



**İNSANSIZ HAVA ARAÇLARI İÇİN MONOKÜLER KAMERA
KULLANILARAK İÇ MEKÂNDAN KONUMLAMA VE HARİTALAMA**

Buğra ŞİMŞEK

**DOKTORA TEZİ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANA BİLİM DALI**

**GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

ŞUBAT 2022

ETİK BEYAN

Gazi Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Buğra ŞİMŞEK
25/02/2022

İNSANSIZ HAVA ARAÇLARI İÇİN MONOKÜLER KAMERA KULLANILARAK İÇ MEKÂNDA KONUMLAMA VE HARİTALAMA

(Doktora Tezi)

Buğra ŞİMŞEK

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Şubat 2022

ÖZET

Günümüzde robotlar ve insansız sistemlere, savunma, sağlık, tarım, madencilik, sürücüsüz araçlar, gezegen keşifleri ve nükleer tesis güvenliği gibi yaygın bir alanda büyük görevler düşmektedir. Gerçekleştirecekleri görevlerde robotik ve insansız sistemlerin kendi konumunu ve çevresini bilmesi operasyonların başarılı bir şekilde sonuçlandırılması için oldukça önemlidir. Dış mekânlarda operasyon gerçekleştirecek insansız sistemler çoğunlukla konum bilgisini elde etmek için Küresel Konum Belirleme Sistemleri uydularından yayınlanan pozisyon ve zaman sinyallerini kullanmaktadır. Fakat kimi durumlarda (sinyal karıştırılması, uydu görüş açısından çıkma, iç mekânda / farklı gezegende görevler gibi) bu sinyalleri kullanmak mümkün olmamaktadır. Bu durumlarda alternatif konumlama ve haritalama metotları kullanılmaktadır. Bu çalışmada düşük maliyetli monoküler RGB kamera kullanılarak görüntü temelli eş zamanlı konumlandırma ve haritalandırma (vSLAM) için hareket bulanıklığına dayanıklı bir çerçeve önerilmiştir. Önerilen çerçeve agresif manevra kabiliyetine sahip dron platformları tarafından gerçekleştirilen hızlı dönme hareketlerinde bile sistemin hareket bulanıklığı kaynaklı takip kaybını uğramasını önleyebilmektedir. Ortaya konulan yeni çerçevede, daha önce vSLAM çerçevelerinde yer almayan hareket bulanıklığı tespit ve azaltma modülleri yer almaktadır. Hareket bulanıklığına dayanıklı vSLAM çerçevesi hem öznitelik temelli hem de doğrudan vSLAM metotları ile denenmiş ve çerçevenin başarısı deneysel olarak kanıtlanmıştır.

Bilim Kodu : 90521

Anahtar Kelimeler : Görüntü temelli eş zamanlı olarak konumlandırma ve haritalandırma, hareket bulanıklığı, öznitelik temelli yöntemler, doğrudan yöntemler

Sayfa Adedi : 228

Danışman : Prof. Dr. Hasan Şakir BİLGE

INDOOR LOCALIZATION AND MAPPING FOR UNMANNED AERIAL VEHICLES
BY USING MONOCULAR CAMERA

(Ph. D. Thesis)

Buğra ŞİMŞEK

GAZİ UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

February 2022

ABSTRACT

Today, robots and unmanned systems have major tasks in such widespread areas such as defence, healthcare, agriculture, mining, self-driving cars, planetary exploration and nuclear facility security. It is very important for the robotic and unmanned systems to be aware of their own location and environment in the tasks they perform, to fulfil their operation needs successfully. Unmanned systems that operate outdoors mostly use position and time signals broadcasted from Global Navigation Satellite Systems to obtain location information. However, in some cases such as signal interference, exiting the satellite line of sight, indoors / a different planet missions, it is not possible to use these signals. In such cases, alternative localization and mapping methods are used. In this study, a motion blur resistant framework is proposed for visual Simultaneous Localization and Mapping (vSLAM) using a low cost monocular RGB camera. The proposed framework can avoid tracking losses due to motion blur, even during rapid rotations performed by aggressively maneuverable drone platforms. In the proposed framework, there are motion blur detection and reduction modules that were not previously included in the vSLAM frameworks. The motion blur-resistant vSLAM framework has been tested with both feature-based and direct vSLAM methods, and the success of the framework has been experimentally proven.

Science Code : 90521

Key Words : Visual simultaneous localization and mapping, motion blur, feature based methods, direct methods

Page Number : 228

Supervisor : Prof. Dr. Hasan Şakir BİLGE

TEŞEKKÜR

Bu teze konu olan çalışma, sınır ötesi operasyonlar kapsamında özellikle mağara, tünel gibi iç mekânlarda görev yaparken şehit veya gazi olmuş askerlerimizin acısının çok taze olduğu bir dönemde başlatılmıştır. Bu açıdan, özellikle bu vatani vatan yapan tüm kahramanlarımıza ve ülkemizin teknolojik ilerlemesine katkı sunan herkese teşekkürü bir borç bilirim.

Eğitim hayatım boyunca her durumda desteklerini esirgemeyen aileme; hedef, hayal ve hayat ortağım eşim Esra ŞİMŞEK'e; gözlerine baktığımda umut dolduğum zeki kızım Meryem Sare'ye yaşamımı güzelleştirdikleri için teşekkür ederim.

Bu teze konu olan çalışmayı gerçekleştirirken desteklerini ve yönlendirmelerini esirgemeyen saygıdeğer hocam Prof. Dr. Hasan Şakir BİLGE'ye, çalışmalarımızın hem akademi hem de sanayi alanında değer bulmasını sağlayan, parçası olmaktan onur duyduğum ASELSAN A.Ş.'ye, bu çalışmanın gerçekleştirilmesine altyapı sağlayan ASELSAN Akademi programına ve programın mimarlarına sonsuz teşekkürler sunarım.

İÇİNDEKİLER

	Sayfa
ÖZET	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
ÇİZELGELERİN LİSTESİ.....	xi
ŞEKİLLERİN LİSTESİ.....	xii
RESİMLERİN LİSTESİ	xv
SİMGELER VE KISALTMALAR.....	xvi
1. GİRİŞ.....	1
2. İNSANSIZ HAVA ARAÇLARI.....	5
2.1. İnsansız Hava Araçlarının Rollerini	5
2.2. İnsansız Hava Araçlarının Sınıflandırılması	10
2.3. İnsansız Hava Araçlarının Tasarımlarına Göre Sınıflandırılması	11
2.4. İnsansız Hava Araçlarının Temel Bileşenleri	13
2.5. Mikro / Nano İnsansız Hava Araçları	14
3. İNSANSIZ SİSTEMLERDE GÖRÜNTÜ TEMELLİ NAVİGASYON .	21
3.1. Eş Zamanlı Konumlama ve Haritalama (SLAM)	21
3.2. Görsel Odometre	24
3.2.1. Öznitelik temelli yaklaşımlar	28
3.2.2. Doğrudan yaklaşımlar	29
3.2.3. Hibrit yöntemler	30
3.3. Görüntü Temelli SLAM (vSLAM)	34
3.3.1. Öznitelik temelli yaklaşımlar	39

	Sayfa
3.3.2. Doğrudan yaklaşımlar	58
3.4. Görsel – Ataletsel Navigasyon (VINS-Visual-Inertial Navigation Systems)	64
3.4.1. Gevşek bağlı – sıkı bağlı sensör füzyonu.....	64
3.5. Performans Karşılaştırmaları	70
3.6. Sonuç.....	74
4. HAREKET BULANIKLIĞI.....	77
4.1. Hareket Bulanıklığı Nedenleri	77
4.2. Homojen Koordinatları Kullanan Merkezi İzdüşüm.....	80
4.3. Kamera Kalibrasyonu.....	84
4.4. Bulanıklığın Tespiti.....	86
4.4.1. Kenar algılamaya dayalı bulanıklık algılama.....	87
4.4.2. Frekans analizine dayalı bulanıklık algılama.....	88
4.5. Hareket Bulanıklığı Modelleri	91
4.5.1. Bir boyutlu doğrusal hareket bulanıklığı modeli	92
4.5.2. Dönel hareket bulanıklık modeli.....	93
4.5.3. Işınsal hareket bulanıklığı modeli	95
4.6. Hareket bulanıklığının düzeltilmesi / azaltılması.....	96
4.6.1. Kör görüntü dekonvolüsyonu.....	97
4.6.2. Kör olmayan görüntü dekonvolüsyonu.....	98
4.6.3. Tek görüntü özelliklerine dayalı bulanıklık düzeltme	100
4.6.4. Çoklu görüntüye dayalı bulanıklık düzeltme	101
5. HAREKET BULANIKLIĞI AZALTMA YÖNTEMLERİ	103
5.1. Lucy-Richardson Algoritması.....	105
5.2. Wiener Filtresi Algoritması.....	106

Sayfa

5.3. Kör Dekonvolusyon Yaklaşımı.....	107
5.4. Görüntü Bulanıklığı Azaltma için Derin Ölçek Yinelemeli Ağ.....	108
5.5. Görüntü Bulanıklığı Azaltma için Yinelemeli Üst Ölçekleme Ağ 114	114
5.6. Tek Fotoğraftan Grafik Temelli Kör Görüntü Bulanıklığı Azaltma	118
5.7. Kör Görüntü Bulanıklığı Azaltma için Basit Bir Minimum Yoğunluk Önceli.....	123
5.8. Görüntü Bulanıklığı Azaltma için Uç Kanallar Yöntemi.....	131
6. ÖNERİLEN METOT VE DENEYSEL ÇALIŞMALAR	139
6.1. Sentetik Hareket Bulanıklığının Öznitelik Tespitine Etkisi.....	143
6.2. Sentetik Hareket Bulanıklığının Öznitelik Eşleştirmesine Etkisi	144
6.3. Mikro / Nano Ölçekteki İnsansız Araçlarda Konumlama.....	147
6.4. vSLAM Tekniklerinde Takip Kaybına Yönelik Çözümler.....	159
6.5. Hızlı Dönüş Hareketi İçeren Veri Kümesinin Oluşturulması	161
6.5.1. Veri kümesinin detaylı analizi	164
6.5.2. Mevcut veri kümelerinin analizi	166
6.6. Önerilen Yeni vSLAM Çerçevesi	170
6.6.1. Hızlı dönüş hareketi içeren veri kümesinde önerilen çerçevenin testleri...	173
6.6.2. Seçilen algoritmaların ani dönüş hareketindeki performansları.....	179
6.7. Hareket Bulanıklığına Dayanıklı vSLAM Çerçevesindeki Performanslar	180
6.7.1. Restore edilen resimlerdeki öznitelik sayılarının karşılaştırılması	184
6.7.2. Eşik değeri seçiminin vSLAM çerçevesinin performansına etkileri.....	186
6.7.3. Hareket bulanıklığına dayanıklı vSLAM çerçevesinin özellikleri.....	187
6.8. Veri Kümesindeki Doğal Hareket Bulanıklığının Öznitelik Tespitine Etkisi.....	189
6.9. AG-Mono: İHA'lar için Monoküler Görüntü Veri Kümesi.....	193
6.9.1. AG-Mono veri kümesinin detaylı analizi.....	196

Sayfa

6.9.2. Önerilen çerçevenin AG-Mono veri kümesindeki performansı.....	198
7. SONUÇ VE ÖNERİLER	203
KAYNAKLAR	207
EKLER.....	225
EK-1.Video görüntüleri	226
ÖZGEÇMİŞ	227

ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 2.1. İHA ve insanlı platformun ilk maliyet karşılaştırması	9
Çizelge 2.2. İHA ve insanlı platformun işletme maliyeti karşılaştırması	9
Çizelge 3.1. Odometrede kullanılan sensörlerin avantaj ve dezavantajları	25
Çizelge 3.2. Görsel odometrede kullanılan kamera türlerinin karşılaştırması.....	27
Çizelge 3.3. Görsel odometre yaklaşımlarının karşılaştırılması	31
Çizelge 3.4. Mevcut veri kümelerine ait özellikler.....	71
Çizelge 6.1. Sensör çeşitlerinin karşılaştırılması	141
Çizelge 6.2. Veri kümelerinin sensör özelliklerinin karşılaştırılması.....	151
Çizelge 6.3. Yavaş (0,8 m/s) hızda vSLAM algoritmalarının performansları.....	157
Çizelge 6.4. Hızlı (1,5 m/s) hızda vSLAM algoritmalarının performansları.....	158
Çizelge 6.5. Bulanıklık azaltma metotlarına ait ortalama odak ölçüm puanları	176
Çizelge 6.6. Önerilen vSLAM çerçevesinin ORB-SLAM2 algoritmasındaki başarısı ..	183
Çizelge 6.7. Önerilen vSLAM çerçevesinin DSO algoritmasındaki başarısı	183
Çizelge 6.8. ORB-SLAM2 algoritması ile tespit edilen öznelik sayıları	186
Çizelge 6.9. Veri kümesinde OÖP lokal minimumları, ağırlıklar ve tercih değerleri	187
Çizelge 6.10. Önerilen vSLAM çerçevesi ile mevcut çerçevelerin karşılaştırılması	188
Çizelge 6.11. AG-Mono veri kümesi pozlandırma süreleri	193
Çizelge 6.12. AG-Mono veri kümesi ortalama OÖP sonuçları	197
Çizelge 6.13. AG-Mono veri kümesi eşik değeri sonuçları.....	198

ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 2.1. İHA ve yer kontrol istasyonuna ait temel bileşenler	13
Şekil 2.2. Mikro/nano İHA’larda monoküler görüntü sistemi	16
Şekil 3.1. Görsel odometre adımları	24
Şekil 3.2. vSLAM akış şeması	35
Şekil 3.3. ORB-SLAM algoritması akış şeması	42
Şekil 3.4. FAST özniteliklerinin tespiti	43
Şekil 3.5. ORB ölçeklendirme piramidi.....	44
Şekil 3.6. İkili öznitelik vektörleri	45
Şekil 3.7 Anahtar nokta, belirlenen ilk ve ikinci pikseller.....	46
Şekil 3.8. Örnek ORB-SLAM takip ve haritalama çıktıları.....	48
Şekil 3.9. Öznitelik temelli vSLAM algoritmalarının kronolojik özeti	48
Şekil 3.10. ORB-SLAM2 akış şeması	49
Şekil 3.11. ORB-SLAM2 algoritmasında yakın (yeşil) ve uzak (mavi) noktalar.....	50
Şekil 3.12. PL-SLAM algoritması modülleri.....	52
Şekil 3.13. CO-SLAM sistem mimarisi	55
Şekil 3.14. SOFT-SLAM öznitelik eşleme ve takip	57
Şekil 3.15. DSO metodu haritası (a) 500 Nokta (b) 2000 Nokta (c) 10000 Nokta	63
Şekil 3.16. Doğrudan metotlara ait harita yoğunluğu karşılaştırması.....	64
Şekil 3.17. RTAB-Map ROS modülünün yapısı	68
Şekil 3.18. RTAB-Map görsel odometre ROS modülü	69
Şekil 3.19. RGB-D kamera kullanımında harita görüntüsü	70
Şekil 3.20. PL-SLAM / ORB-SLAM yörüngelerinin karşılaştırması.....	72
Şekil 4.1. İğne deliği kamera geometrisi	79

Şekil	Sayfa
Şekil 4.2. Bulanıklık seviyesi ile gradyan sonuçlarının karşılaştırılması	87
Şekil 4.3. Bulanıklık seviyesinin frekansa etkisi	89
Şekil 5.1. Görüntü İşleme uygulamalarından farklı evrimsel sinir ağı mimarileri	109
Şekil 5.2. SRN: Ölçek Yinelemeli Sinir Ağı Mimarisi.....	111
Şekil 5.3. Yapay zekâ temelli algoritmalarının ağı mimarileri	116
Şekil 5.4. SIUN metodu ağı mimarisi	118
Şekil 5.5. Net ve bulanık görüntülerin yama bazında minimum pikselleri	127
Şekil 5.6. Bulanıklaşma ile parlak kanal piksel yoğunluğu arasındaki ilişki.....	132
Şekil 6.1. Algılama-planlama-davranış döngüsü	140
Şekil 6.2. Orijinal görüntüde öznitelik çıkarımı	143
Şekil 6.3. Hareket bulanıklığının öznitelik çıkarımına etkisi	144
Şekil 6.4. Hareket bulanıklığının öznitelik eşleştirmeye etkisi.....	146
Şekil 6.5. Veri kümesi oluşturulurken gerçekleştirilen zorlu hareketler	153
Şekil 6.6. vSLAM metotlarının kronolojik sıralaması.....	154
Şekil 6.7. DSO-LDSO ve ORB-SLAM2 nokta bulutu haritaları	155
Şekil 6.8. DSO, LDSO ve ORB-SLAM2 pozisyon ve lokalizasyon sonuçları	156
Şekil 6.9. Deney ortamına ait çizim.....	162
Şekil 6.10. Veri kümesinde odak ölçüm analizi	164
Şekil 6.11. EUROCC-MH05_difficult sekansı odak ölçüm analizi	167
Şekil 6.12. TUM Mono veri kümesi 35.sekansa ait odak ölçüm analizi	169
Şekil 6.13. TUM Mono veri kümesi 14.sekansa ait odak ölçüm analizi	169
Şekil 6.14. Önerilen vSLAM çerçevesi	170
Şekil 6.15. Odak ölçüm puanı eşik değerinin altında kalan görüntüler	173
Şekil 6.16. Seçilen algoritmaların PSNR sonuçları	174

Şekil	Sayfa
Şekil 6.17. Seçilen algoritmaların SSIM sonuçları	175
Şekil 6.18. Orijinal ve restore edilmiş görüntülerin Odak Ölçüm Puanları	177
Şekil 6.19. Restore edilmiş görüntülerde OÖP değişimi	178
Şekil 6.20. 00340 numaralı görüntüye ait odak ölçüm puanları	179
Şekil 6.21. 00807 numaralı görüntüye ait odak ölçüm puanları	180
Şekil 6.22. Seçilen yöntemlerle oluşturulan alt kümeler	181
Şekil 6.23. Seçilen yöntemlerle oluşturulan alt kümeler-2	182
Şekil 6.24. Takip kaybının yaşandığı orijinal görüntü.....	184
Şekil 6.25. Önerilen yöntem ile ORB-SLAM2 algoritmasının tespit ettiği öznitelikler .	185
Şekil 6.26. 340 numaralı görüntüden ORB özniteliklerinin çıkarım sonuçları	190
Şekil 6.27. 825 numaralı görüntüden ORB özniteliklerinin çıkarım sonuçları	190
Şekil 6.28. 340-341 numaralı görüntülerde ORB özniteliklerinin eşleştirmeleri	191
Şekil 6.29. 807-808 numaralı görüntülerde ORB özniteliklerinin eşleştirmeleri	191
Şekil 6.30. 807-808 numaralı görüntülerde Harris özniteliklerinin eşleştirmeleri	192
Şekil 6.31. 807-808 numaralı görüntülerde SURF özniteliklerinin eşleştirmeleri	192
Şekil 6.32. Farklı pozlandırma sürelerinin 20fps görüntü akışındaki gösterimi.....	194
Şekil 6.33. AG-Mono veri kümesine ait deney ortamının şematik gösterimi	195
Şekil 6.34. AG-Mono-V1 görüntü sekansına ait OÖP sonuçları.....	196
Şekil 6.35. AG-Mono-V2 görüntü sekansına ait OÖP sonuçları.....	196
Şekil 6.36. AG-Mono-V3 görüntü sekansına ait OÖP sonuçları.....	197
Şekil 6.37. AG-Mono-V1 görüntü sekansı 87. (a) ve 110. görüntü (b) kareleri.....	199
Şekil 6.38. AG-Mono-V2 görüntü sekansı 83. (a) ve 108. görüntü (b) kareleri.....	200
Şekil 6.39. AG-Mono-V3 görüntü sekansı 83. (a) ve 108. görüntü (b) kareleri.....	201

RESİMLERİN LİSTESİ

Resim	Sayfa
Resim 3.1. ORB-SLAM2 örnek harita çıktısı.....	51
Resim 3.2. PL-SLAM algoritması takip ve haritalama çıktıları	52
Resim 3.3. CO-SLAM yöntemi kullanılan stereo rig yapısı.....	54
Resim 3.4. CO-SLAM 12 kameralı çalışmaya ait görüntüler	56
Resim 3.5. DTAM Algoritması (a) uzaklık (b) regüle edilmiş çıktı (c) takip nokta modeli.....	59
Resim 3.6. LSD SLAM (a) kamera görüntüsü (b) derinlik tahmin haritası (c) çevresel harita.....	60
Resim 3.7. Bazı viSLAM algoritmalarına ait performansların karşılaştırması	67
Resim 4.1. Satranç tahtası kullanılan kalibrasyon örneği	86
Resim 4.2. Bulanıklaştırma çekirdeği (kernel) örneği	90
Resim 6.1. Raspberry Pi kartı ve üzerinde Raspberry Pi v2.1 model kamera	149
Resim 6.2. Veri kümeleri oluşturulurken kullanılan kameralar.....	151
Resim 6.3. Döngü 1 veri kümesine ait döngü kapama işlemi.....	157
Resim 6.4. Paralel iş parçacığı ile örnek bir yörünge'nin çıkarımı	160
Resim 6.5. Veri kümesinde yer alan görüntülerin bir kısmını gösteren resim	163
Resim 6.6. LAP4 yöntemine ait Python programlama dilindeki kodları.....	166
Resim 6.7. EUROCC veri kümesi MH05_difficult isimli veri sekansından görüntü	167
Resim 6.8. TUM Mono veri kümesinde yer alan 35.veri sekansından görüntü	168
Resim 6.9. AG-Mono-V3 görüntü sekansına ait görüntü kümesi	195
Resim 6.10. AG-Mono-V1 görüntü sekansına başarılı dönüş haritası	199

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler

Açıklamalar

m

Metre

db

Desibel

rad

Radyan

sn

Saniye

Kısaltmalar

Açıklamalar

BCP

Parlak Kanal Önceli
(Bright Channel Prior)

CCD

Şarj Bağlantılı Cihaz
(Charge-Coupled Device)

CNN

Evrişimsel Sinir Ağları
(Convolutional Neural Network)

ECP

Uç kanallar Önceli
(Extreme Channels Prior)

FPS

Saniyedeki Görüntü Karesi Sayısı
(Frame Per Second)

GAN

Çekişmeli Üretici Ağlar
(Generative Adversarial Network)

GB

Grafik Temelli Kör Görüntü Bulanıklığı Azaltma
(Graph-Based Blind Image Deblurring)

GPU

Grafik İşlemci Birimi
(Graphics Processing Unit)

GTV

Grafik Toplam Varyasyonu
(Graph Total Variation)

IMU

Ataletsel Ölçüm Birimi
(Inertial Measurement Unit)

Kısaltmalar	Açıklamalar
İHA	İnsansız Hava Aracı
KRBN	Kimyasal, Radyoaktif, Biyolojik ve Nükleer
LSTM	Uzun Kısa Süreli Bellek (Long Short Term Memory)
OÖP	Odak Ölçüm Puanı
PSF	Nokta Yayılım Fonksiyonu (Point Spread Function)
PSNR	Tepe Sinyal Gürültü Oranı (Peak Signal to Noise Ratio)
RDN	Artık Yoğun Ağ (Residual Dense Network)
RGTV	Ağırlıklandırılmış Grafik Toplam Varyasyonu (Reweighted Graph Total Variation)
SIFT	Ölçek Bağımsız Öznitelik Transformu (Scale Invariant Feature Transform)
SIUN	Ölçek Yinelemeli Üst Ölçekleme Ağı (Scale Iterative Upscaling Network)
SFF	Odaktan Şekil Elde Etmek (Shape From Focus)
SRN	Ölçek Yinelemeli Ağ (Scale Recurrent Network)
SSIM	Yapısal Benzerlik Endeksi Metriği (Structural Similarity Index Metric)
viSLAM	Görsel Ataletsel Eş Zamanlı Konumlama ve Haritalama (Visual Inertial Simultaneous Localization and Mapping)
VIO	Görsel Ataletsel Odometre (Visual Inertial Odometry)
vSLAM	Görüntü Temelli Konumlama ve Haritalama (Visual Simultaneous Localization and Mapping)

1. GİRİŞ

İnsansız ve robotik sistemlere son dönemde insan unsuruna yardımcı olarak çeşitli görevler verilmektedir. Günümüzde bu sistemlerin bazı tehlikeli, riskli, insan gücünü aşan görevlerde doğrudan insan yerine kullanılması da mümkün hale gelmiştir. Gelecek yıllarda robotlar ve insansız sistemlerin hemen hemen her alanda yer alacağı görülmektedir. Artan algılama, planlama ve davranış teknolojileri sayesinde artık pek çok görev insansız sistemler tarafından kolayca yerine getirilebilmektedir. Algılama teknolojileri, insansız platformların çevresi ve kendisi hakkında anlamlandırma yapabilmesini ve dış dünyadan alınan verileri planlama amaçlı olarak kullanabilmesini sağlamaktadır. Algılama teknolojileri ile üretilen veriler ne kadar doğru ise o kadar etkili görev ve davranış planlamaları gerçekleştirilebilmektedir. Daha sonra bu davranış hayata geçirilerek insansız sistem tarafından görev icra edilmektedir.

Algılama teknolojileri, insansız platformların hızının, durumunun, pozisyonunun ve görev yaptığı çevrenin özelliklerinin çıkarılmasını sağlamaktadır. Bu teknolojilerin en önemlilerinden bir tanesi de konumlama ve haritalama teknolojileridir. Doğru bir konumlama yapılmadığında, insansız platformun planladığı görevdeki konumlar da yanlış olmaktadır. Benzer şekilde doğru haritalamayı oluşturmak, çevrenin doğru olarak ifade edilmesini ve değerlendirilmesini sağlamaktadır.

İnsansız ve robotik sistemler genellikle dış mekânda konumlama teknolojisi olarak Küresel Konum Belirleme Sistemi (Global Navigation Satellite Systems - GNSS) kullanmaktadır. GNSS uydularından yayınlanan zaman ve konum sinyalleri kullanarak platform kendi konumlamasını gerçekleştirmektedir. Fakat çeşitli yapıların yer aldığı meskûn mahalde ortamlarında GNSS sinyalleri etraftaki bu yapılara çarparak ve farklı yansımalar (çok yolluluk) yaparak insansız sistemlere ulaşmaktadır. Bu durumda insansız sistemler uydulara olan mesafesini tam olarak belirleyememekte ve yanlış konum tahmininde bulunabilmektedirler. İç mekânda görev yapması istenen daha küçük boyutlardaki sistemler ise bu sinyallere erişim sağlayamamaktadır. Benzer şekilde tünel, yer altı ya da farklı gezegenlerde gerçekleştirilecek görevlerde de dünya yörüngesinde yer alan uydulardan yapılan yayınlar kullanılamaz hale gelmektedir. Ayrıca özellikle savunma uygulamalarında uydulardan yayınlanan GNSS sinyalleri düşük güçlü vericiler ile taklit

edilebilmekte ve karıştırma uygulanabilmektedir. Düşman unsurlar tarafından karıştırılan/taklit edilen uydu sinyalleri insansız platformlara farkı bir yerde olduğu bilgisini vererek görevin başarısız olmasına neden olabilmektedir. Yukarıda bahsi geçen nedenler ile GNSS sinyali kullanımına alternatif eş zamanlı konumlama ve haritalama (simultaneous localization and mapping -SLAM) teknolojileri üzerine uzun yıllardır akademik dünyada ve endüstride çalışmalar yapılmaktadır. İnsansız sistemlerin ve özellikle insansız hava araçlarının batarya, ağırlık ve maliyet kısıtları nedeniyle SLAM teknolojilerinde en fazla tercih edilen sensörlerden bir tanesi de görsel sensör olan kameralardır. Yalnızca kameradan elde edilen veriler kullanılarak eş zamanlı konumlama ve haritalama sağlayan teknolojiler görsel eş zamanlı konumlama ve haritalama (visual Simultaneous Localization and Mapping - vSLAM) teknolojileri olarak bilinmektedir.

İnsansız sistemlerdeki teknolojik ilerlemeler sayesinde platform boyutları günden güne küçültülmekte ve pek çok görevde artık daha küçük ve daha az güç tüketimine sahip sistemler kullanılmaktadır. Bu nedenlerle daha düşük hacimli ve hafif kameralara olan ilgi artmaktadır.

Bu teze konu olan çalışmada düşük maliyetli, düşük güç tüketimli, hafif, platform üzerinde kolaylıkla konumlandırılabilir ve her kullanımda kalibrasyon ihtiyacı duyulmayan monoküler RGB kameralar ile gerçekleştirilen konumlama ve haritalama teknolojileri incelenmiştir. Bu tipte kameraların özellikle ilerleyen dönemde pek çok sahada kullanımının artması planlanan mikro / nano ölçekteki insansız sistemlerde de kullanılabilir olması önem taşımaktadır. Çünkü yakın gelecekte bu sistemlerden belirli alanlara daha büyük araçlardan (büyük insansız araçlar, uçaklar ya da helikopterler gibi) yüzlerce, binlerce adet serpilerek, bu sistemlerin alana yayılması ve tekil ya da sürü halinde görev icra etmesi planlanmaktadır. Bu durumda pek çok platformun tekrar kullanımı da mümkün olmayabilir. Bu yüzden en optimum maliyetin ve görev süresinin sağlanması amacıyla bu monoküler kamera tipi seçilmiştir.

Literatürde yer alan vSLAM algoritmalarının performansları genellikle yüksek çözünürlüklü kaliteli kameralar ile oluşturulan veri kümeleri altında değerlendirilmektedir. Fakat yaptığımız çalışmalarda bu algoritmaların düşük kaliteli kameralar ile alınan düşük çözünürlüklü görüntülerde belirlenen yörüngeleri tamamlayamadıkları gözlemlenmiştir. Bu problemin tespiti için uygun fiyatlı, daha düşük kalitede kamera kullanılarak veri

kümeleri oluşturulmuş ve vSLAM algoritmalarının takip kaybına uğramasının nedeninin çoğunlukla hareket bulanıklığı olduğu tespit edilmiştir. Bu nedenle GNSS sinyali alınamayan ortamlarda görüntü temelli konumlandırma ve haritalandırmada en uygun çözümü sağlayabilecek monoküler RGB kameraların, takip kaybına uğramadan çalışabildiği yeni bir vSLAM çerçevesi önerilmiştir.

Bu vSLAM çerçevesi daha önceki çerçevelerde yer almayan ve basit bir konvolüsyon işlemi ile görüntüdeki odak ölçüm skorunun elde edilmesini ve eğer belirlenen eşik değerinin altında kalırsa bu görüntünün bulanıklık azaltma yöntemi ile düzeltilerek takip işlemine devam edilmesini sağlamaktadır. Önerilen bu yeni vSLAM çerçevesi ile:

1. İlk defa odak ölçüm operatörü vSLAM algoritmalarına dâhil edilerek bulanıklık seviyesi tespiti gerçekleştirilmiştir. Bu sayede yüksek hareket bulanıklığı seviyelerine sahip kareler, takip kaybına neden olmadan algılanabilir.
2. Yalnızca bulanıklık düzeyi yüksek kareler düzeltilerek algoritmanın hem haritalama hem de izleme aşamalarında çalışmaya devam etmesi sağlanmıştır.
3. Ek bir donanım ihtiyacı olmadan, mikro / nano ölçekteki insansız istemlerde bile takip kaybına uğramadan görev yönetilmesi amaçlanmıştır.

Önerilen çerçevenin başarımı hem öznitelik temelli hem de doğrudan vSLAM metotlarında 8 farklı bulanıklık azaltma yöntemi ile test edilmiş ve çerçevenin çeşitli işlem hızlarında farklı metotlardaki başarısı sunulmuştur.

Önerilen çerçeve ile GNSS sinyali alınamayan ortamlarda monoküler RGB kamera kullanımında özellikle hızlı öteleme ve dönme hareketlerinde meydana gelen hareket bulanıklığı kaynaklı takip kayıpları önlenmekte ve vSLAM algoritmalarının daha düşük kaliteli kameralarda da konumlama ve haritalama görevlerine devam etmesi sağlanabilmektedir.

Bu tezin ikinci bölümünde genel olarak insansız hava araçlarından, üçüncü bölümünde insansız sistemlerde görüntü temelli navigasyon uygulamalarından bahsedilmiştir. Dördüncü bölümünde hareket bulanıklığı detaylı olarak incelenmiş, beşinci bölümde ise hareket bulanıklığının azaltılmasında kullanılan yöntemler detaylandırılmıştır. Altıncı bölümde gerçekleştirilen deneysel çalışmalar anlatılarak sonuçlar sunulmuştur. Tezin

yedinci ve son bölümünde sonuçlar değerlendirilerek, gelecekte gerçekleştirilebilecek çalışmalar hakkında bilgi verilmiştir.

2. İNSANSIZ HAVA ARAÇLARI

İnsansız hava araçları (İHA) son dönemde savunma, sağlık, tarım, madencilik ve gezegen keşifleri de dahil olmak üzere pek çok alanda kullanılmaktadır. Gün geçtikçe İHA'ların kabiliyetleri ve gerçekleştirebildikleri görev çeşitliliği artmaktadır. Bu durumun çevresel, ekonomik ve diplomatik pek çok nedeni mevcuttur. Bu nedenleri daha iyi analiz edebilmek için, insansız sistemleri insanlı sistemlerle kıyaslamak faydalı olacaktır. Tabi ki farklı operasyonlarda insanlı ve insansız sistemlerin farklı avantajları ve birbirine üstün oldukları durumlar olabilir fakat İHA'ların kullanımının yaygınlaşmasındaki gerekçeleri görmek için bu tip analizlerden önemli bakış açıları kazandırabilirler. İHA'ların savunma amacıyla kullanımı çoğunlukla İngilizce olarak "DDD" (Dull-Dirty-Dangerous / sıkıcı-kirli-tehlikeli) şeklinde ifade edilen askeri bir deyim ile özetlense de bu rollere ek olarak gizli, çevresel, ekonomik ve diplomatik bazı görevlerde de İHA'lar rahatlıkla kullanılabilmektedir (Austin, 2011).

2.1. İnsansız Hava Araçlarının Roller

Sıkıcı Roller

Belirli bir alandaki sürekli gözetim, devriye ve tarama görevleri insan tarafından platformda yer alarak kullanılan hava araçları için oldukça sıkıcı deneyimler olabilmektedir. Pilotların dinlenmeden, sürekli aynı alana bakması ya da kullandıkları platformlara aynı hareketleri yaptırması personelin göreve yoğunlaşmasının ve görev etkinliğinin azalmasına neden olacaktır. İnsansız araçlar belirli bölgenin sürekli olarak gözetlenmesi, aynı mevkiinin taranması gibi görevlerde etkin şekilde görev yürütebilmektedir. Önceden tanımlanmış rotalarda (way point) düzenli hareket gerçekleştirebilmekte ve platform üzerindeki kameralardan, radar sistemlerinden ve diğer sensör sistemlerinden alınan anlık verilerde değişme meydana geldiğinde operatöre bilgi iletebilmektedirler.

Kirli Roller

Kimyasal, radyoaktif, biyolojik ve nükleer (KRBN) maddelere maruz kalınabilecek ortamlarda İHA'ların ve robotların kullanımı, personelin ilgili alanda karşılaşabileceği riskleri önlemektedir. İlgili bölgede kimyasal, radyoaktif, biyolojik ve nükleer (KRBN) ajanların tespiti ve ortamın bu ajanlardan arındırılması açısından İHA'lara ve robotlara görev verilmesi tercih edilebilmektedir.

Tehlikeli Roller

Uçan cisimlere karşı yüksek seviyede korunan alanlarda (radar ve termal algılayıcı sistemlerin ve kameraların kullanıldığı) İHA boyutları açısından insanlı sistemlere (uçak ve helikopter) göre avantajlı olmaktadır. Bu tip görevlerde çoğunlukla uçakları tespit eden ve onlara kilitlenerek imha etmeye çalışan uçaksavar sistemleri alanda savunma amaçlı olarak yer almaktadır. Bu nedenle uçak-helikopter gibi tespiti daha kolay olan insanlı platformların, maruz kalacakları tehlikenin İHA platformlarından daha büyük olması beklenmektedir.

Özellikle askeri personelin zarar görebileceği tehlikeli operasyonlarda İHA'ları ve robotları kullanmak avantaj sağlamaktadır. Askeri personelin arar göre ihtimalinin olduğu durumlarda, insansız platforma zarar gelmesi bu tip tehlikeli operasyonlarda kesinlikle tercih edilmektedir. İHA'ları ve robotları uzaktan kullanan operatörlerin tehlikeli operasyonlarda bile can güvenliği çoğunlukla sağlanabilmektedir. Ormanlık alanlarda ya da doğal afet senaryolarında uçaklar ile alçak uçuş yaparak olaya müdahale etmek bu tip tehlikeli görevlere örnek olarak verilebilir. Bazen bu rollerde uçağın beklenmeyen şekilde hasar alması mürettebatın zarar görmesine neden olabilmektedir. Görevin tehlikeli olduğunu bilmek mürettebatın göreve olan ilgisini de düşürebilmekte ve göreve yönelik duyulan kaygı nedeniyle yanlış uygulamalar yapmasına neden olabilmektedir. Bu tip durumlarda İHA'ların kullanımı tercih edilmektedir.

Gizli Roller

Askeri ve sivil operasyonlarda düşman ya da suçlu tespiti ile personelin önceden uyarılması gereken koşullar olabilmektedir. İHA'ların düşük ses ve termal izleri nedeniyle

dikkat çekmemesi bu tip operasyonlarda fark edilmeden görevin başarıya ulaştırılmasını sağlamaktadır. Benzer şekilde ülkelerin hava sahasında meydana gelen ihlaller kimi zaman diplomatik krizlere hatta çatışmalara neden olmaktadır. Bu tip durumlarda İHA'ların kullanımı fark edilme ihtimalini düşürmektedir.

Araştırma Roller

İnsanlı uçakların insansız olarak küçük ölçekli kopyalarının yapılması ve test amaçlı olarak kullanılması, gerçek çevre koşulları altında daha az maliyetli ve daha az tehlikeli testler yapılmasını mümkün kılmaktadır. Ayrıca gözlemlenen test sonuçlarına göre daha hızlı modifikasyonlar yaparak, sistemin tekrar zaman kaybetmeden testlerini gerçekleştirmek mümkün olabilmektedir. Beklenenden çok daha zorlayıcı hava koşullarında testlerin yapılmasında da bu küçük ölçekli kopyaların kullanımı ile asıl platform üzerinde gerçekleştirilemeyecek testler gerçekleştirilebilmektedir. Gerçekleştirilen bu testlere mürettebatın katılmaması da ayrıca olası kaza durumunda herhangi bir can kaybı ya da yaralanma ihtimalini de ortadan kaldırmaktadır.

Çevresel Açidan Kritik Roller

İHA'ların çevreye vermiş olduğu rahatsızlık seviyesi özellikle sivil görevler düşünüldüğünde uçak ve helikopter gibi platformlara göre oldukça az olmaktadır. Bu durum insansız sistemlere hem platformun kullandığı yakıt miktarı ve emisyon seviyesi açısından çevreyi daha az kirletmesi hem de meydana getirdiği gürültünün daha az olması anlamında üstünlük sağlamaktadır.

Ekonomik Nedenler

İnsansız hava araçlarının, aynı görevlerde kullanılan platform üzerinde insan tarafından kullanılan hava araçları ile karşılaştırıldıklarında kullanılan malzemeler ve platform boyutları değerlendirildiğinde üretim maliyetleri açısından İHA'ların avantajlı oldukları bilinmektedir. Benzer şekilde bakım maliyetleri, hangar maliyetleri ve işletme maliyetleri açısından da İHA'lar daha az maliyetli olmaktadır. Bir İHA ile gerçekleştirilen basit bir görev genellikle bir pilot ve bir gözlemci mürettebatın yer aldığı iki kişilik küçük bir uçak ile gerçekleştirilebilmektedir. Mürettebat ve pilotun platform içerisinde yer alacakları

düşünüldüğünde ayrıca bir hacim gerektiği ortadadır. İHA'larda bu ek hacimlere ihtiyaç duyulmamaktadır. Yapılan çalışmada (Austin, 2011) insanlı hava aracının taşınması gereken kütleler değerlendirildiğinde (kapılar, gövdenin kendisi, camlar vb.) platform ağırlığı 230 kg olurken, aynı görev yaklaşık 10kg ağırlığındaki bir İHA ile gerçekleştirilebilmektedir. Bu ağırlıklara yakıt miktarları eklendiğinde ağırlıkların sırasıyla 750 kg ile 35 kg seviyelerine çıkacağı değerlendirilmektedir.

İlk Maliyetler

İHA'lar açısından değerlendirildiğinde gözetleme sensörleri yaklaşık ağırlığın %3-4'ü, ilgili sensörleri taşımak için gereken motor gücü yaklaşık toplam gücün %2,5'u ve sensörün kapladığı alan tüm platformun %25'i olarak kabaca değerlendirilmektedir (Austin, 2011).

İnsanlı uçakların motor ve platform maliyetleri ise, uçulacak menzil ve ağırlık ile oldukça değişiklik göstermektedir. Sadece keşif amacına yönelik olarak kullanılacak bir uçak düşünülüğünde, İHA'nın maliyeti yaklaşık olarak insanlı uçağın %3 ü olacaktır diye değerlendirilebilir fakat bu durum tam olarak doğru değildir. Çünkü İHA'ların üretiminde hafif olması istenen malzemelerin maliyeti, beklendiği ölçüde ucuza mal edilememektedir. Benzer şekilde İHA'larda yer alan ve kontrolü sağlayan radyo haberleşme sistemleri gibi sistemlere de uçaklarda kontrol amacıyla ihtiyaç duyulmadığı için daha basit haberleşme altyapıları kullanılabilmektedir.

İnsanlı uçakların elektronik parçalarının daha karmaşık ve özellikle uçuş kontrol sisteminin acil durumlar için yedekli olması gibi durumlardan dolayı daha maliyetli olduğu bilinmektedir. Benzer şekilde İHA'larda operatör ile platform arasında arayüz sağlayan kontrol istasyonunun gerekliliği de mevcuttur. Kontrol istasyonu da ayrı bir maliyet olarak değerlendirilmektedir.

Yukarıda yer alan tüm maliyetler değerlendirildiğin ana kalemlerde aşağıdaki gibi bir ortalama yaklaşım yapılabilir:

- İHA'ların platform gövde maliyeti, insanlı uçak maliyetinin %20-40'ı
- İHA kontrol istasyonu, insanlı uçak maliyetinin %20-40'ı

- İHA + İHA kontrol istasyonu İnsanlı uçak maliyetinin %40-80'i

İHA'lar ile insanlı platformların ilk maliyetlerinin karşılaştırılması aşağıda Çizelge 2.1'de verilmiştir.

Çizelge 2.1. İHA ve insanlı platformun ilk maliyet karşılaştırması

İHA Platformu Alt Bileşeni	İnsanlı Uçağa Göre İHA'nın Göreli Maliyeti
İHA'ların platform gövde maliyeti	İnsanlı uçak maliyetinin %20-40'ı
İHA kontrol istasyonu	İnsanlı uçak maliyetinin %20-40'ı
İHA + İHA kontrol istasyonu (Toplam)	İnsanlı uçak maliyetinin %40-80'i

İşletme Maliyetleri

İşletme maliyetleri değerlendirildiğinde, Çizelge 2.2'deki gibi bir genel değerlendirme yapılabilir (Austin, 2011):

Çizelge 2.2. İHA ve insanlı platformun işletme maliyeti karşılaştırması

İHA İşletme Maliyet Kalemi	İnsanlı Uçağa Göre İHA'nın Göreli Maliyeti
İşletme sermayesi faizi	Uçak işletme sermayesinin %40-80 i
Amortisman giderleri	İnsanlı uçağın yaklaşık %30-%60'ı
Hangar maliyetleri ve kontrol istasyonu idamesi	İnsanlı uçağın ihtiyacının %20'si
Personel(operatör) giderleri	Pilot ve mürettebat maliyetlerinin %40-80 i
Yakıt giderleri	İnsanlı uçak yakıt maliyetinin %5'i
Bakım maliyetleri*	İnsanlı uçak bakım maliyetlerinin %20'si
Sigorta maliyetleri**	İnsanlı uçağın sigorta maliyetlerinin %30'u
İHA işletme maliyetleri (Toplam)	İnsanlı uçağın işletme maliyetlerinin %40'ı

* Hava mürettebatına yerdeki personelden daha fazla ücret ödenir, daha erken emekli olurlar ve düzenli, daha titiz bir tıbbi muayene sürecinden geçmektedir.

**** Hava mürettebatının can güvenliğini sağlamak ve platformda meydana gelen hasarları temin etmek açısından insanlı uçaklarda riskin fazla olması nedeniyle sigorta maliyetleri yüksektir.**

Yukarıda bahsi geçen nedenler İHA'lara olan ilginin nedenlerini göstermektedir. Bir İHA tasarlanırken, bahsi geçen rollerden birine ya da birkaçına uygun şekilde tasarlanabilmektedir. Bazı İHA'lar sadece tek amaç için (sadece keşif gözetleme gibi) tasarlanırken, bazıları birçok amaca hizmet edecek şekilde çeşitli kabiliyetler ile donatılabilmektedir. Bu nedenle şimdiye kadar farklı birçok İHA sınıflandırması yapılmıştır. Aksiyon mesafesi göz önüne alındığında aşağıdaki gibi bir sınıflandırma yapmak mümkündür (Carrillo, López, Lozano ve Pégard, 2012).

2.2. İnsansız Hava Araçlarının Sınıflandırılması

Yüksek irtifa uzun süreli dayanıklılık (High-altitude long-endurance - HALE)

15000 metre irtifada ve 24 saatten fazla uçuş yapabilen sistemler bu sınıfta yer almaktadır. Uzun menzilde keşif ve gözetleme operasyonlarında çoğunlukla kullanılmaktadır.

Orta irtifa uzun süreli dayanıklılık (Medium-altitude long-endurance -MALE)

Maksimum 24 saat dayanıklılık ile 5000–15000 metre irtifa arasında uçabilen sistemlerdir. HALE tipi insansız hava araçlarına benzer misyonlarda kullanılabilirler. Fakat bu sistemler daha yakın mesafelerde görev icra etmektedirler.

Orta Menzilli veya Taktik İHA (Tactical UAV -TUAV)

HALE ve MALE platformlarından daha küçük ve daha sade sistemlerdir. 100 kilometre ile 300 kilometre arasındaki mesafelere yapılan uçuşlarda görev almaktadırlar.

Yakın Menzilli İHA

100 km çalışma menzilinde görev alırlar. Keşif, hedef belirleme, havaalanı güvenliği, elektrik hatları muayenesi ve trafik izleme gibi operasyonlarda tercih edilmektedirler.

Mini İHA (MUAV)

Ağırlığı 20 kg mertebesinde olan bu tip İHA'lar, yaklaşık olarak 30 km menzillerde görev yapmaktadır.

Mikro İHA (MAV)

Kanat açıklığı 150 milimetreden fazla olmayan İHA'lar bu sınıfa girmektedir. Binaların yer aldığı şehir içi ortamlarda da görev alabilen İHA'lardır.

Nano Hava Araçları (NAV)

Yaklaşık 10 milimetrelilik boyutlara sahip, sürü halinde radar karıştırmak gibi amaçlarda kullanılan, ayrıca oldukça kısa menzillerde gözetleme yapabilen araçlardır. Bu araçlar üzerinde olabildiğince küçük, kompakt tahrik sistemleri ve kameralar barındırmaktadır.

2.3. İnsansız Hava Araçlarının Tasarımlarına Göre Sınıflandırılması

İnsansız hava araçları yapılarına / tasarımlarına göre dört ana kategoride incelenmektedir (Carrillo ve diğerleri, 2012).

Sabit Kanat

Uzun mesafeli, uzun menzilli ve yüksek irtifa gerektiren operasyonlarda kullanılmak üzere planör benzeri, kanatları hareket etmeyen ve süzülme hareketi gerçekleştirebilen sistemlerdir. Sivil uygulamalarda meteorolojik tahminler ve afet ile mücadele gibi amaçlarla kullanılmaktadır. Askeri uygulamalarda ise yukarıda bahsi geçen keşif ve gözetleme uygulamaları için uygun fonksiyonel özellikler göstermektedir. Bu sistemler GPS sinyallerini ve ataletsel navigasyon sistemlerini kullanarak uygun konumda uçuş gerçekleştirmektedir.

Döner kanat

Döner kanatlı insansız hava araçları, üzerinde pervane ve pervaneyi döndüren motor sistemlerini kullanarak uçabilen ve çoğunlukla süzülme hareketi gerçekleştiremeyen araçlardır. Döner kanatlı İHA'ların en bilindik örneği helikopterler ve dronlardır. Helikopterler aslında iki rotor bulunduran uçaklar olarak değerlendirilmektedir. Dron platformlarında ise ihtiyaca ve kaldırılacak yüke göre genellikle 4, 6 ve 8 adet rotor kullanılabilir. Bu sistemlerin en avantajlı tarafı havada asılı kalabilme yetenekleridir. Sadece dikey olarak kaldırma kuvveti üreterek havada oldukları gibi kalabilmektedirler. Bilindiği üzere dronların kullanımı gerek sivil gerekse askeri uygulamalarda oldukça geniştir. Tarım, fotoğrafçılık ve gezegen keşfi gibi sivil alanlarda bile örneklerini görmek mümkündür.

Çırpan Kanat

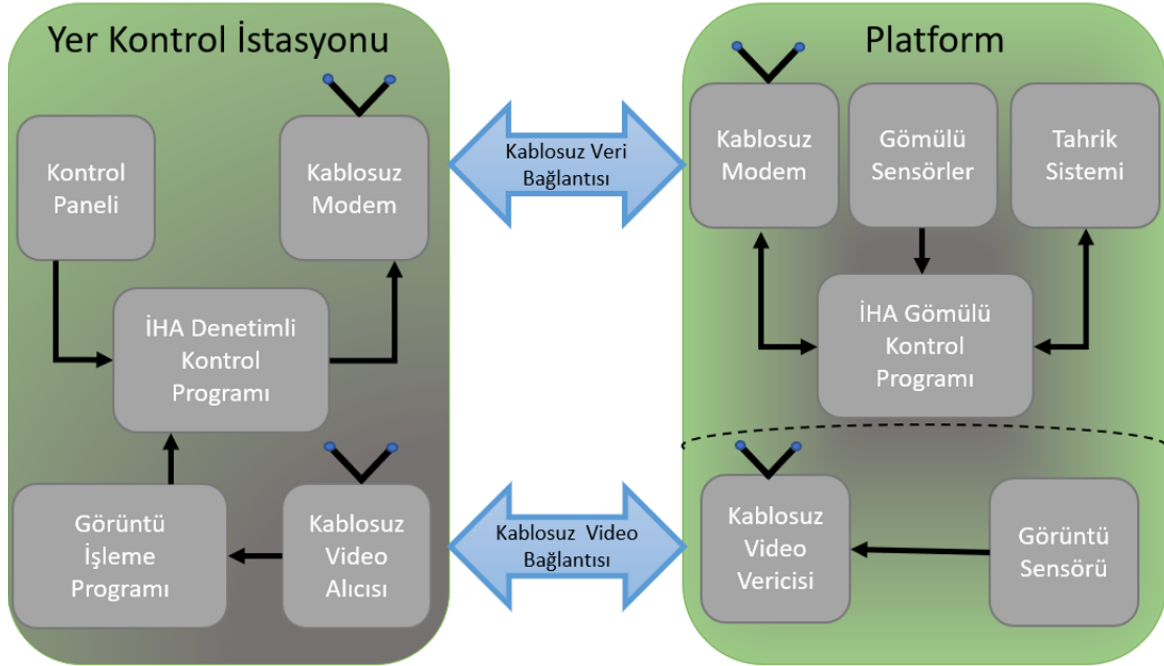
Çırpan kanatlı İHA'lar halen geliştirme aşamasında olup olgunluk seviyesinin yüksek olduğu bazı örnek ürünler mevcuttur. Bu platformlar oldukça küçük görev yükleri taşıyabilmektedir ve kısa mesafelerde görev almaktadırlar. Bu platformların avantajları dikey iniş kalkış yapabilmesi ve az güç gereksinimine sahip olmasıdır. Üzerinde oldukça hafif mikro kamera taşıyan versiyonları da mevcuttur. Süzülme hareketi ile güç verimliliği, ani manevralar ile atıklık ve ince tüylü kanat tasarımları ile sessizlik özelliklerini sağlayabilmektedir. Bu platformların kontrolü klasik aerodinamik kurallar ile açıklanamadığı için, kontrol algoritmalarını geliştirmek oldukça zordur.

Balonlar

Balonlar havadan daha hafif insansız hava araçları olarak, uzun menzillerde görev yapabilen, içindeki helyum gazı sayesinde uçabilen araçlardır. Bu araçların kalkış anında ek bir güç harcamasına gerek duyulmadığı için bu güç yükselen balonu yönlendirmekte kullanılabilir. Genellikle GPS tabanlı navigasyon sistemi kullanılan bu platformlarda, binalara yakın olarak şehir içinde ve yerel alanlarda da görevler gerçekleştirilebilir. Bu platformlar içinde helyum gazlı balon içerirler. Aynı zamanda denizanası gibi biyo-esinlenmeli yapısıyla kollarını kullanarak yükselme sağlayan platformlar da mevcuttur.

2.4. İnsansız Hava Araçlarının Temel Bileşenleri

Aşağıdaki Şekil 2.1’de İHA platformunun üzerinde ve yer kontrol istasyonunun yer alan temel bileşenler gösterilmiştir.



Şekil 2.1. İHA ve yer kontrol istasyonuna ait temel bileşenler

Şekil 2.1’de platform üzerinde gömülü işlemci, algılama ve tahrik bileşenleri ve yer istasyonu ile bağlantı sağlayan veri ve video linkleri gösterilmektedir. Platform üzerinde olmayan kısımda ise bir yer istasyonu bilgisayarı ve bir uçuş simülatörü kontrol paneli kablosuz bağlantı ile insansız hava platformuna bağlı konumda yer almaktadır. Operatör kontrol kumandası ile platformu kontrol ederken, komutlar kablosuz olarak platforma iletilir. Yer kontrol istasyonunda hem platformun durum tahmini ile ilgili veriler değerlendirilirken hem de görüntü işleme programları görüntüden anlamlı bilginin çıkarımı amacıyla kullanılmaktadır. İnsansız platformlarda ve özellikle hava araçlarında batarya kullanımını en aza indirmek ve en uzun uçuş süresini sağlamak amacıyla platform üzerinde en az işlem yapılacak şekilde optimizasyon gerçekleştirilmektedir. Bu nedenle platform üzerinde yüksek hesaplama maliyetine neden olacak her türlü işlemin eğer yapılabiliyorsa, kontrol istasyonunda gerçekleştirilmesi önem arz etmektedir. Özet olarak bir İHA platformunda yer alması gereken temel bileşenler aşağıdaki gibi sıralanabilir (Carrillo ve diğerleri, 2012):

- Tahrik sistemi
- Algılama sistemi
- Kablosuz iletişim bağlantıları
- Gömülü İşlemci ile
 - ✓ Sensör verilerinin okunması.
 - ✓ İletişimin yönetimi.
 - ✓ Kontrol algoritması.
- Batarya

2.5. Mikro / Nano İnsansız Hava Araçları

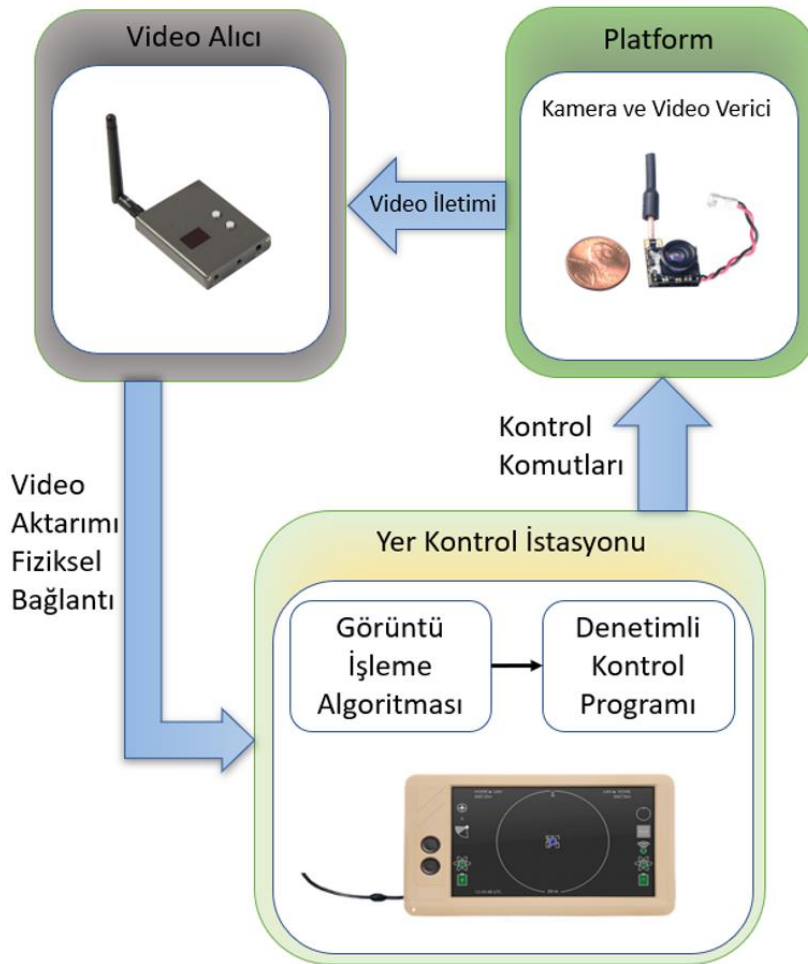
İHA'lar insanların hayati tehlikesinin en aza indirilmesi ve personelin görev alamayacağı ortamlarda görev icra etmesi bakımından tüm dünyada silahlı kuvvetler tarafından tercih edilen unsurlardandır. Son dönemde teknoloji yarışında İHA platformlarının boyutlarının azaltılması ve bu küçük platformların daha çevik hareket edebilmesi hedefine yönelik çalışmalar mevcuttur. Buradaki temel amaç hem sistemin fark edilebilirliğinin azaltılması hem de personel tarafından kolayca taşınabilirliğinin sağlanmasıdır. Tüm bu amaçlarla yapılan çalışmalar mikro / nano ölçekte insansız hava araçlarının ortaya çıkmasına neden olmuştur. Mikro / nano insansız hava araçları oldukça küçük, binalara fark edilmeden pencerelerden girebilen ve bu personelden önce keşif yaparak personeli koruyan insansız araçlardır. Mikro / nano ölçekteki insansız araçların boyutları 5-10 santimetre seviyesinde ve ağırlıkları ise 100 gramın altındadır. Bu ölçüler böceklerden büyük fakat kuşlardan küçük bir skalaya işaret etmektedir. Askeri kullanımının yanı sıra bu araçların sivil kullanımı da mevcuttur (Atcioglu, 2016).

Mikro / nano insansız sistemlere yönelik çalışmalar, 1900'lü yılların başlarında mikro sistemlerin (Micro-Electro-Mechanical Systems - MEMS) çalışılmasıyla başlamıştır. Daha sonra bu sistemlere ait mekanik, kontrol, haberleşme ve sensör çalışmaları ile süreç ilerlemiştir. Bu aşamada var olan pek çok sistemin küçültülme çalışmaları gerçekleştirilmiştir. Son 10 yıldır artık bu teknolojinin çıktılarını yavaş yavaş görmek mümkün olmaya başlamıştır. Mikro işlemci teknolojisindeki gelişmeler daha küçük hacimde daha fazla transistörün bir araya getirilmesini sağlamıştır. Benzer şekilde uzaktan kontrol teknolojileri de daha uygun fiyatlı ve hızlı kontrol sağlayan bir seviyeye ulaşmıştır

(Everaerts, 2006). Mikro / nano insansız hava araçlarının geliştirilmesi konusu ABD Savunma Bakanlığı ileri araştırma projeleri ajansı (Defense Advanced Research Projects Agency - DARPA) fonlarının bu konuya destek vermesi ile ivme kazanmıştır. DARPA'nın fonladığı projelerden elde edilen çıktılar takım ve birlik seviyesinde keşif ve gözetleme, arama ve kurtarma operasyonlarında savaş sahasının durumsal farkındalığının oluşturulmasında kullanılmıştır (Kendrick, 2014). İlerleyen çalışmalarda araştırmacılar doğadan esinlenerek, kanat çırparak uçan canlıları taklit etmiş ve bu canlıların uçuş aerodinamiğini çözmeye çalışmışlardır. Bu çalışmalarda platformun fark edilebilirliğini düşürebilmek amacıyla kuşların uçuş hareketlerine olabildiğince sadık kalmışlardır. Amerikan Savunma Bakanlığı'nın yol haritasına göre (Department of Defence [DoD], 2013) 2035 yılı ve sonrasında sahada mikro / nano insansız araçların kullanımının artması beklenmektedir. Bu nedenle pek çok savunma ve havacılık firmasının bu alana ilgisi yoğunlaşmış durumdadır. Hem platform tasarımları hem farklı platformların kontrolü anlamında ciddi çalışmalar sürdürülmektedir. Piyasaya ticari/askeri ürün olarak sürülen platformların mevcut durumda adet fiyatları 60.000 USD - 100.000 USD seviyelerindedir. Bu fiyatlar sahada binlerce platformun bir arada kullanılmasını bugün şartlarıyla desteklememektedir. Bu durum bu alanda yapılacak çalışmaların ve araştırmaların bu hedeflere ulaşılincaya kadar gündemde olacağını göstermektedir.

Mikro/nano insansız araç teknolojisi ile ilgili çözülmesi için üzerinde çalışılan en büyük zorluklar ise yeterli uçuş süresi sağlayacak batarya teknolojisi ve otonom uçuşu destekleyecek algılama ve planlama teknolojilerinin bu sistemler için geliştirilmesi olarak görülmektedir. Bu sistemlerin mevcut uçuş süreleri 7-15 dakika olarak gözlemlenmektedir. Bazı platformlar ise deneme aşamasında doğrudan kablolu olarak kullanılmaktadır. Çoğunlukla platformun bir yer istasyonu üzerinden kontrolü söz konusudur. Platform kamerasından aktarılan video görüntü yer istasyonunda değerlendirilmekte ve operatör tarafından platformun kontrolü gerçekleştirilmektedir. Bu alanda yapılan ilk otonom / yarı otonom sistem geliştirme çalışmaları da yine benzer şekilde yer istasyonu üzerinden daha güçlü donanımlar ile gerçekleştirilmektedir. İnsansız platformlar operasyonlarını sürdürürken, konumlama yapmak için genellikle GNSS sinyallerini kullanmaktadır. GNSS uydularından alınan sinyallerin platforma ulaşma süresi ve her uydudan iletilen kodun çözümü ile platformun uydulara uzaklığı tespit edilmektedir. Böylece uydulara olan uzaklık verisinden platformun dünya üzerindeki konumu da bulunmuş olur. Fakat meskûn mahalde ya da GNSS sinyallerinin alınamadığı iç mekânlarda konum bilgisinin üretilmesi

için alternatif yöntemlere ihtiyaç duyulmaktadır. Bu yöntemlerde ise, platform üzerinde konumlandırılmış sensörler marifetiyle konum çıkarımında veri kaynağı olarak kullanılmaktadır. Ayrıca mikro / nano platformlar için GNSS tabanlı bir konumlandırma yeterince hassas olamamaktadır. GNSS temelli navigasyona alternatif olarak bu tip platformlarda sensörler kullanılarak pozisyon tahmini ve konumlandırma gerçekleştirilir. Mikro / nano insansız araçlarda genellikle daha hafif, daha kompakt, daha az güç tüketen ve yüzlerce platformun bir arada operasyon yapmasını sağlayacak şekilde ucuz sensörlerin kullanımı oldukça önemlidir. Bu nedenle hem boyut olarak küçük hem de harcanan güç açısından avantajlı olan aynı zamanda çevrenin özellikleri hakkında da detaylı bilgi verebilen sensör olarak monoküler RGB kameralar ön plana çıkmaktadır.



Şekil 2.2. Mikro/nano İHA’larda monoküler görüntü sistemi

Şekil 2.2’de gösterilen sistemde, mikro / nano boyutlardaki platform üzerinde olabildiğince küçük bir RGB kamera yer almaktadır. Bu kameraya ait bir verici yine platform üzerinden

video iletimini kablosuz olarak gerçekleştirmektedir. Özellikle savunma uygulamalarında bu yayının olabildiğince korunaklı şekilde iletilmesi için çeşitli şifreleme yöntemleri de kullanılmaktadır. İletilen video bir video alıcı sayesinde, yer kontrol istasyonuna ulaşmaktadır. Yer kontrol istasyonunda görüntü işleme algoritmaları koşturularak videonun yorumlanması ve işlem amacına özgü verilerin çıkarımı gerçekleştirilmektedir. Eğer bu veriler platformun poz tahmininde kullanılacaksa, yer istasyonu üzerinde konumlama ve haritalama yapılmaktadır.

Monoküler görüntüleme sistemlerinde, ardışık görüntüler dizisinden, hareket ve 3 boyutlu yapılar hakkında bilgi tek kamera kullanılarak çıkarılabilmektedir. Burada sadece platformun gerçekleştirdiği hareketi içeren ana matris (essential matrix) E ya da hareket ile birlikte kamera içsel parametrelerini de içeren temel matris (fundamental matrix) F kullanarak iki görüntü arasındaki göreceli hareket kestirilmektedir. E matrisi sadece gerçek dünyadaki pozisyon ile ilgili verileri kullanırken (X matrisi ile ifade edilen), F matrisi gerçek dünyadaki koordinatlar ile görüntü düzlemindeki koordinatları arasındaki ilişkiyi içeren bir matristir. Bu matrisler kullanılarak her iki ardışık görüntü arasındaki öznitelik ya da pikselin konumu izlenerek x_1 - x_2 bu izlenen noktanın 3 boyutlu projeksiyonu gerçekleştirilmektedir. Kullanılacak olan matrisin (E ya da F) ayrıştırılması döndürme ve öteleme hareketi için olası 4 farklı çözüm sağlar. Bu çözümlerden yalnızca bir tanesi fiziksel olarak mümkündür. Çünkü sadece bir tanesi pikselin ya da özneliğin kameranın önünde bulunduğu duruma ait olup, 3 boyutlu dünyada yeniden oluşturulması mümkündür. Bir kez dönme ve öteleme hareketinin doğru kombinasyonu biliniyorsa, nesnelerin 3 boyutta tekrar oluşturulması üçgenleme yöntemi ile gerçekleştirilmektedir. Dönme hareketinde tahmin eşsizdir ve projeksiyon açısından herhangi bir soruna neden olmamaktadır. Fakat öteleme hareketi ve yapının derinliğini bulmak, ancak skaler bir faktöre kadar kurtarılabilir. Bunun nedeni hareket temel matrisinin (E matrisi) rankının 2 olmasıdır. Hareket temel matrisi sadece hareketin yönünün çıkarılmasını sağlamaktadır fakat harekete ait büyüklük bilgisi yer almamaktadır. 3 boyutlu yapıyı gerçeğe uygun ve hassas bir şekilde yeniden oluşturabilmek için, art arda gelen görüntüler arasında yeterince geniş bir taban çizgisi (yani geniş mesafeden aynı sahnenin çekimi) ve farklı açılardan yapının görüntülenmesi gerekmektedir. Fakat özellikle dronlarda bu gereksinimleri gerçekleştirmek oldukça zordur. Dron platformlarına yerleşik kameralardan alınan görüntülerde, iki ardışık görüntü arasında oldukça küçük hareketlerin gerçekleştirildiği gözlemlenmektedir. Ayrıca platform yörüngesi yoğunlukla kameranın gördüğü yönde

ilerleme hareketini içerecektir. Burada iki görüntü arasındaki temel çizgi fazla olana kadar alınan görüntülerin atlanması ve fazla olduğunda görüntülerin kullanılması akla gelebilir fakat dron hareketlerinin hızı ve pozisyon değişiklikleri çoğunlukla aynı sahneyi sürekli görecektir açıyı sağlayacak şekilde olmamaktadır.

Monoküler kameralar kullanıldığında 3 boyutlu yapıların tekrar oluşturulması yalnızca araç hareket ettiği müddetçe mümkün olabilmektedir. Dron platformlarının normal operasyonunda yapıyı tekrar oluşturacak şekilde aynı objelerin etrafında uzun süre gezinmesi ve dokuyu ortaya çıkarması pek operasyon ortamında pek makul olmamaktadır. Bu nedenle özellikle monoküler kamera kullanan dron platformları önünde yer alan yapının her özelliğini değerlendirememekte ve sadece kameradan elde edilen veri ile çarpışma önleme sağlanamamaktadır.

Bir diğer sorun olan ölçek belirsizliği ise genellikle ek sensörlerin (lazer mesafe ölçer vb.) kullanımı ile çözülmeye çalışılmaktadır. Dron platformunda ek sensörlerin kullanılmasının hem ağırlık hem de batarya açısından maliyetinin iyi belirlenmiş olması gerekmektedir. Ayrıca kullanılan sensörlerin ölçüm frekanslarının ve hassasiyetlerinin birbirleri ile uyumlu olması ve kalibrasyon sürelerinin batarya ömrünün fazlaca azalmasına neden olmayacak şekilde hızlı olması gerekmektedir. Özellikle mikro / nano platformlarda mümkün olan en az sensör ile operasyonun yönetilmesi elzemdir. Bazı durumlarda ortama işaretlerin yerleştirilmesi ve bu işaretlerin tanınarak derinliğin çıkarılması da mümkündür. Fakat dron platformları çoğunlukla keşif-gözetleme görevlerinde görev alarak ortamın genel hatlarıyla haritasının çıkarılması ve ortam hakkında bilgi sahibi olunması amacıyla kullanıldıkları için, çoğunlukla bilinmeyen ve işaret yerleştirilemeyecek ortamlarda kullanılmaktadırlar.

Stereo görüntüleme sistemlerinde ise herhangi bir ortamın ya da yapının tekrar oluşturulması için platformun hareket etmesine gerek olmamaktadır. İki farklı kameradan iki farklı görüş açısı ile aynı zaman diliminde alınan görüntüler kullanılarak objelere ait uzaklık bilgisi de çıkarılabilmektedir. Stereo kameralarda da taban çizgisinin genişliği gereksinimi önemlidir. Özellikle iki kameranin birbirinden olabildiğince uzak konulmasının avantajları mevcuttur. Fakat pek çok dron platformu buna pek uygun değildir.

Stereo sistemlerin tek kareden (iki farklı kamera ile) derinlik elde etmesinin oldukça önemli bir avantaj olduğunu bilinmektedir. Bununla birlikte bazı dezavantajlar da mevcuttur. Özellikle stereo tabanlı görüntü işleme algoritmaları monoküler tabanlı olanlara göre daha yüksek işlem kapasitesi ve daha fazla kaynak gerektirmektedir. İki farklı kameraya sahip olduğunuz düşünülünce tek kamera için gereken her işlemin (filtreleme, düzeltme ve gürültü önleme gibi) eş zamanlı olarak iki kamera için de gerçekleştirilmesi gerekmektedir. Aynı zamanda tanınan objelerin ya da çıkarılan özniteliklerin sağ ve sol kamerada eşleştirilmesi ve aynı olduklarının tespiti bir diğer gereken işlemdir. Kamera senkronizasyonunun ve görüntülerin alındığı veri yolunun dolmasının önlenmesi gibi iki kamera kullanmanın zorluklarının da kısıtlı kaynaklar ile aşılması gerekmektedir.

Stereo kameralar seçildiğinde platformdaki kamera kaynaklı ağırlığın monoküler kameralar ile karşılaştırıldığında en az iki kat artacağını bilinmelidir. Ayrıca hacimsel olarak da daha fazla yere ihtiyaç duyulmaktadır. İnsansız hava araçlarında kaldırılacak yükün artırılması, belirli bir noktaya ulaştığında mevcut sistemin yükü kaldıramamasına sebep olmaktadır. Böyle durumlarda daha fazla yük kaldırabilecek daha güçlü motorların seçimi, daha güçlü motorları besleyecek daha yüksek güç sağlayabilen bataryanın seçimine neden olmaktadır ve birbirini takip eden bir silsile şeklinde artan gereksinimler platformun boyutlarının büyümesine yol açmaktadır. Özellikle mikro/nano ölçekteki insansız hava araçlarında olabildiğince hafif gövde ve sensörlerin seçimi platformların havada kalma süresini arttırmaktadır (Carrillo ve diğerleri, 2012).

3. İNSANSIZ SİSTEMLERDE GÖRÜNTÜ TEMELLİ NAVİGASYON

3.1. Eş Zamanlı Konumlama ve Haritalama (SLAM)

Eş zamanlı lokalizasyon ve haritalama, 30 yıldır robotik ve insansız sistemler alanında aktif araştırma alanı durumundadır (Zaffar, Ehsan, Stolkin ve Maier, 2018). SLAM teknolojileri otomatik uzaktan kontrollü robotik ve insansız sistemlerin otonom insansız sistemlere dönüştürülmesi için gereken algılama (perception) teknolojileri arasında yer almaktadır. Bu teknolojiler algılama-planlama ve davranış döngüsünde otonom sistemleri meydana getirmektedirler. Günümüzde sağlıktan tarıma, madencilikten sürücüsüz araçlara, gezegen keşiflerinden nükleer çalışmalara kadar oldukça geniş bir alanda robotlara rastlamak mümkün hale gelmiştir (Ben-Ari ve Mondada, 2018). Otonom olarak operasyon gerçekleştirebilmek için robotların çevreyi tanımlaması ve ilgili çevrede nerede olduğunu bilmesi gerekmektedir. Bu problem eş zamanlı lokalizasyon ve haritalama problemi olarak bilinmektedir. SLAM genel olarak aşağıda belirtilen teknolojilerden oluşmaktadır:

- Haritalama (Burgard, Hebert ve Bennewitz, 2016)
- Özellik çıkarımı ve Eşleştirme (Bay, Tuytelaars ve Van Gool, 2006; Oliva ve Torralba, 2006)
- Konum Tahmini (Montemerlo ve Thrun, 2003; Grisetti, Stachniss ve Burgard, 2005)
- Döngü Kapatma (Williams, Cummins, Neira, Newman, Reid ve Tardós, 2009)

SLAM kullanılarak anlık bir harita oluşturulduğunda, konum tahmini aşamasından bağımsız olarak, bu haritanın engelden kaçınma, obje tanıma ve görev planlama gibi amaçlar için kullanılması da mümkün olabilmektedir (Bailey ve Durrant-Whyte, 2006). İlk yapılan çalışmalarda, SLAM sistemleri uzaklık sensörleri (akustik ve lidar) kullanılarak tasarlanmış olan sistemler olarak karşımıza çıkmaktadır. Bu sensörler derinlikle ilgili güvenilir veriler vermekle birlikte ortamın özellikleri hakkında bilgi vermemektedir. İlerleyen fazlarda ise görsel sensörler (kameralar) kullanılarak ortam özelliklerinin daha öncelikli olduğu fakat derinlik verisinin yeterince güvenilir olmadığı uygulamalar ortaya çıkmıştır. Bu iki sensörün avantajlarının birleştirilmesine yönelik olarak bu alanda en son yapılan çalışmalar incelendiğinde hem uzaklık hem de görsel bilgilerin bir arada kullanıldığı RGB-D sensör ve stereo kameraların daha çok tercih edildiğini görülmektedir.

(Cadena, Carlone, Carrillo, Latif, Scaramuzza, Neira ve Leonard, 2016). Özellikle uzaklık ölçümünün önemli olduğu çarpışma kaçınma ya da kamikaze dalış gibi operasyonlarda hem nesnenin tanınması hem de derinlik bilgisinin detaylı şekilde üretilmesi açısından bu yöntemler önem taşımaktadır. Bununla birlikte bu kameraların boyutları standart monoküler RGB kameralardan büyüktür ve daha ağırdırlar. Güç tüketimi olarak da değerlendirildiğinde, daha fazla güç harcadıkları bilinmektedir.

Akustik ve LİDAR Sensörler

Akustik sensörler renk, parlama, saydamlık gibi, kameraların çalışmasını olumsuz etkileyen ortam özelliklerinden etkilenmemektedir. Dolayısıyla karanlık ortamlarda kullanıma uygun sensörlerdir. Bu sensörler, özellikle çarpışmadan kaçınma ve güvenlik amaçlı kullanılmaktadır. Bununla birlikte yüksek miktarda toz, kir ve rutubetli ortamlarda sensör okumasında farklılıklar olabilmektedir. Ayrıca bu sensörlerin güç tüketimleri birkaç mili-watt ile birkaç watt aralığında değişmektedir. Bu açıdan oldukça düşük güç tüketimine sahip sensörlerdir.

LIDAR sensörler ise saniyede 1 milyon ışın gönderip alabilmektedir. Bu ışın demetleri sayesinde 3B haritalama yapmayı sağlayan nokta bulutu oluşturulabilmektedir. LIDAR sensörler oldukça güvenilir ($\sim \pm 2$ cm hassasiyet) derinlik bilgisi sağlamaktadır. Bununla birlikte bu sensörlerin ağırlık ve güç açısından dezavantajları mevcuttur (Zaffar ve diğerleri, 2018). LIDAR sensörler, sürücüsüz araçlar, insansız kara araçları ve bacaklı robotlar gibi uygulamalarda oldukça yaygın şekilde kullanılmaktadır. Bu uygulamalarda özellikle geniş görüş açısına sahip ya da 360 derecelik LIDAR sensörler tercih edilmektedir. Bu uygulamaların nispeten daha büyük bataryalara sahip olmaları LIDAR sensörlerin güç ihtiyacının karşılanmasında kilit rol oynamaktadır. Bununla birlikte insansız hava araçlarında (İHA) batarya kısıtları nedeniyle çoğunlukla tercih edilmemek ile birlikte, tercih edildiğinde tek yönlü ve sayıca daha az lazer demeti ileten daha az güç harcayan LIDAR sensörler tercih edilmektedir. LIDAR sensörler, ayrıntılı ve hassas haritalama ve konumlamanın gerektiği çalışmalara (maden, jeoloji vb.) daha uygundur.

Monoküler Kamera

Monoküler kameralar çoğunlukla 3 renk sağlayan standart RGB kameralardır. SLAM'da sadece monoküler kamera kullanıldığı durumda oldukça basit bir donanım altyapısı yeterli olmaktadır (Royer, Lhuillier, Dhome ve Lavest, 2007). Günümüzde akıllı cep telefonlarında da yer alan kameralar bile konumlama ve haritalama amacıyla kullanılabilir. Bununla birlikte statik resimlerden, bilinmeyen ortamlarda derinlik bilgisi doğrudan hesaplanamadığı için bu sensörler hassas mesafe ölçüm uygulamalarına uygun olmamaktadır. Bununla birlikte, daha önceden boyutları bilinen objelerin olduğu ortamlarda, derinlik çıkarımı belirli bir hassasiyete kadar gerçekleştirilebilmektedir. Kameraların pasif sensör olması ve güç harcamalarının az olması sadece kamera (camera-only) kullanılan yöntemlerin seçilmesinde etkilidir. Kamera edinme maliyetleri, çözünürlük, görsel özellikler, depolama vb. özelliklere göre çeşitlilik arz etmektedir. Daha düşük çözünürlüklü, kullanımı kolay kameralar için maliyet 10\$ (standart RGB gibi) seviyelerindeyken, yüksek özelliklere sahip 2500\$ (olay kameraları gibi) maliyet seviyesinde kameralar da mevcuttur.

RGB-D Kamera

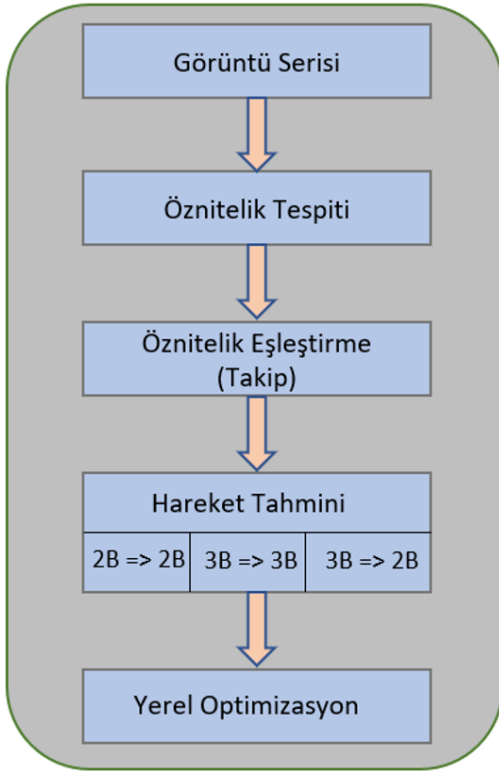
RGB-D kameralar uzaklık temelli sensörler ile görsel sensörlerin güçlü yönlerinin bir araya getirilmesi avantajını sağlamaktadır. RGB-D kameralar monoküler kameralar ile kızılötesi verici ve alıcıları bir araya getirmektedir. Her piksel için hem renk hem de derinlik bilgisi bu sensörler sayesinde alınabilmektedir. RGB-D kameralar kompakt yapıda fakat genelde geniş yüzeyli kameralardır.

Stereo Kamera

Stereo kamera, iki ayrı gözün birbirine olan mesafesine benzer şekilde yerleştirilmiş kameralar içermektedir. Aynı pozun iki farklı açıdan görüntülenmesi ile derinlik bilgisi elde edilebilmektedir. Stereo kameralar da pasif sensörlerdir. Monoküler kameralarda meydana gelen ölçeklemede kayma (scale drift) hatası bu kameralarda derinlik bilgisi doğrudan alındığı için ortadan kalkmaktadır (Post, 2020). Stereo kameralarda da edinme maliyeti çözünürlük, algılayabildiği uzaklık gibi etkenlere bağlı olarak değişiklik göstermektedir.

3.2. Görsel Odometre

Odometre zamana bağlı olarak sensör pozisyonunda ardışık olarak meydana gelen değişimlerin kullanılarak pozisyon tahmini yapılmasına verilen isimdir. Kamera temelli odometre ise (görüntü temelli odometre, visual odometry - VO) sensör olarak kameraları kullanmaktadır. Genel olarak görsel odometre adımları aşağıdaki Şekil 3.1’de verilmiştir.



Şekil 3.1. Görsel odometre adımları

Görsel odometrede tek kameradan veriler alındığı gibi birden çok kameradan alınan veriler de anlamlandırılabilir. Aşağıda lokalizasyon tespitinde faydalanan sensörlerin avantaj ve dezavantajları aşağıdaki tabloda gösterilmektedir (Aqel, Marhaban, Saripan ve Ismail, 2016).

Çizelge 3.1. Odometrede kullanılan sensörlerin avantaj ve dezavantajları

Sensör/teknoloji	Avantaj	Dezavantaj
Tekerlek odometresi	- Konum/Yönelim belirlemede kolaylık - Kısa sürede duyarlılık, yüksek örnekleme hızı - Düşük fiyatlı çözüm	- Teker (çark) kayması nedenli konumda sürüklenme. - Zamanla yığılma hatası - Sayısal diferansiyel işlemleriyle yapılan hız kestirimlerinin neden olduğu ek gürültü.
Ataletsel Navigasyon Sistemleri (INS)	3 eksenli ivmeölçer ve jiroskop kullanarak konumlama ve yönelme - Kesinti müdahalelerine maruz kalmama	- Konumda sürüklenme (konum kestirimleri 2.derece integral işlemi gerektirir) - Uzun süredeki sürüklenme hataları
GPS/GNSS	- Bilinen hata değerleri ile kesin konumlama sağlama - Zamanla hata yığılmalarının olmaması	- İç mekân, sualtı ve kapalı alanlarda kullanılamama. - RF müdahalelerinden etkilenme.
Ultrasonik sensör	- Sensörden hedefe skalar mesafe ölçüsü sunma - Uygun fiyatlı çözüm	- Yansıyan sinyal dalgasının malzemeye ve yüzeydeki çıkıntılarının yönelimine bağlı olması - Başka sensörlerin müdahalelerinden etkilenmesi - Düşük açısal çözünürlük ve tarama hızı

Çizelge 3.1. (devam) Odometrede kullanılan sensörlerin avantaj ve dezavantajları

Sensör/teknoloji	Avantaj	Dezavantaj
Lazer sensör	• Sonar sensöre benzemesiyle beraber daha yüksek duyarlılığa ve tarama hızına sahip olması • Tek noktadan mesafe ölçümü veya mesafeler dizisinden mesafe ölçümü	• Yansıyan sinyal dalgasının malzemeye ve yüzeydeki çıkıntıların yönelimine bağlı olması • Pahalı çözüm
Optik Kamera	• Resimlerin büyük anlam taşıyan bilgiyi barındırması • Yüksek konumlama duyarlılığı sağlaması • Uygun fiyatlı çözüm	• Görüntü işleme ve veri ayıklama tekniği gerektirmesi • Görüntü işlemede yüksek hesaplama maliyeti

Görsel Odometre'nin yaygın olarak tanınması 2004 yılındaki NASA'nın Mars görevinde sağlanmıştır. "Görsel odometre" ismi de tekerlek odometresi ile benzerliğinden dolayı ilk defa bu yılda verilmiştir (Nistér, Naroditsky ve Bergen, 2004b). Tekerlek odometresinde tekerin her dönüşünden ilerleme tespiti yapıldığı gibi görsel odometrede de piksellerdeki zamana bağlı değişim dikkate alınarak yer değiştirme tespit edilmektedir.

Görsel Odometre tekniği GPS, INS, tekerlek odometresi ve sonar sistemlere göre oldukça hassastır. Bu teknikte göreceli pozisyon hata aralığı %0,1 ile %2 arasındadır (Scaramuzza ve Fraundorfer, 2011). Ayrıca görsel odometrede, tekerleğin bazı yüzeylerde boş dönmesi ile ilgili bazı problemler de olmayacağı için tek başına tekerlek odometresinden daha güvenilir olarak değerlendirilmektedir.

Görsel Odometre iç mekânlarda başarıyla uygulanmasına rağmen, dış mekân uygulamalarında hala çözülme bekleyen pek çok problem mevcuttur. Maruz kalınan direkt gün ışığı, gölgeler ve dinamik değişimler hala üzerinde çalışılan konular olarak gözlenmektedir. Görsel odometre konusunda ana problemlerin hesaplama maliyeti, ışık ve görüntünün alınma farklılıkları olduğunu söylemek mümkündür (Gonzalez, Rodriguez, Guzman, Pradalier ve Siegwart, 2013; Nagatani, Ikeda, Ishigami, Yoshida ve Nagai, 2010; Nourani-Vatani ve Borges, 2011; Yu, Pradalier ve Zong, 2011). İç mekânlarda ise ışığın heterojen dağıldığı, dokuların ve yakalanacak özelliklerin olmadığı bölgelerde ve statik ya da dinamik objelerin gölgelerinin bulunduğu ortamlarda hata oranlarının yüksek çıktığı bilinmektedir (Gonzalez, Rodriguez, Guzman, Pradalier ve Siegwart, 2012; Nourani-Vatani ve Borges, 2011). Görsel Odometre aynı zamanda araç sürüş asistanlığı için de kullanılmakta olup LIDAR sistemlerden daha maliyet verimli bir çözüm sunmaktadır. Görsel Odometrede monoküler (Royer ve diğerleri, 2007; Gonzalez ve diğerleri, 2013; Nourani-Vatani ve Borges, 2011; Yu ve diğerleri, 2011; Gonzalez ve diğerleri, 2012; Lovegrove, Davison ve Ibanez-Guzmán, 2011; Martinez, 2013; Pathak, Moro, Fujii, Yamashita ve Asama, 2019; Jiang, Xiong, Chen ve Lee, 2014), stereo (Nistér, Naroditsky ve Bergen, 2006; Howard, 2008; Azartash, Banai ve Nguyen, 2014; Golban, Istvan ve Nedeveschi, 2012; Soltani, Taghirad ve Ravari, 2012; Siddiqui ve Khatibi, 2015; Mouats, Aouf, Sappa, Aguilera ve Toledo, 2014; McManus, Furgale ve Barfoot, 2013; Jiang, Xu ve Liu, 2013), RGB-D (Fabian ve Clayton, 2014; Steinbrücker, Sturm ve Cremers, 2011; Fang ve Zhang, 2015), omnidirectional (Scaramuzza ve Siegwart, 2008; Bunschoten ve Krose, 2003; Scaramuzza ve diğerleri, 2011, Tardif, Pavlidis ve Daniilidis, 2008) kameralar gibi görsel sensörlere uygun şekilde çalışabilecek algoritmalar yer almaktadır. Bu üç kamera türünün avantajları ve dezavantajları aşağıda yer alan Çizelge 3.2’de listelenmiştir (Aqel ve diğerleri, 2016).

Çizelge 3.2. Görsel odometrede kullanılan kamera türlerinin karşılaştırması

VO kamera türleri	Avantaj	Dezavantaj
Monoküler	- Düşük maliyet ve kolay yerleştirme - Hafiflik: küçük robotik uygulamalar için iyi - Kolay kalibrasyon	Ölçeklemedeki belirsizlik durumu

Çizelge 3.2. (devam) Görsel odometrede kullanılan kamera türlerinin karşılaştırması

VO Kamera Türleri	Avantaj	Dezavantaj
Stereo	- Görüntünün ölçek ve derinlik bilgilerinin kolay alınması 3 boyutlu görsel sağlaması	- Monoküler kameradan daha pahalı ve kalibrasyonunun daha uğraştırıcı olması - Stereo kamera baseline'ı sahnenin kameradan uzaklığından küçük olduğunda monoküler kamera gibi çalışır. - Arayüzleme ve senkronlamada zorluk
Her yönlü Omnidirectional	- Çok geniş görüş alanı (FOV) (360 dereceye kadar FOV) - Dünya yapılarının temiz 3 boyutlu modellerini oluşturabilme - Dönme değişmezliği (dönmede eşdeğerlik)	- Karmaşık sistem - Birden fazla kameranın kalibrasyonu ve senkronizasyonu - Yüksek bant genişliği gerekliliği - Pahalılık

Görsel Odometre’de öznitelik temelli (feature based), doğrudan (appearance-based / direct) ve hibrit (öznitelik ve doğrudan bir arada) yaklaşım olmak üzere 3 yaklaşım mevcuttur.

3.2.1. Öznitelik temelli yaklaşımlar

Öznitelik temelli yöntemler NASA’nın Mars’a göndermiş olduğu Roverlarda (insansız kara aracı) kullanılmasının ardından popülerlik kazanmıştır (Maimone, Cheng ve Matthies, 2007; Johnson, Goldberg, Cheng ve Matthies, 2008).

Monoküler

Öznitelik temelli yaklaşımlarda görüntülerde yer alan özniteliklerin (corner, line and curves) ardı ardına gelen resim karelerindeki durumlarına göre hareketin tahminini gerçekleştirmektedir. Önceki resim karesi (frame) ile yeni gelen resim karesinde yer alan özelliğin Öklid farkı (Euclidean distance) hesaplanarak değişim hesaplanmaktadır.

Stereo

Stereo görsel odometre uygulandığında, ilk kameradan elde edilen resim karesinden çıkarılan öznitelikler, ikinci kameradan elde edilen aynı sıradaki resim karesinde karşılık gelen anahtar noktalarla eşleştirilmektedir. Böylece uzayda noktaların kameraya göre 3 boyutlu konumu elde edilmektedir. Kamera hareketi, elde edilen özniteliklerin her iki kameradan alınan resim karelerindeki yer değişimine göre tahmin edilmektedir. Elde edilen bir dizi öznitelik noktasının art arda gelen iki resim karesindeki eşleşmesinin yapılması için öznitelik tanımlayıcılar (feature descriptors) arasında en yakın komşu çiftleri (nearest neighbour pairs) belirlenmektedir.

Eğer stereo kameradan sahnedeki özelliğe olan mesafe, stereo kamera temel çizgisinden (iki kameranın birbirine olan uzaklığı - baseline) çok fazla ise (40 kat gibi) stereo kamera ile haritalama gerçekleştirilememektedir. Yakın mesafelerde tespit edilen öznitelikler kullanılarak gerçekleştirilen haritalama performansı daha yüksekken, uzak mesafeler sadece yörünge kestiriminde bu anahtar noktalar kullanılmaktadır.

3.2.2. Doğrudan yaklaşımlar

Bu yöntemlerde görünümdeki değişimler takip edilerek konum değişimi elde edilmektedir. Burada piksel yoğunluk (intensity) değerine odaklanılarak piksellerin hareketleri gözlemlenmektedir. Kameranın hareket hızı optik akış (optical flow) kullanılarak tahmin edilebilmektedir. Optik akış algoritmalarında görüntüde odaklanılan piksel ve komşularının yoğunluk değerlerini takip ederek ardı ardına gelen resimlerde piksel yer değiştirmeleri bulunmaktadır. Ardı ardına gelen resim karelerinin zaman bilgisini kullanılarak hareketli platformun ya da hareketin hızı tespit edilebilmektedir.

Doğrudan yaklaşımlarda şablon eşleştirme (template matching) metodu da oldukça kullanılan yöntemlerdendir. Bu yöntemde resim karesinde bir şablon belirlenmekte ve bu şablonun art arda gelen karelerde ne şekilde değişeceği izlenmektedir. Yer değişimleri ve dönme hareketleri bu şablonlar üzerinden belirlenmektedir. Bazı şablon karşılaştırma konusunda ön plana çıkan algoritmalar olarak (Yoo, Hwang, Kim, Ki ve Cha, 2014; Brunelli, 2009) bilinmektedir. Doğrudan yaklaşımlara (Gonzalez ve diğerleri, 2013; Nourani-Vatani ve Borges, 2011; Yu ve diğerleri, 2011; Gonzalez ve diğerleri 2012; McManus ve diğerleri, 2013) çalışmaları örnektir.

3.2.3. Hibrit yöntemler

Öznitelik temelli yaklaşımlar öznitelik sayısının fazla olduğu özellikle meskûn mahallerde daha başarılı sonuçlar sağlamaktadır (Brunelli, 2009; Gonzalez ve diğerleri, 2012). Bununla birlikte dokusuz ve özniteliklerin az bulunduğu bölgelerde ya da tek bir dokunun (tekrarlı doku) bulunduğu bölgelerde bu algoritmaların verimi düşmektedir.

Doğrudan yaklaşımlarda ise başarı, düşük özellik bulunan ve dokusuz bölgelerde öznitelik temellilere kıyasla daha yüksektir (Kicman ve Narkiewicz, 2013; Nourani-Vatani ve Borges, 2011). Dokusuz bölgelerde oldukça büyük bir şablon belirlenerek bu şablondaki değişimlere göre daha hassas ve başarılı sonuçlar elde edilmektedir.

Bazı senaryolarda bu iki ayrı yöntemin kombinasyonunun en iyi sonuç verdiği tespit edilmiştir. Bu yöntemlerde hem piksel yoğunluk değeri hem de belirlenen bazı öznitelikler takip edilmektedir (Scaramuzza ve Siegwart, 2008).

Görsel Odometrede Kısıtlar

Genel olarak görsel odometrenin kısıtlayıcı yönleri olarak hesaplama maliyeti, ışık ve resim karesinin elde edildiği ortamın özellikleri sayılabilmektedir. Ortamda kameranın doğrudan gün ışığına maruz kalması, gölgeler, resimde meydana genel bozulmalar (hareket bulanıklığı gibi) ve dokusuz bölgeler kısıtlayıcı etkiler olarak değerlendirilmektedir (Nourani-Vatani ve Borges, 2011; Yu ve diğerleri, 2011; Gonzalez ve diğerleri, 2012; Pathak ve diğerleri, 2019).

Ölçeklemede Belirsizlik

Monoküler kamera kullanılarak gerçekleştirilen odometre çalışmalarında sistemin ölçeklemede belirsizlikten (scale uncertainty) olumsuz olarak etkilendiği bilinmektedir. Bu tez çalışmasında gerçekleştirilen deneylerde monoküler kameranın kullanıldığı algoritmalarda ölçek sorununun 30 metrelik bir koridorda 9 metreye kadar kaydığı gözlemlenmiştir. Stereo kameralarda temel çizgi uzaklığı bir referans teşkil etmektedir. Bu yüzden stereo kameralarda ölçekleme doğru bir şekilde yapılabilmektedir. Monoküler kameraların kullanıldığı durumda ise ilerleyen bölümlerde bahsedilecek olan döngü kapama algoritmalarının kullanılması ve global harita optimizasyonunun gerçekleştirilmesi gerekmektedir. Görsel odometre çalışmalarının haritalama özellikleri kazanarak ve gelişerek vSLAM yaklaşımlarına dönüşmektedir. Görsel Odometre’de öne çıkan yaklaşımlar ve karşılaştırmalı gösterimi aşağıdaki Çizelge 3.3’te verilmiştir (Aqel ve diğerleri, 2016).

Çizelge 3.3. Görsel odometre yaklaşımlarının karşılaştırılması

Referans	Kamera Türü	Yaklaşım	VO Kestirim Duyarlılığı	Kısıtlar
Gonzalez ve diğerleri, 2012	2 monoküler kamera: yerdeğişimi için yere bakan kamera ve yönlendirme için öne bakan kamera	Doğrudan yaklaşım (NCC: Normalized Cross Corelation şablon eşleme)	Hata toplam hareket mesafesinin %3’ünden az ve 8 derecelik ortalama yönlenme hatasından az	1,5 m/s’nin üzerinde hızlarda gölge ve bulanıklık nedeniyle yanlış eşleşme Düz olmayan yüzeylerde ölçek değişimleriyle başa çıkamama

Çizelge 3.3. (devam) Görsel odometre yaklaşımlarının karşılaştırılması

Referans	Kamera Türü	Yaklaşım	VO Kestirim Duyarlılığı	Kısıtlar
Van Hamme, Goeman, Veelaert ve Philips, 2015	Monoküler kamera	Öznitelik tabanlı yaklaşım (zemindeki özelliklerin izlenmesi için ters perspektif projeksiyonu ve Kalman filtresi)	%8,5'dan büyük translasyon (öteleme) hatası (800m için)	İlgili segmentlerdeki yol ortamının düzlemsel olmayışı nedeniyle tahmini bazı yörünge segmentleri üzerinde önemli dönme eğilimi
Scaramuzza ve Siegwart, 2008	Çok yönlü kamera	Hibrit yaklaşım (translasyonu tahmin etmek için SIFT özellik noktalarını karadan izlenirken, görselin görünüm benzerliğinin ölçümü (NCC, Manhattan ve Euclid mesafesi) aracın dönüşünü tahmin etmek için kullanıldı.)	Kat edilen mesafenin %2'sinden daha küçük 5 derece ortalama yönlendirme hatası	Düzlemsel hareket varsayımını ihlal eden yol tümseklerinden kaynaklanan kaçınılmaz görsel kilometre sayacı sürüklenmesi ve sapması
Nistér ve diğerleri, 2004b	Stereo kamera	Öznitelik tabanlı yaklaşım	380 m'nin üzerinde kat edilen mesafenin %1,63'lük hatası	Belirtilmemiştir

Çizelge 3.3. (devam) Görsel odometre yaklaşımlarının karşılaştırılması

Referans	Kamera Türü	Yaklaşım	VO Kestirim Duyarlılığı	Kısıtlar
Howard, 2008	Stereo kamera	Öznitelik tabanlı yaklaşım (Öznitelik eşleştirme ve sınır içi değer algılaması için stereo aralık verilerini kullanma)	380 m'nin üzerinde kat edilen mesafede %0.25'lik hata	Kendi gölgesinin yanlış eşleşmelere yol açması Bitki örtülü bir çevrede etkili çalışmaması
Nourani-Vatani ve Borges 2011	Monoküler kamera	Doğrudan yaklaşım (entropiye dayalı en iyi şablonu seçen NCC çok şablonlu eşleme)	Toplam hareket mesafesinin %5'inden küçük hata 5 derece ortalama istikamet hatası	Pürüzlü yüzeylerde ölçek değişimi ile ilgili eksiklik Sistemin güneşli/gölgeli alan konusunda etkisiz olması
Yu ve diğerleri, 2011	Monoküler kamera	Doğrudan yaklaşım (NCC döndürülmüş şablonlu eşleme)	%1.38 mesafe hatası ve 2,8 derece istikamet hatası	Görsel ölçek çeşitliliğinde etkisiz olması (gölgeler ve bulanıklık)
Nagatani ve diğerleri, 2010	Telecentric kamera	Doğrudan yaklaşım (NCC şablon eşleşmesi)	İç mekân testinde %3'ün altında hata 0,4 m/s hızda %1,5 hata (100 m'lik yörüngede)	Zemin çeşitliliğinden kamera yüksekliğini kestirememesi

Çizelge 3.3. (devam) Görsel odometre yaklaşımlarının karşılaştırılması

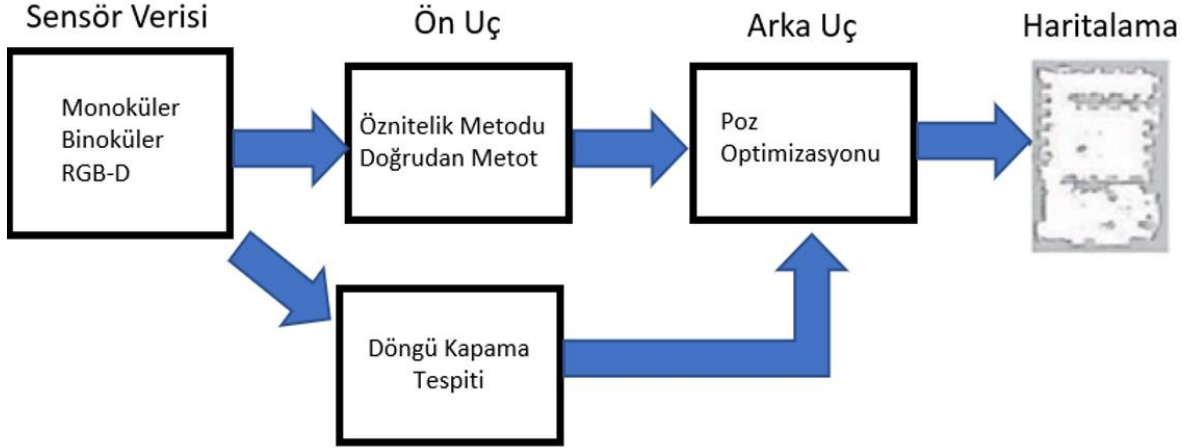
Referans	Kamera Türü	Yaklaşım	VO Kestirim Duyarlılığı	Kısıtlar
Zhang, Zhao, Xiong, Ma ve Gao, 2014	Monoküler kamera	Öznitelik tabanlı yaklaşım (Lucas Kanade Tomasi (LKT) kullanarak özelliklerin takibi)	Kat edilen mesafenin %1'inin altında hata	Karmaşık zemin koşullarında görüntü ölçeği belirsizliği, örneğin gevşek zemin döşemeleri

vSLAM ile görsel odometre birbiri ile oldukça ilişkilidir. Çünkü her iki teknik de sensor pozisyonlarını çeşitli dış dünya verilerinden tahmin ederek çalışmaktadır. Bu konuda incelenilecek araştırmalar (Fraundorfer ve Scaramuzza, 2012; Engel, Sturm ve Cremers, 2013) mevcuttur. vSLAM ile VO arasındaki fark global harita optimizasyonu modülüdür. Bu konuda (Davison, 2003) çalışması incelenebilir. Görsel odometre yöntemlerine global harita optimizasyonu özelliğinin eklenmesi ile vSLAM yöntemlerine geçiş yapılmış olur.

3.3. Görüntü Temelli SLAM (vSLAM)

Görsel eş zamanlı lokalizasyon ve haritalama (vSLAM) ise daha önce bahsi geçen monoküler, stereo, RGB-D gibi kameraların kullanıldığı SLAM tekniğine verilen addır. vSLAM'de amaç anlık olarak harita oluştururken, eş zamanlı olarak kamera pozunu tahmin etmektir (Mur-Artal ve Tardós, 2017). Bu teknikte yalnızca görsel sensörlerden faydalanarak lokalizasyon ve haritalama gerçekleştirilmektedir. Görsel sensörler kullanılarak, robot bağımsız şekilde kendi pozisyonunu tahmin edebilmekte ve eş zamanlı olarak bulunduğu çevrenin haritasını da çıkarabilmektedir (Mur-Artal ve Tardós, 2017). Fakat vSLAM diğer sensörlerin kullanıldığı SLAM metotlarına göre teknik olarak hesaplama yoğunluğu nedeniyle daha zorlu bir metottur (Taketomi, Uchiyama ve Ikeda, 2017). Aşağıda vSLAM genel konseptini gösteren Şekil 3.2 verilmiştir. Ön uçta (front-end) görsel odometre yöntemleri kullanılırken, arka uçta (back-end) platforma ait geçmişe

dair yörünge ile yörünge birleştirme yapılır (Kümmerle, Grisetti, Strasdat, Konolige ve Burgard, 2011).



Şekil 3.2. vSLAM akış şeması

İlk vSLAM uygulamaları monoküler kamera kullanılarak 2000’li yılların başlarında ortaya konulmuştur. Bu uygulamalar öznitelik noktalarını (feature points) takip etmek ve haritalandırmak üzerine kurulmuştur. Bu nedenle bu uygulamalara öznitelik temelli (feature based) uygulamalar ismi verilmektedir (Taketomi ve diğerleri, 2017). Daha detaylı haritalamaların yapıldığı ve resimlerin ön işleme gerek duymadan takip edilebileceği bir diğer yöntem ise doğrudan yöntem (direct approach) olarak bilinmektedir. Doğrudan yöntemde anlık gelen resim ile bir sonraki resim piksel yoğunluğu açısından karşılaştırılarak takip ve haritalama işlemi gerçekleştirilmektedir (Taketomi ve diğerleri, 2017). Doğrudan metotlar özniteliklerin oldukça az olduğu ya da herhangi takip edilecek bir dokunun bulunmadığı ortamlar için (texture-less or feature-less environments) tüm resmin bir sonraki resim ile karşılaştırıldığı ortamlarda doğru çalışmayabilmektedir (Taketomi ve diğerleri, 2017). Fakat bu alanlardaki başarıyı performans olarak öznitelik temelli yöntemlere göre daha iyi sonuçlar vermektedir.

RGB-D kameraların kullanıldığı yöntemlerde ise hem derinlik hem de renkli görüntü bilgisi bir arada yer almaktadır. Benzer şekilde stereo kameralar kullanılarak hem iç hem de dış mekânlarda haritanın oluşturulması ve ek olarak engelden kaçınma, yol planlaması gibi uygulamaların gerçekleştirilmesi mümkündür. vSLAM uygulamalarında temel unsurlar ve gerçek zamanlı çalışma için gerekli modüller aşağıdaki gibidir:

- Başlatma (Initialization)
- Takip (Tracking)
- Haritalama (Mapping)
- Yeniden Konumlama (Relocalization)
- Global Harita Optimizasyonu (Global map optimization)

Başlatma

Kamera yörünge tahmini ve 3 boyutlu ortamın oluşturulması için belirli bir koordinat sistemi tanımlayarak başlangıç yapılması gerekmektedir. Başlangıçta global bir koordinat sistemi oluşturularak, kameranın ilk anda tespit ettiği özellikler (öznitelik ya da pikseller) bu koordinat sisteminde gösterilmektedir. Bu aşamadan sonra tespit edilen özellikler devamlı olarak takip edilerek hem takip hem de haritalandırma işlemleri yerine getirilmektedir.

Takip

İlk oluşturulan haritaya aktarılan öznitelikler ya da art arda gelen resim karelerinde yer alan pikseller takip edilerek kameranın bulunduğu yerin tahmini gerçekleştirilmektedir.

Haritalandırma

Kamera ile ilk defa görüntülenen bilinmeyen bölgeler nokta bulutu şeklinde haritaya eklenerek, harita 3 boyutlu olarak genişletilmektedir.

Yeniden Konumlama

Takip işlemi, kameranın hızlı hareketleri ya da bazı düzensizliklerden (disturbance) dolayı doğru şekilde yapılamadığında yeniden konumlama yapmak gerekmektedir. Bu tip hareketlerin ardından kamera daha önce görmüş olduğu öznitelikleri / pikselleri tekrar yakalayabilirse, kamerayı / platformu daha önce bu özelliklerin görüldüğü yere yeniden konumlayabilmektedir. vSLAM algoritmalarının gerçek zamanlı olarak çalışabilmesi için yeniden konumlama olmazsa olmazdır. Aksi halde vSLAM algoritması takipteki ilk

kayıptan (tracking loss) itibaren çalışmasını doğru şekilde sürdürememektedir. Yeniden konumlamanın yer almadığı vSLAM algoritmaları genellikle akademik çalışmalar olmakla birlikte, bu tarz çalışmalar gerçek dünyada kullanılamamaktadır.

Global Harita Optimizasyonu

Haritalar genellikle kamera hareketinin doğru tahmin edilememesinden kaynaklı birikmiş tahmin hataları içermektedir. Bu hataları (scale drift gibi) en aza indirmek için global harita optimizasyonu yapılmaktadır. Bu aşamada tüm haritanın bütünü ele alınmakta ve harita yeniden işleme tabi tutulmaktadır. Bazı kamera hareketlerinin ardından yeniden başlangıç noktası civarına gelindiğinin farkına varılırsa daha önce yapılmış olan kümülatif hatanın tespiti mümkün olmaktadır.

Elde edilen bu hatanın giderilmesi için döngü kapatma (loop closure) algoritmaları kullanılmaktadır. Döngünün tespiti için elde edilen yeni görüntü kareleri eskileri ile karşılaştırılmaktadır. Eğer birbirine çok yakın nitelikte özellikler mevcutsa, buradan daha önce geçildiğine dair bir sonuca varılmaktadır. Bu durumda konum ve harita yeni duruma göre güncellenmektedir.

Demet Dengeleme

Harita üzerinde kamera pozunu ve haritayı optimize etmek için kullanılmaktadır. Bu optimizasyon tahmin hatalarını oldukça verimli şekilde önlemektedir (Triggs, McLauchlan, Hartley ve Fitzgibbon, 1999). Eğer küçük ortamlarda kullanılacaksa döngü kapatma algoritması olmaksızın da kullanılabilir. Çünkü kümülatif hata küçük ortamlarda daha azdır.

Özet olarak; vSLAM algoritmaları başlatma, takip, haritalama, yeniden konumlama ve global harita optimizasyonu aşamalarından oluşmaktadır (Taketomi ve diğerleri, 2017). vSLAM de görüntü temelli odometre gibi öznitelik temelli metotlar ve doğrudan metotlar olmak üzere 2 genel gruba ayrılmaktadır.

Uzun bir süre vSLAM alanına öznitelik tabanlı yöntemler hâkim olurken, son yıllarda doğrudan ve yoğun formülasyonlar olmak üzere bir dizi farklı yaklaşım popülerlik

kazanmıştır. Tüm bu metotların altında, girdi olarak gürültülü ölçümler alan ve bilinmeyen (Y) gizli model parametreleri (3B dünya modeli ve kamera hareketi) için bir X tahmincisi hesaplayan olasılıklı bir model vardır. Tipik olarak, gerçek ölçümleri elde etme olasılığını maksimize eden model parametrelerini bulan bir Maksimum Olabilirlik (Maximum Likelihood) yaklaşımı kullanılmaktadır:

$$X^* := \operatorname{argmax}_x P(Y|X) \quad (3.1)$$

Öznitelik temelli yöntemler iki temel adımda ilerlemektedir. İlk olarak, ham sensör ölçümleri, ilgili noktaların görüntü koordinatlarının hesaplanması gibi genel problemin bir kısmını çözerek bir ara temsil oluşturmak için önceden işlenmektedir. İkinci olarak, hesaplanan ara değerler, geometri ve kamera hareketini tahmin etmek için olasılıksal bir modelde gürültülü ölçümler (Y) olarak yorumlanmaktadır. İlk adımda tipik olarak, seyrek bir dizi anahtar noktanın çıkarılması ve eşleştirilmesi gerçekleştirilmektedir. Bununla birlikte, yoğun, düzenli optik akış biçiminde bağlantılar oluşturmak gibi başka seçenekler de mevcuttur. Ayrıca, çizgi veya eğri parçaları gibi diğer geometrik şekillerin parametrik temsillerini çıkaran ve eşleştiren yöntemler de bu grupta yer almaktadır.

Doğrudan yöntemler, bu ön hesaplama adımını atlamaktadır ve gerçek sensör değerlerini (belirli bir zaman periyodu boyunca belirli bir yönden alınan ışık) bir olasılık modelinde Y ölçümleri olarak doğrudan kullanmaktadır.

Pasif görüş durumunda, sensör fotometrik ölçümler sağladığından, doğrudan yaklaşım fotometrik hatayı optimize etmektedir. Öznitelik temelli yöntemler ise geometrik bir hatayı optimize etmektedir, çünkü önceden hesaplanmış değerler noktasal konumlar veya akış vektörleri geometrik niceliklerdir. Derinlik kameraları veya lazer tarayıcılar (doğrudan geometrik miktarları ölçen) gibi diğer sensör modeller için doğrudan formülasyonlar da geometrik bir hatayı optimize edebilmektedir.

vSLAM yöntemlerinin harita yoğunlukları karşılaştırıldığında yoğun ve seyrek olarak iki değerlendirme yapmak mümkün olabilmektedir. Seyrek yöntemler yalnızca seçili bir dizi bağımsız noktayı (geleneksel olarak köşeler) kullanmakta ve bu noktaları yeniden oluşturmaktadır. Yoğun yöntemler 2 boyutlu görüntü alanındaki tüm pikselleri kullanarak

ve tüm pikselleri yeniden oluşturmaktadır. Ara yaklaşımlar (yarı yoğun) tüm yüzeyi yeniden yapılandırmaktan kaçınarak, ancak yine de (büyük ölçüde bağlantılı ve iyi kısıtlı) bir alt kümeyi kullanıp ve yeniden yapılandırmaktadır (Engel, Koltun ve Cremers, 2017).

Bununla birlikte, kullanılan görüntü bölgesinin kapsamı dışında, daha temel ve sonuç olarak ortaya çıkan bir fark, önceden bir geometrinin eklenmesinde yatmaktadır. Seyrek formülasyonda, komşuluk kavramı yoktur ve kamera pozları ve öz değerleri (intrinsic parameters) göz önüne alındığında geometri parametreleri (anahtar nokta konumları) koşullu olarak bağımsızdır. Diğer yandan yoğun (veya yarı yoğun) yaklaşımlar, kullanılan görüntü bölgesinin bağlantılılığından yararlanmaktadır. Bir geometriyi önceden formüle etmek için, geometrinin pürüzsüz olduğu kabul ederek işlem yapmak gerekmektedir. Aslında, yalnızca pasif görüşten gözlemlenebilir yoğun bir dünya modeli yapmak için böyle bir öncel zorunlu olarak gereklidir. Genel olarak, bu öncelik doğrudan ek bir log-olasılık enerji terimi (Newcombe, Lovegrove ve Davison, 2011; Stühmer, Gumhold ve Cremers, 2010) şeklinde formüle edilmektedir (Engel ve diğerleri, 2017).

Seyrek ve Öznitelik Temelli yaklaşımlar, bir dizi anahtar nokta eşleşmesinden 3 boyutlu geometriyi tahmin eden ve böylece önceden bir geometri olmadan geometrik bir hata kullanan en yaygın kullanılan formülasyonlardır. Örnekler arasında, en bilindik yöntemler olarak monoSLAM (Davison, Reid, Molton ve Stasse, 2007), PTAM (Klein ve Murray, 2009) ve ORB-SLAM2 (Mur-Artal ve Tardós, 2017) yöntemleri yer almaktadır.

Seyrek ve Doğrudan yaklaşımlar, geometrik bir önceli dahil etmeden, doğrudan görüntülerde tanımlanan bir fotometrik hatayı optimize etmektedirler. Bu formülasyonu kullanan çalışma olarak DSO (Engel ve diğerleri, 2017) algoritması ön plana çıkmaktadır. DSO yöntemi, genişletilmiş Kalman filtresine dayanan çalışmalarının aksine, doğrusal olmayan bir optimizasyon çerçevesi kullanmaktadır.

3.3.1. Öznitelik temelli yaklaşımlar

Özellik temelli yaklaşımlar filtre tabanlı ve demet dengeleme tabanlı olmak üzere iki alt alanda incelenmektedir.

MonoSLAM

İlk monoküler vSLAM 2003 yılında MonoSLAM ismiyle ortaya konulmuştur (Davison, Reid, Molton ve Stasse, 2007). MonoSLAM filtre tabanlı algoritmalar arasında gösterilen örneklerden biridir. Kamera hareketi ve 3 boyutlu bilinmeyen ortamın haritasının oluşturulması Genişletilmiş Kalman Filtresi (Extended Kalman Filter- EKF) kullanılarak oluşturulmuştur. 6 serbestlik derecesinde kamera hareketleri ve öznitelik noktalarının (feature points) 3 boyutlu pozisyonları EKF de durum vektörü olarak kullanılmıştır. Başlatma safasında bilinen bir obje kullanılarak harita oluşturulmaya başlanmaktadır. Global koordinat sisteminde ilk öznitelikler bu şekilde oluşturulmaktadır. Bu yöntemin dezavantajı hesaplama maliyetinin, ortamın büyüklüğü ile orantılı olarak artmasıdır. Büyük alanlarda, durum vektörünün boyutu öznitelik noktalarının sayısının artmasından dolayı oldukça artmaktadır. Bu nedenle gerçek dünya uygulamalarına uygulanması zordur (Taketomi ve diğerleri, 2017).

PTAM (Paralel Tracking and Mapping)

MonoSLAM yönteminde hesaplama maliyetinden kaynaklı olarak karşılaşılan sorunun çözümü için PTAM yöntemi geliştirilmiştir. PTAM de takip ve haritalama CPU üzerinde farklı threadlerde gerçekleştirilmektedir. Farklı iş parçacıkları kullanıldığı için takip ve haritalama birbirini etkilemeden, herhangi bir veri paylaşımına ihtiyaç duymadan eş zamanlı olarak çalışabilmektedir. Kamera pozisyonunun takibi anlık olarak yapılırken, haritalamada özelliklerin 3 boyutlu gösterimi için hala hesaplama maliyeti gerekmektedir. PTAM demet dengelemenin ilk defa real time vSLAM ile birlikte kullanıldığı uygulamadır. PTAM'den sonra çıkan SLAM algoritmaları bu çoklu iş parçacığı (multi-thread) yaklaşımını benimsemiştir. PTAM'da harita başlatılırken beş nokta algoritması kullanılmaktadır (Nistér, 2004a).

Takip aşamasında haritalanmış noktalar çevrenin dokusunda yer alan öznitelikler kullanılarak 2 boyuttan 3 boyuta geçirilir. PTAM de öznitelik noktaları kullanılarak üçgenleme gerçekleştirilir ve kamera pozisyonu tahmin edilir. Bu açıdan PTAM ilk öznitelik temelli yaklaşım olarak karşımıza çıkmaktadır. Bir resim karesi anahtar resim karesi (keyframe) olarak seçilir ve bu kareden büyük farklılığın olduğu başka bir kareye

geçiş olacağı zaman bu yeni kare de anahtar resim karelerine eklenmektedir. Resim kareleri arası farklılık daha kesin bir üçgenleme için kullanılmaktadır.

Demet dengeleme işlemi bazı resim kareleri girdi olarak kullanarak lokal olarak ve tüm resim kareleri ile global olarak uygulanmaktadır. PTAM algoritmasında takip işleminde “randomized tree” temelli arama kullanılmaktadır (Williams, Klein ve Reid, 2007). Filtre temelli haritalama ile demet dengeleme temelli haritalama literatürde karşılaştırılmıştır (Strasdat, Montiel ve Davison, 2012).

Buna göre;

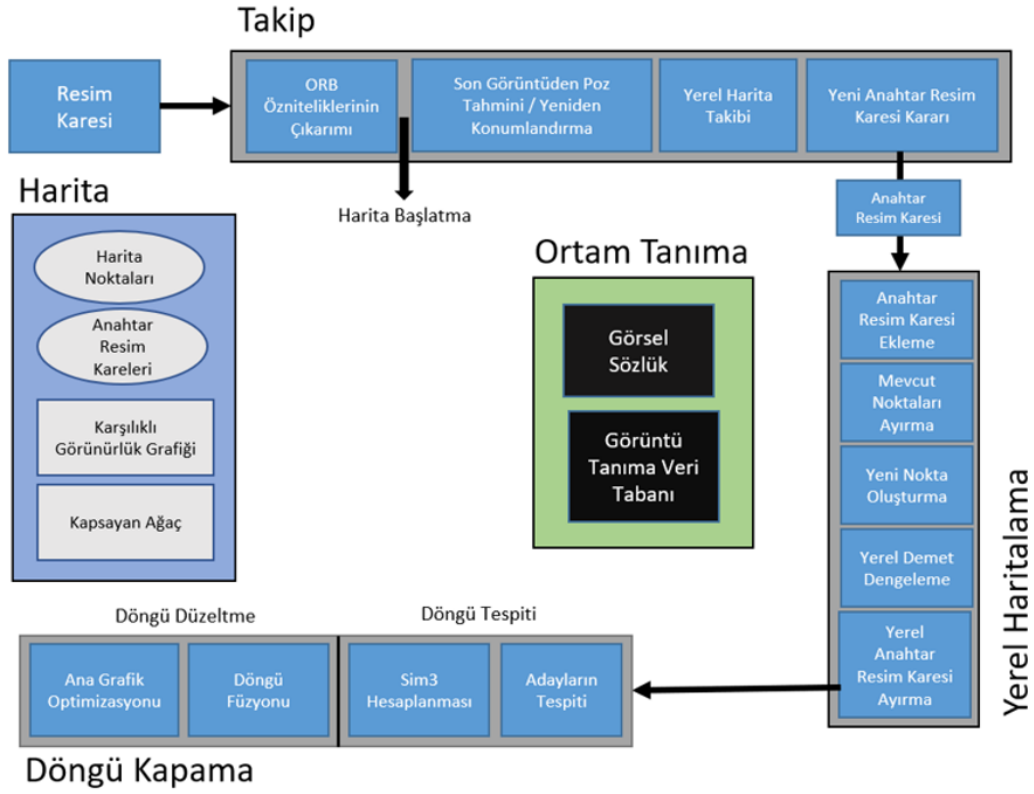
- Öznitelik temelli yaklaşımlarda özellik noktalarının sayısı arttıkça vSLAM algoritmalarının yörünge tahmini ve haritalama performanslarının arttığı gözlenmiştir.
- Performans karşılaştırması yapıldığında demet dengeleme yapılan algoritmaların filtre tabanlılara göre çok daha iyi performans sağladığı gözlenmiştir.

vSLAM- Global Harita Optimizasyonu

Stereo kameralar kullanıldığında koordinat sisteminde ölçeklendirme (scale) doğru bir şekilde belirlenebilmektedir. Fakat monoküler kamera kullanımında ölçeklendirmede bir miktar kayma meydana gelebilmektedir. Bu durumda harita optimizasyonu ve döngü kapatma yapmak gerekmektedir. PTAM yönteminin geliştirilmiş versiyonu olan ORB-SLAM metodu demet dengeleme, kapalı döngü algılama ve harita optimizasyonu özellikleriyle öne çıkmaktadır. Özellik temelli yaklaşımlar incelendiğinde en olgunlaşmış algoritmanın ORB-SLAM ve ORB-SLAM2 algoritması olduğu görülmektedir.

ORB-SLAM

ORB-SLAM algoritmasında çoklu iş parçacığı (multi-thread) kullanılarak takip, haritalama ve kapalı döngü algılama, poz optimizasyonu ve demet dengeleme bir arada ve uyumlu şekilde çalışmaktadır (Mur-Artal, Montiel ve Tardos, 2015). ORB-SLAM aynı zamanda açık kaynak kodlu olduğundan kolayca denenip geliştirmeye açıktır. ORB-SLAM yöntemine ait akış şeması Şekil 3.3’te gösterilmiştir.



Şekil 3.3. ORB-SLAM algoritması akış şeması

ORB-SLAM yöntemi, ORB özniteliklerinin tespitini gerçekleştirerek, bu öznitelikleri konumlama ve haritalama işleminde kullanmaktadır.

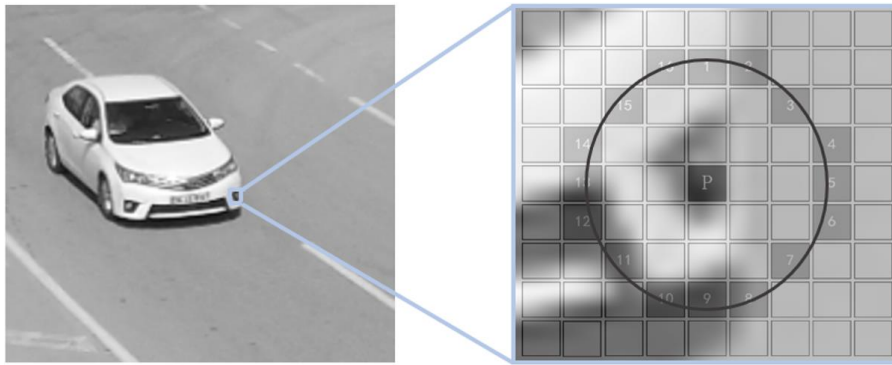
ORB (Oriented FAST and Rotated BRIEF)

ORB öznitelikleri kullanılarak gerçek zamanlı olarak oldukça hızlı hesaplamalar ve karşılaştırmalar yapılabilmektedir. ORB, FAST köşe tespiti altyapısını kullanırken, BRIEF öznitelik tanımlayıcısını temel almaktadır. ORB, OpenCV laboratuvarlarında SIFT ve SURF özniteliklerine göre daha performansı yüksek ve uygulanabilir bir alternatif olarak geliştirilmiştir. SIFT ve SURF patentli algoritmalar olduğu için ORB açık ve kullanımı ücretsiz olarak tasarlanmıştır. ORB'un literatüre en önemli katkıları aşağıdaki gibidir:

- FAST'e hızlı ve doğruluğu yüksek yönelim bileşeninin eklenmesi
- Yönlendirilmiş BRIEF özelliklerinin verimli hesabı, varyans ve korelasyon analizi
- Dönel değişmezlik altında BRIEF özelliklerinin kullanımı

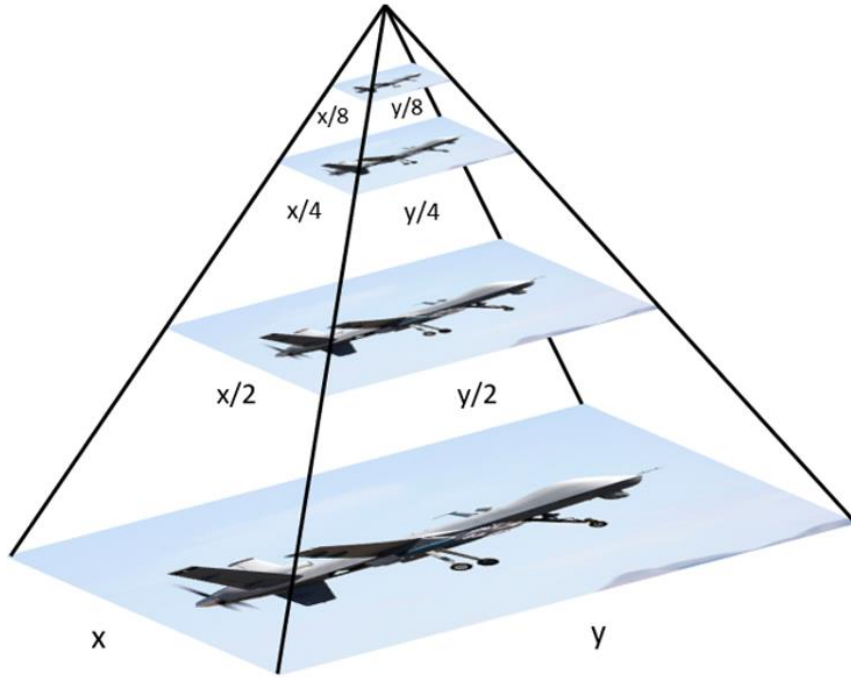
FAST (Features from Accelerated and Segments Test)

Aşağıdaki Şekil 3.4’de FAST öz niteliklerinin belirlenmesinin gösterildiği bir örnek yer almaktadır. Burada gösterilen p noktasının, çevre pikselleri ile yoğunluk (intensity) farkı araştırılmaktadır. Belirli bir yarıçaptaki çevre piksellerde gösterilen 1,5,9,13 piksellerine (4 yöndeki temel pikseller) öncelikli olarak bakılmaktadır. Bu noktalarda p ye göre yoğunluk farkı varsa aynı yarıçaptaki ‘n’ nokta da yoğunluk farkı var mı diye kontrol edilir. Çemberdeki pikseller “p’den daha açık”, “p’den daha koyu” veya “p’ye benzer” şeklinde üç sınıfa ayrılmaktadır. 8 pikselden fazlası p’den daha koyu veya daha parlaksa, p anahtar nokta olarak seçilir. Burada farklı sayıdaki noktaları ifade eden $n \geq 12$ olmak zorundadır. Böylece FAST algoritması en temel olarak kenarların konumları hakkında bilgi vermektedir.



Şekil 3.4. FAST öz niteliklerinin tespiti

Ancak, FAST öz nitelikleri bir yönlendirme bileşenine ve çok ölçekli yapıya sahip değildir. ORB algoritması, öz nitelik tespitinde çok ölçekli bir görüntü piramidi (Şekil 3.5) kullanmaktadır. Görüntü piramidi, tek bir görüntünün, tümü görüntünün farklı çözünürlüklerdeki sürümleri olan görüntü dizilerinden oluşan çok ölçekli bir temsildir. Piramitteki her seviye, görüntünün bir önceki seviyeden daha düşük örneklenmiş düşük ölçekli sürümünü içermektedir. ORB piramidinde, görüntüdeki anahtar noktaları algılamak için FAST algoritmasını kullanılmaktadır. ORB, her ölçek seviyesinde anahtar noktaları tespit ederek, farklı bir ölçekte bu noktaları etkili bir şekilde konumlandırmaktadır. Bu yönüyle, ORB kısmi ölçek değişmezdir.



Şekil 3.5. ORB ölçeklendirme piramidi

Anahtar noktaların yeri belirlendikten sonra, o anahtar nokta etrafındaki yoğunluk seviyelerinin nasıl değiştiğine bağlı olarak her bir anahtar noktaya sola veya sağa bakma gibi bir yön atanmaktadır. ORB, yoğunluk değişikliğini tespit etmek için yoğunluk merkezini kullanmaktadır. Yoğunluk merkezi, bir köşenin yoğunluğunun merkezinden dengelendiğini varsayar ve bu vektör bir yönlendirmeyi belirlemek için kullanılabilir. İlk olarak, bir yamanın momenti şu şekilde tanımlanmaktadır:

$$m_{pq} = \sum_{x,y} x^p y^q I(x,y) \quad (3.2)$$

Bu momentler ile, yamanın "kütle merkezi" şu şekilde bulunabilmektedir:

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (3.3)$$

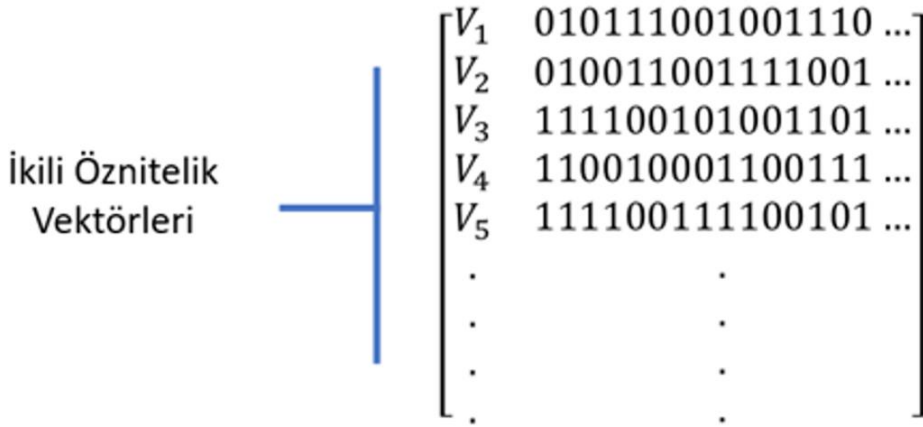
Köşenin merkezi O'dan, -OC ağırlık merkezine doğru bir vektör oluşturulur. İlgili yamanın yönelimi (orientation of the patch) aşağıdaki gibi ifade edilmektedir:

$$\Theta = \text{atan2}(m_{01}, m_{10}) \quad (3.4)$$

Yamanın yönü hesapladıktan sonra, onu kanonik bir forma döndürebilmekte ve ardından tanımlayıcıyı hesaplayarak, böylece bir miktar dönme değişmezliği elde edilmektedir.

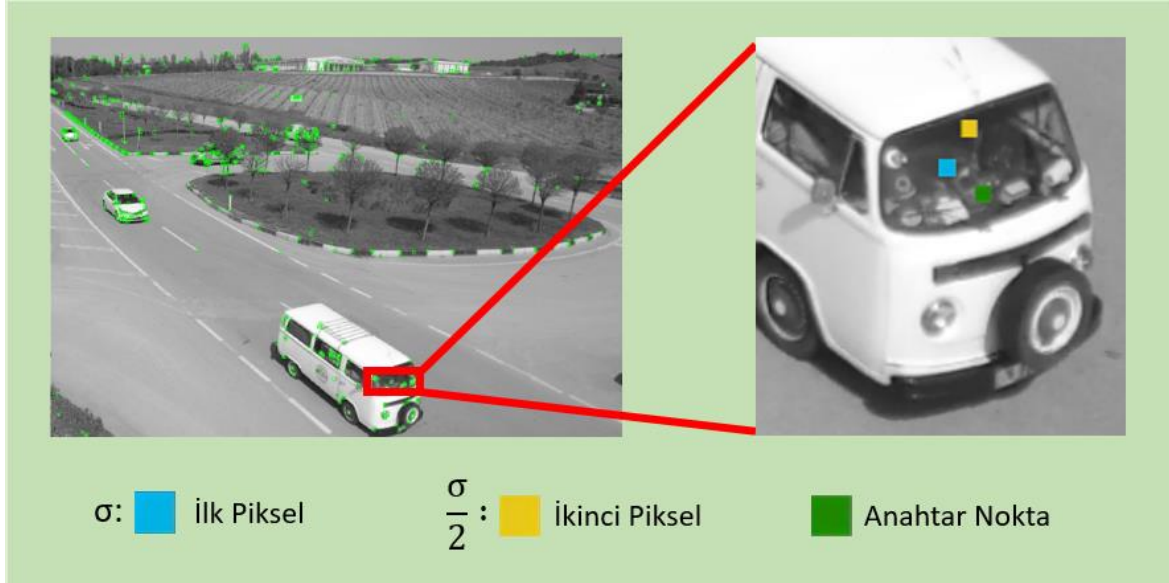
BRIEF (Binary Robust Independent Elementary Features)

BRIEF yöntemi görüntü yamalarında yer alan öznitelikleri alarak onları bir nesneyi temsil edecek şekilde ikili öznitelik vektörüne (aynı zamanda ikili özellik tanıtıcısı) çevirmektedir. BRIEF kullanıldığında her anahtar nokta (keypoint) 128-512 bitten oluşan bir öznitelik vektöründe (Şekil 3.6) temsil edilmektedir.



Şekil 3.6. İkili öznitelik vektörleri

BRIEF, tanımlayıcının yüksek frekanslı gürültüye duyarlı olmasını önlemek için bir Gauss çekirdeği kullanılarak görüntü yumuşatılmaktadır. Ardından BRIEF, anahtar noktanın çevresinde tanımlanmış bir çevrede rastgele bir piksel çifti seçmektedir. Piksel etrafında tanımlanan komşuluk, birkaç piksel genişliği ve yüksekliği olan kare şeklinde bir yama olarak bilinmektedir. Rastgele çiftteki ilk piksel, standart sapma veya sigma yayılımı ile anahtar nokta etrafında ortalananmış bir Gauss dağılımından çizilmektedir. Rastgele çiftteki ikinci piksel, standart sapma veya iki sigma yayılımı ile birinci pikselin etrafında ortalananmış bir Gauss dağılımından çizilmektedir. Bu durumda eğer ilk piksel ikinciden daha parlaksa, karşılık gelen bite '1' değeri atanır. Tam tersi durumda '0' değeri ilgili bite atar. Bu durum Şekil 3.7'de gösterilmektedir.



Şekil 3.7 Anahtar nokta, belirlenen ilk ve ikinci pikseller

Devamında BRIEF rastgele bir piksel çifti seçmekte ve onlara değer atamaktadır. 128 bitlik vektör boyutu için, bir anahtar nokta için bu işlemi 128 kez kısaca tekrarlanmaktadır. BRIEF, bir görüntüdeki her bir anahtar nokta için bunun gibi bir vektör oluşturmaktadır. Bununla birlikte, BRIEF aynı zamanda dönme hareketine göre değişmez değildir, bu nedenle ORB rBRIEF (döndürmeye duyarlı BRIEF) kullanmaktadır. ORB, BRIEF'in hız özelliğini kaybetmeden döndürmeye karşı duyarlılık fonksiyonu eklemektedir.

Yumuşatılmış bir görüntü yaması 'p' olmak üzere, ikili bir test 'τ' şu şekilde tanımlanmaktadır:

$$\tau(p;x,y)=\begin{cases} 1 & :p(x)<p(y) \\ 0 & :p(x)\geq p(y) \end{cases} \quad (3.5)$$

Burada $p(x)$, bir x noktasındaki p 'nin yoğunluğudur. Öznitelik, n tane ikili testin bir vektörü olarak aşağıdaki gibi tanımlanmaktadır:

$$f(n)=\sum_{1 \leq i \leq n} 2^{i-1} \tau(p;x,y) \quad (3.6)$$

BRIEF'in eşleştirme performansı, birkaç dereceden fazla düzlem içi dönüşler için keskin bir şekilde düşmektedir. ORB, kilit noktaların yönelimine göre BRIEF'i yönlendirmek için

bir yöntem önermektedir. Herhangi bir özellik setinin (x_i, y_i) konumundaki n ikili testlerinin ifadesinde $2 \times n$ matrisine ihtiyacımız bulunmaktadır:

$$S = \begin{pmatrix} x_1, x_2, \dots, x_n \\ y_1, y_2, \dots, y_n \end{pmatrix} \quad (3.7)$$

Yama oryantasyonunu θ ve karşılık gelen rotasyon matrisi R_θ olmak üzere S 'nin yönlendirilmiş bir S_θ versiyonunu:

$$S_\theta = R_\theta S \quad (3.8)$$

Böylece, yönlendirilen BRIEF operatörü Eş. 5.9'daki hale gelmektedir:

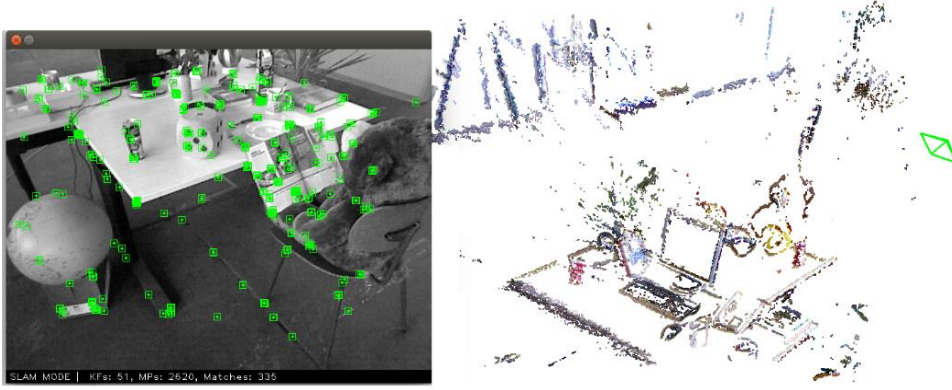
$$g_n(p, \theta) = f_n(p) | (x_i, y_i) \in S_\theta \quad (3.9)$$

Daha sonra açığı $2\pi/30$ (12 derece) artışlarla ayırklaştırır ve önceden hesaplanmış BRIEF modellerinin bir arama tablosunu (lookup table) oluşturmaktadır. Anahtar nokta oryantasyonu θ görünümlemler arasında tutarlı olduğu sürece, doğru nokta kümesi S_θ tanımlayıcıyı hesaplamak için kullanılmaktadır. ORB, rBRIEF algoritmasını aşağıdaki gibi kullanmaktadır:

- 1) Her test, tüm eğitim yamalarına üzerinde çalıştırılmaktadır.
- 2) Testler, T vektörünü oluşturarak ortalama 0,5'ten uzaklıklarına göre sıralanmaktadır.
- 3) Açgözlü (greedy) arama gerçekleştirilmektedir.
- 4) İlk testi R sonuç vektörüne koyularak T 'den çıkarmaktadır.
- 5) T 'den sonraki test alınır ve R 'deki tüm testlerle karşılaştırılmaktadır. Mutlak korelasyonu bir eşikten büyükse, elenmektedir; yoksa R 'ye eklenmektedir.
- 6) R 'de 256 test olana kadar önceki adımı tekrarlamaktadır. 256'dan az varsa eşiği yükseltmekte ve yeniden denemeye etmektedir.

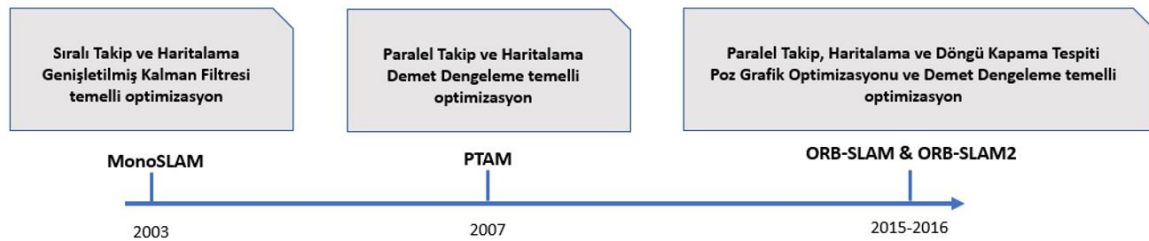
rBRIEF, yönlendirilen BRIEF'e göre varyans ve korelasyonda önemli bir gelişme göstermektedir.

ORB-SLAM Algoritması, grafik işlemci birimi (GPU) ya gereksinim duymadan gerçek zamanlı olarak performans gösterebilmektedir. Güçlü yeniden konumlandırma özelliği ile takip kaybı meydana gelmesi durumunda dahi, daha önce gördüğü özellikleri kullanarak yeniden konumlama gerçekleştirebilmektedir. ORB-SLAM yöntemine ait takip ve haritalama çıktıları Şekil 3.8’de verilmektedir.



Şekil 3.8. Örnek ORB-SLAM takip ve haritalama çıktıları

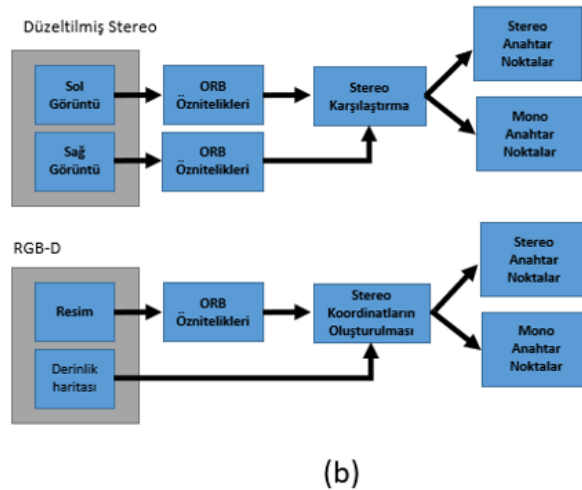
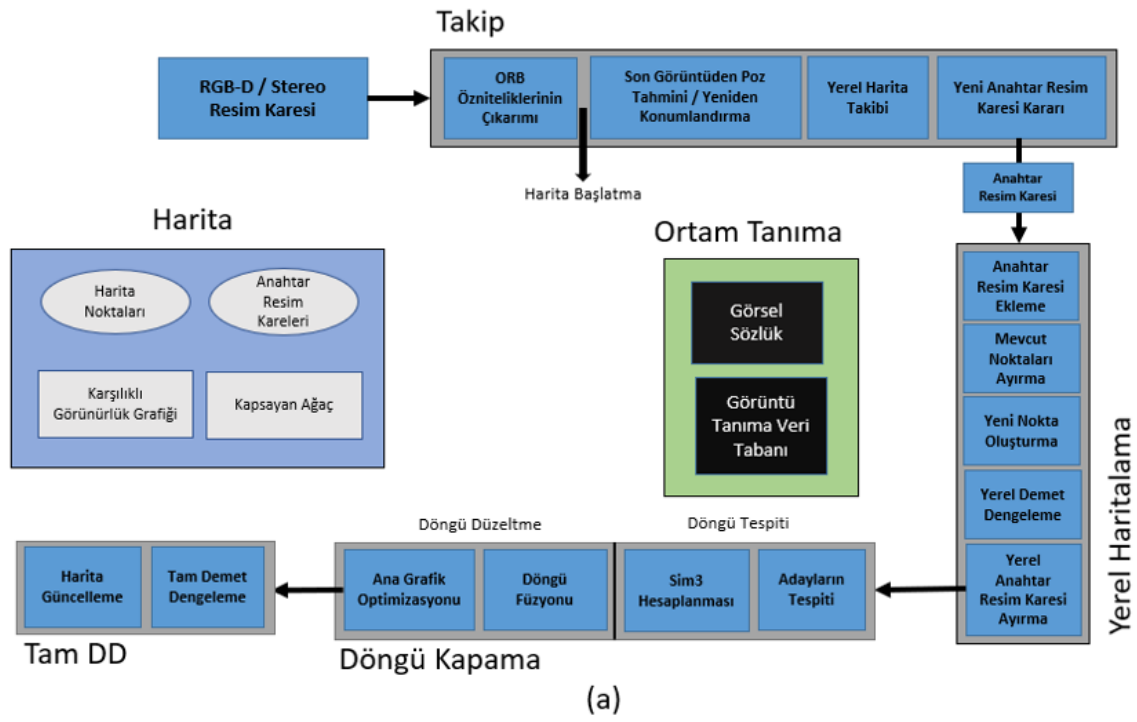
Anahtar resim karesi (keyframe) ve demet dengeleme bazlı yaklaşımların filtre bazlı yaklaşımlardan daha güvenilir olduğu literatürde ortaya konulmuştur (Soltani ve diğerleri, 2012). Bu nedenle ORB-SLAM yönteminde demet dengeleme yaklaşımı kullanılmaktadır. Bütün resim kareleri ve tüm öznitelik noktaları ile işlem yapmak pek tercih edilebilir bir yöntem değildir. Bu nedenle, uygulama bazında optimum öznitelik noktası ile hesaplama kompleksliğini en aza indirmek önem arz etmektedir. Öznitelik temelli vSLAM algoritmalarının kronolojik özeti aşağıda yer alan Şekil 3.9’da (Servières, Renaudin, Dupuis ve Antigny, 2021) verilmiştir.



Şekil 3.9. Öznitelik temelli vSLAM algoritmalarının kronolojik özeti

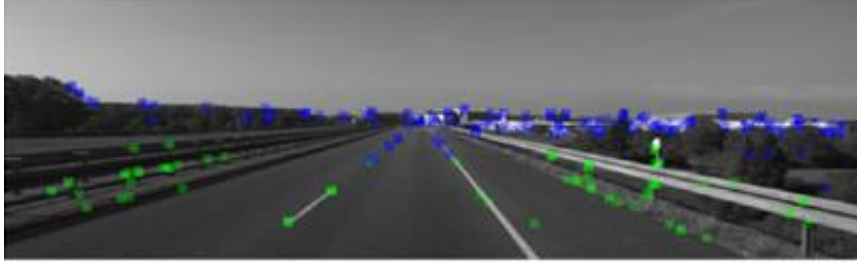
ORB-SLAM 2

ORB-SLAM'in, ORB-SLAM 2 adıyla stereo ve RGB-D kameralar kullanarak aynı yetenekleri sergileyen yeni bir versiyonu da akademik dünyaya sunulmuştur. Monoküler kamera kullanılarak gerçekleştirilen uygulamalarda uzaklık kestirimi kesin olmayan üçgenleme yöntemleriyle yapılmaktadır (Mur-Artal ve Tardós, 2017). Stereo ve RGB-D kameralarda ise uzaklık bilgisi doğrudan görsel sensörden alınmaktadır. Ayrıca monoküler kameradan kaynaklı bazı ölçeklendirme hataları da bu kameralarda engellenebilmektedir. ORB-SLAM2 yöntemine ait akış şeması ve ilgili modüller Şekil 3.10'da verilmiştir.



Şekil 3.10. ORB-SLAM2 akış şeması

Stereo ve RGB-D kameralarda kameranın temel çizgi (baseline) mesafesinin 40 katına kadar olan kısmı yakın, mesafenin bu sınırı aştığı yerler ise uzak bölge olarak adlandırılmaktadır. Üçgenleme ve pozisyon tahmininde yakın noktalar alınarak daha güvenilir sonuçlar elde edilmektedir. Uzak noktalar ise hareketin oryantasyonunun tespitinde (Şekil 3.11) fayda sağlamaktadır (Mur-Artal ve Tardós, 2017).



Şekil 3.11. ORB-SLAM2 algoritmasında yakın (yeşil) ve uzak (mavi) noktalar

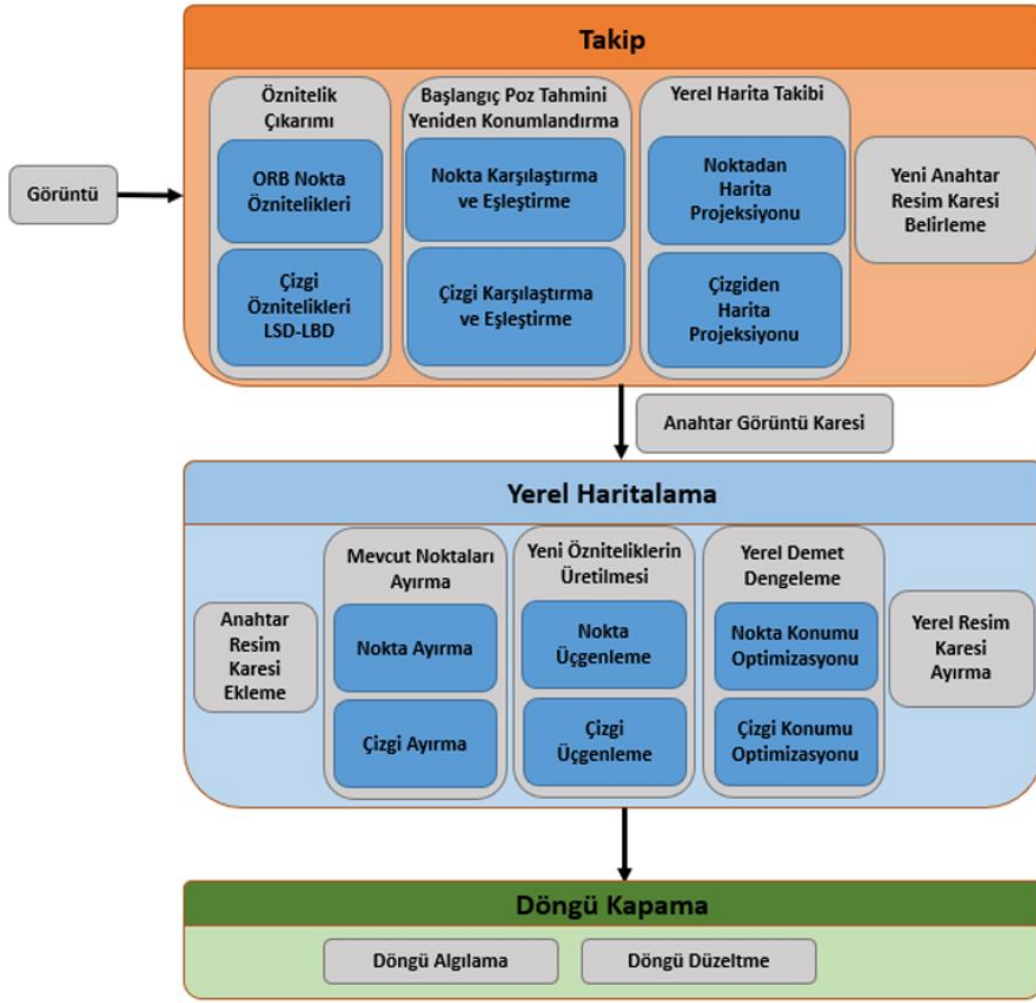
Stereo ve RGB-D kamera kullanmak, daha ilk resim karesini elde edilmesinden itibaren nesnelerin uzaklık bilgisini yüksek hassasiyette sağlamaktadır. Monoküler kamera kullanılan yaklaşımda ise uzaklık konusunda sadece ölçek olarak bize bir fikir sunulmaktadır. ORB-SLAM2 (Mur-Artal ve Tardós, 2017), Stereo LSD SLAM (Engel, Schöps ve Cremers, 2014) ile karşılaştırıldığında doğruluk ve işlem zaman açısından birçok durumda daha yüksek performans göstermektedir. İlgili çıktılar karşılaştırmalı olarak performans ile ilgili başlıkta paylaşılmaktadır. ORB-SLAM2 algoritmasından elde edilen nokta bulutu haritaları ORB-SLAM den (monoküler kamera sürümü) farklı olarak daha yoğun şekilde ve renkli olarak elde edilebilmektedir. Fakat bu renklendirilmiş harita gerçek zamanlı olarak elde edilmesi işlemci açısından zorlayıcı olabilmektedir. ORB-SLAM2 algoritmasının çıktıların başka programlar yardımıyla harita olarak yeniden oluşturulması da mümkün olabilmektedir. Bu konsept, askeri uygulamalarda bir iç mekânın keşfedilmesi ve ardından detaylı haritasının oluşturularak bu harita üzerinden operasyon yürütülmesi şeklindeki bir senaryoya daha uygun görülmektedir. ORB-SLAM2 algoritmasına ait yeniden oluşturulmuş bir harita çıktısı Resim 3.1’de verilmektedir (Mur-Artal ve Tardós, 2017).



Resim 3.1. ORB-SLAM2 örnek harita çıktısı

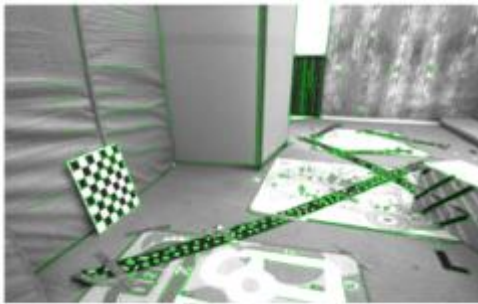
PL – SLAM (Point and Line SLAM)

Monoküler ve stereo sürümleri mevcut olan PL-SLAM hem nokta hem de çizgilerin özelliklerini çıkararak takip eden ve harita oluşturan bir SLAM çeşididir. Bu yöntem ORB SLAM ile kurgu olarak çok benzerdir. Noktalar, ORB özellik tespiti ile bulunurken, çizgiler LSD-LBD yöntemleri ile bulunmaktadır. Bu nedenle ORB-SLAM yöntemine LSD (line segment detector) ve LBD (line band descriptor) yöntemlerinin eklendiği bir vSLAM metodu olarak değerlendirilmektedir (Pumarola, Vakhitov, Agudo, Sanfeliu ve Moreno-Noguer, 2017). PL-SLAM yöntemine ait modüller aşağıda yer alan Şekil 3.12’de gösterilmektedir.



Şekil 3.12. PL-SLAM algoritması modülleri

Harita yapısı ORB-SLAM algoritmasına göre daha yoğundur. ORB-SLAM algoritmasında kullanılan FAST köşe tanıma algoritması sadece noktaları bulduğu için bu algoritmanın haritası da sadece nokta bulutu bir haritadan oluşmaktadır. PL-SLAM algoritmasında bu haritalara çizgi özellikleri de (Resim 3.2) dâhil edilmiştir (Pumarola ve diğerleri, 2017).



Resim 3.2. PL-SLAM algoritması takip ve haritalama çıktıları

Bu yöntemde döngü tamamlama ana grafik (essential graph) yöntemi kullanılarak hızlandırılmıştır. Her anahtar resim karesi (keyframe) sadece 100 öznitelik taşıyacak şekilde, öznitelik sayısı az tutulmuştur. Yerel haritalandırma (local mapping) ve döngü tamamlama paralel iki iş parçası (thread) olarak eş zamanlı olarak gerçekleştirilmektedir. Bu adımlar gerçekleştirildiğinde PL-SLAM'in ORB-SLAM algoritmasından pek çok durumda daha hızlı çalıştığı ve hassas sonuçlar verdiği ilgili yayında belirtilmektedir. İlerleyen kısımlarda bu karşılaştırmalara yer verilmiştir.

CO-SLAM

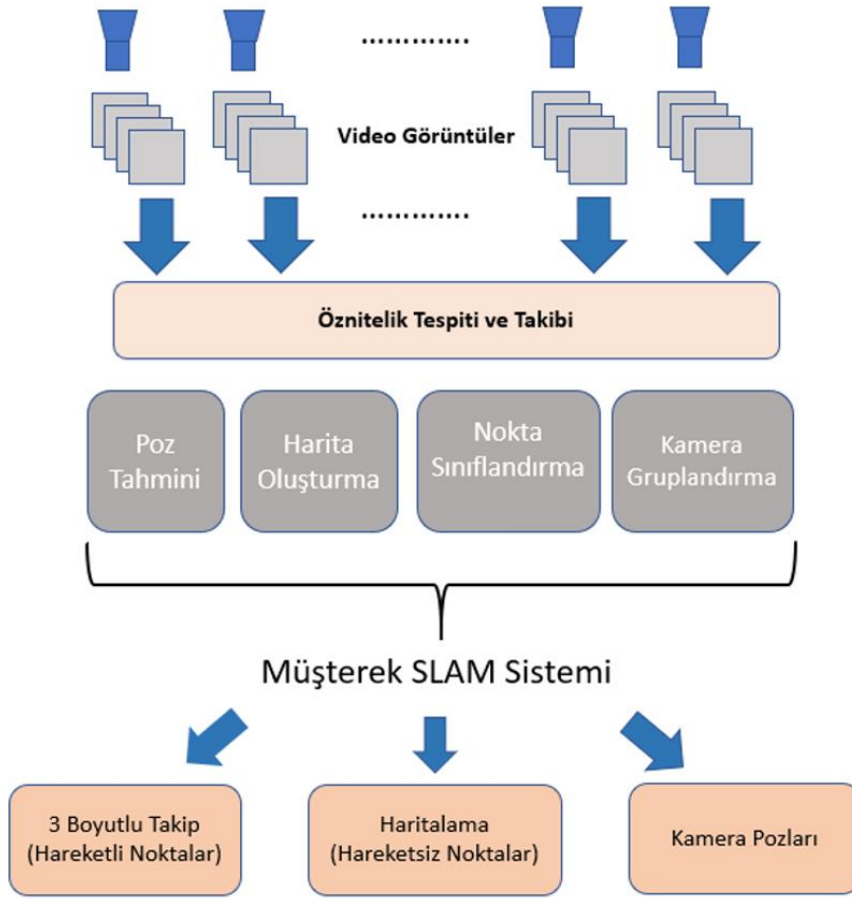
Bu yöntem birden çok kamera ile dinamik ortamda haritalama ve konumlama yapılmasını amaçlamaktadır. Kameralar farklı platformlara birbirinden bağımsız olarak hareket edebilecek şekilde yerleştirilmiştir. Tüm kameralardan elde edilen veriler ile global bir harita ve arka plan ve ön plan verileri elde edilebilmektedir. İlgili yöntem statik ortamlarda ise daha verimli ve başarılı çalışmaktadır. Akademik dünyadaki pek çok algoritma tek kamera ile ve statik ortamda çalışmaktadır. Bununla birlikte gerçek dünyada pek çok hareketli nesnenin bulunduğu bilinmektedir. Bu algoritmalar hareketli objeler ile karşılaşıldığında başarısız olmaktadır. Hareketli nesneleri de değerlendiren SLAM algoritmalarının bir diğer avantajı ise çarpışmadan kaçınma ve yörünge planlamaya imkân tanımlarıdır. Statik kabul edilen dünyada bu işlemleri gerçekleştirmek hatalara neden olmaktadır. Bu algoritma bir platform üzerinde iki veya daha fazla kamera bulunduran (Cremers, 2017) ve stereo kamera kullanan (Nister, 2004b, Paz, 2008) sistemlerden birbirlerinden farklıdır. Bu algoritma beden farklı parçalarına kameralar takmak suretiyle giyilebilir sistemler, her birinde tek kamera bulunan robot takımları gibi uygulamalarda kullanılabilmektedir. CO-SLAM her resim karesi (frame) için statik ve dinamik harita noktalarını tespit ederek harita oluşturmakla birlikte, doğru olmaya tespitleri sonraki frameelerde tespit ederek haritadan silmektedir. Algoritma takımın aynı bölgeye doğru ilerlediği ve görüntüde çakışma (overlap) olduğu durumlarda haritanın kalitesini arttırmaya yönelik çalışırken takım kendi içinde birden fazla gruba da ayrılabilir durumdadır (Zou ve Tan, 2012). Bu çalışmada stereo rig kullanılarak stereo kamera yapısı oluşturulmuştur. Stereo rig yapısı aşağıdaki Resim 3.3'de gösterilmiştir. Rigte yer alan kameraların birbirilerine olan uzaklığın haritalara oldukça etkilediği görülmüştür. Uzaklığın kısa olması harita kalitesini düşürmektedir. Bu problemin çözülmesi için (Paz, 2008, Cremers, 2017) numaralı çalışmalarda kamera mesafeleri değiştirilerek harita kalitesinin arttırılması

amaçlanmıştır. Fakat bu çalışmalardaki katkılar CO-SLAM ile karşılaştırılınca yetersiz kalmaktadır. Ayrıca CO-SLAM algoritmasında her kameranın bağımsız olarak hareket etme kabiliyeti de bulunmaktadır.



Resim 3.3. CO-SLAM yöntemi kullanılan stereo rig yapısı

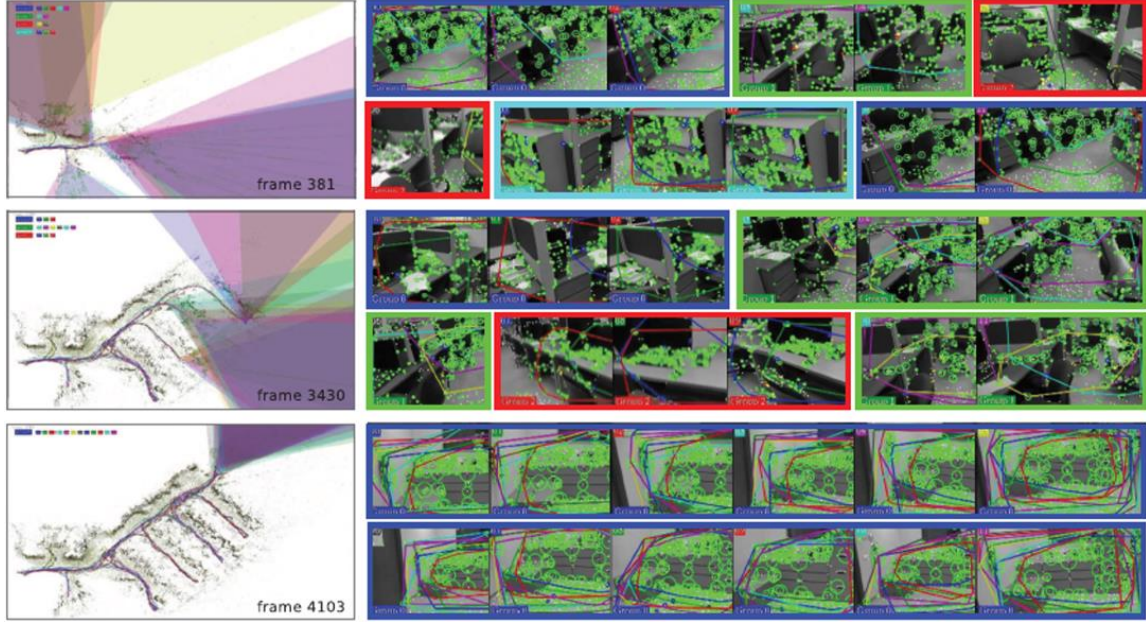
Literatürde dinamik çevrelerde SLAM için daha çok LIDAR ve radar verileri kullanılmaktadır. CO-SLAM bu yöntemlere alternatif (daha düşük hassasiyet ile) bir yöntem sunmaktadır. Böylece pasif, kompakt ve enerji verimli olan kameraların kullanımı mümkün olmaktadır. CO-SLAM metoduna ait sistem mimarisi Şekil 3.13’de verilmiştir.



Şekil 3.13. CO-SLAM sistem mimarisi

Her vSLAM uygulaması gibi CO-SLAM’de de kullanım öncesi kameraların kalibrasyonu gerekmektedir. CO-SLAM öznelik noktalarını (feature points) Kanade-Lucas-Tomasi (KLT) tespit yöntemini kullanarak yakalamaktadır. Hem öznelik noktalarının bulunması hem de bu noktaların takibinde KLT takibi başarılı sonuçlar vermektedir. CO-SLAM bileşenleri kamera pozu tahmini (camera pose estimation), harita oluşturma (map building), nokta sınıflandırması (point classification), ve kamera gruplama (camera grouping) dır. CO-SLAM yönteminde başlangıçta tüm kameraların aynı başlangıç sahnesine bakması gerekmektedir. Başlangıcın ardından kamera pozu tahmini bileşeninde, 3 boyutlu harita noktalarının dağılımından 2 boyutlu görüntü özelliklerinin elde edilmektedir. Intra kamera ve inter kamera poz tahmini gerçekleştirilmektedir. Bu aşamada eş zamanlı olarak yeni elde edilen noktalar harita oluşturma bileşeni tarafından haritaya eklenmektedir. Nokta sınıflandırması aşamasında dinamik ve statik notların ayrımı gerçekleştirilmektedir. Bu aşamada özellikle hareketli objeler tanınmaktadır. Kamera gruplama bileşeni ise aynı gruptaki aynı sahneyi görüntülemekte olan kameraları aynı grupta birleştirilmekte aksi durumlarda ise gruplara dağılma işlemi gerçekleştirilmektedir.

Bu gruplar tekrar birleştirilebilir (grouping and splitting, merging) ve tekrar ayrılabilir durumdadır. Bu bileşenler kullanılarak 12 kameranın kullanıldığı, her kameranın görüş açısının ve ilgili framelerin gösterildiği Resim 3.4 gösterilmektedir (Zou ve Tan, 2012).

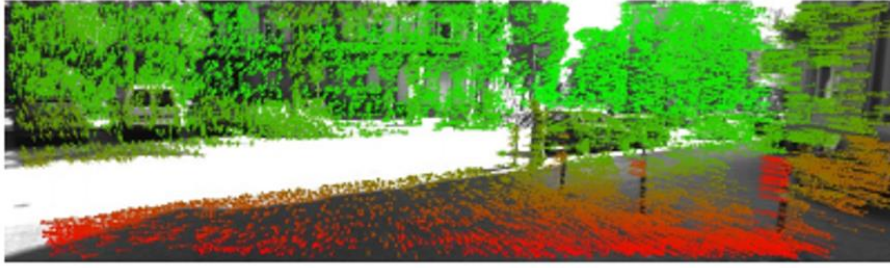


Resim 3.4. CO-SLAM 12 kameralı çalışmaya ait görüntüler

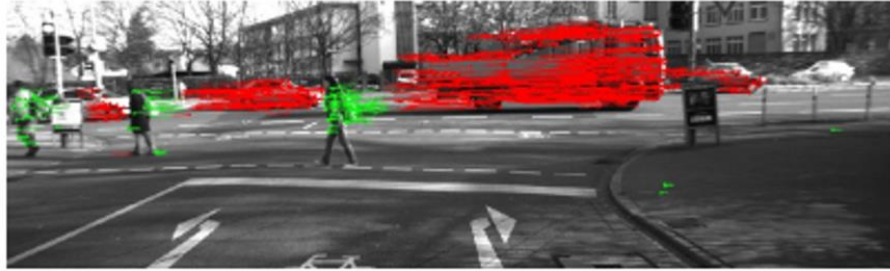
CO-SLAM video verileri üzerinden offline olarak çalışmaktadır. Hesaplamalar merkezi bir bilgisayar üzerinden gerçekleştirilmektedir. Bu sistemin dağıtık olarak uygulanması ve online hale getirilmesi gerekmektedir. Fakat konsept gösterim fazında bu tarz bir çalışmanın gösteriminin gerçekleştirilebileceği değerlendirilmektedir.

SOFT-SLAM (Feature selection and tracking - SLAM)

Stereo kamera tabanlı görsel odometre kullanan bu yöntem, resmin alt parçalara bölünerek, her bir alt parçadan seçilen sınırlı sayıda özelliğin akan resim karelerinde (frame) takibi yöntemiyle çalışmaktadır (Cvišić, Ćesić, Marković ve Petrović, 2018). Dönme ve öteleme hareketlerinin tahmininde iki farklı poz tahmin algoritması kullanması (sırasıyla 5 nokta (five point algorithm) ve 3 nokta (three point algorithm)) diğer algoritmalarından ayrılan yönlerindendir. SOFT-SLAM algoritmasında köşe (corner) ve damla (blob) öznitelikleri tespit edilerek (Şekil 3.14) bu özniteliklerin takibi gerçekleştirilmektedir (Cvišić ve diğerleri, 2018).



(a) Öznitelik Eşleştirme (2 resim karesi arasında, hareketli kamera)



(b) Öznitelik Takibi (5 resim karesi arasında, sabit kamera)

Şekil 3.14. SOFT-SLAM öznitelik eşleme ve takip

Öznitelik takibi yapılırken, her öznitelik bir ID, yaş (age), resimdeki şu anki konum bilgisi (refined current position), öznitelik gücü (feature strenght), ait olduğu sınıf (belonging class) ve başlangıç tanımlayıcısı (initial descriptor) ile ifade edilmektedir.

Her özniteliğin ilgili resim karesindeki (frame) ve bir sonraki resim karesindeki piksel pozisyonu takip edilir ve bu şekilde takip edilen öznitelik bir sonraki karede de yer alıyorsa yaşı (age) bir arttırılmaktadır. Bu özniteliğin yeni pozisyonu da kaydedilerek bir sonraki iterasyonda kullanılmak üzere işleme devam edilmektedir. Bu öznitelik canlı olduğu sürece aynı öznitelik tanımlayıcısı (feature descriptor) ile işleme devam edilmektedir.

Aynı öznitelik tanımlayıcısı kullanmak aynı özniteliğin başlangıçtan itibaren takibini sağlamaktadır. İlgili öznitelikte birden çok fazla değişiklik olduğu durumda bu özniteliğin takibi bırakılır ve yanlış öznitelik tespiti olarak kabul edilmektedir. Bu yöntemin her resim karesinde bir tanımlayıcı tanımlayarak yapılan karşılaştırmalardan çok daha iyi performans gösterdiği görülmüştür. Bu tarz bir öznitelik tanımlayıcısının dezavantajı ise kaymaya (drift) yatkın olmasıdır. Bu nedenle başlangıçta bulunan ve takip edilen noktanın kolaylıkla kaybedilmesine de açıktır.

SOFT-SLAM algoritması KITTI (Geiger, Lenz, Stiller ve Urtasun, 2013) veri kümesi üzerinde çok başarılı sonuçlar elde etmektedir. KITTI'nin başarı listesinde (leaderboard) oldukça üste seviyelerde yer almaktadır. Bu tablo performans karşılaştırma ile ilgili bölümde verilmiştir. Görsel odometreye ek olarak ilgili metotta IMU entegrasyonu da gerçekleştirilebilmektedir. SOFT-SLAM algoritmasının IMU ile desteklenen sürümü, temel sürüme göre kayma hatalarının giderildiği bir sürüm olarak daha yüksek performans göstermektedir (Cvišić ve diğerleri, 2018).

3.3.2. Doğrudan yaklaşımlar

Doğrudan (Direct) metotlar, özniteliksiz (featureless) metotlar olarak da isimlendirilmektedir. Direkt metotlarda giriş olarak alınan resim herhangi bir ön işlem (köşe bulma, çizgi bulma, detector, descriptor vb.) geçirilmemektedir. Öznitelik temelli yöntemler ile doğrudan yaklaşımlar arasındaki bir diğer temel fark ise öznitelik temelli yaklaşımlarda işlem iki ayrı adımda yürütülmektedir. Özniteliklerin çıkarılması ile eşleştirmesi ve bu özniteliklerin takibi ile SLAM probleminin çözülmesi bu iki ayrı adımı oluşturmaktadır. Doğrudan yöntemlerde tek adımda SLAM probleminin çözülmesi gerçekleştirilmektedir. Doğrudan metotlar direkt olarak resim formasyonunu ve sensör ölçümlerini (kameralarda renk bilgisi) kullanır (Cremers, 2017). Ayrıca bu iki yöntemde maliyet fonksiyonları da birbirinden farklıdır. Öznitelik temelli yaklaşımlarda ilgili resmin, alakalı 3 boyuttaki noktalarının geometrik yeniden oluşturma hatası (geometric reprojection error) hesaplanırken, doğrudan yöntemlerde SLAM hesaplamasında maliyeti tüm resimlerdeki renk ve parlaklık kaynaklı fotometrik hataya dayandırılır. Piksellerin gri skaladaki değerini (gray-scale value) kullanarak kamera pozunun bulunmasını sağlar. Resmin tüm piksel bilgisi kullanılarak saniyede yüzlerce resim karesi (frame) işlenebilmektedir. Dolayısıyla işlem hızı bakımından öznitelik çıkarma süresi kaybı bu yöntemlerde gerçekleşmemektedir. Buradaki temel mantık aynı nesneye ait pikselin farklı resim karelerinde de yaklaşık benzer grilik seviyesinde olacağı kabulü ile çalışmaktadır. Öznitelik temelli yaklaşımlara göre bir farkı da fotometrik hatalar minimize edilerek kamera hareketleri tahmin edilmektedir.

DTAM (Dense Tracking and Mapping)

Harita başlangıçta anlık stereo ölçüm ile başlatılmaktadır. Kamera hareketleri sentetik olarak harita üzerinde ortam ile birlikte oluşturulmaktadır. Herhangi bir çıkarım yapmadan stereo kameradan aldığı derinlik verisini kullanılmaktadır. DTAM (Newcombe ve diğerleri, 2011) özniteliklerin kaybolmasından etkilenmemektedir. Fakat her bir pikselin mesafesi stereo kamera ile ölçüldükten sonra global optimizasyon gerçekleştirildiği için hesaplama yükü bu algorithmada fazladır. Yoğun haritaların oluşturulması ve yoğun olarak piksellerin takibinin gerçekleştirilmesi için resim karelerinde yer alan neredeyse tüm verinin kullanılması gerekmektedir. Haritaların oluşturulmasında ters derinlik metodu kullanılmaktadır. Yani en uzaktaki bölgeler sanki sonsuz gibi kabul edilir ve bu değerin matematiksel tersi alındığında '0' değeri elde edilmektedir. Bu yöntemde en uzaktaki bölgeler yok edilir ve yakında bulunan cisimler daha ayrıntılı olarak elde edilmektedir. Aşağıda yer alan şekilde bu durum gösterilmiştir. Resim 3.5'te b şıkında yer alan elde edilen aynı maliyet hacminde (cost volume) regule edilmiş çözümü (regularised solution) yer almaktadır. Bu çözümün gerçek dünya görseline yakın bir şekilde 3 boyutlu model oluşturduğunu görülmektedir (Newcombe ve diğerleri, 2011).

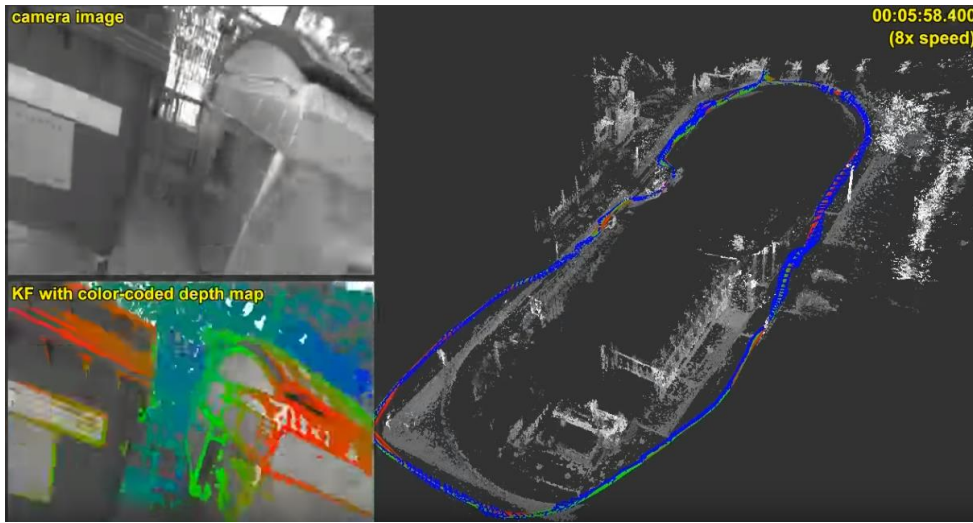


Resim 3.5. DTAM Algoritması (a) uzaklık (b) regüle edilmiş çıktı (c) takip nokta modeli

LSD SLAM (Large Scale Direct Monocular SLAM)

Bu metot doğrudan metotlar arasında en ön plana çıkan metotlardandır (Engel ve diğerleri, 2014). Resim kareleri diğer doğrudan metotlarda olduğu gibi herhangi bir işleme tabi tutulmadan doğrudan alınır ve resimde yer alan her pikselin gri skaladaki değerleri doğrudan kullanılmaktadır. Resim karelerinde yer alan piksellerin hareketine bağlı olarak takip sağlanmaktadır. Anlık duruma göre ORB-SLAM e oranla daha sık noktalardan meydana gelen, Resim 3.6'da gösterilen yoğun bir harita oluşturmaktadır. Bu yöntemin

dezavantajı, dokusu olmayan ortamlarda (textureless areas) sonucun doğru verilememesidir. Aslında bu durum tüm vSLAM algoritmalarının ortak dezavantajı olarak bilinmektedir. Bununla birlikte, LSD-SLAM yönteminde nokta bulutu haritasının yoğun olmasından kaynaklı olarak işlem gücü gereksinimi daha fazla olabilmektedir. Görüntüde yer alan ortam özelliklerinden ortamın derinliği çıkarılmaya çalışılmaktadır. Başlangıçta her piksel için bir rastgele derinlik değeri atanmaktadır. Kamera pozisyonu bu derinliklere göre tespit edilmektedir. İki resim karesi arasındaki hareket takip edilerek pozisyon güncellenmektedir. Derinlik bilgisi de benzer şekilde art arda devam eden görüntülere değerlendirilerek anlık tahminlere göre güncellenmektedir. 7 serbestlik derecesinde (DoF) poz grafiği optimizasyonu kullanılarak geometrik haritalama gerçekleştirilmektedir. Aşağıda LSD SLAM yöntemine ait kamera görüntüsü, görüntü üzerinde farklı renklerde derinlik tahmini ve ortaya çıkan çevresel harita Resim 3.6 gösterilmektedir (Engel ve diğerleri, 2014).



Resim 3.6. LSD SLAM (a) kamera görüntüsü (b) derinlik tahmin haritası (c) çevresel harita

DSO: Direct Sparse Odometry

DSO, doğrudan odometre yöntemleri arasında en bilindik yöntemlerdendir. DSO (Engel ve diğerleri, 2017) yöntemi, kamera pozları, kamera yapıları ve geometri parametreleri (ters derinlik değerleri) dahil olmak üzere ilgili tüm model parametreleri için tam olasılığı ortaklaşa optimize eden tek tam doğrudan yöntemdir. Optimizasyon, (Leutenegger, 2015)'den esinlenerek, eski kamera pozlarının yanı sıra kameranın görüş alanından çıkan

noktaların marjinalleştirildiği kayar bir pencerede gerçekleştirilmektedir. Mevcut yaklaşımların aksine, DSO yöntemi lens zayıflaması, gama düzeltmesi ve bilinen pozlandırma süreleri dahil olmak üzere fotometrik kamera kalibrasyonunun tüm avantajlarından yararlanmaktadır. DSO metodu, geçmiş görüntü karelerinin yer aldığı bir pencerede, fotometrik hatayı, fotometrik kalibrasyona sahip model kullanarak sürekli optimize etmektedir. Diğer doğrudan yöntemlerin aksine DSO’da ilgili tüm parametreler (kamera içsel parametreleri, kamera dışsal parametreleri ve ters derinlik değerleri) belirli bir penceredeki seyrek demet dengelemenin fotometrik eşdeğerini etkin bir şekilde gerçekleştirmektedir. Bununla birlikte diğer doğrudan yöntemlerin, geometrik temsili (3 boyutlu noktaların referans çerçevede ters derinlik olarak ifade edilmesi) korunmaktadır.

Fotometrik hata (E_{pj}), noktanın ters derinliği (dp), kameranın içsel parametrelerine (c) ilgili karelerin pozlarına (T_j ve T_i) ve parlaklık transfer fonksiyonu parametrelerine (a_i, b_i, a_j, b_j) bağlıdır. DSO yönteminin ele aldığı fotometrik hata her görüntü karesi ve her nokta için aşağıdaki gibidir:

$$H_{\text{fotometrik}} = \sum_{i \in F} \sum_{p \in P_i} \sum_{j \in \text{obs}(p)} E_{pj} \quad (3.10)$$

Burada i tüm F çerçeveleri üzerinde, p i çerçevesindeki tüm P_i noktaları üzerinde ve j, p noktasının görünür olduğu tüm çerçeveler $\text{obs}(p)$ üzerinde çalışmaktadır. Anahtar görüntü karelerinin belirlenmesinde ve yeni anahtar görüntü karelerinin eklenmesinde ORB-SLAM2 (Mur-Artal ve Tardós, 2017) yöntemindeki gibi aşağıdaki 3 kriteri kullanmaktadır.

1. Değişiklikleri görüntülemek: Anahtar görüntü karelerinin belirlenmesi için ortamın eski görüntü karelerinde yer almayacak kadar değiştiği gözlemlenmektedir. Bunu sağlamak için, optik akışın ortalama kare (mean square) değeri kaba takip esnasında hesaplanmaktadır. Ortalama kare değerinin Eş 3.10’daki formüle göre bulunmaktadır.

$$f := \left(\frac{1}{n} \sum_{i=1}^n \|p - p'\|^2 \right)^{\frac{1}{2}} \quad (3.11)$$

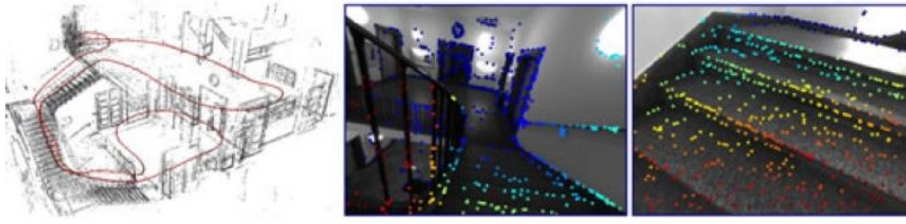
2. Kamera ötelemesi, daha fazla anahtar görüntü karesinin alınmasını gerektiren (f küçük olsa bile) tıkanmalara ve bozulmalara neden olmaktadır. Bu, dönme olmaksızın ortalama akışla ölçülmektedir, p_t ötelenmiş nokta konumu ($R=I_{3 \times 3}$) olmak üzere

$$f := \left(\frac{1}{n} \sum_{i=1}^n \|p - p_t'\|^2 \right)^{\frac{1}{2}} \quad (3.12)$$

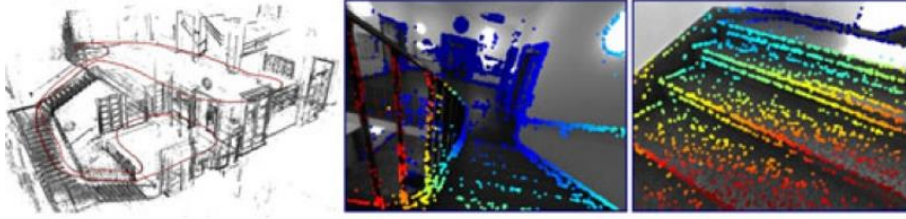
3. Kameranın pozlandırma süresi önemli ölçüde değişirse, yeni bir ana kare çekilmektedir. Bu, iki görüntü karesi arasındaki bağıl parlaklık faktörü (a) ile ölçülmektedir.

$$a := \log(e^{a_j - a_i} t_j t_i^{-1}) \quad (3.13)$$

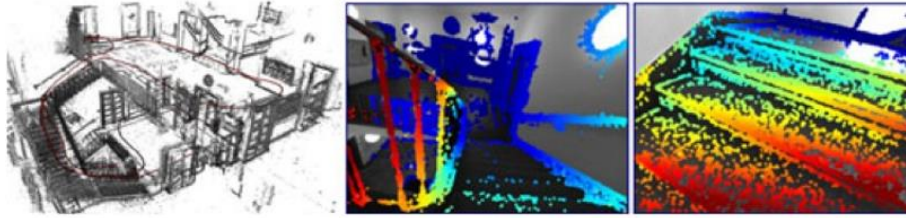
DSO metodunda, harita oluşturulurken kullanılacak nokta sayısı değiştirilebilmektedir. Nokta sayısına göre nokta yoğunluğu değişmektedir. Aşağıdaki Şekil 3.15'te nokta sayısının 500, 2000 ve 10000 nokta olduğu üç farklı duruma ait harita ve aktif noktalar gösterilmiştir (Engel ve diğerleri, 2017).



(a) Nokta Sayısı N=500 adet



(b) Nokta Sayısı N=2000 adet



(c) Nokta Sayısı N=10000 adet

Şekil 3.15. DSO metodu haritası (a) 500 Nokta (b) 2000 Nokta (c) 10000 Nokta

Doğrudan yaklaşım, fotometrik gürültüye karşı daha sağlamdır ve iyi kalibre edilmiş veriler üzerinde üstün doğruluk sağlamaktadır. Bu, on yıldan fazla bir süredir öznitelik temelli yaklaşımların egemenliğinden sonra doğrudan formülasyonların son zamanlarda yeniden canlanmasının ana açıklamalarından biridir.

SVO: Semi Direct Visual Odometry

Bu yöntemdeki yarı doğrudan yaklaşım ile sadece başlatma ve anahtar resim karelerinin belirlenmesinde öznitelikler kullanılmaktadır (Forster, Zhang, Gassner, Werlberger ve Scaramuzza, 2016). Çıkarılan öznitelikler karşılaştırma ve takip aşamalarında kullanılmamaktadır. Bu nedenle öznitelik temelli metotlara göre daha hızlıdır. Hibrit bir metot olarak değerlendirilmektedir. Yukarıda bahsi geçen doğrudan yaklaşım temelli yöntemlerin harita yoğunluğuna sıralaması Şekil 3.16'da özet olarak gösterilmektedir (Servières ve diğerleri, 2021).



Şekil 3.16. Doğrudan metotlara ait harita yoğunluğu karşılaştırması

3.4. Görsel – Ataletsel Navigasyon (VINS-Visual-Inertial Navigation Systems)

Uzun yıllardır, ataletsel navigasyon sistemleri çeşitli platformların, özellikle GPS olmayan su altı, iç mekân, kanyon ya da başka gezegenler gibi ortamlarda, 6 serbestlik derecesinde pozisyon tahmininde kullanılmaktadır. Ataletsel navigasyon sistemleri çoğunlukla, ataletsel ölçüm birimi (inertial measurement unit - IMU) tabanlı olarak ivme değeri ve açısal hızların elde edilmesi temeline dayanmaktadır. Tasarım ve üretim sektörlerindeki gelişmeler sayesinde oldukça ucuz ve hafif mikro elektro mekanik (MEMS) IMU'ların kullanımı günümüzde yaygınlaşmıştır. Bu ürünler kullanılarak mobil cihazlar, hava araçları, arttırılmış ve sanal gerçeklik uygulamaları ve otonom sürüş teknolojilerinde hassaslığı yüksek pozisyon tahminleri gerçekleştirilebilmektedir. Fakat uzun süreli kullanımlarda yüksek hızdaki IMU ölçümleri gürültülerden fazlaca etkilenmektedir. Son teknolojiye sahip ve taktiksel seviye IMU'lar bu gürültülerden etkilenmemektedir fakat bu ürünlerde de pek çok uygulama için oldukça pahalı ve boyut olarak büyük oldukları görülmektedir. Öte yandan kameralar, küçük, hafif, enerji verimli sensörler olarak ataletsel navigasyon sistemlerini hassasiyet yönünden besleyebilmektedir. Bu sayede görsel ataletsel navigasyon sistemleri ortaya çıkmıştır. Görsel ataletsel navigasyon sistemlerinde, IMU ölçümleri ile kameradan alınan resim kareleri kullanılarak poz tahmini gerçekleştirilmektedir (Delmerico ve Scaramuzza, 2018).

3.4.1. Gevşek bağlı – sıkı bağlı sensör füzyonu

Gevşek bağlı sistemler kamera ve IMU verilerinden elde edilen sonuçların sonradan birleştirildiği sistemler iken, sıkı bağlı sistemlerde IMU verileri ile kameradan elde edilen görüntüler birlikte değerlendirilerek sonuç oluşturulmaktadır.

Gevşek bağlı sistemlerde hesaplama maliyeti daha az iken hassasiyet de düşüktür. Sıkı bağlı sistemlerde tek işlem ile iki sensörden gelen veri birleştirildiği için hesaplama maliyeti yüksektir fakat daha hassas sonuçlar elde edilmektedir.

Görsel Odometre ile vSLAM arasında haritanın optimize edilmesi işlemi fark olarak bulunmaktadır. Görsel ataletsel sistemlerde de bu durum geçerlidir. Görsel ataletsel odometre (VIO) herhangi bir harita optimizasyonu yapmazken görsel ataletsel SLAM (viSLAM) yöntemleri haritanın en son haline getirilmesi ve global olarak hataların en aza indirilerek optimum haritanın elde edilmesi aşamasını da içermektedir. Bu nedenle görsel ataletsel SLAM yöntemlerinin sonuçları daha hassas ve ortaya çıkan harita gerçek zamanlı uygulamalarda da kullanılabilir. Bununla birlikte viSLAM algoritmaları, VIO algoritmalarından daha kompleks ve hesaplama maliyetleri daha yüksektir. Genel bilgi olarak yer alması açısından literatürde ön plana çıkan bazı görsel ataletsel SLAM algoritmalarına dair genel bilgiler aşağıdaki gibidir:

MSCKF (Multi-state constraint Kalman Filter): Çok durumlu kısıtlama Kalman filtresi, birçok modern, tescilli VIO sisteminin temelini oluşturmaktadır. Orijinal MSCKF algoritması, belirli bir görüntü özelliğini gözlemleyen tüm kamera pozları arasındaki geometrik değerlendirmeyi gerçekleştirerek 3B özellik konumunun bir tahminini korumaya gerek kalmadan ifade eden bir ölçüm modeli önermektedir. Birkaç çeşitli sürümü bulunmaktadır.

OKVIS (Open Keyframe-based Visual-Inertial SLAM): Keyframe ve optimizasyon temelli kayan pencere (sliding window) yaklaşımı kullanılmaktadır. Ana kare pozlarının kayan bir penceresinde doğrusal olmayan optimizasyon kullanılmaktadır. Maliyet fonksiyonu, görsel işaretler için ağırlıklı yeniden projeksiyon hatalarının ve ağırlıklı atalet hatası terimlerinin bir kombinasyonu ile formüle edilmektedir. Ön uç, özellikleri bulmak için çok ölçekli bir Harris köşe dedektörü kullanır ve ardından çerçeveler arasında veri ilişkilendirmesi yapmak için bunlar üzerinde BRISK tanımlayıcılarını hesaplamaktadır. OKVIS'in monoküler VIO için optimize edilmediği ve stereo konfigürasyon kullanarak üstün performans gösterdiği bilinmektedir.

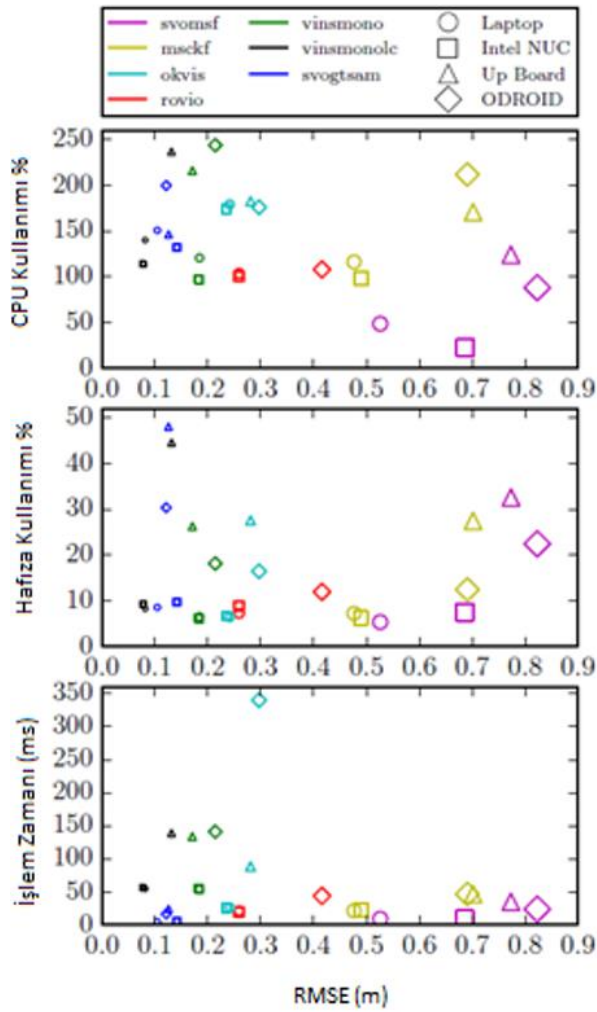
ROVIO (Robust Visual Inertial Odometry): Birkaç yenilik öneren genişletilmiş Kalman Filtresine (EKF) dayalı bir görsel-atalet durum tahmincisidir. 3B konumları robot merkezli yatak vektörleri ve mesafeleri ile parametrelendirilen FAST köşe özelliklerine ek olarak, bu özelliklerin etrafındaki görüntü akışından çok seviyeli yamalar oluşturulmaktadır. Yama özellikleri izlenir, IMU tarafından tahmin edilen harekete dayalı olarak değerlendirilir ve güncelleme adımıyla yenilik terimleri olarak fotometrik hatalar

kullanılır. Genişletilmiş Kalman Filtresi kullanılarak 3D işaretler (landmark) ve resim parçalarındaki (image patch) özellikler takip edilmektedir.

VINS-Mono: OKVIS'e benzer şekilde, gürbüz köşe özelliklerini izleyen, doğrusal olmayan optimizasyon tabanlı bir kayar pencere tahmincisidir. IMU ile Kamera başlangıç aşamasından sonra sıkı bağlı (tightly-coupled) sensör füzyonu yöntemine göre ortak bir çıktı üretmektedir. Döngü tamamlama algoritmasının başarısı ön plana çıkmaktadır.

SVO+MSF (Multi-Sensor Fusion): Genişletilmiş Kalman Filtresine ait genel bir mimaridir. Bu mimari farklı sensörlerden gelen verilerden durum tahminini gerçekleştirmeyi sağlamaktadır. Semi-Direct Visual Odometre ise hesaplama maliyeti düşük, FAST köşe yakalama algoritmasını kullanarak özellikleri çıkaran ve bu özelliklerin çevresindeki fotometrik hatayı minimize eden bir odometre çeşididir. SVO+MSF yönteminde kamera, IMU verisini destekleyici olarak kullanılmaktadır.

SVO+GTSAM: SVO+MSF yaklaşımına back-end de iSAM2 grafik optimizasyonu eklenmiş versiyonudur. Yukarıda bahsi geçen algoritmaların 4 farklı donanım üzerindeki başarımlar ve RMS hataları (metre cinsinden) aşağıdaki Resim 3.7'de verilmiştir (Delmerico ve Scaramuzza, 2018).



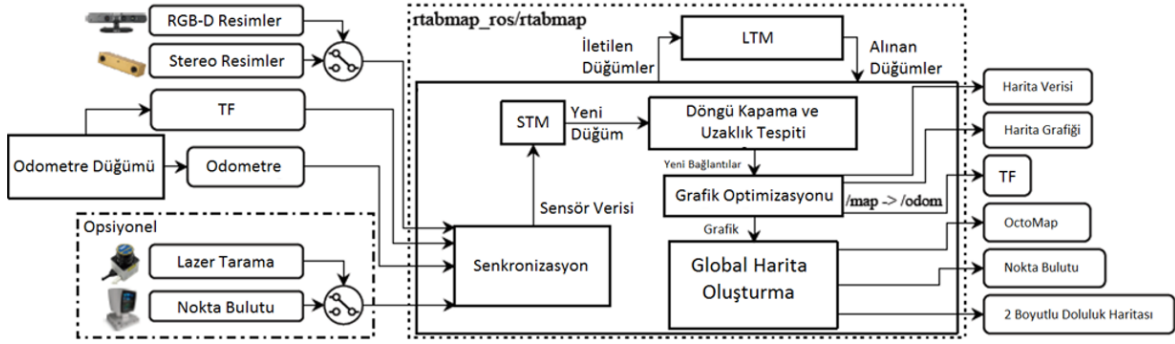
Resim 3.7. Bazı viSLAM algoritmalarına ait performansların karşılaştırması

Görsel ataletsel yöntemlerde kamera-IMU kalibrasyonu ve senkronizasyonu oldukça önemli bir yere sahiptir. Bu konuda (G. Huang, 2019) çalışmasından detaylı bilgi edinmek mümkündür.

RTAB-Map SLAM Kütüphanesi

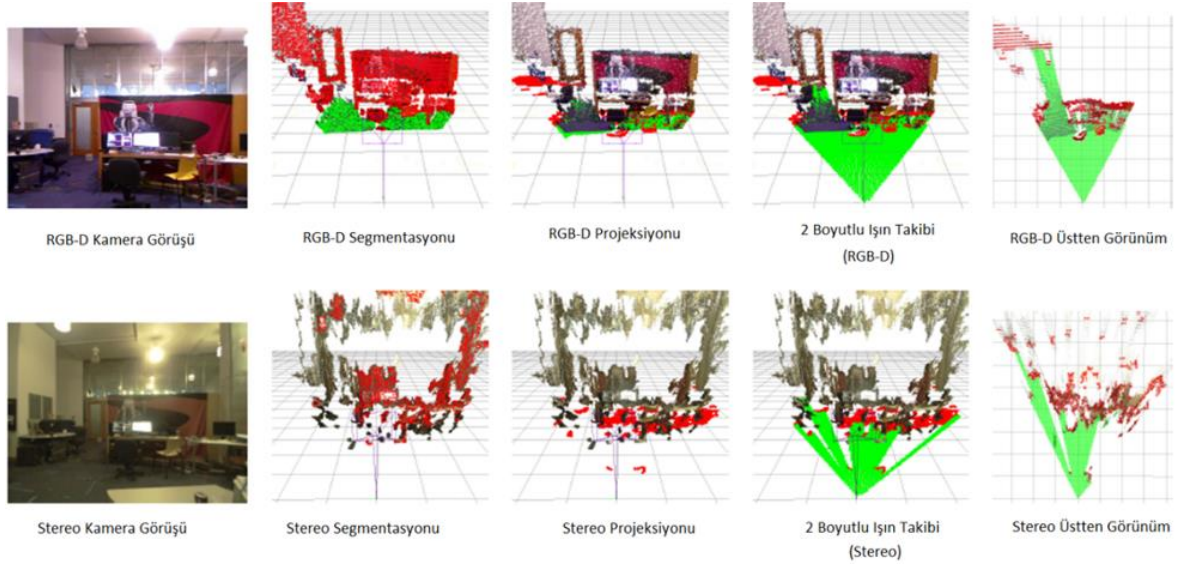
RTAB-Map kütüphanesi 2013 yılından bu yana açık kaynak kodlu olarak geliştirilen ve genişletilen bir kütüphanedir. İlgili kütüphane hem LIDAR tabanlı hem de görsel SLAM algoritmalarını kapsayan şekilde geliştirilmiştir. Görünüm temelli (appearance based) olarak döngü tamamlama algoritması ve uzun zaman kullanımının önünü açan hafıza yönetim eklentileri sayesinde uzun süre gerçek zamanlı olarak kullanılabilir. Ayrıca döngünün tespitinde kullanılan kendileri tarafından geliştirilen graph optimizasyon algoritması (OctoGraph) sayesinde performansın ciddi oranda artırıldığı ifade

edilmektedir. İlgili makalede geniş alanların haritalamasında da teknik sorunlar yaşanmadan başarı ile büyük ve tek bir haritanın oluşturulabildiği yer almaktadır. Aşağıda RTAB-Map ROS modülünün yapısı gösterilmiştir (Labbé ve Michaud, 2019).



Şekil 3.17. RTAB-Map ROS modülünün yapısı

RTAB-Map yapısında yukarıda verilen şekilde de görüleceği gibi, lazer sensör ve nokta bulutu haritası da giriş olarak kullanılabilmektedir. RTAB-Map de geçici hafıza (working memory) ile uzun zamanlı (long-term) hafıza birbirinden ayrılmaktadır. Aşağıdaki şekilde verilen Şekil 3.18’de odometre ile ilgili olan ROS düğümüne (node) tekerlek encoder ve IMU da entegre edilebilmektedir (Labbé ve Michaud, 2019).



Şekil 3.19. RGB-D kamera kullanımında harita görüntüsü

RTAB-MAP kütüphanesi altında FOVIS, Viso2, DVO, OKVIS, ORB-SLAM2, MSCKF ve Google TANGO projesinden oluşan 7 ayrı algoritma entegre durumdadır (Labbé ve Michaud, 2019).

3.5. Performans Karşılaştırmaları

Önceki bölümlerde genel olarak SLAM teknikleri, görsel odometre, görsel SLAM, görsel ataletsel SLAM konularında ön plana çıkan yöntemler incelenmiştir. Bu bölümde daha önce bahsi geçen yöntemlerin birbirlerine göre performans karşılaştırmaları yine akademik dünyada kabul gören çalışmalar üzerinden özetlenmektedir.

Görsel SLAM ve görsel odometre yöntemlerinin performanslarının denendiği bazı çok bilindik veri kümeleri vardır. Bu veri kümeleri aşağıda listelenmiştir. Bu veri kümeleri dışında üniversitelerin ve araştırma gruplarının oluşturduğu veri kümelerine de bazı çalışmalarda rastlanmaktadır. Herhangi bir video kaydı almadan, uygun platform ve fiziki ortam oluşturularak ROS ortamında da mevcut algoritmalar denenerek sonuçlar elde edilebilmektedir (Zhao, Xu, Bu, Jiang ve Han, 2019).

Çizelge 3.4. Mevcut veri kümelerine ait özellikler

Veri kümesi	Yıl	Ortam-Çevre	Kullanılan Sensör Tipi
KITTI (Geiger ve diğerleri, 2013)	2013	Dış Mekân	Çoklu Kamera – IMU
ICL (Handa, Whelan, McDonald ve Davison, 2014)	2014	Simülasyon	RGBD
TUMRGBD (Sturm ve diğerleri, 2012)	2012	İç Mekân	RGBD
TUMMono (Engel, Usenko ve Cremers, 2016)	2016	İç Mekân	Monoküler Kamera
EUROC (Burri, Nikolic, Gohl, Schneider, Rehder, Omari ve Siegwart, 2016)	2016	İç Mekân	Stereo Kamera- IMU
TUMVI (D. Schubert, Goll, Demmel, Usenko, Stückler ve Cremers, 2018)	2018	İç / Dış Mekân	Stereo Kamera- IMU

Görsel Odometre ve SLAM Algoritma Karşılaştırmaları

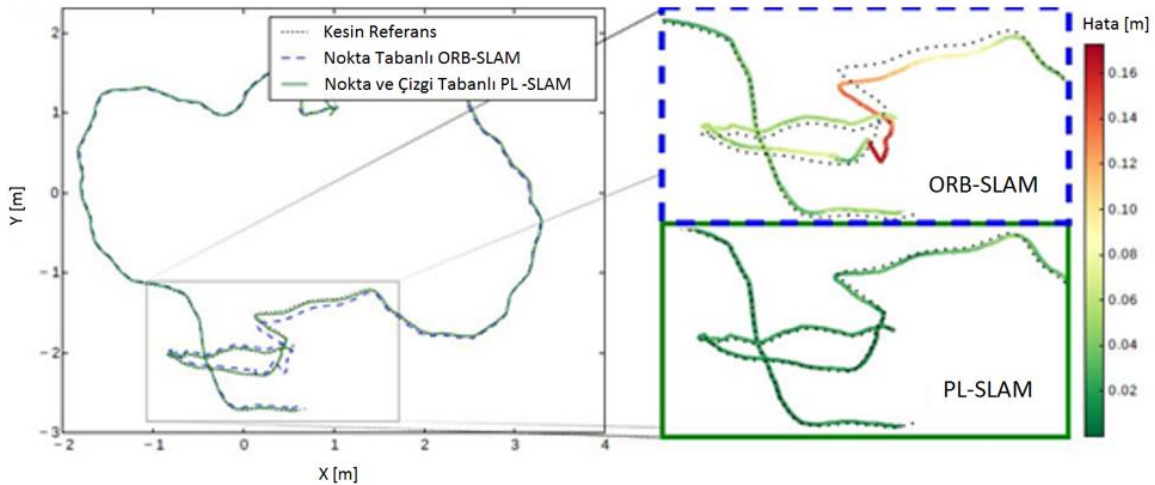
Görsel Odometre ve Görsel SLAM teknikleri hem kendi içlerinde (odometre-odometre, vSLAM-vSLAM) hem de birbirleriyle (odometre-vSLAM) performans açısından karşılaştırılabilmektedir. Görsel Odometre ile görsel SLAM algoritmalarının farkı daha önce de bu raporda belirtildiği üzere global harita optimizasyonu özelliğidir. Dolayısıyla vSLAM metotlarında en son ortaya çıkan global haritadan hataların ortadan kaldırılmış olduğunu ya da en aza indirilmiş olduğunu varsayabilirken, odometre çalışmalarında görsel sensörlerden kaynaklanan birikmiş hatanın ortadan kaldırılması için bir işlem basamağı bulunmaz. Bu nedenle iki yöntemin birbirleriyle karşılaştırılması durumunda ortaya çıkan son haritanın karşılaştırılması pek adil görülmemektedir. Fakat bu yöntemler arasında yörünge tahmini performans karşılaştırması yapılabilir. Aşağıda literatürde yer alan performans ve başarımlar karşılaştırmaları listelenmiştir.

ORB-SLAM, PL-SLAM ve Doğrudan Yaklaşım Kullanan Yöntemler

PL-SLAM algoritması ORB-SLAM algoritmasına çizgi özelliklerinin (line features) eklendiği ve bu özelliklere göre takip ve haritalama yapan bir yöntemdir. Elde edilen harita görüntüleri daha önce ilgili başlıkta verilmiştir. Literatürdeki performans karşılaştırmaları incelendiğinde PL-SLAM yönteminin işlem zamanı olarak ORB-SLAM'den çok daha maliyetli olduğu gösterilmektedir.

Hem haritalama adımlarında hem de ortaya özniteliklerin yakalanması ve takibi aşamalarında ORB-SLAM yönteminin daha hızlı çalıştığı gözlenmektedir (Pumarola ve diğerleri, 2017). Pek çok incelemede ORB-SLAM algoritmasının işlem yükünün düşük olduğu fakat elde edilen haritanın da seyrek nokta haritası şeklinde olduğu konusu işlenmiştir. Ayrıca ORB-SLAM algoritması öznitelik temelli olduğu için dokusuz (featureless) bölgelerde başarısı düşük bir yöntemdir.

PL-SLAM metodunun akademik dünyaya katkısı ise ORB-SLAM'e göre daha iyi bir harita oluşturması ve daha iyi bir yörünge tahmini (trajectory estimation) yapabilmesidir. Bu konudaki sonuçlar PL-SLAM çalışmasında aşağıda Şekil 3.20'de gösterilmiştir.



Şekil 3.20. PL-SLAM / ORB-SLAM yörüngelerinin karşılaştırması

PL-SLAM ile ORB-SLAM2 in KITTI veri kümesindeki yörüngelerindeki RMS hata değerleri hem dönme hareketinde hem de öteleme hareketindeki başarıları da incelenmiştir.

Bu değerlerden KITTI veri kümesindeki kayıt serisinde ORB-SLAM2 algoritmasının daha az hatalı çalıştığı ilgili çalışmada paylaşılmıştır (Pumarola ve diğerleri, 2017).

Birçok vSLAM algoritmasında dokusuz alanlarda (textureless area) performans ciddi oranda düşmektedir. ORB-SLAM gibi nokta tespiti ve takibi yapan algoritmalarda çoğu zaman dokusuz bölgelerde öznitelik bulunamadığı için kopmalar meydana gelebilmektedir. PL-SLAM gibi uygulamalar çizgileri de öznitelik olarak kabul ettiği için ve dokusuz alanlarda dahi yapısal çizgileri (tavan ve taban çizgileri gibi) kullanarak bu alanlarda performans olarak başarımı daha yüksektir (Pumarola ve diğerleri, 2017). Akademik dünyadaki çalışmalar arasında öznitelik temelli yaklaşımlar ile doğrudan yaklaşımların karşılaştırıldığı çalışmalar da mevcuttur. Daha önceki başlıklarda bahsi geçen PL-SLAM, ORB-SLAM, PTAM, LSD-SLAM algoritmaları TUM RGB-D (Sturm, Engelhard, Endres, Burgard ve Cremers, 2012) veri kümesinde denenmiştir. Bu çalışmada RGB-D kamera kullanıldığında bilinen yörüngeye göre en az hata yapılan yöntemin PL-SLAM olduğu, onu PTAM yönteminin takip ettiği ilgili çalışmada gösterilmektedir. Fakat yörünge tahminin karşılaştırmasında işlem zamanı ile ilgili bir bilgi verilmemiştir (Pumarola ve diğerleri, 2017).

Öznitelik temelli metotlarda ORB-SLAM2, doğrudan yaklaşımlarda ise LSD-SLAM metodunun en çok kullanılan metotlardan olduğu bilinmektedir. Bu iki yöntemin karşılaştırıldığı çalışmada KITTI (Geiger ve diğerleri, 2013) veri kümesi üzerinde öteleme (göreceli ve mutlak) ve dönme performansları irdelenmiştir. Bu çalışmaya göre ORB-SLAM2 metodunun stereo kamera ile kullanımında LSD-SLAM'e göre ciddi performans farkı gözlemlenmiştir. ORB-SLAM2 yöntemi 10 ayrı sekans içeren veri kümesinden 7 tanesinde LSD-SLAM metodunu geride bırakmaktadır. Aynı çalışmada EUROC (Burri ve diğerleri, 2016) veri kümesinde de ORB-SLAM2 metodu 3 ayrı sekansın hepsinde LSD-SLAM metodundan yüksek performans sağladığı gösterilmektedir. Bu durum ilgili çalışmada detaylarıyla gösterilmiştir (Mur-Artal ve Tardós, 2017).

ORB-SLAM ve RTAB-MAP Karşılaştırması

ORB-SLAM ile RTAB-MAP algoritmasının aynı senaryoda karşılaştırıldığı çalışmada (Ragot, Khemmar, Pokala, Rossi ve Ertaud, 2019) gidilen toplam mesafe ve ORB-SLAM ve RTAB-MAP'in yörünge tahminleri karşılaştırılmıştır. ORB-SLAM2'nin stereo ve

RGB-D kamera kullandığı durumların RTAB-MAP yönteminde RGB-D kamera kullanımına göre alınan mesafe tahmininde daha iyi performans gösterdiği sonucuna ulaşılmıştır. Karşılaştırmada temel kabul edilen kesin referans (ground truth) VICON kamera sistemi kullanılarak çıkarılmıştır. Bu sonuçlar ilgili çalışmada (Ragot ve diğerleri, 2019) gösterilmiştir.

KITTI veri kümesi üzerinde gerçekleştirilen değerlendirmeler sonucunda bir başarı listesi ortaya çıkmış durumdadır. Literatür taramasında yer alan bazı algoritmaların başarı sırası ilgili çalışmada (Labbé ve Michaud, 2019) verilmiştir. Bu başarı sırası dönme ve öteleme hareketlerinde meydana gelen hatalara göre belirlenmiştir (Labbé ve Michaud, 2019).

Doğrudan Yaklaşım Kullanan Yöntemlerin Kendi Arasında Karşılaştırmaları

Doğrudan yaklaşım kullanan yöntemlerden olan DSO ve SVO odometre yöntemleri ile LSD-SLAM algoritması özellikleri bakımından karşılaştırılmıştır. Bu karşılaştırmada beklendiği üzere harita optimizasyonu kapsamına giren yeniden konumlama (relocalization) ve döngü kapama (loop closure) aşamaları vSLAM aşamaları DSO da yer almazken, SVO da ise sadece yeniden konumlama (relocalization) aşaması yer almaktadır. Bunun dışındaki karşılaştırmalarda ise iki yöntemin LSD-SLAM'den performanslarının daha yüksek olduğu gösterilmiştir (Y. Chen, Zhou, Lv ve Deveerasetty, 2018). Genel olarak SLAM yöntemlerinde harita ile ilgili performans kaygıları, araştırmacıları daha kullanılabilir haritalar elde etmeye yönlendirdiği için yörünge tahmininin verimliliğinin azaldığı gözlenmektedir. Bu azalama araştırmacıların hesaplama maliyetini düşürmek için daha basit özellik tanıma ve takip algoritmalarına yöneldiklerini göstermektedir (Y. Chen ve diğerleri, 2018).

3.6. Sonuç

Bu bölümde genel olarak SLAM, Görsel Odometre (VO), Görsel SLAM (vSLAM), Görsel- Ataletsel Navigasyon Sistemleri (VINS) konularında literatürde ön plana çıkan yöntemlerden incelenmiştir.

Görsel Odometre, platformların almış olduğu mesafeyi ve bu mesafeyi alırken çizdikleri rotayı elde etmeyi planlayan ve bu rota üzerinden pozisyon tahmini yapmayı sağlayan,

kamera temelli bir odometre çeşididir. Bu yöntemde haritalama işlemi gerçekleştirilse bile, haritanın global olarak optimizasyonu yapılmadığı için harita kullanılabilir bir çıktı olarak elde edilememektedir.

Görsel SLAM metotları ise içerisinde hem odometre hem de global harita optimizasyonu içeren yöntemlerdir. Bu yöntemlerde eklemeli olarak artan hatalar, döngü tamamlama yöntemleri ile ortadan kaldırılmaya çalışılmaktadır. Görsel odometre yöntemleri SLAM algoritmalarında yörünge çıkarımında kullanılmaktadır.

Görsel Ataletsel Navigasyon Sistemleri kamera ile IMU'ların ortak kullanıldığı sistemlerdir. Burada her iki sensör tipinden gelen veriler ayrı ayrı değerlendirilip elde edilen iki sonuç ortaklanırsa gevşek bağlı, her iki sensörden gelen veriler birlikte değerlendirilip ortak sonuç oluşturulursa sıkı bağlı sistemler olarak adlandırılmaktadır. Görsel ataletsel SLAM yöntemleri, IMU'nun uzun süreli görevlerdeki biriken hata seviyesinden ve uzun süren kalibrasyon zamanından olumsuz şekilde etkilenmektedir. Performans karşılaştırması başlığında, bahsi geçen yöntemlere ait performans karşılaştırmaları paylaşılmıştır. Literatürde daha çok algoritmaların doğruluk oranlarıyla ilgili çalışmalar yapıldığı gözlenmiş ve bu çalışmalar ile ilgili sonuçlar kullanılmıştır. Görsel ataletsel navigasyon sistemleri için ise açık kaynak kodlu yöntemlerin performans karşılaştırması ilgili konu başlığı altında paylaşılmıştır.

Performans karşılaştırmaları incelendiğinde sadece monolitik kamera kullanarak elde edilen haritaların ve yörünge (trajectory) tahminlerinde çok fazla kaymaya (drift) maruz kaldığı görülmektedir. Stereo ve RGB-D kameraların kullanıldığı yöntemlerin bu kaymalar açısından daha hata oranı düşük yöntemler olduğu görülmektedir. Görsel ataletsel yöntemler ile görsel yöntemlerin RTAB-MAP altyapısında gerçekleştirilen karşılaştırmalarında bu yöntemlerin stereo ve RGB-D kamera kullanan görsel yöntemlere göre pek fazla üstünlükleri gözlemlenmemiştir.

4. HAREKET BULANIKLIĞI

İster insansız sistemlerin üzerinde konumlandırılınsın ister insan tarafından çekim yapılıyor olsun standart RGB kameralar kullanıldığında meydana gelebilecek görüntü bozulmalarından bir tanesi de hareket bulanıklığıdır. Hareket bulanıklığı, pozlandırma süresi boyunca kamera ile çekilen nesne arasındaki göreceli hareketten kaynaklanan bulanıklıktır.

Kullanılan İHA platformlarının boyutları küçüldükçe ve insan unsuru ile birlikte hatta insan unsuru yerine kullanımları arttıkça, platformda harcanan güç seviyesini azaltmak ve İHA kısımlarına karşı maliyeti en düşük seviyede tutmak önemli bir hale gelmiştir. Bu nedenle hem elektronik / mekanik tasarımların, hem de sensör boyutlarının azaltılması ve düşük güç tüketimli, hafif ve ucuz alternatiflerin kullanılması önem arz etmektedir. Mikro / nano insansız araçların kullanımının artması ile tek kullanımlık platformlar daha fazla tercih edilecektir. Bu platformlar belirli bir bölgeye toplu şekilde serpilerek o bölgenin tamamında otonomi seviyesi yüksek şekilde operasyon gerçekleştirecek ve birçok platformun tekrar kullanımı/ geri dönemi beklenmeyecektir. Bu durum ayrıca maliyet faktörünü en aza indirmenin önemini göstermektedir.

Günümüzde pek çok görev, akıllı telefonlarda da yer alan, piyasada pek çok muadili bulunan standart monoküler RGB kameralar kullanılarak yerine getirilebilir. Bu kameralar daha kapsamlı ve yüksek güçlü yüksek kaliteli kameralardan ucuzdurlar ve bekleneneği üzere elde edilen görüntüler daha düşük kalitelidir. Ayrıca bu kameralar daha az hassastırlar ve özellikle dış etkilere daha çok etkilenirler. Standart RGB kameraların hızlı platformlar üzerinde kullanımı, hareket kaynaklı olarak bulanıklığına sebep olabilir. Bu kısımda monoküler RGB kameraların maruz kalabileceği görüntü bulanıklıkları ele alınmaktadır.

4.1. Hareket Bulanıklığı Nedenleri

Standart RGB monoküler kameralar kullanıldığında, 2 farklı bulanıklık meydana gelebilmektedir. Bunlar optik bulanıklık ve hareket bulanıklığıdır. Optik bulanıklık, daha çok kamera ile pozlandırma süresinde merceğin uygun mesafede

konumlandırılmamasından meydana gelmektedir. Çoğunlukla yakınlaştırma ve uzaklaştırma işlemi esnasında pozlandırma gerçekleştirildiğinde meydana oluşmaktadır. Hareket bulanıklığı ise ya çekim yapan kameranın ya da görüntüsü alınan nesnenin hareketi nedeniyle görüntü karesinin bulanıklaşması olarak ortaya çıkmaktadır. Optik bulanıklık genellikle tüm görüntü karesini etkilerken, hareket bulanıklığı kameranın hareketiyle ortaya çıkan bulanıklık tipinde tüm görüntü karesini, sadece nesnenin hareket ettiği bulanıklık tipinde ise görüntünün bir kısmını etkilemektedir. Hareket bulanıklığının şiddetini etkileyen üç önemli faktör vardır. Bunlar nesne ya da platformun hareket hızı, pozlandırma süresi ve kamera ile nesne arasındaki mesafedir.

Kadrajda yer alan nesnenin ya da kameranın hızı arttıkça, hareket bulanıklığının seviyesi artmaktadır. Daha uzakta yer alan nesneler aynı hızda olsalar bile, daha yakında yer alan nesnelerden daha az bulanıklığa neden olmaktadır (Loughborough University, 2011).

Pozlandırma süresi arttıkça, görüntüde hareket bulanıklığının oluşması ihtimali artmaktadır. Fakat bu sürenin bütünüyle 0 milisaniye gibi bir değere yakınsaması da görüntünün kalitesini etkilemektedir. Pozlandırma süresinin parlaklık üzerine doğrudan bir etkisi vardır. Uzun pozlandırma zamanı daha parlak resim elde etmemizi sağlarken, pozlandırma sırasında hareket oluşabileceği için hareket bulanıklığı kaynağı olarak görülmektedir. Kısa pozlandırma süresi ise daha koyu resimler elde edilmesine neden olmaktadır. Bu nedenle pozlandırma süresinde optimum bir değer belirlenmesi önemlidir.

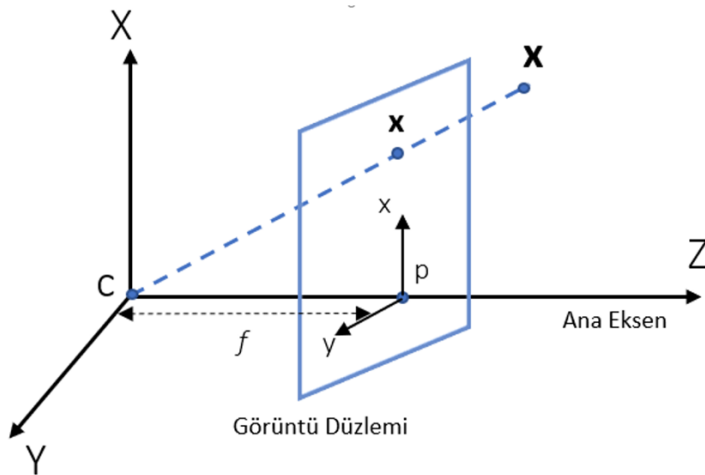
Pozlandırma süresinin resimde meydana gelen kayma miktarına etkisini gerçek bir örnek üzerinde incelemek için şu örnek açıklayıcı olacaktır. İç mekânlarda görev yapmak üzere tasarlanan insansız hava araçlarından olan bir platformun açıklanan maksimum hızı yaklaşık 18 km/s olup, pozlandırma zamanı 1/400 s olan bir kamera kullanıldığında platform pozlandırma sonlandırılana kadar görüntü karesinde yaklaşık 13mm kayma görülebilmektedir.

İnsansız hava araçlarındaki kameralarda platformun hareketi çeşitli hareket bulanıklığı tiplerine neden olmaktadır. Platform hareketinde ortaya çıkan hareket bulanıklığı, tüm görüntü karesini etkileyen ve görüntü işleme uygulamalarında görüntünün kullanımını daha zorlaştıran bir bulanıklıktır. Platformun ileri yönde veya geniş şekilde dönüş yönündeki hareketlerinde tek yönlü (directional) hareket bulanıklığı ortaya çıkmaktadır.

Platform yuvarlanma (roll) eksenini boyunca hareket ederek dönüşler gerçekleştirirken meydana gelen hareket ileri yönde oluşan hareket bulanıklığına dikey ek bir hareket bulanıklığı daha meydana gelmektedir. Bu durumda ortaya çıkan hareket bulanıklığı köşegen (diagonal) vektör şeklinde ortaya çıkmaktadır. Platformun hareketi sapma ekseninde olursa (yaw axis) daha kompleks bir bulanıklık olan dairesel (circular) hareket bulanıklığı oluşmaktadır (Sieberth, Wackrow ve Chandler, 2016).

Kamera Modeli

İnsansız hava aracı platformlarında görüntüleme sistemleri, platform üzerinde yer alan sensör ile çevreye ve platformun kendisine ait bilgilerin çıkarılması amacıyla kullanılmaktadır. Görüntüleme sensörleri pasif sensörler oldukları için güç tüketimi, maliyet ve boyutlar bakımından oldukça avantajlıdır. Diğer sensörlerin olumsuz etkilendiği bazı koşullarda da kimi zaman kullanılabilirler. Fakat görüntüleme sensörleri kullanılırken bir matematiksel model oluşturulmaktadır. Bu model sayesinde 3 boyutlu dünyada tanımlanan objeler ve öznitelikler 2 boyutlu dünyaya aktarılabilir. Uzaydaki 3 boyutlu noktaların bir düzlem üzerine merkezi izdüşümünü koordinat sisteminin merkezi kabul edilerek $Z = f$ koordinatında yer alan düzlem görüntü düzlemi ya da odak düzlemi olarak isimlendirilmektedir. İğne deliği kamera modelinde (pinhole camera model) uzayda $X = (X,Y,Z)^T$ ile ifade edilen noktanın görüntü düzlemi ile projeksiyon merkezi doğrultusunda kesiştiği yer x harfi ile gösterilmektedir.



Şekil 4.1. İğne deliği kamera geometrisi

Yukarıda Şekil 4.1’de iğne deliği kamera geometrisi verilmiştir. Şekilde C kamera merkezini, p ana eksen üzerindeki ana noktayı göstermektedir. Kamera koordinat sisteminin merkezine yerleştirilmiş, görüntü düzlemi ise kameranın önünde yer almaktadır. Oluşan üçgene benzer üçgen kuralları diğer noktalar $(X,Y,Z)^T$ için de uygulandığında $\left(f\frac{X}{Z}, f\frac{Y}{Z}, f\right)^T$ noktaları ile görüntü düzleminde gösterilebilir. Bu durumda gerçek dünyadan görüntü düzlemine merkezi izdüşüm dönüşümü aşağıdaki gibi olmaktadır.

$$(X,Y,Z) \mapsto \left(f\frac{X}{Z}, f\frac{Y}{Z}\right)^T \quad (4.1)$$

Bu dönüşüm, 3 boyutlu bir Öklid uzayından (R_3) bir 2 boyutlu Öklid uzayına (R_2) çevrim işlemidir. Bu projeksiyonda yer alan C merkezine kamera merkezi ya da optik merkez adı verilmektedir. Bu merkezden çıkarak görüntü düzlemini kesen doğruya ana eksen (principal axis) ya da ana ışın (principal ray) ismi verilmektedir. Ana eksenin görüntü düzlemi ile kesiştiği noktaya ana nokta (principal point) ismi verilirken ve ‘p’ harfi ile gösterilmektedir. Kamera merkezi boyunca yer alan ve görüntü düzlemine paralel olan düzleme kameranın ana düzlemi denir (Hartley, 2003).

4.2. Homojen Koordinatları Kullanan Merkezi İzdüşüm

Gerçek dünya koordinatları ile görüntü noktaları eğer homojen vektörler olarak gösterilirse, merkezi iz düşüm homojen koordinatlar arasında basitçe lineer bir şekilde ifade edilebilmektedir. Bu durumda, Eş. 4.1’deki eşitlik matris çarpımı olarak Eş. 4.2’deki gibi yazılabilmektedir (Carrillo ve diğerleri, 2012):

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4.2)$$

Yukarıdaki denklemde yer alan 3 x 4 boyutlarındaki matris köşegen matrisi $\text{diag}(f, f, 1)$ ve 3x3 boyutlarında bloğa bölünmüş $[I | 0]$ bir birim vektöre ‘0’lardan oluşan bir kolon vektörü eklenilmesi ile çarpım ifadesi olarak ‘ $\text{diag}(f, f, 1) [I | 0]$ ’ şeklinde ifade edilebilmektedir. Burada X matris notasyonu dünyadaki gerçek koordinatlar olmak üzere homojen 4 boyutlu vektör $(X,Y,Z,1)^T$ ile ifade edilmektedir. Ayrıca x görüntü noktası

homojen 3 boyutlu bir vektör olarak ve P ise 3x4 boyutlarında homojen bir projeksiyon matrisi olarak gösterilmektedir. Eş. 4.2, daha sade bir şekilde aşağıdaki gibi (Eş. 4.3) yazılabilmektedir.

$$x=PX \quad (4.3)$$

Bu durumda merkezi projeksiyonun iğne deliği kamera modeli için kamera matrisi de aşağıdaki şekilde tanımlanabilmektedir (Hartley, 2003).

$$P = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (4.4)$$

Ana Nokta'ya ait Ofset Denkleminin gösterildiği Eş. 4.1 kameranın koordinat sisteminin başlangıç noktasında olduğunu kabul etmekte ve görüntü düzlemi ana noktaya yerleştirilmektedir. Pratikte bu durum öyle olmayabilir, daha genel bir ifade ile denklem aşağıdaki şekilde ifade edilebilmektedir.

$$(X,Y,Z)^T \mapsto \left(f \frac{X}{Z} + p_x, f \frac{Y}{Z} + p_y\right)^T \quad (4.5)$$

Bu denklemde p_x ve p_y ifadeleri ana noktanın koordinatlarını göstermektedir. Bu denklem homojen koordinatlarda yazılırsa;

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} fX+Zp_x \\ fY+Zp_y \\ Z \end{bmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4.6)$$

Buradan K matrisi aşağıdaki gibi seçilirse;

$$K = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \quad (4.7)$$

En sade haliyle;

$$K = [I|0]q \quad (4.8)$$

şeklinde ifade edilmektedir. Eş. 4.8'deki ifadede K matrisi kamera kalibrasyon matrisi olarak adlandırılmaktadır. Kamera koordinat sisteminin merkezine yerleştirilmiş olduğunun varsayıldığını ifade etmek için, $(X,Y,Z,1)^T$ ifadesi q harfi ile ayrıca gösterilmektedir. Kameranın ana eksen, koordinat sisteminde Z eksenindedir ve q noktası bu koordinatları ifade etmektedir. Bu koordinat sistemine kamera koordinat çerçevesi (camera coordinate frame) adı verilmektedir.

Dışsal Parametreler, kamera koordinat çerçeveleri dışında, başka bir koordinat sistemi çerçevesinde (dünya koordinat çerçevesi- world coordinate frame) yer alan gerçek dünyaya ait noktalardır. Dünya koordinat çerçevesi ile kamera koordinat çerçevesi arasında öteleme ve dönme hareketi ile açıklanabilecek bir ilişki mevcuttur. $\tilde{X} \in R^{3 \times 1}$ gerçek dünya koordinat çerçevesindeki homojen olmayan bir noktayı gösterirken, $\tilde{q}$ bu noktanın kamera koordinat çerçevesinde aynı noktayı gösteriyor olsun. Bu durumda $\tilde{q} = R(\tilde{X} - \tilde{C})$ eşitliği mümkün olmaktadır. Burada $\tilde{C}$ kameranin dünya koordinat çerçevesindeki merkezi ve $R \in R^{3 \times 3}$ rotasyon matrisini ifade etmektedir. Bu eşitlik homojen koordinatlarda aşağıdaki gibi yazılabilmektedir.

$$q = \begin{bmatrix} R & -R\tilde{C} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} R & -R\tilde{C} \\ 0 & 1 \end{bmatrix} X \quad (4.9)$$

ve;

$$x = KR[I|\tilde{C}]X \quad (4.10)$$

eşitliği ortaya çıkmaktadır. Buradan, iğne deliği kamera modelinin 9 serbestlik derecesi olduğu çıkarılabilmektedir. K'da üç, R'de üç ve $\tilde{C}$ 'de üç olmak üzere toplamda 9 serbestlik derecesi vardır. Bu eşitlikteki R ve $\tilde{C}$ kamera oryantasyonu ve pozisyonunu gerçek dünyadaki koordinatlar ile bağlayan kameranin 'dışsal parametreleri' (extrinsic parameters) ve 'dış oryantasyonu' (exterior orientation) dur. Gerçek dünyadaki koordinatlardan, görüntüye çevrim yapıldığında $\tilde{q} = (R\tilde{X} + t)$ ifadesi yazılarak;

$$P = K[R|t] \quad (4.11)$$

Kamera matrisi ifadesi bulunur. Eş. 4.10 kullanılarak;

$$t = -R\tilde{C} \quad (4.12)$$

ifadesi elde edilir.

İçsel Özellikler, iğne deliği kamera modeli, görüntü koordinatlarının ekseninde her iki yönde eşit ölçeklerde türetildiğini varsaymaktadır. Şarj bağlantılı cihaz (Charge-coupled device-CCD) kameraları kare olmayan piksellere sahip olabilir. Eğer görüntü koordinatları pikseller ile ölçülüyorsa, kare olmayan pikseller eksenin her iki yönünde eşit ölçek yapısını bozmaktadır. Özellikle birim uzaklık başına piksel sayısı görüntü koordinatlarında x ve y yönlerinde sırasıyla m_x ve m_y ise dünya koordinatlarından piksel koordinatlarına dönüşüm K matrisindeki eşitliğin soldan ek $\text{diag}(m_x, m_y, 1)$ faktörü ile çarpılarak gerçekleştirilmektedir. Böylece kalibrasyon matrisinin genel formu aşağıdaki gibi olur.

$$K = \begin{bmatrix} a_x & 0 & x_0 \\ 0 & a_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.13)$$

Burada $\alpha_x = f m_x$ ve $\alpha_y = f m_y$ kameranin piksel cinsinden odak uzunluğunu x ve y yönlerinde ifade etmektedir. Benzer şekilde, $\tilde{x}_0 = (x_0, y_0)$ piksel boyutu cinsinden ana nokta koordinatlarını ifade eder. Burada $x_0 = m_x p_x$ ve $y_0 = m_y p_y$ dir. Böylece CCD kameralar 10 serbestlik derecesinde olmaktadır.

$$K = \begin{bmatrix} a_x & s & x_0 \\ 0 & a_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.14)$$

Eklenen s parametresi, eğiklik parametresi olup, normal kameralar için 0 olarak alınmaktadır (Hartley, 2003).

4.3. Kamera Kalibrasyonu

Kamera kalibrasyonu aşamasında, 3 boyutlu dünyadan 2 boyutlu görüntüye geçiş için lazım olan kamera projeksiyon matrisini tahmin edilmektedir. En basit şekliyle 3 boyutlu bir X noktasına karşılık gelen 2 boyutlu görüntüdeki x noktasının bilinmeyen kamera dönüşümü P altında bulunmasıdır. $X_i \leftrightarrow x_i$ arasındaki bağlantı örnekleri yeterince elde edildiğinde P matrisi elde edilebilmektedir. Kameranın içsel özelliklerini ifade eden K matrisi, elde edilen P matrisinden RQ 'nun ayrışması ile çıkarılmaktadır (Hartley, 2003; Carrillo ve diğerleri, 2012).

$X_i \leftrightarrow x_i$ eşleştirmeleri değerlendirildiğinde, 3 boyutlu X_i noktaları ve 2 boyutlu x_i noktalarının ilişkileri tespit edilebilmektedir. 3×4 boyutlarında P kamera matrisi $x_i = PX_i$ işlemi ile her i indisi için bulunmalıdır. $x_i = PX_i$ işlemi homojen koordinatlar içermektedir. Bu nedenle x_i ve PX_i birbirine orantılı olmalıdır. Bu nedenle bu iki ifadenin vektörel çarpımı $x_i \times PX_i = 0$ olmaktadır. P matrisinin 3 satırı p_1^T , p_2^T ve p_3^T olmak üzere;

$$PX_i = \begin{bmatrix} p_1^T X_i \\ p_2^T X_i \\ p_3^T X_i \end{bmatrix}; \quad x_i \times PX_i = \begin{bmatrix} y_i p_3^T X_i - \omega_i p_2^T X_i \\ \omega_i p_1^T X_i - x_i p_3^T X_i \\ x_i p_2^T X_i - y_i p_1^T X_i \end{bmatrix} \quad (4.15)$$

Her $X_i \leftrightarrow x_i$ eşleştirmesi için bir ilişki vardır.

$$\begin{bmatrix} 0^T & -\omega_i X_i^T & y_i X_i^T \\ \omega_i X_i^T & 0^T & -x_i X_i^T \\ -y_i X_i^T & x_i X_i^T & 0^T \end{bmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = 0 \quad (4.16)$$

Her p_i^T , P matrisinin i . satırını ifade eden, bir vektördür. Eş. 4.15'de mevcut olan 3 denklem lineer bağımsız olduğu için sadece iki satır seçilerek işlem gerçekleştirilebilmektedir.

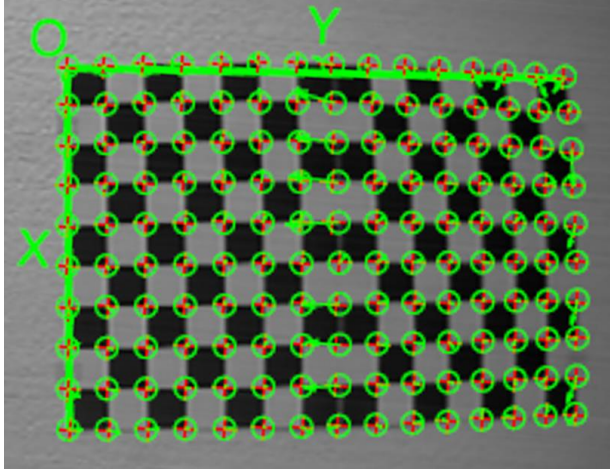
$$\begin{bmatrix} 0^T & -\omega_i X_i^T & y_i X_i^T \\ \omega_i X_i^T & 0^T & -x_i X_i^T \end{bmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = 0 \quad (4.17)$$

Böylece 11 bilinmeyenli 2 bağımsız denklem elde edilir. Buradan ‘n’ adet ilişki için (en az 6 adet ilişki değerlendirilmeli) $2n \times 12$ boyutlu A matrisi elde edilmektedir. Projeksiyon matrisi P, $A_p=0$ denklemi çözülerek bulunur. Burada p, P matrisinin elemanlarını içeren vektörü ifade etmektedir.

Satranç Tahtası Kullanarak Kalibrasyon

En yaygın kalibrasyon yöntemi, 3 boyutlu referans çerçevede koordinatları bilinen belirli sayıda noktaya sahip olan bir nesneyi kameranın önüne koyarak görüntü alınması işlemidir. Aslında, uygun şekilde karakterize edilebilen herhangi bir obje kalibrasyon amacıyla kullanılabilir fakat en çok tercih edilen düzenli desenlere sahip olan satranç tahtalarıdır. İşaretleyiciler ile kaplı kutu gibi 3 boyutlu nesneler ile de kalibrasyon yapmak mümkündür. Fakat satranç tahtası desenleriyle kalibrasyon yapmak oldukça kolaydır (Carrillo ve diğerleri, 2012).

Bu yöntemde satranç tahtasının kamera tarafından en az iki pozisyonda ve oryantasyonda görülmesi gerekmektedir (Z. Zhang, 1999). Algoritma satranç karelerinin köşelerini hesaplayarak, projektif dönüşümü n farklı görüntü noktasından belirli bir ölçek faktörüne kadar çıkarmaktadır. Bu yöntem ile içsel ve dışsal kamera parametreleri çıkarılır ve ışınal lens bozulmaları modellenmektedir. Bu prosedür, kapalı formlu bir çözüm içerir ve maksimum olabilirlik kriterine (maximum likelihood criterion) dayalı doğrusal olmayan iyileştirme uygulanmaktadır. Bu kalibrasyon metodunun en kolay uygulamasına MATLAB altyapısı kullanılarak ulaşılabilmektedir (Bouguet, 2011). İndirilen yazılım sayesinde kolayca istenilen kameranın içsel ve dışsal parametreleri çıkarılabilmektedir. Aşağıda kamera kalibrasyonundan bir örnek Resim 4.1’de verilmiştir.



Resim 4.1. Satranç tahtası kullanılan kalibrasyon örneği

Bizim çalışma kapsamında kullanmış olduğumuz kamera olan Raspberry v2 kamera için 640 x 480 pikselde kalibrasyon matrisi ve bozulma parametreleri aşağıdaki şekilde elde edilmiştir.

$$\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 518.695 & 0 & 332.380 \\ 0 & 518.757 & 229.837 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.18)$$

$$[k_1 \quad k_2 \quad p_1 \quad p_2 \quad k_3] = [0.24 \quad -0.358 \quad -0.009 \quad 0.005 \quad 0.029] \quad (4.19)$$

4.4. Bulanıklığın Tespiti

Bulanıklığın tespit edilmesinde görüntü işleme metotlarının kullanılması ile daha doğru ve güvenilir tespitlerin yapılmasına başlanmıştır. İnsan gözü çoğu zaman bariz şekilde hareket bulanıklıklarının ayırdına varabilmektedir fakat bulanıklık tespitinde kullanılan çeşitli metrikler ile karşılaştırıldığında, insan algısının bulanıklık tespitinde her zaman metrikler ile uyumlu sonuçlar vermediği görülmektedir. Bu nedenle daha hassas tespitler için ilgili metriklerin kullanımı önem arz etmektedir.

Bir resmin bulanık olup olmadığı ya da bulanıklık seviyesini tespit etmek oldukça zorlu bir problemdir. Çoğunlukla aynı pozisyondan elde edilmiş keskin bir görüntü ile bulanıklık seviyesi tespit edilebilmektedir. Bu keskin görüntü karesi, referans görüntü olarak

adlandırılmaktadır. Referans görüntü olmadan bulanıklık, kenar algılama ve frekans analizi yöntemleri ile tahmin edilebilmektedir.

4.4.1. Kenar algılamaya dayalı bulanıklık algılama

Bulanıklık seviyesi az olan keskin görüntülerde kenarlar belirgin olarak görülürken, bulanık resimlerde kenarların bulunduğu pikseller komşu piksellerle daha fazla benzerlik göstermektedir. Bu durum bulanıklığın algılanması amacıyla kullanılmaktadır. Özellikle gri seviyede, kenarların olduğu bölgeler komşu piksellerden yoğunluk değeri bakımından ayrışmaktadır. Gradyan hesabı yapıldığında kenar bölgelerinde zıtlığın yüksek seviyede olduğu görülmektedir. Fakat bulanıklık seviyesi arttıkça, zıtlık azalır ve gradyan daha geniş bir bölgede daha düz bir şekilde ortaya çıkmaktadır. Kenarlar bulanıklık nedeniyle ortadan kalkmış ve komşu pikseller ile bir girişim meydana gelmiştir. Bu durumda kenar algılama az tespit edilebilen ya da hiç tespit edilemeyen sonuçlar vermektedir (Narvekar ve Karam, 2011). Bu durum Şekil 4.2’te gösterilmiştir.



Şekil 4.2. Bulanıklık seviyesi ile gradyan sonuçlarının karşılaştırılması

Bulanıklık arttıkça gradyan sonucunun düz bir eğime ulaştığı görülmektedir. Bulanıklık tespitinde kenar algılama temelli yaklaşımda ilk adım gradyan yönünün hesaplanmasıdır.

Bulunan gradyan, kenarın hangi tarafının parlak, hangi tarafının koyu olduğu tespit edilmesinde kullanılmaktadır. Ardından kenar algılama algoritması kullanılarak sadece kenar pikseller elde edilmektedir. Ardından, gerçek bulanıklık değerinin ifadesi olan kenar yayılım değeri bulunmaktadır. Gradyanın lokal ekstrema değerinden her iki yönde piksel sayarak bu değer tespit edilmektedir. Lokal ekstrema değerinden aşağı yönde gradyan azalır ve lokal ekstremadan yukarı yönde yoğunluk (intensity) artmaktadır. Artan ve azalan bölgede piksel sayımı gerçekleştirilerek yayılım değeri bulunmaktadır. Burada elde edilen toplam piksel değeri kenar yayılımını ifade etmektedir. Son olarak ise resim kalitesinin hesaplanmasıdır. Bu kalite metriği tüm kenarların ortalama kenar yayılma değeri ve küçük bir insan denek grubundan elde edilen öznel derecelendirmeleri ifade eden bir parametre kullanılarak gerçekleştirilmektedir (Ong, Lin, Lu, Yang, Yao, Pan ve Moschetti, 2003). Bu örnek kenar algılama temelli yaklaşımın bulanıklık tespitine yönelik en temel örneklerinden birisidir (Sieberth, Wackrow ve Chandler, 2016).

Bahsi geçen çalışmadaki eksiklik, rotasyonlarından bağımsız olarak tüm kenarların değerlendirilmeye alınmasından kaynaklanmaktadır. Eğer bulanıklığa dik yönünde bir kenar varsa, bu tespit için bir ipucu sağlayabilmektedir. Bununla birlikte eğer bulanıklık ile aynı yönde bir kenar varsa bu kullanışlı bir sonuç sağlayamayacaktır. Bu yöntemin daha geliştirilmiş bir versiyonu olan (Narvekar ve Karam, 2011), yaklaşımında ise görüntü alt görüntülere ayrılarak kenarlara ait piksel sayımı gerçekleştirilmektedir. Eğer kenarlar bulanıklık tespitinde kullanılamayacak ise sayım işlemine devam edilmemektedir (Sieberth ve diğerleri, 2016).

4.4.2. Frekans analizine dayalı bulanıklık algılama

Bulanıklık algılama konusunda uygulanabilecek alternatif bir yaklaşım ise görüntünün frekans analizinin yapılmasıdır. Bir resim frekans bandında 2 boyutlu bir vektör ile ifade edilebilmektedir (Rahtu, Heikkilä, Ojansivu ve Ahonen, 2012). Bulanık resimlerde yüksek frekanslar çoğunlukla ya hiç bulunmamaktadır ya da çok az miktarda yer almaktadır. Genellikle keskin pozlanmış resimlerin frekans bandında daha geniş yayıldığı görülürken, yüksek bulanıklığa sahip resimlerde yüksek frekansların bulunmadığı (Şekil 4.3) görülmektedir (Sieberth ve diğerleri, 2016).



Şekil 4.3. Bulanıklık seviyesinin frekansa etkisi

Görüntü uzamsal (spatial) alandan frekans alanına dönüştürüldüğünde genlik ve faz olarak iki bileşen ile incelemek mümkün olur. Bulanıklık hem büyüklüğü hem de fazı etkilemektedir. Bir FFT yaklaşımının temel zayıflığı, görüntüyü dönüştürmek için hesaplama süresidir. FFT yaklaşımında uzamsal (spatial) alandan frekans alanına geçiş ve sonra tekrar uzamsal alana geri dönüş filtre temelli yöntemlere nazaran oldukça uzun sürebilmektedir. Bu dönüşüm işlemleri filtre uygulamaktan yaklaşık iki kat daha uzun sürmektedir. Ayrıca frekans alanında ışınsal (radyal) yayılım ile analiz yapmak uzamsal alanda işlem yapmaktan çok daha zordur. Bir görüntünün frekans alanına dayalı bulanıklık tespiti, görüntünün yüksek frekanstaki bileşenlerinin sayısı ve büyüklüğünün analizine dayanmaktadır. Bu yöntemlerin kullanımı son dönemde pek fazla tercih edilmemektedir. Kenar algılama ve frekans analizine dayalı yöntemleri birleştiren ve “kenar keskinliği bulma” yöntemi olarak bilinen yöntemler de mevcuttur (Crete, Dolmiere, Ladret ve Nicolas, 2007). Kenar algılama ve frekans analizine dayalı klasik yöntemlerin gerçek zamanlı olarak kullanımının işlem süresi nedeniyle mümkün olmadığı görülmektedir.

Bulanıklık tespiti genelde orijinal resim ile bulanık resmin karşılaştırılması şeklinde gerçekleştirilmektedir. Bu aşamada referans olarak kullanılacak orijinal resmin de bazı gürültüler içermesi tespiti zorlaştırabilmektedir. Fakat aynı resmin tamamen gürültüsüz olarak elde edilmesi de mümkün olmadığı için bu tarz durumlarda görel karşılaştırmaların kullanılması tercih edilmektedir.

Nokta yayılım fonksiyonu

Bir görüntüdeki bulanıklığın azaltmanın \ düzeltmenin (deblur) en kolay yolu, tam olarak bulanık görüntü ile aynı sahneyi gösteren orijinal bir keskin görüntüyü kullanmaktır. Bu keskin görüntü aracılığıyla resimde bulanıklaşmaya neden olan hareket nokta yayılım fonksiyonu (point spread function - PSF) tahmin edilebilmektedir. Eğer keskin bir resim elimizde yoksa o zaman gerçekleştirilen hareketin tahmin edilmesi gerekmektedir. Hareket tahmininde kameranın görüntü yakalarken izlemiş olduğu yolun tahmini gerçekleştirilebilmektedir. Bu hareket çok basit, temel bazı hareketler olabileceği gibi, oldukça karmaşık da olabilmektedir. Ayrıca gerçekleştirilen hareketin hızı sabit ya da değişken olabilmektedir. Hareket tahmini de benzer şekilde bir nokta yayılım fonksiyonu ile ifade edilmektedir. Nokta yayılım fonksiyonuna “resimde meydana gelmiş olan bozulmanın ifadesi” denilmektedir. Eğer PSF büyük ise resimde meydana gelmiş olan bulanıklaşma da büyüktür. PSF yalnızca uniform olan yani resmin her yanına yayılmış olan hareket bulanıklığının ifadesi olarak kullanılabilir.

Bulanıklaştırma çekirdeği (kernel)

Daha kompleks bulanıklıklarda 4 boyutlu (üç uzaysal yön ve bir zaman boyutu) bulanıklık çekirdeği (kernel) kullanılmaktadır. Bulanıklık çekirdeği ise pozlandırma zamanı esnasındaki hareketin yörüngesini vermektedir. Buradaki yörüngenin uzun ve kompleks olması kameranın nesneye göre uzun ve kompleks bir hareket yaptığının ifadesidir. Çekirdeklerde farklı hızlar, farklı grilik seviyesinde ifade edilmektedir.



Resim 4.2. Bulanıklaştırma çekirdeği (kernel) örneği

4.5. Hareket Bulanıklığı Modelleri

Hareket bulanıklığının modellenmesi elde etmek istediğimiz ideal olan keskin resim ile elimizde olan bulanık resim arasındaki modelin ortaya çıkarılması olarak adlandırılabilir. İdeal resim, ani olarak pozlandırılmış olan, hiçbir bulanıklığın meydana gelmesine fırsat verilmeyen resimdir. Bu durum pratikte mümkün olmamaktadır. Pozlandırma süresi, kamera ile sahne arasında göreceli bir hareket olduğu zaman hareket bulanıklığına neden olan geçici bir kutu filtresi tanımlamaktadır. Bu filtre önemli yüksek frekanslı uzamsal detayları ortadan kaldırır.

Bulanıklaştırma homojen olduğunda bozucu hareket etmeni uzamsal olarak (her pikseli içerecek şekilde) konvolüsyon işlemi ile frekans alanında ise nokta çarpımı ile ifade edilmektedir (Üzer, 2010).

$$z=u*h[x,y] = \int u(x-s,y-t)h(s,t)dsdt \quad (4.20)$$

Ayrık durum için;

$$z=\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} u(i-k, j-l)h(k,l) \quad (4.21)$$

Burada u ideal görüntü olup, h konvolüsyon çekirdeği veya nokta yayılma fonksiyonu (PSF) ve z bulanık görüntüdür. PSF, görüntüdeki her (x, y) piksel konumunun tek noktadan maruz kaldığı ışığı tanımlayan bir enerji yoğunluğu fonksiyonudur. PSF h, aşağıdaki enerji tasarrufu kısıtlamasını karşılamalıdır:

$$z=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)dxdy=1 \quad (4.22)$$

Ayrık durum için;

$$z=\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} h(i,j)dxdy=1 \quad (4.23)$$

Basit model sadece homojenlik koşulu altında kullanılabilir. Eğer kamera odak uzaklığı kısaysa ya da kamera optik ekseninde önemli derecede hızlı dönüş hareketi

yapıyorsa resimde bulanıklığın yoğunluğu değişmektedir. Başka bir deyişle, sahne derinliği ve kamera ile sahnenin görelî hareketinin kompleks bir fonksiyonu haline gelmektedir. Uzamsal olarak değişken olmayan PSF bu tip kompleks bir hareketi modellemek için yeterli olmamaktadır. Uzamsal değişken bulanıklığı ifade edebilmek için daha genel lineer işlemleri kullanmak gerekmektedir.

$$z=u\hat{*}h[x,y]=\int u(x-s,y-t)h(x-s,y-t;s,t)dsdt \quad (4.24)$$

Bu denklemde u ideal bir görüntüdür, h PSF'dir ve z bulanık görüntüdür. Görüleceği üzere homojen bulanıklık için olan denklem resim koordinatları x ve y ile PSF in değişmediği durumundaki özel halidir. Yukarıda yer alan Eş. 4.24'deki işlem uzay bağımlı konvolüsyon olarak ifade edilmektedir. Bu denkleme gürültünün de eklendiği durumda formüller aşağıdaki gibi olmaktadır.

$$z=u\hat{*}h[x,y]+n(x,y) \quad (4.25)$$

$$z=u\hat{*}h[x,y]+n(x,y) \quad (4.26)$$

Hareket bulanıklığı, temel olarak, deklanşör hızı tarafından belirlenen pozlandırma süresi boyunca görüntüleme sensörü ile sahne arasındaki bağıl hareketin sonucudur. Bu nedenle bulanıklık hareket tipinin değişimi ile değişebilmektedir. Bu hareket öteleme, dönme, ölçeklendirmede ani değişim ya da tüm bunların kombinasyonu şeklinde olabilmektedir. Bundan dolayı tüm hareket tipini içeren tek bir modelleme yapılması mümkün olmamaktadır. Gerçekleşen hareketin tipine göre hareket bulanıklığının modeli doğrusal, döneî ve ışınsal olarak incelenmektedir.

4.5.1. Bir boyutlu doğrusal hareket bulanıklığı modeli

Optik eksene dik düzlemde 0-t_{poz} ile ifade edilen pozlandırma periyodu aralığında hareket meydana geldiği zaman meydana gelen hareket yönündeki uzamsal bağımsız 1 boyutlu dikdörtgen dürtü şeklindeki PSF (Üzer, 2010) Eş. 4.27'deki gibi ifade edilir:

$$h(x,y)=\begin{cases} \frac{1}{L} & -\frac{L}{2} \leq x \leq \frac{L}{2} \\ 0 & \text{aksi durumda} \end{cases} \quad (4.27)$$

Bu ifadede PSF yoğunluğu L harfi ile ifade edilir ve görelî hız ile pozlandırma süresinin çarpımına eşittir. Bu görelî hareket pozlandırma periyodu boyunca sahnenin yatay eksenî ile bir açî yaptığında;

$$h(x,y) = \begin{cases} \frac{1}{L} & \sqrt{x^2+y^2} \leq L \text{ ve } \frac{x}{y} = -\tan \phi \\ 0 & \text{aksi durumda} \end{cases} \quad (4.28)$$

ϕ değeri ‘0’ kabul edilirse, denklemin ayrık versiyonu yaklaşık olarak Eş. 4.29’daki gibi elde edilmektedir;

$$h(i,j) = \begin{cases} \frac{1}{L} & \text{eğer } n_1=0, |n_2| \leq \left\lfloor \frac{L-1}{2} \right\rfloor \\ \frac{1}{L} \left\{ (L-1) - 2 \left\lfloor \frac{L-1}{2} \right\rfloor \right\} & \text{eğer } n_1=0, |n_2| \leq \left\lfloor \frac{L-1}{2} \right\rfloor \\ 0 & \text{aksi durumda} \end{cases} \quad (4.29)$$

4.5.2. Dönel hareket bulanıklık modeli

Pozlandırma süresi boyunca ($0-t_{poz}$) sahne ile kamera arasında dönme hareketi olduğunda dönel hareket kaynaklı bulanık görüntüler kaydedilmektedir. Bu hareket homojen olmayan bir bulanıklık meydana getirmektedir. Dönel hareketin bulanıklık yörüngesi dairesel yaylardır ve bu yayların yarıçapına bağlı olarak bulanıklık değişmektedir. Bu nedenle rotasyonel hareketler, görüntülerde uzaysal olarak değişken bulanıklığa neden olmaktadır. Döndürme hareketiyle bulanık görüntü, aşağıdaki gibi formüle edilmektedir (Wenying, Jingxin ve Yi, 2008; Üzer, 2010);

$$z = \frac{1}{T} \int_0^T u(x-x_0(t), y_0(t)) dt \quad (4.30)$$

burada $x_0(t) = r \cos(\omega t)$, $y_0(t) = r \sin(\omega t)$, $r = \sqrt{x^2 + y^2}$ ve ω dönme hareketine ait açısal hızdır. Görüntü düzlemi kutupsal koordinat sisteminde temsil ediliyorsa, dönme hareketinin neden olduğu bulanık görüntü aşağıdaki gibi formüle edilmektedir;

$$z(r, \Theta) = \frac{1}{T} \int_0^T u(r, \Theta - \omega t) dt \quad (4.31)$$

burada (r, θ) görüntü noktasının kutupsal koordinatıdır. Yukarıdaki denkleminin parametreleri aşağıdaki şekilde değiştirildiğinde;

$l = r\theta$ $s = r\omega t$, $N_r = r\omega T$ ve alt indis r olmak üzere;

$$z_r(l) = \frac{1}{N_r} \int_0^{N_r} u_r(l-s) ds \quad (4.32)$$

Ayrık durum için;

$$z_r(i) = \frac{1}{N_r} [u_r(i) + u_r(i-1) + \dots + u_r(i-N_r+1)] \quad (4.33)$$

$i = 1, 2, \dots, M_r$ ve M_r bulanıklaşan dairesel yaydaki piksel dizilerinin periyodu olarak PSF ayrık biçimde aşağıdaki şekilde yazılmaktadır:

$$h_r(i) = \begin{cases} \frac{1}{N_r} & 1 \leq i \leq N_r \\ 0 & N_r \leq i \leq M_r \end{cases} \quad (4.34)$$

Sadeleştirme işlemi ile bulanık görüntü;

$$z_r(i) = \sum_0^{N_r} u_r(m) h_r(m) = u_r(i) * h_r(i) \quad (4.35)$$

Burada $h_r(i)$, görüntüyü rotasyonel olarak bulanıklaştıran PSF'dir, $N_r = r\omega T$ bulanıklık yoğunluğudur ve $\theta = \omega T$, T pozlama süresi boyunca bulanıklık açısını göstermektedir. Bulanıklık yoğunluğu N_r ve yarıçap r arasında bir korelasyon bulunmaktadır. Bulanıklık yoğunluğu yarıçapın değişimi ile değişmektedir. Bulanıklığın yoğunluğu, θ sabit ve yarıçap değişken iken bulanıklık açısının hesaplanmasıyla elde edilebilmektedir. Bulanıklık yoğunluğu belirlendiğinde, PSF elde edilmektedir.

Yukarıda bahsedildiği gibi, bulanıklığın yoğunluğu r yarıçapı ile değişmektedir. Bu nedenle dönme merkezini bulmak ve ardından dairesel bulanıklık yayları boyunca pikselleri elde etmek önemlidir. Bulanığın merkezini, görüntünü merkezinde (Wenying ve

diğerleri, 2008) ve öteleme olmadığını ve kameranın açısal hızı sabit kabul eden (Klein ve Drummond, 2005) çalışmalar mevcuttur (Üzer, 2010).

4.5.3. Işınsal hareket bulanıklığı modeli

Işınsal (merkezden yayılan) bulanıklık bir çeşit uzamsal değişken hareket bulanıklığıdır. Bu bulanıklık bazen görüntü alma sürecinde, görüntü sensörünün yakınlaştırma işlemi sırasında nesneye doğru hızla hareket ederken de oluşabilmektedir. Bu tip bulanıklıkta merkezden dışarı doğru uzaklaştıkça artan bir bulanıklık söz konusudur. Kamera nesneye yaklaşırken, kameranın görüş alanını tanımlayan üçgen gittikçe küçülür fakat objede bir değişiklik yoktur. Eğer kameranın pozlandırma süresi, göreceli hız ve kamera ile ilgilenilen obje arasındaki mesafe ile karşılaştırılabilir derecede ise, elde edilen görüntü radyal olarak bulanıklık içerecektir. Bazı görüntü tabanlı kilitlenme sistemine sahip füzelerde ve dron üzerinden alınan fotoğraflarda bu tip bulanıklıklar görülebilmektedir (Üzer, 2010).

Literatürde radyal bulanık görüntüler üzerine birkaç (Webster ve Reeves, 2007; Boracchi, Foi, Katkovnik ve Egiazarian, 2008) çalışma mevcuttur. Radyal bulanıklığa ait matematiksel modelleme (Webster ve Reeves, 2007) çalışmasında önerilmiştir. Uzamsal olarak değişmez bulanıklığın, uzamsal değişen şekilde yeni koordinat sistemine dönüşümü temeline dayanan modele göre radyal bulanıklığın ifadesi Eş. 4.36'da verilmiştir.

$$z_{\theta}(\phi) = \int_0^T u_{\theta}(r(\phi, t)) dt \quad (4.36)$$

Buradaki formüle göre her verilen θ değerine göre bulanıklık tek boyutlu bir hal almaktadır. Burada u_{θ} keskin resim, z_{θ} radyal gürültülü resim, r obje yüzeyine ait radyal koordinat ve θ obje ve sensör yüzeyindeki r ve ϕ ile ifade edilebilen her noktanın açısal koordinatıdır. Bir diğer formül ise ilgilenilen nesne ile kamera arasındaki mesafeyi gösteren aşağıdaki ifadedir. İlgili denklemde x ifadesi yatay uzamsal koordinattaki mesafeyi temsil etmektedir. Kamera hızı v ile gösterilmektedir. Bu ifade yukarıda yer alan formülde θ ifadesinin değiştirilmesi ile elde edilmektedir (Üzer, 2010).

$$r = x \tan(\phi), \quad \frac{dx}{dy} = -v \quad (4.37)$$

4.6. Hareket bulanıklığının düzeltilmesi / azaltılması

Şarj bağlı cihaz (charge coupled device- CCD) ve karşılaştırmalı metal oksit yarı iletken (complementary metal oxide semiconductor- CMOS) teknolojisinin görüntü elde etmek amacıyla sensörlerde kullanılmasıyla birlikte görüntü zenginleştirme çalışmaları önemli bir araştırma konusu olmuştur. Bu şekilde sağlık ve savunma gibi pek çok alanda görüntü kalitesinin artırılması hatta süper çözünürlük uygulamalarının kullanıldığı (Wang, Fevig ve Schultz, 2008) ve hatta insansız hava araçlarından elde edilen resimlerin mobil cihazlarda görüntü zenginleştirilmesinin yapıldığı çalışmalar bile mevcuttur (González, Patricio, Berlanga ve Molina, 2019). Süper çözünürlükte genel mantık birbiri üzerine binen resimlerdeki ortak alanları tespit ederek bu alanların birleştirilmesi yoluyla daha yüksek çözünürlüklü yeni resimler elde etmektedir. Görüntü bulanıklığının düzeltilmesi (image deblurring) uygulamalarında ise aynı resmin bulanıklık nedeniyle kalitesinde meydana gelen azalmanın, resmin restore edilmesi yoluyla arttırılmaya çalışılmaktadır.

Hareket bulanıklığı görüntü kalitesini azalttığı için görüntü kullanılabilirliğini attırmak amacıyla görüntü bulanıklığı azaltma / önleme (deblurring) yöntemleri kullanılmaktadır. Görüntü temelli takip ve arama kurtarma- keşif gözetleme uygulamalarında bu tip netleştirme metotlarından faydalanan çalışmalar mevcuttur. Kamera ile sahne arasında meydana gelen göreceli hızlı hareketler genellikle hareket bulanıklığına neden olmaktadır. Gerçekten de bu tür bozulmalar, kenar algılama (Canny, 1986), köşe algılama (Harris ve Stephens, 1988) ve ölçekte değişmez öznitelik dönüşümü (Scale-Invariant Feature Transform - SIFT) (Lowe, 2004) gibi görüntü temelli takip içeren birçok öznitelik çıkarma algoritmasının başarısız olmasına neden olabilmektedir (Sieberth ve diğerleri, 2016).

Bulanıklık düzeltme, meydana gelen bulanıklığın ortadan kaldırılması ve keskin resmin elde edilmesine yönelik olarak gerçekleştirilen işlemlerdir. Bulanıklık, görüntü işleme literatüründe genel olarak bulanıklığa neden olan nokta yayılım fonksiyonu ile keskin resmin konvolüsyonu olarak ifade edilmektedir. Netleştirme veya bulanıklık düzeltme işlemi (deblurring) ise “deconvolution” işlemi olarak tanımlanmaktadır (Jia, 2007). Dekonvolüsyon işlemi herhangi bir ek bilgi olmadan sadece bulanık görüntü bilinerek yapılırsa işlemin adı kör görüntü dekonvolüsyonu (blind image deconvolution) eğer ek bilgi kaynakları gerekiyorsa kör olmayan görüntü dekonvolüsyonu (non-blind image

deconvolution) ismini almaktadır. Bulanık görüntüye dair elimizde ek bilgi bulunup bulunmamasına göre hareket bulanıklığını düzeltmek kör görüntü dekonvolüsyonu ve kör olmayan görüntü dekonvolüsyonu olmak üzere ikiye ayrılmaktadır.

4.6.1. Kör görüntü dekonvolüsyonu

Kör dekonvolüsyon işleminde giriş olarak sadece bulanık görüntü kullanılmaktadır. Herhangi bir sensörden elde edilmiş ek bir bilgi söz konusu değildir. Dekonvolüsyon işlemi bulanık resim ile elde edilen tahmini hareketi ifade eden vektörün dekonvolüsyonunun alınması ile keskin resmin elde edilmesini amaçlamaktadır. Kör görüntü dekonvolüsyonu algoritmaları klasik hareket bulanıklığı giderme algoritmaları olarak değerlendirilebilmektedir. Bu yöntemlerde nokta yayılım fonksiyonu tahmini yapılırken daha sonra bu fonksiyon ile bulanık görüntü arasında dekonvolüsyon işlemi gerçekleştirilmektedir. Bu yöntemle en uygun klasik metotlar Wiener dekonvolüsyon (Tukey, 1952), ve Richardson-Lucy dekonvolüsyon (Lucy, 1974; Richardson, 1972) yöntemleridir. Bu yöntemler yıllar önce ortaya konulmuş olmasına rağmen, basit ve verimli olduğu için pek çok uygulamada kullanılmaya devam edilmektedir. Her iki yöntem de olasılık teorisindeki Bayes teoremini kullanmaktadır (Sieberth ve diğerleri, 2016).

$$P(A||B)=\frac{P(B|A)P(A)}{P(B)} \quad (4.38)$$

Bu teoremde $P(A)$, A'nın olasılığı; $P(B)$, B'nin olasılığı; $(A||B)$, B doğru iken A'nın olasılığı; $P(B|A)$, A doğru iken B'nin olasılığıdır. Resimde meydana gelen bulanıklık Eş. 4.39'daki şekilde ifade edilir.

$$h = f \otimes g \quad (4.39)$$

Bu eşitlikte f orijinal görüntü, g nokta yayılım fonksiyonu ve h ise bulanıklık nedeniyle bozulmuş görüntüdür. Kullanılan operatör ise $(\otimes)$ konvolüsyon operatörüdür. Richardson metodunda f , g ve h 'nin ayrık olasılık-frekans fonksiyonları oldukları varsayılmıştır. Bu durumda “ f ” ifadesi Bayes kuralı kullanılarak iteratif olarak bulunabilmektedir. Bunun için bir başlangıç değerinin (f_0) kullanılması gerekmektedir (Richardson, 1972). Buradaki eksik yan ise bulanık görüntünün gürültüden arındırılmış kabul edilmesidir. Bulanık

görüntüde bir gürültü mevcutsa, restore edilmiş görüntüde de bazı bozulmalar olabilmektedir. Ayrıca bu metotlar kullanılırken aslında gerçekte olmayan yapay halkalar (ringing artifacts) meydana gelebilmektedir. Bu halkalar görüntüdeki kenarlara bitişik yüksek kontrastlı çizgiler olarak meydana gelmektedir. Bu halkalar genel olarak görüntü işleme uygulamalarını olumsuz olarak etkilemektedir. Özellikle hedef tespit algoritmalarının bu halkalardan olumsuz etkilendiği bilinmektedir. Yapay halkaların ortadan kaldırılması/azaltılması için tam değişim düzenlileştirme (total variation regularization) (Dey, Blanc-Féraud, Zimmer, Kam, Olivo-Marin ve Zerubia, 2004) ve gradyan seyreklik kısıtı (gradient sparsity constraint) (Levin, Fergus, Durand ve Freeman, 2007) kullanan yöntemler mevcuttur.

Geleneksel yöntemlerin hareket tahmini ve dekonvolüsyon işlemlerini hızlandırmaya yönelik olarak da bazı çalışmalar mevcuttur. Bu yöntemlerde genellikle ek bazı bilgiler kullanılarak PSF tahmini gerçekleştirilmektedir. Alfa maskesi kullanarak PSF tahmini işlemini kullanan (Jia, 2007) çalışması ve resmin doğal istatistiklerinden faydalanan dekonvolüsyon aşamasını hızlandıran (Fergus, Singh, Hertzmann, Roweis ve Freeman, 2006) yöntemleri bilindik yöntemlerdendir.

Wiener (Tukey, 1952) ve Richardson-Lucy dekonvolüsyon (Lucy, 1974; Richardson, 1972) gibi olasılıksal hesaplamalara dayanan (L. Sun, Cho, Wang ve Hays, 2014) ve Extreme Channels Prior (ECP) (Yan, Ren, Guo, Wang ve Cao, 2017) gibi yöntemler de mevcuttur. Bu tip çalışmalarda resimde gürültü mevcut olsa bile netleştirme sonrasında görüntü işleme uygulamalarında kullanım güvenilirliğini arttırmayı amaçlamaktadır. Benzer şekilde 'Hyper-Laplacian Önceli Kullanan Hızlı Görüntü Dekonvolüsyonu' gibi kör dekonvolüsyon metotları (Krishnan ve Fergus, 2009) veya (Michaeli ve Irani, 2014) çalışmaları da mevcuttur. Bu yöntemlerin dezavantajı ise özellikle düşük kapasiteli donanımlarda ve gerçek zamanlı uygulamalarda bellek kullanımının yüksek olması ve hesaplama zamanının uzun olmasıdır.

4.6.2. Kör olmayan görüntü dekonvolüsyonu

Kör olmayan dekonvolüsyon yöntemlerinde bulanık resim ile birlikte resimdeki bulanıklık ile ilgili bazı ek bilgiler de mevcuttur. Ek bilgiler resimlerde üst üste gelen bölgeler (Agrawal, Xu ve Raskar, 2012), hassas IMU ölçümleri (Joshi, Szeliski ve Kriegman, 2008)

kamera objektifini hızlı şekilde açıp-kapatan kameralar ile görüntü almak (Raskar, Agrawal ve Tumblin, 2006) (flutter shutter) ve renk kanalına bağlı maruz kalma süreleri (Lelégard, Delaygue, Brédif ve Vallet, 2012) olabilmektedir (Sieberth ve diğerleri, 2016).

Tüm bu yöntemlerde kullanılan ek bilgiler ile bulanıklık çekirdeği ortaya çıkarılmaya çalışılmaktadır. Ortaya çıkarılan bulanıklık çekirdeği daha sonra netleştirme işleminde kullanılmaktadır. Üst üste binen farklı görüntülerin kullanılması yönteminde ayrıca süper çözünürlüklü görüntüler elde etmek de mümkün olabilmektedir (Raskar ve diğerleri, 2006).

Statik bir arka plana sahip görüntülerde hareketli objeden PSF üreten bir diğer yöntemde mevcuttur (Agrawal, Xu ve Raskar, 2012,). PSF e ayrık fourier dönüşümü uygulanarak elde edilen frekans uzayında yanlış pozlanmış ve bilgi içermeyen frekanslar tespit edilebilmektedir. Farklı pozlandırma zamanlarında alınan resimlerde bilgi içermeyen frekanslar diğer dönüştürülmüş PSF'lerin frekansları ile doldurulabilmektedir. Daha sonra elde edilen birleşik PSF frekans alanından uzamsal alana çevrilerek netleştirme amacıyla dekonvolüsyon işleminde kullanılmaktadır. Bu yöntemde yer alan kabuller ise aynı pozun farklı pozlanma zamanlarında aynı açıdan alınması ve arka planın hareketli olmamasıdır.

Fourier dönüşümünü kullanan bir diğer çalışmada (Jung, Kim ve Ko, 2009) bir resimde yer alan aynı bölgeye ait tüm frekanslar bir araya getirilerek birleştirilmiştir. Bu metot resimde olmayan frekansları elde etmeyi sağlamaktan ziyade var olan frekansların bir araya getirilmesi üzerine kurgulanmıştır. Bu yöntemde aynı sıralamada birbirine tam olarak hizalanmış olarak bir araya getirilen resimler kullanılmıştır. Bu nedenle gerçek bir hareket bulanıklığından çok hizalanmış resimlerde düzeltme yapmaya yönelik bir yöntem olarak karşımıza çıkmaktadır. Hizalanmamış resimler ile herhangi bir test gerçekleştirilmemiştir.

Daha doğru PSF tahmini gerçekleştirmek için birden çok kamera kullanan (Li, Yu ve Chai, 2008) yöntemi de mevcuttur. Bu yöntemde bir düşük çözünürlüklü ve bir yüksek çözünürlüklü kamera paralel olarak işlem yapmaktadır. Düşük çözünürlüklü kamera yüksek hızda (fps olarak), yüksek çözünürlükte olan düşük hızda görüntü almaktadır. Bu iki kameradan alınan görüntüler birlikte kullanılarak görüntünün bulanıklığının giderilmesi sağlanmaktadır. Benzer şekilde değişik pozlandırma sürelerinde görüntüler alarak oluşan

PSF'leri inceleyen ve video akarken en uygun hızı belirleyerek bulanıklığı ortadan kaldıran çalışma (Agrawal, Xu ve Raskar, 2009) da mevcuttur (Üzer, 2010).

Bu çalışmalara ek olarak jiroskop kullanarak kameralar tarafından ölçülemeyen dönüş yönünü ve açısını ölçen (Klein ve Drummond, 2004) çalışmasında kenar tespiti yaparak poz tahmini de gerçekleştirilebilmektedir. Bulanıklığın tespiti ve azaltılması konusunda ek sensörlerin ya da donanımların kullanıldığı pek çok çalışma mevcuttur. Platform boyutları küçüldükçe ek donanım ile bu problemlerin aşılması zorlaşmaktadır. Bu nedenle son dönemde üzerinde çalışma yapılan tek resimde bulanıklık önleme çalışmaları oldukça önem kazanmıştır.

Bulanıklığın düzeltilmesinde kullanılan görüntü sayısına göre düzeltme çeşitleri “tek görüntü özelliklerine dayalı” ve “çoklu görüntüye dayalı” olmak üzere ikiye ayrılmaktadır.

4.6.3. Tek görüntü özelliklerine dayalı bulanıklık düzeltme

Kamera ve sahne arasındaki bağıl hareket tam olarak biliniyorsa, PSF analitik olarak hesaplanabilmektedir. Ancak, bunu başarmak için ekstra sensörler gereklidir. Yakalanan görüntünün kendisini kullanarak bulanıklık tahmini literatürde hala en yaygın senaryo olmasına rağmen gerçekten zor bir problemdir. Hareket hakkında herhangi bir ek bilgi olmadan, tek bir görüntü üzerinden bulanıklık tahmininin yapılmasına “tek görüntü üzerinde kör dekonvolüsyon” (single image blind deconvolution) denilmektedir. Bu tipteki bulanıklığı ortadan kaldırma yöntemi en zor yöntemlerden biridir. Bununla birlikte bu tipten bir algoritma ek ölçüm ve donanım gerektirmediği için küçük platformlarda özellikle mikro /nano insansız araçlarda oldukça kullanışlı olabilmektedir. Tek bir görüntüde pek çok nedenle (kamera sallanması, az ışık, tüketici kameralarındaki uzun pozlandırma süreleri) bulanıklık meydana gelebilmektedir. Tek resim üzerinden bulanık tahmini için limitli hareket tipleri ve istatistiksel metotlar yardımcı olarak kullanılmaktadır.

Geleneksel yöntemlerde PSF genellikle bir alçak geçiren filtre olarak modellenir ve görüntünün bozulmasına düzgün hareketin neden olduğu varsayılmaktadır. Fakat bu yöntemler bazen yoğun bulanık görüntü durumunda başarısız olur. Yine de tek görüntü üzerinde kamera sallantılarında yoğun ve karmaşık bulanıklık seviyelerinde bile başarılı olan (Fergus ve diğerleri, 2006) gibi çalışmalar mevcuttur. Fakat bu çalışmada da resim

gradyanının kesin bir dağılıma sahip olması bir ön kabul olarak seçilmiştir. Bu açıdan bakıldığında yine manuel girdiler söz konusudur. Bulanıklığı maksimum posteriori (MAP) problemi olarak gören ve başlangıç olarak kaba bir bulanıklık çekirdeği atayarak PSF parametrelerini iteratif olarak tahmin etmeye çalışan yöntemler de mevcuttur.

4.6.4. Çoklu görüntüye dayalı bulanıklık düzeltme

Bulanıklık hakkında daha fazla bilgi elde etmek için geliştirilen bir diğer yöntem ise birden fazla görüntü kullanmaktır. Böylece bir ek görüntüler yer alan bilgiler de işleme dahil edilmiş olur. Bu yöntemde bulanık bir görüntü kümesi giriş olarak verilirken genelde bir tek bulanık olmayan resim çıktı olarak alınmaktadır (Bascle, Blake ve Zisserman, 1996). Bu çalışmaların ilk örneklerinde görüntü kümesinde yer alan resimlerin aynı yönde ve aynı hızda hareket yaptıkları bilinerek hareket bulanıklığı tahmini gerçekleştirilmektedir. Bu durumda meydana gelen bulanıklık yoğunluğu da hareketin büyüklüğü ile orantılı olmaktadır. Fakat bu kabul her zaman doğru değildir. Çünkü görüntüler arasında farklı yönde ve büyüklükte hareketlerin gerçekleştirilmesi olasılığı oldukça yüksektir. Gauss piramidi kullanarak resmi alt örneklere ayırarak bulanıklık tahmini yapan (Agrawal ve diğerleri, 2009) ve görüntü kümesinde yer alan her bir görüntü için pozlandırma süresini değiştirerek yönü aynı ama büyüklükleri farklı PSF elde eden yöntemler de mevcuttur (Rav-Acha ve Peleg, 2005).

5. HAREKET BULANIKLIĞI AZALTMA YÖNTEMLERİ

Bu bölümde görüntü bulanıklığının azaltılmasında / önlenmesinde kullanılan klasik, öncel (prior) tabanlı ve yapay zekâ temelli algoritmalar incelenmiştir. Performansıya ön plana çıkan algoritmalar detaylandırılarak, önerilen görüntü bulanıklığına dayanıklı vSLAM çerçevesinin testlerinde kullanılmıştır. Bulanıklaştırma çekirdeğinin önceden bilinip bilinmemesine bağlı olarak, görüntü bulanıklığı azaltma metotları iki kategoriye ayrılmaktadır.

Kör olmayan görüntü bulanıklığının azaltılması (Hansen, Nagy ve O'leary, 2006), onlarca yıldır üzerinde çalışılan bir konu olup, bilinen bulanıklık çekirdeği ile gizli keskin görüntüyü restore etmeyi sağlamaktadır. Restore edilmiş görüntüyü elde edilmesi esnasında, gerçekleştirilen dekonvolüsyon işlemi gürültüye karşı kararsız olduğu için; az miktarda gürültü bile ciddi bozulmalara yol açmaktadır. Bu sorunu çözmek için klasik yöntemler arasında Wiener filtresi (Tukey, 1952) ve Richardson-Lucy algoritması (Lucy, 1974; Richardson, 1972) bulunmaktadır. Son yıllarda, kör olmayan görüntü bulanıklaştırma bir optimizasyon problemi olarak modellenmiştir. Toplam varyasyon (TV) (Rudin, Osher ve Fatemi, 1992) ve seyrek önceller (Levin ve diğerleri, 2007; Krishnan ve Fergus, 2009) gibi birçok ünlü görüntü önceli, görüntü restorasyonundaki hatalı pozlamayla başa çıkmak ve görüntü düzenleme yapmak için ortaya konulmuştur.

Hemen hemen tüm geleneksel enerji optimizasyon temelli yöntemler (Hirsch, Schuler, Harmeling ve Schölkopf, 2011; Krishnan, Tay ve Fergus, 2011; Whyte, Sivic ve Zisserman, 2014) dinamik bulanıklığını iteratif olarak azaltma yöntemine odaklanmışlardır. Bu iterasyonlarda, verimli şekilde düşük ölçeklendirilmiş resimlerde, ölçeklendirme arttırılarak, orijinal ölçeğe ulaşılmaktadır. Öğrenme temelli olarak ortaya çıkan ilk yöntemlerde bulanıklık çekirdeğini (blur kernel) tahmin eden (J. Sun, Cao, Xu ve Ponce, 2015; Schelten, Nowozin, Jancsary, Rother ve Roth, 2015) ve kör olmayan dekonvolüsyon (J. Zhang, Pan, Lai, Lau ve Yang, 2017a) gerçekleştirilen ya da resim hakkında öncül bilgilere (Sreehari, 2016; K. Zhang, Zuo, Gu ve Zhang, 2017b) sahip olunan sinir ağları içeren yöntemler mevcuttur. Fakat bu ilk yöntemler, hala yinelemeli-optimizasyon çerçevesi altında, varsayımsal modeller kullanan algoritmalarlardır.

Kör görüntü bulanıklık azaltma (Almeida ve Almeida, 2009; Levin, Weiss, Durand ve Freeman, 2011) önceden bilinen bir bulanıklık çekirdeği bilgisi olmadan gizli bir keskin görüntüyü restore etmektir. Geleneksel gradyan tabanlı önceller doğal görüntülerde başarısız olma eğilimindedir (Levin ve diğerleri, 2011), çünkü Fourier alanında çoğunlukla düşük frekanslı bulanık görüntüler yer almaktadır. Düzenleştirme (regularization) ve optimizasyonun ilerlemesiyle, bu sorunu çözmek için çerçevicik (framelet) tabanlı öncel (Cai, Ji, Liu ve Shen, 2011), l_0 -norm tabanlı öncel (Pan, Hu, Su ve Yang, 2016a; L. Xu, Zheng ve Jia, 2013) , seyrek kodlamalı öncel (L. Sun, Cho, Wang ve Hays, 2013) , düşük sıralı öncel (Ren, Cao, Pan, Guo, Zuo ve Yang, 2016) ve karanlık kanal önceli (Pan, Sun, Pfister ve Yang, 2016b) gibi yöntemler örnek gösterilebilir. Ancak, her birinin kendi eksiklikleri vardır. Çerçevicik (framelet) dayalı öncel (Cai ve diğerleri, 2011), heterojen bulanıklaştırma senaryolarını işlemek için yeterince genel olmayan, elle hazırlanmış çerçevelerin işlevlerine dayanmaktadır. l_0 -norm önceli de (Pan ve diğerleri, 2016a; L. Xu ve diğerleri, 2013) doğası gereği kombinasyoneldir (dolayısıyla dışbükey değildir) ve l_1 -norm'da dışbükey gevşemesi, optimum performans için hassas parametre ayarlaması gerektirmektedir. Seyrek kodlamaya dayalı öncel (L. Sun ve diğerleri, 2013), eğitim seti ile hedef görüntü arasında pratikte doğru olmayabilecek istatistiksel benzerliği varsaymaktadır. Düşük sıralı öncel (Ren ve diğerleri, 2016), genel olarak $O(N^3)$ karmaşıklığına sahip olan maliyetli tekil değer ayrıştırma (SVD: singular value decomposition) çözümünü gerektirmektedir. Karanlık kanal önceli (Pan ve diğerleri, 2016b), l_0 -norm ve doğrusal olmayan karanlık kanal hesaplamasını birleştiren kapsamlı bir önceldir, bu nedenle de yüksek karmaşıklık içermektedir.

Doğal görüntülerde önceller (priors), görüntü elde edilirken (pozlandırma zamanı içinde) ortaya çıkan bulanıklık ve benzeri bozuklukları bastırmak ve kaliteyi artırmak için tasarlanmıştır. Bu yöntemlere yukarıda bahsedilenlere ek olarak toplam varyasyon (Chan ve Wong, 1998), ağır-kuyruklu gradyan önceli (Shan, Jia ve Agarwala, 2008), hiper-Laplacian önceli (Krishnan ve Fergus, 2009) gibi yöntemler de eklenebilir. Bu yöntemlerin kullandığı ölçeklendirme yaklaşımına kabadan inceye yaklaşım ismi de verilmektedir. Bu yaklaşımda bir piramit içinde farklı ölçeklerde, kademeli restorasyon gerçekleştirilmektedir. Bunlara ek olarak frekans alanında uygulanan limitli çalışmalar da mevcuttur (Delbracio ve Sapiro, 2015).

Daha sonrasında ortaya çıkan öğrenme temelli yöntemler ile uçtan uca, çoklu katman yapısıyla daha verimli sonuçlar ortaya konulmuştur. Evrimsel Sinir Ağları (CNN: Convolutional Neural Network) kullanarak özellikle resimde yer alan yazıları restore eden (Hradiš, 2015) ve Çekişmeli Üretici Ağlar (GAN: Generative Adversarial Networks) kullanan (Kupyn, 2018) resim hareket bulanıklığını azaltan / yok eden yöntemler (deblurring) de geliştirilmiştir.

Görüntüde yer alan bulanıklığın şiddetini ve yönünü tahmin eden metotlar (J. Sun, Cao, Xu ve Ponce, 2015) ve birden çok CNN'i bir araya getirerek, kabadan inceye yaklaşımına uygun şekilde, yinelemeli bir optimizasyon gerçekleştiren yöntemler bulunmaktadır (Schuler, Hirsch, Harmeling ve Schölkopf, 2015). Ayrıca dekonvolüsyon çekirdeğini tahmin eden başka bir çalışma da ilgili veri kümesinde başarılı olmuştur (Chakrabarti, 2016). Çoklu katman kullanarak, görüntünün farklı ölçeklerdeki halinin derin bir ağ ile ilerlemeli şekilde görüntü restorasyonunu gerçekleştiren (Nah, Hyun Kim ve Mu Lee, 2017) çalışması da bir diğer örnek çalışmadır.

5.1. Lucy-Richardson Algoritması

Gürültü türüne bağlı olmaması ve yinelemeli bir algoritma olması gibi birçok nedenden dolayı en iyi bilinen bulanıklaştırma algoritmalarından biridir. Bu algoritma ile gerçekleştirilen işlem genel olarak Eş. 5.1'deki gibi formülize edilmektedir (Lucy, 1974, Richardson, 1972):

$$f^{n+1} = f^n H\left(\frac{g}{Hf^n}\right) \quad (5.1)$$

Burada, f^{n+1} bir önceki f^n 'den gelen yeni yaklaşım, g kameradan alınan bulanık görüntü, n yineleme sayısıdır. H ile ifade edilen fonksiyon nokta yayılım fonksiyonudur. İlk yinelemede, f^n bulanık görüntü g ile aynıdır. Poisson istatistiğine göre bulanık görüntü ile tahmin edilen görüntü arasındaki farkı en aza indirilmektedir. Bu algoritmanın dezavantajı, sonuç olarak restore edilmiş görüntüdeki yapay halka oluşumları ve yineleme nedeniyle işlemin fazla zaman almasıdır.

5.2. Wiener Filtresi Algoritması

Görüntü bulanıklığı azaltma amacıyla kullanılan en yaygın tekniklerden biri Wiener filtrelemedir. Wiener filtresi, doğrusal hareket veya odaklanmamış optiklerin neden olduğu görüntülerdeki bulanıklığı gidermek için büyük bir yeteneğe sahiptir. Bulanık görüntü, kötü örneklemenin bir sonucu olarak görülebilmektedir. Görüntüdeki her piksel, kameranın önündeki tek bir sabit nokta için yoğunluk değeri içermelidir. Ne yazık ki, kamera hareket ettirilirse veya deklanşör hızı çok yavaşsa, belirli bir piksel, kameranın hareket çizgisi boyunca bulunan noktalardan gelen yoğunlukların bir karışımı olmaktadır (Singh, 2013). Wiener filtresi, ek gürültüyü ortadan kaldırdığı ve aynı anda bulanıklığı tersine çevirdiği için ters filtreleme ve gürültü yumuşatma arasında en uygun dengeyi sağlamaktadır. Wiener Filtresi Algoritması'nın çalışması Eş. 5.2. – Eş. 5.5. denklemlerinde açıklandığı gibi çalışmaktadır (El-Henawy, Amin ve Kareem Ahmed, 2014).

$$b(x,y)=h(x,y)*s(x,y)+n(x,y) \quad (5.2)$$

Bu denklemde $s(x, y)$ bilinmeyen keskin görüntü, $h(x, y)$ lineer ölçekte nokta yayılım fonksiyonudur. $n(x, y)$, $s(x, y)$ 'den bağımsız büyüklüğü bilinmeyen ek gürültü ve $b(x, y)$ gözlenen bulanık görüntüdür. Wiener ters konvolüsyon filtresi $g(x, y)$ 'yi bulur ve aşağıdaki Eş 5.3'de ifade edildiği gibi $s(x, y)$ 'yi tahmin etmek için kullanır.

$$\hat{s}(x,y)=g(x,y)*b(x,y) \quad (5.3)$$

Burada $\hat{s}(x, y)$, ortalama kare hatasını en aza indiren bir $s(x, y)$ tahminidir. Filtre, frekans alanında çalışır. İlk adım, Eş 5.4.'de ifade edildiği gibi $g(x, y)$ 'nin frekans alanı versiyonunu hesaplamaktır.

$$G(u,v) = \frac{H(u,v) W(u,v)}{|H(u,v)|^2 W(u,v) + N(u,v)} \quad (5.4)$$

$G(u,v)$, g 'nin Fourier dönüşümü; $H(u,v)$, h 'nin Fourier dönüşümü; $W(u,v)$, $s(x,y)$ 'nin ortalama spektral güç yoğunluğu; $N(u,v)$, $n(x,y)$ ile ifade edilen gürültünün ortalama güç spektral yoğunluğudur. Son olarak, filtreleme, Eş.5.5'de ifade edildiği gibi frekans alanında gerçekleştirilmektedir.

$$\hat{S}(u,v)=G(u,v)B(u,v) \quad (5.5)$$

Burada $\hat{S}(u,v)$ tahmini keskin görüntünün ve $B(u,v)$ gözlemlenen bulanık görüntünün Fourier dönüşümüdür. Wiener algoritmasının ters filtrelemeye dayalı olduğu görülmektedir. Bu algoritmanın bilinen dezavantajları şunlardır (El-Henawy ve diğerleri, 2014).

- H nokta yayılım fonksiyonu bilinmediği durumda belirli bir görüntü için tahmin edilebilir, ancak iyi bir H tahmini yapabilmek için çok sayıda deneme ve çaba gerekir.
- Bazı durumlarda (gürültünün çok yüksek olduğu durumda) başarısız olur çünkü sinüs fonksiyonu bazı x ve y değerlerinde 0'a eşittir.

5.3. Kör Dekonvolüsyon Yaklaşımı

Kör dekonvolüsyon yaklaşımı, nokta yayılım fonksiyonunun doğru bir tahminini gerçekleştirmeyi amaçlamaz. Bunun yerine, bozulma fonksiyonunun bir başlangıç tahmini hesaplanmaktadır. Daha sonra sistem, ortalama kare hatasını en aza indiren doğru bir tahmin elde edilene kadar yinelemeli bir kör dekonvolüsyon süreci gerçekleştirmektedir. Kör dekonvolüsyon için iki temel yaklaşım vardır. Birincisi izdüşüm tabanlı kör ters evrişim olarak adlandırılır, ikincisi ise maksimum olabilirlik restorasyonudur. İlk yaklaşım, nokta yayılma fonksiyonunun (PSF) ilk değer atamasını ve ardından gerçek görüntünün ilk tahminini yapmaktadır. Daha sonra, önceden tanımlanmış bir yakınsama kriteri karşılanana kadar süreç tekrarlanmaktadır. Bu yaklaşımın avantajları, destek boyutundaki yanlışlıklara karşı dayanıklı olması ve gürültüye karşı duyarsız olmasıdır. Bu yaklaşımın dezavantajları, benzersiz olmaması ve uygun olmayan yerel minimumlar başlatılırsa hatalara yol açabilmesidir. İkinci yaklaşım, PSF ve kovaryans matrisi gibi parametreleri tahmin etmek için maksimum olasılığı kullanmaktadır. PSF tahmini benzersiz olmadığından, yaklaşım boyut, simetri ve diğer parametreler gibi parametreleri dikkate alabilmektedir. Bu yaklaşımın avantajları, düşük hesaplama karmaşıklığıdır. Yaklaşım, keskin görüntünün bulanıklığını, gürültüsünü ve güç spektrumlarını elde edebilir. Bu yaklaşımın dezavantajı, tahmini maliyet fonksiyonunun yerel minimumlarına yakınsamasıdır (El-Henawy ve diğerleri, 2014).

5.4. Görüntü Bulanıklığı Azaltma için Derin Ölçek Yinelemeli Ağ

Görüntü bulanıklığı azaltma, bilgisayarla görme ve görüntü işleme alanlarında uzun zamandır çalışılan, kamera odağı, hareket, titreşim gibi nedenlerden kaynaklanan bulanıklığın azaltılması ve resimde yer alan kenar yapılarının ortaya çıkarılarak keskin bir görüntü elde edilmesi metodudur.

Tek resim kullanılarak bulanıklığın azaltılması, oldukça zorlu bir problemdir. Geleneksel metotlarda çoğunlukla bulanıklık seviyesinin modellenmesi (düzgün/dağınık vb.) ve resim hakkında önceden elde olan verileri kullanan (Chan ve Wong, 1998; Cho ve Lee, 2009; Goldstein, 2012; Pan, Hu, Su ve Yang, 2014) yöntemler yer almaktadır. Bu yöntemlerin ortak özelliği yinelemeli olarak parametrelerin değiştirildiği daha yüksek hesaplama maliyetli yöntemler olmasıdır. Ayrıca, kullanılan bulanıklık modelleri çoğunlukla gerçek dünyadaki karmaşık bulanıklıklar ile uyuşmamaktadır.

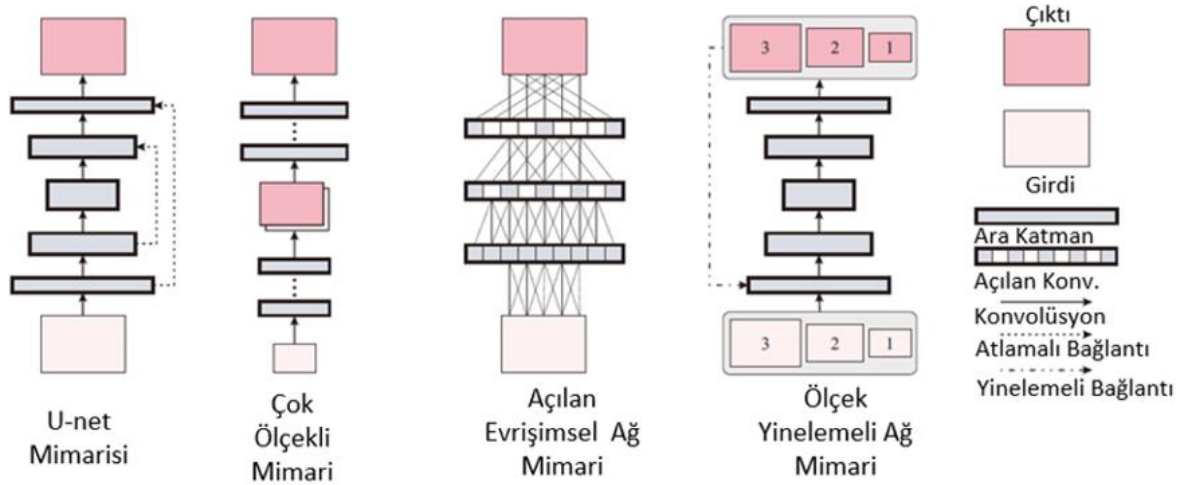
Öğrenme temelli yöntemlerde ise ilk ortaya konulan çalışmalarda, görüntü dışındaki verilerin de kullanılabilmesi için, geleneksel metotlardaki bazı modüller öğrenilen parametreler ile değiştirilmiştir (Schuler ve diğerleri, 2015; Xiao, Wang, Heidrich ve Hirsch, 2016). Daha sonraki yöntemlerde ise video (Nah ve diğerleri, 2017) ve görüntüler (Su, Delbracio, Wang, Sapiro, Heidrich ve Wang, 2017, Wieschollek, 2017) için uçtan uca öğretilebilir ağ yapısının kullanımına başlanmıştır. Çok ölçekli evrimsel sinir ağının kullanıldığı (Nah ve diğerleri, 2017) çalışmada oldukça başarılı sonuçlar elde edilmiştir. Çalışmada bulanık görüntünün çok kaba bir ölçeğinden başlanarak, tam çözünürlüğe ulaşılan kadar, görüntünün restorasyonu aşamalı olarak gerçekleştirilmektedir. Bu yaklaşım aslında geleneksel metotlardaki çoklu yinelemeli yapıya oldukça benzerdir.

Ölçek yinelemeli ağda (scale-recurrent structure) ve çok ölçekli yöntemlerde, her ölçekte aynı bulanıklık probleminin çözülmesi hedeflendiği için, çözücü ve karşılık gelen parametreler genellikle aynıdır. Gerçekleştirilen çalışmalarda, her ölçek için farklı parametrelerin kullanımı kısıtlayıcı olmayan çözüm uzayında farklı problemlere neden olabilmektedir. Diğer bir sorun, girdi olan görüntülerin, farklı çözünürlükte olduğu durumda, eğer parametrelerin her ölçek için değiştirilmesine izin verilirse, öğrenme aşamasında çözüm tek bir çözünürlüğe ya da hareket ölçeğine yakınsayabilmektedir. Derin ölçek yinelemeli ağ (SRN) (Tao, Gao, Shen, Wang ve Jia, 2018) yönteminde ise ağ

ağırlıkları ölçekler arasında paylaştırılarak bu sorun bertaraf edilmiştir. Bu yaklaşım ile SRN yönteminde;

1. Eğitilebilir parametre sayısı, önemli ölçüde azaltılmıştır. Aynı eğitim verileri ile çalışılsa bile, veri sayısı ile parametre sayısının çarpımından gelen işlem yükü azaltılmış olur.
2. SRN yapısında ayrıca her ölçekte görüntü restorasyonu gerçekleştirilerek, resmin daha kaliteli bir versiyonu oluşturulur.

SRN yönteminde yer alan kodlayıcı-çözücü ResBlock ağı, SRN yöntemine kadar olan benzer yapılardan daha üstün performans gösteren ve öğrenme fizibilitesini arttıran bir ağıdır. SRN çalışmasındaki parametre sayısının azaltılması ve her ölçekte görüntü restorasyonu ile uçtan uca bir görüntü bulanıklık azaltımı çalışmasının gerçekleştirilmesi öğrenme verimliliğini 4 katına (öğrenme zamanı benzer restorasyon algoritmalarına göre 4'te 1 ine düşmüştür) çıkarmıştır. Eğitilebilir parametre sayısı 3'te 1'ine düşürülmüştür. Görüntü işleme uygulamalarında kullanılan farklı evrişimsel sinir ağı mimarileri aşağıdaki Şekil 5.1'de verilmiştir (Tao ve diğerleri, 2018).



Şekil 5.1. Görüntü İşleme uygulamalarından farklı evrişimsel sinir ağı mimarileri

Bulanıklık azaltma (deblurring) yöntemlerinde hem restorasyon kalitesinin artırılması hem de farklı durumlara uyum sağlamaya yönelik, birçok çalışma yapılmış, farklı uygulamalar önerilmiştir (Shan ve diğerleri, 2008). Doğal görüntü öncelleri (priors), görüntü elde edilirken (pozlandırma zamanı içinde) ortaya çıkan bulanıklık ve benzeri bozuklukları

bastırmak ve kaliteyi artırmak için tasarlanmıştır. Bu yöntemler toplam varyasyon (Chan ve Wong, 1998), seyrek görüntü önceli (Levin ve diğerleri, 2009), ağır-kuyruklu gradyan önceli (Shan ve diğerleri, 2008), Hiper-Laplacian önceli (Krishnan ve Fergus, 2009), l_0 -norm gradyan önceli (L. Xu, Zheng ve Jia, 2013) gibi yöntemlerdir. Bu yöntemlerin birçoğu kabadan inceye bir yaklaşımı takip etmektedir. Bu yaklaşımda bir piramit içinde farklı ölçeklerde, kademeli restorasyon gerçekleştirilmektedir. Bunlara ek olarak frekans alanında uygulanan limitli çalışmalar da mevcuttur (Delbracio ve Sapiro, 2015).

SRN- Ağ Mimarisi

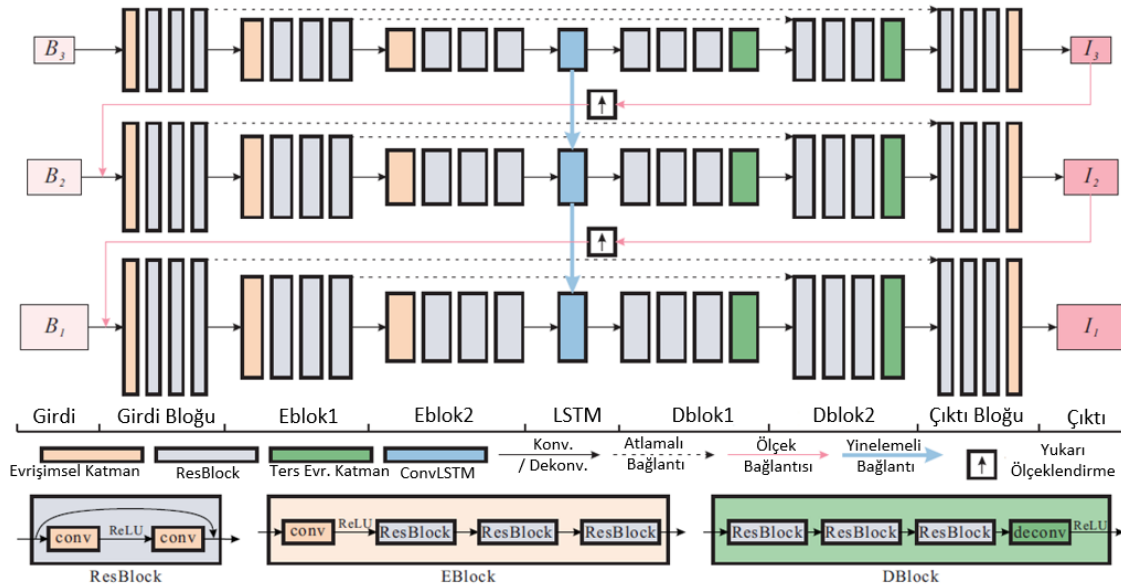
SRN çalışmasında SRN-DeblurNet adıyla bir mimari önerilmiştir. Bu mimaride, girdi olarak alınan bulanıklık görüntüler, alt örnekleme yoluyla düşük ölçeklere çevrilir. Her katmanda daha keskin, daha az bulanık görüntüler elde edilmektedir. Resmin tam çözünürlüğüne kadar restorasyon kademeli olarak gerçekleştirilmektedir. Tam çözünürlükteki keskinliği yüksek görüntü, istenilen restore edilmiş görüntü çıktısı olur. SRN ağını aşağıdaki eşitlik ile ifade edebiliriz:

$$\mathbf{I}^i, \mathbf{h}^i = \text{NET}_{\text{SR}}(\mathbf{B}^i, \mathbf{I}^{i+1\uparrow}, \mathbf{h}^{i+1\uparrow}; \theta_{\text{SR}}) \quad (5.6)$$

Bu denklemde i , ölçek indeksidir ve $i=1$ için en ince ölçeği ifade eder. $\mathbf{B}^i$ ve $\mathbf{I}^i$ ifadeleri, i 'inci ölçek için sırasıyla bulanık ve o restore edilmiş gizli görüntüleri ifade etmektedir. NET_{SR} ise ilgili yöntemde önerilen ölçek yinelemeli ağı ifade etmektedir. Bu denklemde, θ_{SR} SRN'in öğrenme parametresidir. Ağ yinelemeli olduğunda, gizli durum $\mathbf{h}^i$ ölçek arasında bilgi aktarmaktadır. Gizli durum, önceki daha kaba ölçekten, görüntü yapılarını ve çekirdek bilgilerini yakalamaktadır. Operatör $(.)^\uparrow$ ise, öznitelikleri veya görüntüleri $(i+1)$. ölçekten, (i) . ölçeğe aktarmak için kullanılan operatördür.

Bu denklemden, ağ yapısının oldukça esnek olduğu görülebilir. Örneğin, yinelemeli ağ çeşitli formlarda olabilir. Vanilla RNN ya da uzun-kısa süreli hafıza (LSTM) ya da kapılı tekrarlayan birim (GRU: gated recurrent unit) ağ çeşidi olarak seçilebilir. İkinci olarak, $(.)^\uparrow$ operatörü yerine de olası başka operatörler de seçilebilir. Bu operatör yerine dekonvolüsyon (ters evrişim) katmanı, alt-piksel evrişim (konvolüsyon) katmanı ya da

görüntü yeniden boyutlandırma da seçilebilmektedir. SRN ağı mimarisinde, her ölçekte, ağın optimum verimlilik için, uygun şekilde tasarlanması gerekmektedir.



Şekil 5.2. SRN: Ölçek Yinelemeli Sinir Ağı Mimarisi

Kodlayıcı- Çözücü ResBlock'ları

Kodlayıcı-çözücü ağı, önce girdi verilerini aşamalı olarak daha küçük uzamsal boyutlara ve daha fazla kanala sahip öznitelik haritalarına çeviren (kodlayıcı tarafında) ve daha sonra bu haritaları girdinin şekline çeviren (çözücü tarafında) simetrik CNN yapılarını ifade etmektedir. Öznitelik haritaları arasındaki atlama bağlantıları, farklı seviyelerdeki bilgilerin bağlanmasını sağlamaktadır. Ayrıca gradyen yayılımından faydalanabilmekte ve yakınsamayı hızlandırabilmektedir. Kodlayıcı, evrişim (convolution) katmanlarının birkaç tanesini içerebilmektedir. Çözücü tarafında ise bir dizi ters evrişim katmanı (deconvolution layers) veya yeniden boyutlandırma kullanılmaktadır (Liu, Yeh, Tang, Liu ve Agarwala, 2017; Delbracio ve Sapiro, 2015). Ağ derinliğini arttırmak için her seviyeden sonra ek evrişim katmanları eklenmektedir.

Kodlayıcı-kod çözücü yapısının birçok görüş görevinde etkili olduğu kanıtlanmıştır (Liu ve diğerleri, 2017; Delbracio ve Sapiro, 2015; Tao, Gao, Liao, Wang ve Jia, 2017; N. Xu, Price, Cohen ve Huang, 2017). Ancak, aşağıdaki hususlar göz önüne alındığında, doğrudan kodlayıcı-kod çözücü ağını kullanmak bulanıklık azaltma işlemleri için en iyi seçim

değildir. İlk olarak, bulanıklık giderme görevi için, algılayıcı alanın şiddetli hareketi işlemek için büyük olması gerekir, bu da kodlayıcı/kod çözücü modülleri için daha fazla seviye eklenmesine neden olmaktadır. Ancak bu strateji, çok sayıda ara özellik kanalı ile parametre sayısını hızla arttırdığından pratikte önerilmemektedir. Ayrıca, orta öznitelik haritasının uzamsal boyutu, yeniden yapılandırma ve uzamsal bilgiyi tutmak için çok küçük olacaktır.

İkinci olarak, kodlayıcı/kod çözücü modüllerinin her seviyesine daha fazla evrişim katmanı eklemek ağın yakınsamasını (her seviyede düz evrişim ile) yavaşlatacaktır. Bu nedenle önerilen SRN yapısında, içinde gizli durumlar (hidden states) içeren yinelemeli modüllerin kullanımı gerekli olmuştur.

Kodlayıcı/kod çözücü ResBlock

Kodlayıcı-kod çözücü ağlarını SRN çerçevesine uyarlamak için çeşitli değişiklikler yapılmıştır. İlk olarak, artık öğrenme bloklarını kullanarak, kodlayıcı/kod çözücü modüllerini iyileştirilmiştir (He, Zhang, Ren ve Sun, 2016). SRN metodunda gerçekleştirilen deneylere dayanarak, ResNet (He ve diğerleri, 2016) çalışmasındaki orijinali yerine ResBlocks kullanımı seçilmiştir. Şekilde de gösterildiği üzere önerilen Kodlayıcı ResBlock'ları (EBlock'lar), bir evrişim katmanı ve ardından birkaç farklı ResBlock içermektedir.

Evrişim katmanı 2 adımdan oluşur. Önceki katmanın çekirdek sayısını iki katına çıkarmaktadır ve öznitelik haritalarını yarı boyuta küçültmektedir. Her bir ResBlock'u 2 evrişim katmanı içerir. Ayrıca, tüm evrişim katmanları aynı sayıda çekirdeğe sahiptir. Çözücü ResBlock'ları (DBlocks), EBlock'lara simetriktir.

SRN çerçevesi, ResBlock'larının ardından bir dekonvolüsyon katmanı içermektedir. Dekonvolüsyon katmanı, özellik haritalarının uzamsal boyutunu iki katına çıkarmak ve kanalları yarıya indirmek için kullanılmaktadır.

Ölçek yinelemeli yapıda, ağ içinde tekrarlayan modüller tasarlanmıştır. Derin video süper çözünürlük uygulamasının (Tao ve diğerleri, 2017) stratejisine benzer şekilde, ardışık ölçekleri bağlamak için gizli durum için darboğaz katmanına evrişim katmanları

eklenmektedir. Son olarak, her evrişim katmanı için 5×5 boyutunda büyük evrişim çekirdekleri kullanılmaktadır. Değiştirilen ağ şu şekilde ifade edilir:

$$f^i = \text{NET}_E(B^i, I^{i+1\uparrow}; \theta_E) \quad (5.7)$$

$$h^i, g^i = \text{ConvLSTM}(h^{i+1\uparrow}, f^i; \theta_{\text{LSTM}}) \quad (5.8)$$

$$I^i = \text{NET}_D(g^i; \theta_D) \quad (5.9)$$

Burada NetE ve NetD kodlayıcı ve kod çözücü CNN'lerdir, θ_E ve θ_D ise sırasıyla bu CNN'lere ait parametrelerdir. NET_E ve NET_D 'de sırasıyla EBlock'ların ve DBlock'ların 3 aşaması kullanılmaktadır.

θ_{LSTM} , ConvLSTM'deki parametreler kümesidir. Gizli durum h^i , bir sonraki ölçeğe aktarılan ve ince ölçek problemine fayda sağlayan ara sonuç ve bulanıklık desenleri hakkında faydalı bilgiler içermektedir. SRN yöntemi 3 ölçek içerir. Burada $(i + 1)$ ölçeği, i. ölçeğinin yarısı büyüklüğündedir. Kodlayıcı/kod çözücü ResBlock ağı için, şekilde gösterildiği gibi 1 InBlock, 2 EBlock, ardından 1 Convolutional LSTM bloğu, 2 DBlock ve 1 OutBlock vardır. InBlock, 32 kanallı bir öznitelik haritası üretir. OutBlock, önceki öznitelik haritasını girdi olarak alır ve çıktı görüntüsü oluşturmaktadır. Her EBlock/DBlock içindeki tüm evrişim katmanlarının çekirdek sayıları aynıdır. EBlock'lar için çekirdek sayısı sırasıyla 64 ve 128'dir. DBlock'lar için bunlar 128 ve 64'tür. EBlock'larda ve evrişim katmanlarında, evrişim katmanı için adım boyutu 2, diğerleri ise 1'dir. Doğrultulmuş lineer birim (ReLU), tüm katmanlar için etkinleştirme işlevi olarak kullanılır ve tüm çekirdek boyutları 5'e ayarlanmaktadır.

Kayıplar

Her ölçek için, ağ çıktısı ile temel gerçek (ground truth) arasında (bilineer enterpolasyon kullanılarak aynı boyuta alt-örneklenmiş) Öklid kaybı kullanılmaktadır.

$$L = \sum_{i=1}^n \frac{\kappa_i}{N_i} |I^i - I^*_i|_2^2 \quad (5.10)$$

I^i ve I^i_* sırasıyla i ölçeğinde ağ çıktısı ve temel gerçektir. $\{\kappa_i\}$ her ölçeğin ağırlıklarıdır. Yapılan SRN çalışmasında $\kappa_i = 1.0$ olarak belirlenmiştir. N_i , I^i 'deki normalize edilecek elementlerin sayısıdır.

5.5. Görüntü Bulanıklığı Azaltma için Yinelemeli Üst Ölçekleme Ağı

Görüntü Bulanıklığı Azaltma için Yinelemeli Üst Ölçekleme Ağı (SIUN) (Ye, 2020) metodu, resim karelerinin ölçekleri (scales) arasında ağırlık paylaşımının avantajlarını kullanmaktadır. Ayrıca metodun farklı resimlere uygulanabilmesi için daha esnek eğitim ve tahmin için farklı iterasyonlar gerçekleştirmek için altyapı da sağlamaktadır. Her bir ağ katmanında çözünürlüğü arttırmak yerine süper çözünürlük (super resolution) yapısı getirilerek, detayları içeren bir restore edilmiş resim elde edilmektedir.

Mevcut bulanıklık azaltma algoritmaları, daha önce de bahsedildiği üzere geleneksel iteratif optimizasyon algoritmaları ve öğrenme temelli algoritmalar olmak üzere ikiye ayrılmaktadır.

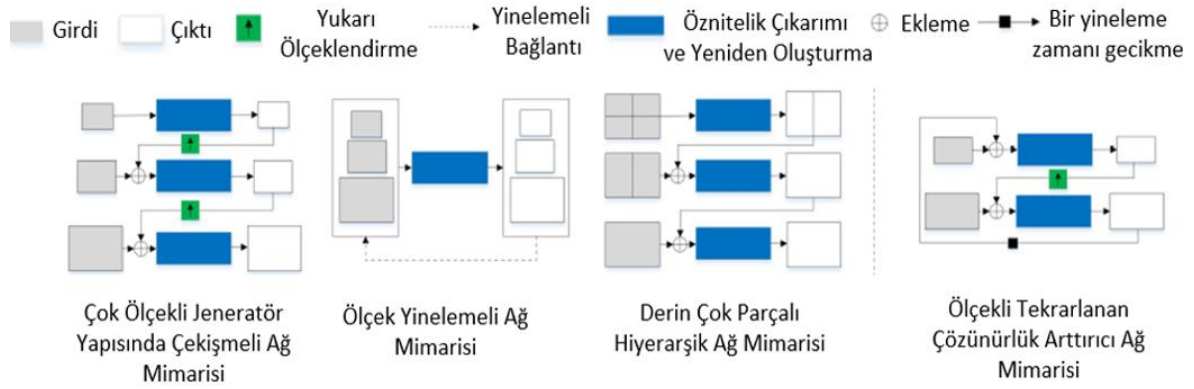
Geleneksel metotlar, bulanıklık problemini, bulanıklığın karakterini ortaya çıkarmaya ve bazı öncül varsayımlar ile sonuca varmaya çalışmaktadırlar. Fakat gerçek dünya senaryolarında, meydana gelen bulanıklık çoğunlukla, varsayım modelinden daha kompleksir.

Karmaşık doğal bulanıkları çözmek için, öğrenme temelli yöntemler önerilmiştir. Bu alanda (Schuler, Burger, Harmeling ve Schölkopf, 2013) çalışmasında 2 adımlı bir yöntem önerilmektedir. Bu çalışmada ikinci adımda doğal resimler içeren geniş bir veri kümesinde sinir ağı kullanılarak öğrenme sağlanmıştır. (J. Zhang ve diğerleri, 2017a) çalışmasında evrimsel ağ öğretilerek, çoklu bir çerçevede bulanık resimlerin dekonvolüsyonu ile iteratif olarak restorasyonu sağlanmıştır. Bir diğer çalışmada (L. Xu, Ren, Liu ve Jia, 2014) 2 modüllü, derin evrimsel ağ geliştirilmiştir. Bu çalışmalar öğrenme tabanlı metotlar arasında klasik öğrenme yöntemleri içeren metotlar olarak değerlendirilmektedir.

İlerleyen dönemlerde daha gelişmiş ve karmaşık ağ modelleri ortaya konulmuş, kör bulanıklık azaltma (blind deblurring) için uçtan uca derin öğrenme metotları ortaya

konulmuştur. Çok ölçekli jeneratör yapısında çekişmeli ağ kullanılan önerilen metotta (Nah ve diğerleri, 2017) daha ince ölçekte görüntü bulanıklık azaltma, daha kalın ölçekteki öznetelikler tarafından desteklenmektedir. Böylece görünür olmayan resim, $\frac{1}{2}k$ ($k=2$) ölçeğinden, orijinal boyutlarına ulaşmaya kadar adım adım oluşturulmaktadır. Fakat bu basamaklı yapı, ağ boyutunu ve derinliğine istenmeyen derecede arttırmaktadır. Bu durumu çözmek için, (Tao ve diğerleri, 2018) uzun kısa süreli bellek (LSTM: long short term memory) ile ölçekli tekrarlayan sinir ağları (SRN: scale recurrent network) kullanılarak ölçekler arası ağırlık paylaşımı sağlanmıştır. Bu şekilde ağ parametreleri, önemli ölçüde azaltılabilmektedir. Bu yöntemlerde daha fazla işlem gücü gereksinimi mevcuttur. İşlem gücünü azaltmaya yönelik ortaya konulan diğer çalışmada (H. Zhang, Dai, Li ve Koniusz, 2019), derin çok parçalı hiyerarşik ağ (DMPHN: deep multi-patch hierarchical network) yapısı kullanılmıştır. İlgili çalışmada çok parçalı hiyerarşide, resmin çoklu parçaları giriş olarak alınıp, üst seviyelerde tek bir restore edilmiş resim oluşturulmaktadır. Bahsi geçen öğrenme temelli metotlar, en ön plana çıkan, son teknoloji metotlar olup hem öğrenme hem de tahmin sürecinde katman sayısı belli olan, çoklu ölçekli çerçeve kullanmışlardır.

Ölçekli tekrarlanan çözünürlük arttırıcı ağda (Scale-iterative Upscaling Network- SIUN) ise, küçük ölçeklerden başlanarak iteratif olarak, her iterasyonda ölçek arttırımı yapılarak, tüm çözünürlükte resim restorasyonu gerçekleştirilene kadar devam edilmektedir. Bu sistemde de SRN'e (Tao ve diğerleri, 2018) benzer şekilde ağ parametreleri ve katmanlar arası paylaşımı azaltılmaktadır. Belirlenmiş sayıda katman bulunmadığı için, eğitim ve tahmin aşamalarında daha esnek bir yapı mevcuttur ve farklı çözünürlüklerdeki resimler için kullanılabilir. Ayrıca yöntemde yer alan süper çözünürlük mimarisi sayesinde, resimde yer alan daha fazla yapının çıkarılması da mümkün hale gelmiştir. Süper çözünürlük yapısı ile, özellikle yüz ve yazı içeren resimlerde çözünürlük iyileştirmeleri daha etkili olmaktadır. Bahsi geçen yöntemlere ait mimarilerin görselleri aşağıda Şekil 5.3'te verilmiştir (Ye, 2020):



Şekil 5.3. Yapay zekâ temelli algoritmalarının ağ mimarileri

CNN temelli çok ölçekli mimarilerde, özellikle ardışık ölçeklemeli yapılar kullanıldığı için ağ büyüklüğü ve derinliği çok fazla büyümektedir. Bu nedenle LSTM temelli SRN modeller oluşturulmuş ve ağırlıkların ölçeklendirmeler arasında paylaşımı sağlanabilmektedir. Bu yöntemlerde ise hesaplama maliyetleri ciddi oranda artmaktadır. Hesaplama maliyetlerini azaltmak amacıyla ortaya konulan DMPHN yöntemlerinde ise bulanıklığın azaltımına oldukça küçük bir ölçek ile başlanmaktadır. Bu yöntemlerde de ağ tasarımı bir kez yapıldığında, konfigürasyon değişiklikleri mümkün olmamaktadır. Özellikle farklı bulanıklık seviyelerine sahip görüntülerin yer aldığı veri kümelerine uygulanamamaktadır. Bu durumu ortadan kaldırmak için çoklu katman yapısı ve iteratif yöntemlerin birleştirildiği SIUN yöntemi ortaya konulmuştur.

SIUN yönteminde “kodlayıcı çözücü ağı” ve “süper çözünürlük” ağı isimli iki farklı ağ mevcuttur.

Kodlayıcı-Çözücü Ağı

Kodlayıcı çözücü mimarisi, bir çeşit sinir ağı tasarım desenidir. Sinir ağı tabanlı makine çevirisi ve diziden diziye tahmin uygulamalarında yaygın olarak kullanılmaktadır. Son dönemlerde bilgisayar ile görme uygulamalarında da başarı göstermektedir. Kodlayıcı-çözücü ağları, isminden de anlaşılacağı üzere kodlayıcı ve çözücü isimlerinde iki ayrı simetrik parçadan oluşmaktadır. Bu bölümlerden kodlayıcı, resimden resme görevlerde, giriş görüntüsünü öznitelik uzayında haritasını oluşturmaktadır. Çözücü kısmı ise öznitelik uzayındaki haritayı giriş olarak alarak çıkış görüntüsünün elde edilmesini sağlamaktadır. Yani kodlayıcı-çözücü parçaları, öznitelik uzayıyla, resim uzayı arasında geçiş yapılan

bölümlerdir. SIUN yönteminde kodlayıcı-çözücü bölümleri, görüntüdeki yapıların restorasyonunu gerçekleştirecek şekilde özelleştirilmiştir.

Süper Çözünürlük Ağı

Ölçekte alt-örnekleme (down-sampling) görüntü işleme algoritmalarında kolay işlem yapılmasını sağlayan, resimde yer alan bazı özelliklerin çıkarılması yoluyla resmin sadeleştirilmesini sağlayan bir yöntemdir. Alt-örnekleme gerçekleştirildiğinde görüntüde yer alan bulanıklığın büyüklüğü de azaldığı için bulanıklık azaltma (deblurring) algoritmalarında daha kolay işlem yapılabilir. Fakat, alt-örnekleme gerçekleştirilen görüntülerde bazı görüntü detayları kaybolmaktadır. Alt-örnekleme gerçekleştirilmiş resimlere, tekrar üst örnekleme (up-sampling) yapıldığında kaybolan bu detayların, tekrar oluşturulması mümkün olamamaktadır. SIUN yönteminde bu problemin çözümü için, üst-örnekleme katmanı yerine süper çözünürlük katmanı eklenmiştir. Bu katman, düşük çözünürlükteki görüntünün yüksek çözünürlükte tekrar oluşturulmasını sağlamaktadır. Böylece, görüntüde yer alan yüksek frekans bilgileri bir üst ölçek seviyesine aktarılmış olmaktadır. Ayrıca tüm resimde yer alan detayların restorasyonu için, Artık yoğun ağ (RDN: Residual dense network) kullanılmıştır. RDN sayesinde hem yerel hem de global özniteliklerin restorasyonu gerçekleştirilebilmektedir.

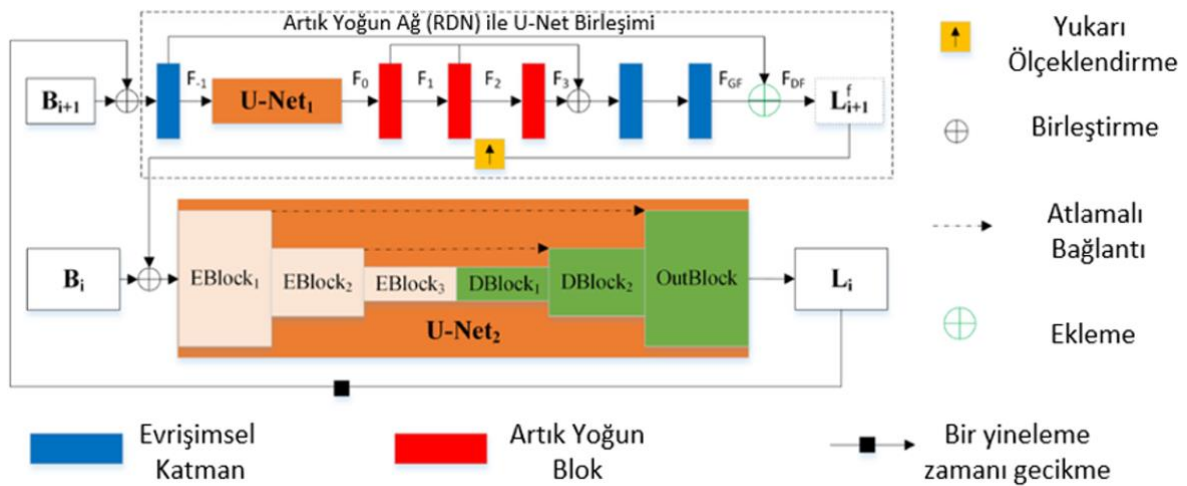
Müfredat Öğrenme (Curriculum Learning)

Bu öğrenme yönteminde öğrenme esnasında, görevin zorluğu kademeli olarak arttırılmaktadır. Müfredat öğrenme yönteminde, ağ performansının çeşitli görevlerde pozitif olarak etkilendiği gözlemlenmiştir (Elman, 1993, Bengio, 2009). SIUN yönteminde ise ölçekte alt-örnekleme gerçekleştirilmiş olan görüntü, orijinal ölçeğine getirilene kadar kademeli olarak restore edilmektedir. SIUN da kullanılan müfredat öğrenme yönteminin uygulaması ölçek artımı yapılırken görüntülerde kademeli restorasyon yapılması şeklindedir.

Ağ yapısı

SIUN metodunda ağ yapısı, iki seviyelidir. Birinci seviyede, U-Net (U-Net 1) ile birleştirilmiş, değiştirilmiş bir RDN yapısı kullanılmaktadır. U-Net regresyon kapasitesi

arttırılmış bir kodlayıcı çözücü yapısıdır (Ronneberger, 2015). İlk seviyede, daha küçük ölçeklerdeki görüntülere bulanıklık azaltma işlemi gerçekleştirilmektedir. Bu seviyenin çıktısı, bir üst seviyeye ölçeklenerek, ikinci seviyeye girdi olarak verilmektedir. Böylece görüntü boyutu iki katına çıkarılırken, bulanıklık seviyesi de düşürülmüş olmaktadır. İkinci seviyede de diğer U-net (U-Net2) modülü yer almaktadır. İkinci seviyede daha büyük ölçekte görüntüler ile çalışılmaktadır. Bu aynı zamanda bulanıklık seviyesinin de arttırılması anlamına gelmektedir. Fakat ilk seviyede gerçekleştirilen bulanıklık azaltımı sayesinde daha büyük ölçekteki görüntülerde de daha keskin görüntü restorasyonu gerçekleştirilebilmektedir. Bu iki seviye arasındaki işlemler, resmin orijinal ölçeğine ulaşılan kadar sürekli olarak gerçekleştirilmektedir. Resmin orijinal ölçeğine erişildiğinde ise final görüntü çıktısı elde edilmiş olur (Ye, 2020). SIUN metodu ağ mimarisi Şekil 5.4'te detaylı olarak gösterilmiştir.



Şekil 5.4. SIUN metodu ağ mimarisi

5.6. Tek Fotoğraftan Grafik Temelli Kör Görüntü Bulanıklığı Azaltma

Kör görüntü bulanıklık azaltma, yani bulanıklaştırma çekirdeği bilgisi olmadan bulanıklık azaltımı, son derece zorlayıcı bir problemdir. Problem iki kısımda çözülmektedir:

1. Bulanık görüntüden bir bulanıklık çekirdeği tahmin etmek
2. Verilen tahmini bulanıklık çekirdeğini kullanarak, hedef, bulanıklığı az olan görüntüyü elde etmek için bulanık girdiyi dekonvolüsyon işlemine tabi tutmak.

Grafik temelli (GB) yöntem, bir görüntü yamasını ağırlıklı bir grafik üzerinde bir sinyal olarak yorumlayarak grafik tabanlı bir kör görüntü bulanıklık azaltma algoritmasıdır.

İlk olarak, hedefin güçlü gradyanlarını koruyan ancak ayrıntıları düzelten bir taslak olan bir iskelet görüntüsünün, bulanıklık çekirdeğini doğru bir şekilde tahmin etmek için kullanılabileceğini ve benzersiz bir iki modlu (bi-model) kenar ağırlık dağılımına sahip olduğu vurgulanmaktadır.

Ardından, bulanıklığa sahip bir görüntüye ait parça (patch) verildiğinde iki modlu bir kenar ağırlık dağılımını verimli bir şekilde ortaya koyan yeniden ağırlıklandırılmış bir grafik toplam varyasyonu (RGTV: reweighted graph total variation) tasarlanmaktadır.

Ayrıca, RGTV'yi grafik frekans alanında analiz etmek için, bir grafik “ l_1 -Laplacian düzenleyicisi” olarak temsil eden yeni bir ağırlık fonksiyonu ortaya konulmuştur. Bu, gürültü ve bulanıklığa karşı sağlamlık, güçlü parçalı pürüzsüz (PWS: strong piecewise smooth) filtreleme ve keskinlik geliştirme dahil olmak üzere arzu edilen özelliklerle öncekinin bir grafik spektral filtreleme yorumuna imkân sağlanmaktadır.

RGTV ile kör bir görüntü bulanıklığı azaltma hedefinin en aza indirilmesi, dışbükey olmayan türevlenemez bir optimizasyon problemine dönüşmektedir. RGTV için yeni grafik spektral yorumunu kullanarak, dönüşümlü olarak iskelet görüntüsü ve bulanıklık çekirdeğini çözen verimli bir algoritma tasarlanmıştır. Özellikle Gauss bulanıklığı için, hızlandırılmış grafik spektral filtreleme kullanarak kör Gauss bulanıklık giderme için başka bir hızlandırma stratejisi ortaya konulmuştur. Son olarak, hesaplanan bulanıklık çekirdeği ile, hedef görüntüyü restore etmek için en son kör olmayan görüntü bulanıklık azaltma algoritmaları kullanılabilmektedir.

Görüntü bulanıklığı, odak dışı pozlandırma veya pozlandırma süresi boyunca nesneler ve kamera arasındaki hareketlerden kaynaklanan yaygın bir görüntü bozulmasıdır. Bulanıklaştırma işlemi genellikle şu şekilde modellenmektedir:

$$b = x \otimes k + n \quad (5.11)$$

Eşitlikte b gözlemlenen bulanık görüntü, x gizli keskin (orijinal) görüntü, k bulanıklık çekirdeği, n gürültü ve $\otimes$ evrişim operatörüdür. Tersten ele alındığında, görüntü

bulanıklığını azaltma/yok etme işlemi, bulanık görüntü b 'den keskin/net görüntü x 'i kurtarmaktır. Bulanıklaştırma çekirdeğinin (k) bilinip bilinmemesine bağlı olarak iki farklı görüntü bulanıklık azaltma /yok etme kategorisi vardır. Bunlar, kör olmayan görüntü bulanıklık azaltımı (Hansen ve diğerleri, 2006) ve kör görüntü bulanıklık azaltımı (Almeida, 2010; Levin ve diğerleri, 2011) olarak sınıflandırılmaktadır. GB (Bai, 2018) (Graph-Based Blind Image Deblurring) algoritmasında hem gizli görüntü x 'in, hem de bulanık çekirdek k 'nin bilinmediği ve yalnızca bulanık görüntü b verildiğinde geri yüklenmesi gerektiği kör görüntü bulanıklığını azaltma sorununa odaklanılmıştır.

Kötü pozlanmışlığın üstesinden gelmek için, kör görüntü bulanıklık azaltma metodunda, görüntü netliğini destekleyen ve bulanıklığı cezalandıran bir öncel (prior) tasarlamak önemlidir. Bununla birlikte, doğal görüntülerin geleneksel radyan temelli önceller başarısız olma eğilimindedir (Levin ve diğerleri, 2011), çünkü bunlar genellikle Fourier alanında çoğunlukla düşük frekanslı bulanık görüntüleri tercih eder. Son zamanlarda, bu problemle başa çıkmak için birçok karmaşık görüntü öncelleri önerilmiştir. l_0 -norm tabanlı öncel (Pan ve diğerleri, 2016a), düşük sıralı öncel (Ren ve diğerleri, 2016) ve karanlık kanal önceli (Pan ve diğerleri, 2016b) gibi örnekler vermek mümkündür.

Bu öncellerin yanı sıra, grafiksel sinyal işlemenin (GSP: graph signal processing) ilerlemesiyle (Shuman, 2013), farklı görüntü uygulamaları için grafik tabanlı önceller tasarlanmıştır (Hu, 2015, Kheradmand, 2014). Pikselleri, düğümler şeklinde modelleyerek, pikseller arası benzerlikleri yansıtan ağırlıklı kenarlara sahip düğümlerin oluşturduğu grafiklerde, görüntüler grafikler üzerinde sinyaller olarak yorumlanabilir.

GB metodunda, kör görüntü bulanıklık azaltımı için grafik tabanlı bir öncel önerilmiştir. Spesifik olarak, doğal görüntüyü doğrudan hesaplamak yerine, bir iskelet görüntüsü kullanılmaktadır. İskelet görüntüsü, hedef görüntünün güçlü gradyanlarını koruyan ancak ayrıntıları düzelten parçalı pürüzsüz (PWS: a piecewise smooth) bir temsildir. İlgili yöntemde bu temsilin, bulanıklık çekirdeğini tahmin etmek için yeterli olduğu savunulmaktadır (Bai, 2018).

İlgili yöntemde, bulanık görüntü parçalarının aksine, iskelet görüntü parçası için bir grafiğin kenar ağırlıklarının benzersiz (unique) iki modlu dağılıma sahip olduğu gözlemlenmiştir.

Bu nedenle, bulanık bir görüntü parçası verilen iki modlu dağılımı desteklemeden önce yeniden ağırlıklandırılmış bir grafik toplam varyasyonu (RGTV: reweighted graph total variation) önerilmektedir. Böylece, RGTV'nin avantajlarını, grafik toplam varyasyonu (GTV) (Elmoataz, 2008, Berger, 2017) ve düğüm alanındaki grafik Laplacian düzenleyici (Pang, 2017) gibi önceki grafik düzgünlüğü öncelleriyle yan yana konulup, analiz edilmektedir. RGTV'yi grafik frekans alanında analiz etmek için, RGTV'nin l_2 -Laplacian düzenleyicisine (Pang, 2017) benzer şekilde bir grafik l_1 -Laplacian düzenleyici olarak ifade edilebilmesi için yeni bir grafik ağırlık fonksiyonu tanımlanmaktadır. Bunu yapmak, RGTV'nin gürültüye / bulanıklığa karşı sağlamlık ve güçlü PWS filtrelemesi gibi istenen özelliklere sahip bir düşük geçişli grafik spektral filtresi olarak yorumlanabileceği anlamına gelmektedir.

RGTV'nin grafik spektral yorumuna dayanarak, dönüşümlü olarak iskelet görüntü ve bulanıklık çekirdeğini çözen verimli bir algoritma tasarlanmıştır. Ayrıca, özellikle Gauss bulanıklığı için, hızlandırılmış grafik spektral filtreleme kullanarak kör Gauss bulanıklık giderme için başka bir hızlandırma stratejisi önerilmektedir (Hochreiter, 1997).

Son olarak, tahmini bulanıklık çekirdeği k ile, (Pan ve diğerleri, 2016, Krishnan ve Fergus, 2009, Zoran, 2011) gibi kör olmayan bir bulanıklık giderme yöntemini kullanarak bulanık görüntü çözülmektedir. GB metodunun katkıları aşağıdaki gibi özetlenebilmektedir:

1. Bulanık bir gözlemden bir iskelet görüntü parçasını restore etmek için iki modlu bir ağırlık dağılımını destekleyen RGTV önceli grafik tabanlı bir görüntü önerilmektedir, böylece daha sonra uygun bir bulanıklık çekirdeği kolayca türetilenmektedir.
2. GTV / RGTV'nin bir grafik l_1 -Laplacian düzenleyicisi olarak ifade edilebilmesi için bir grafik ağırlık fonksiyonu önerilmektedir. Öncel daha sonra arzu edilen spektral özelliklere sahip bir alçak geçiren grafik filtresi olarak yorumlanabilir. GTV'nin grafik frekans yorumunu sağlayan ilk çalışma GB metodunu ortaya koyan çalışmadır (Bai, 2018).
3. RGTV'nin spektral yorumuna dayanarak, dışbükey olmayan türevlenemez optimizasyon problemini dönüşümlü olarak çözmek için verimli bir algoritma ortaya konulmuştur. Böylece iskelet görüntüsü ve bulanıklık çekirdeği için çözülecek iki alt problemin kapalı form çözümleri mevcuttur.

4. Spesifik olarak Gauss bulanıklığı için, hızlandırılmış grafik spektral filtreleri kullanarak kör Gauss bulanıklık giderme için başka bir hızlandırma stratejisi önerilmiştir (Hochreiter, 1997).

GB metodunda grafik tabanlı öncel RGTV, yüksek oranda veriye uyarlanabilir ve hızlı grafik spektral filtreleri kullanılarak RGTV'nin bir grafik l_1 -Laplacian düzenleyicisi olarak yeni bir yorumu yoluyla verimli bir şekilde çözülebilmektedir.

Grafiğe Dayalı Görüntü Önceli

GSP (Shuman, 2013), grafiklerle tanımlanan düzensiz veri çekirdekleri üzerindeki sinyalleri incelemek için ortaya çıkan bir alandır. Pikselleri, pikseller arası benzerlikleri yansıtan ağırlıklı kenarlara sahip düğümler olarak modelleyerek, görüntüler (veya görüntü parçaları) grafik sinyalleri olarak yorumlanabilmektedir.

Görüntü restorasyonu için, grafik Laplacian düzenleyici (Pang, 2017) ve GTV (Elmoataz, 2008, Berger, 2017) gibi grafik tabanlı görüntü öncelleri, farklı ters problemler (bozuk görüntüden, orijinal keskin görüntünün elde edilmesi) için tasarlanmıştır.

1) Grafik Laplacian Önceli: (Hu, 2015) çalışmasında, birleşik bir çerçevede istenen grafik sinyalini ve benzerlik grafiğini optimize ederek sıkıştırılmış bir JPEG PWS görüntüsünün kodlanması için bir şema tasarlanmıştır. Yapılan (Pang, 2017) çalışmada ise sürekli etki alanındaki grafik Laplacian düzenlemesini analiz etmiş ve görüntü gürültü giderme için görüşler sağlamıştır. Çift stokastik bir Laplacian grafikte, kör olmayan görüntü bulanıklaştırma (Kheradmand, 2014) için bir çerçeve geliştirilmiştir.

2) Grafik Toplam Varyasyon Önceli, (GTV: Graph Total Variation Prior) hem arzu edilen PWS koruma özelliklerine hem de dışbükeyliğe sahip olan p-Dirichlet enerji öncellerinden biridir. Manifold işlemede (Elmoataz, 2008) çalışmasında ayrık p-Dirichlet enerjisini analiz etmiştir. Grafik sinyallerinin doğrusal olmayan çok katmanlı gösterimi için (Hidane, 2013) çalışmasında, GTV kullanılmıştır. Grafiklerde ikili kısıtlı GTV düzenlemesi (Couprie, 2013) çalışmasında önermiştir. (Berger, 2017) çalışmasında grafik düğümlerinin bir alt kümesinde alınan gürültülü örneklerden düzgün bir grafik sinyali elde etmek için GTV kullanılmıştır.

GB çalışmasında, kör görüntü bulanıklaştırma problemini çözmeden önce yeni bir RGTV önerilmektedir. Önerilen yöntem ile:

- i) Literatürde grafik spektral alanında GTV ve RGTV'nin analizini ilk defa gerçekleştirilmiştir
- ii) RGTV'nin spektral yorumuna dayanarak, dışbükey olmayan optimizasyon problemini çözmek için verimli bir alternatif yinelemeli algoritma tasarlanmıştır.
- iii) Özellikle Gauss bulanıklığı için, hızlandırılmış grafik spektral filtreleme kullanarak kör Gauss bulanıklaştırma için bir hızlandırma stratejisi ortaya konulmuştur (Susnjara, 2015).

5.7. Kör Görüntü Bulanıklı Azaltma için Basit Bir Minimum Yoğunluk Önceli

Karanlık kanal önceli, aşırı kanal önceli ve yerel maksimum gradyan önceli gibi karmaşık metotlarda bulunan hesaplama maliyetlerini azaltmak ve bu metotlarda yer alan alt problemleri çözmek amacıyla Kör görüntü Bulanıklı Azaltma için Basit Bir Minimum Yoğunluk Önceli (PMP) (Wen, 2020) metodunda literatürde ilk olarak yerel minimum piksellerden önce basitleştirilmiş bir seyreklik, yani yama bazında minimum piksellerin (PMP: patch-wise minimal pixels) kullanımını önermektedir. Net görüntülerin PMP'si, bulanık olanlardan çok daha seyrek olduğundan net ve bulanık görüntüleri ayırt etmede oldukça etkili olmaktadır. Bu yöntemde, bulanıklık gidermede PMP'nin seyrekliğinden verimli bir şekilde yararlanmak için yeni bir algoritma tasarlanmıştır. Yeni algoritma, doğrudan yarı kuadratik bölme algoritmasını kullanmak yerine, maksimum bir posterior (MAP) çerçevesi altında PMP'ye seyreklik indüklemeyi iteratif bir şekilde uygulamaktadır. Bu sayede, hesaplama açısından çok daha verimli olurken, mevcut algoritmalarda titiz olmayan yaklaşım çözümlerinden kaçınır. İlgili yöntem bulanıklaştırmayı azaltma kalitesi, sağlamlık ve hesaplama verimliliği gibi açılardan, mevcut yöntemlerden daha avantajlı görülmektedir (Wen, 2020).

Kör bulanıklaştırma problemini iyi pozlanmış hale getirmek için, görüntü önceli ve bulanık çekirdek modeli kullanımı, en etkili yöntemlerin çokca tercih ettiği yöntemlerdendir. İyi geliştirilmiş görüntü öncelleri arasında (Chan ve Wong, 1998; Fergus, 2006) gradyan seyreklik önceli, normalleştirilmiş seyreklik önceli (Krishnan ve diğerleri, 2011), yama

önceli (L. Sun ve diğeleri, 2013) , grup seyrekliği önceli (Lin, 2018), yoğunluk önceli (Pan ve diğeleri, 2014) , karanlık kanal önceli (Pan ve diğeleri, 2016b; Pan, Sun, Pfister ve Yang, 2017), uç kanal önceli (Yan ve diğeleri, 2017), gizli yapı önceli (Bai, 2019), yerel maksimum gradyan önceli (Chen, 2019), sınıfa özel öncel (Anwar, 2018) ve derin bir ağ kullanmadan önce öğrenilmiş görüntü (Li, 2018) gibi örnekler mevcuttur. Bu arada, bulanık çekirdek modelleri, çoklu homografilerden (Hu, 2014b; L. Zhang, Zhou ve Huang, 2016), derinlik varyasyon modelinden (L. Xu ve Jia, 2012; Sheng, 2019), düzlem döndürme modelinden (J. Sun ve diğeleri, 2015) ve ileri hareket modelinden (Zheng, 2013) gelen bulanıklığa sahip tek biçimli olmayan modeli içermektedir. Bu yöntemlerin çoğu, maksimum bir posterior (MAP: maximum a posterior) çerçevesi altında görüntü önceli ve bulanık çekirdek modelinden yararlanmaktadır (Wen, 2020).

Genel olarak, ilgili bulanıklık giderme sorunları dışbükey olmadığı için (Wen, 2019), görüntü önceli ve/veya çekirdek modelinden yararlanmak için düzenleştirmeyi dahil etmek, iyi bir yerel çözüm elde etme olasılığını etkin bir şekilde artırmaya yardımcı olmaktadır. Ek olarak, hüristik kenar seçimi (heuristic edge selection) de MAP'e yardımcı olmanın etkili bir yolu olarak görülmektedir.

Görüntü gradyan seyrekliği popüler ve yaygın olarak kullanılan bir öncel olsa da görüntü yoğunluğu veya gradyan tabanlı önceller, daha önceden ortaya konulan (Pan ve diğeleri, 2014; Pan ve diğeleri, 2016a; Yan ve diğeleri, 2017; Chen, 2019) görüntü gradyanları ile birlikte kullanıldığında iyi bir tamamlayıcı etkinlik göstermektedir. Doğal görüntüler için tasarlanan önceller ve modeller genellikle metin görüntüleri (Cho, Wang ve Lee, 2012), yüz görüntüleri ve düşük ışıklı görüntüler (Hu, 2014a) gibi belirli görüntülere genelleme yapmadığından, aynı anda birden fazla önceli kullanmak hem doğal hem de doğal görüntülerde tatmin edici performans elde etme potansiyeline sahip olmaktadır. Gelişmiş önceller umut verici performans gösterse de iki sınırlama vardır:

- i) Birden çok öncelin birlikte kullanılması, karşılık gelen algoritmaları hesaplama açısından pahalı hale getirmektedir.
- ii) Karşılık gelen algoritmalarındaki bu öncellerle ilgili alt problemler açık bir şekilde çözülemediğinden, bu tür öncellerin kapasitesinden en iyi şekilde yararlanılmasını sınırlayan titiz olmayan yaklaşık çözüm yaygın olarak kullanılmaktadır. Bu noktada bazı sınırlamalar bulunmaktadır.

Bu sınırlamaları çözmek için PMP metodunun katkıları aşağıda verilmiştir. PMP, yerel minimum piksellerin bir koleksiyonudur. Sezgisel olarak, bulanıklaştırma işlemi görüntü pikselleri üzerinde yumuşatıcı bir etkiye sahip olduğundan, bulanıklaştırma işleminden sonra yerel minimum piksellerin yoğunluğu artmaktadır. Sonuç olarak, net görüntülerin PMP'si, bulanık olanlardan çok daha seyrek olmaktadır. PMP metriği, doğrudan yoğunluk önceli (direct intensity prior) kadar basit olup, net ve bulanık görüntüleri ayırt etmede etkili olmaktadır. Karanlık kanal öncelinin kullanıldığı (dark channel prior) (Pan ve diğerleri, 2016) çalışmada bahsedilen yöntemin verimli hesaplamayı kolaylaştıran bir basitleştirmesi şeklinde olmaktadır.

İkinci olarak, MAP çerçevesi altında PMP'nin seyrekliliğinden yararlanmak için yeni bir algoritma önerilmektedir. Yeni algoritma, MAP bulanıklaştırma işleminde gizli (orijinal, net) görüntünün PMP'sinde seyrekliliği indüklemeyi esnek bir şekilde uygulamaktadır. (Pan ve diğerleri, 2016; Yan ve diğerleri, 2017; Chen, 2019) gibi doğrudan yarım karesel bölgelendirmeye dayalı artırılmış MAP formülasyonlarını çözen mevcut algoritmalarla karşılaştırıldığında, önerilen algoritma maddedeki hesaplama verimliliğini büyük ölçüde geliştirmektedir.

Son yıllarda, tek görüntü kör bulanıklaştırma, görüntüler ve bulanıklık çekirdekleri üzerinde çeşitli etkili öncellerin yardımıyla çok ilerleme kaydetmiştir (Jia, 2014). Çoğu çalışma, varyasyonel Bayes ve MAP çerçevelerine dayanmaktadır. Tipik olarak, bu tür bir kör bulanıklaştırma yönteminin genellikle iki adımı vardır. İlk olarak, bulanıklık çekirdeği, MAP çerçevesi altındaki bulanık gözleminden tahmin edilmektedir. Ardından, tahmini bulanıklık çekirdeğine dayalı olarak, gizli keskin görüntü, kör olmayan evrişim yöntemleriyle, örneğin (J. Zhang, Zhao, Xiong, Ma ve Gao, 2014, Zoran, 2011) ile tahmin edilmektedir.

Saf MAP yöntemi doğal görüntülerde başarısız olabileceğinden, doğal görüntülerin ve seçilimin göze çarpan kenarları için istatistiksel öncellerden yararlanmak en son teknolojinin (state-of-the-art) başarısının anahtarıdır. Doğal görüntülerde, gradyan seyrekliliği önceli, en yaygın olarak kullanılan önceldir. Ancak (Levin, Weiss, Durand ve Freeman, 2009) çalışmasındaki gibi, pratikte, MAP çerçevesinde gradyan seyrekliliği önceli kullanan yöntemler, net görüntülerden ziyade bulanık görüntüleri tercih etmektedir. Bu sınırlama, açık keskin kenar takibi (Cho ve Lee, 2009; Cho, Paris, Horn ve Freeman, 2011)

veya buluşsal kenar seçimi (L. Xu ve Jia, 2010) gibi tekniklerle hafifletilebilmektedir. Ancak, bu tür tekniklerin gizli (keskin) görüntülerde güçlü kenarların olduğu varsayımı her zaman karşılanmayabilmektedir.

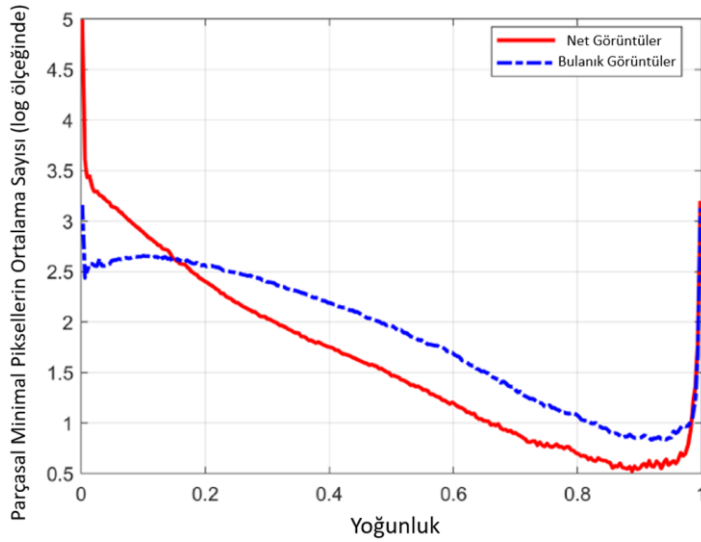
MAP'i canlandırmak için tasarlanmış, normalleştirilmiş seyreklik önceli (Krishnan ve diğerleri, 2011), dahili parça tekrarlaması (Michaeli ve Irani, 2014) ve doğrudan yoğunluk önceli (Pan ve diğerleri, 2014) gibi çalışmalar mevcuttur. Bu tip yöntemler doğal veya belirli (daha önceden seçilmiş) görüntüler için etkili olmasına rağmen, genellikle hem doğal hem de belirli görüntülerde tatmin edici performans sağlayamamaktadır. Karanlık kanal önceli (Pan ve diğerleri, 2016a) ve önceden öğrenilen verilere dayalı öncel (Li, 2018) hem doğal hem de seçilen görüntülerde performans açısından tatmin edici sonuçlar elde edebilmektedir. Fakat, ilgili optimizasyon algoritmaları hesaplama açısından pahalı olmaktadır. PMP metodunda ise bu hesaplama maliyetleri düşürüldüğü belirtilmektedir.

Son zamanlarda, veriye dayalı yöntemler de güçlü derin öğrenme tekniklerinin, yardımıyla başarılı olmuştur. Örneğin, (J. Sun ve diğerleri, 2015) çalışmasında tekdüze olmayan hareket bulanıklığını tahmin etmek ve ortadan kaldırmak için bir evrimsel sinir ağı (CNN) kullanmıştır. (Nah ve diğerleri, 2017), bulanıklık çekirdeği üzerinde herhangi bir varsayım olmaksızın, gizli görüntüyü uçtan uca bir şekilde kurtarmak için çok ölçekli bir CNN önermektedir (Wen, 2020).

Ek olarak, uzamsal olarak değişken yinelemeli sinir ağı ve ölçek yinelemeli ağ (J. Zhang, Pan, Ren, Song, Bao, Lau ve Yang, 2018; Tao ve diğerleri, 2018) yöntemleri bulanıklığı gidermek için tasarlanmıştır. Ayrıca, (Kupyn, 2018) çalışmasında koşullu üretken çekişmeli ağlarına (GAN) dayalı bulanıklık giderme için uçtan uca öğrenilmiş bir yöntem önermektedir. Video bulanıklığı azaltmak için uçtan uca CNN modeli (Su ve diğerleri, 2017) çalışmasında ele alınmaktadır. Bu veriye dayalı yöntemler çeşitli senaryolarda olumlu performans gösterebilse de başarıları büyük ölçüde eğitim verileri ile test verileri arasındaki tutarlılığa bağlı olmaktadır. Bu, genelleme yeteneklerinin sınırlandırılmasına yol açmaktadır.

Minimal piksellerin ortalama sayısı bulanık ve net görüntüler için aşağıdaki Şekil 5.5'te verilmiştir. İlgili şekilde 5000 adet doğal görüntü üzerinde net ve bulanık görüntülerin yama bazında minimum pikselleri için yoğunluk histogramı (20×20 piksel parça boyutu

ile) ile net görüntülerin PMP'sinin (0,9 gibi bir eşiğin altında) Hiper Laplacian dağılımını takip ettiğini gösterilmektedir ve net görüntülerin PMP'si bulanık görüntülerin PMP'sinden çok daha seyrek olmaktadır.



Şekil 5.5. Net ve bulanık görüntülerin yama bazında minimum pikselleri

Parçasal Minimal Pikseller (Patch-Wise Minimal Pixels)

PMP, örtüşmeyen parçalar (yamalar) üzerindeki yerel minimal piksellerin bir koleksiyonudur. Bir $I \in \mathbb{R}^{m \times n \times c}$ görüntüsünün, $r \times r$ yama boyutuna sahip P adet kesişmeyen parçalara bölündüğünde; $P = \frac{m}{r} \cdot \frac{n}{r}$ olmaktadır. PMP ifadesi;

$$P(I)(i) = \min_{(x,y) \in \Omega_i} \left(\min_{c \in \{r,g,b\}} I(x,y,c) \right) \quad (5.12)$$

olarak yazılabilmektedir. $i = 1, 2, \dots, P$ için Ω_i i-inci parçanın piksel konumlarının indeks kümesini gösterir. $P(I) \in \mathbb{R}^P$ olmaktadır ve parça-şeklinde I 'nın yerel minimum piksellerini içermektedir.

Bulanık görüntülerin PMP'sinin doğal net görüntülerden çok daha az seyrek olduğu Şekil 5.5'ten görülebilmektedir. VGG1 veri setinden 5000'den fazla doğal görüntü üzerinden net görüntüler ve bunların bulanık karşılıkları arasındaki PMP'nin histogram istatistiğini karşılaştırılmıştır. Bulanık görüntüler, veri kümesindeki orijinal keskin görüntülerden, bulanık çekirdekleri kullanılarak üretilmiştir (Levin, Weiss, Durand ve Freeman, 2009).

Net doğal görüntülerin PMP'sinin, bulanık görüntülerden önemli ölçüde daha fazla sıfır öğeye sahip olduğu görülebilmektedir. Bir eşiğin (örneğin, 0.9) altındaki net görüntülerin PMP'si, bir Hiper Laplacian dağılımını takip etmekte ve seyreklik özelliği göstermektedir. PMP'nin bu seyreklik özelliği, net görüntüleri bulanık olanlardan ayırt etmek için doğal bir ölçüm sağlamaktadır. Bu özelliğe dayanarak, ilgili yöntemde önerilen algoritma, bulanıklık azaltma işleminde daha doğru çekirdek tahmini elde etmek için PMP 'nin bu özelliğini kullanmaktadır.

$P(B)$ ve $P(I)$ sırasıyla bulanık ve net görüntülerin PMP'si olmak üzere;

$$P(B) \geq P(I) \quad (5.13)$$

olmaktadır. Bu özellik bulanıklaştırma işleminin PMP değerlerini artırdığını gösterir, bu da bulanık görüntülerin PMP'sinin net muadillerinden daha az seyrek olmasına yol açmaktadır. Genel formülde, işlem kolaylığı açısından $c = 1$ alındığında,

$$P(I): \mathbb{R}^{m \times n} \Rightarrow \mathbb{R}^P \quad (5.14)$$

olmaktadır. Buradan transpoz işlemi ile,

$$P^T(z): \mathbb{R}^P \Rightarrow \mathbb{R}^{m \times n} \text{ her } z \in \mathbb{R}^P \quad (5.15)$$

Buradan;

$$I_p: P^T(P(I)) = I \circ M \quad (5.16)$$

$\circ$ operatörü element bazında Hadamard çarpımını göstermektedir. Burada $M \in \mathbb{R}^{m \times n}$, I'nin PMP alt kümesine karşılık gelen ikili maskedir. Eğer $I(i,j)$, parçadaki bir minimal pikselse $M(i,j)=1$ olmaktadır. Diğer durumlar için $M(i,j)=0$ olmaktadır.

Mevcut Yoğunluk Seyreklik Öncelleri ile Karşılaştırma

Yoğunluk seyrekliği, kör bulanıklaştırma için (Pan ve diğerleri, 2014; Pan ve diğerleri, 2016b)'da da dikkate alınmıştır. Bununla birlikte, önerilen PMP, aşağıda açıklanan (Pan ve diğerleri, 2014; Pan ve diğerleri, 2016b)'da ele alınan yoğunluk metriklerinden temel olarak farklıdır. Yoğunluk seyrekliğinden yararlanan en yakın çalışma, (Pan ve diğerleri, 2016b)'da ele alınan karanlık kanal metriğidir. Bir görüntü için, $I \in \mathbb{R}^{m \times n \times c}$, karanlık kanal;

$$D(I)(i,j)=\min_{(x,y) \in \Omega_{i,j}} \left(\min_{c \in \{r,g,b\}} I(x,y,c) \right) \quad (5.17)$$

$i = 1, 2, \dots, m$ ve $j = 1, 2, \dots, n$ için, burada $\Omega_{i,j}$ (i, j) inci pikselde ortalananmış parçanın piksel konumlarının indeks kümesini belirtmektedir. Buradan görüleceği üzere, karanlık kanal, $m \times n$ boyutunda bir çıktı ile evrişim benzeri bir şekilde hesaplanırken, önerilen PMP, vektör çıktısı $P(I) \in \mathbb{R}$ olan örtüşmeyen parçalar üzerinde hesaplanmaktadır. Sonuç olarak, önerilen PMP metodu, karanlık kanal öncelinden çok daha basittir, böylece daha verimli algoritma tasarımını kolaylaştırmaktadır. Öte yandan önerilen PMP metriği, doğrudan yoğunluk metriği kadar basittir, ancak Şekil 5.5'te de gösterildiği gibi net ve bulanık doğal görüntüleri ayırt etmede etki göstermektedir.

PMP Kullanarak Bulanıklık Azaltma Algoritması (Seyreklik Düzeltmesi)

PMP'nin seyreklik düzenlemesini geleneksel MAP çerçevesine esnek bir şekilde dahil ederek verimli bir algoritma ortaya konulduğunda bulanıklık azaltımı mümkün olabilmektedir. Bu algoritma, yarı karesel bölme algoritmasının bir çeşididir. MAP formülasyonu aşağıdaki gibidir:

$$\min_{(k,l)} L(k \otimes I, B) + \gamma G(k) + \mu R(I) \quad (5.18)$$

burada γ ve μ pozitif ağırlık parametreleridir. L ise, $k \otimes I$ 'in bulanık görüntü B ile tutarlı olmasını kısıtlayan bir veri doğruluğu terimidir. Problemi iyi bir şekilde ortaya koymak için, G ve R , sırasıyla bulanıklık çekirdeği ve gizli görüntü ile ilişkili ceza fonksiyonlarıdır. Doğal görüntülerin gradyanı seyrek olduğu için, $R(I)$ genellikle ∇I 'nin (I 'nin gradyanının)

l_0 -norm cezası olarak seçilir. Ek olarak, l_2 norm kazancı olarak hem kayıp fonksiyonu L hem de çekirdek cezası da seçilmektedir.

$$\min_{(k,l)} \|k \otimes I - B\|_2^2 + \gamma \|k\|_2^2 + \mu \|\nabla I\|_0 \quad (5.19)$$

l_2 -norm sadece Gauss gürültüsü için optimal değildir, aynı zamanda hızlı Fourier dönüşümü (FFT) yoluyla ilgili alt problemlerin hızlı hesaplanmasını kolaylaştırdığı için verimli algoritmaların geliştirilmesine de olanak tanımaktadır. Gizli (orijinal, net) görüntünün PMP'sinin seyrekliğinden daha fazla yararlanmak için, örneğin, $r \times r$ bir yama boyutu için $P(I) \in \mathbb{R}^P$, yukarıdaki formülün kısıtlanmış bir şekli aşağıdaki gibi ifade edilmektedir:

$$\min_{(k,l)} \|k \otimes I - B\|_2^2 + \gamma \|k\|_2^2 + \mu \|\nabla I\|_0 \quad (5.20)$$

$$P(I)(i) \sim p(x), i \in \{1, \dots, P\} \quad (5.21)$$

Burada $p(x)$, daha önce de belirtildiği üzere, 0.9 gibi bir eşiğin altındaki x için Hiper Laplacian dağılımının bir olasılık yoğunluk fonksiyonudur. Kısıtlı formül, görüleceği üzere düzgün ve dış bükey değildir. MAP benzeri amaç fonksiyonlarını çözmede mevcut çoğu bulanıklaştırma algoritmasına benzer şekilde, bulanıklık çekirdeğini ve gizli görüntüyü dönüşümlü olarak güncelleyerek çözmek için Eş. 5.16'da yer alan verimli bir algoritma önerilmiştir. Önerilen algoritmada, Eş. 5.16'daki kısıtlama, gizli orijinal görüntü alt probleminde $P(I)$ üzerinde yinelemeli seyreklik indükleme yoluyla yaklaşık olarak uygulanmaktadır. MAP çerçevesinde PMP'nin seyrekliğini daha rahat kullanabilmek için Eş. 5.16 formülünün doğal bir alternatifi olarak aşağıdaki formül kullanılabilir:

$$\min_{(k,l)} \|k \otimes I - B\|_2^2 + \gamma \|k\|_2^2 + \mu \|\nabla I\|_0 + \alpha \|P(I)\|_0 \quad (5.22)$$

Bu formülde α pozitif bir ağırlık parametresidir ve son terim, gizli görüntünün PMP'sinde seyreklik indükleme elde etmek için l_0 -norm cezasını kullanır. Bu formülasyon, (Pan ve diğerleri, 2014; Pan ve diğerleri, 2016; Pan ve diğerleri, 2017)'deki algoritmalara benzer alternatif bir şekilde yarı ikinci dereceden bölme algoritması ile doğrudan çözülebilmektedir. Bununla birlikte, çözüm için önerilen algoritmanın çözme için

doğrudan yarı ikinci dereceden bölme algoritmasından daha üstün olduğunu PMP metodu sayesinde gösterilmiştir.

5.8. Görüntü Bulanıklığı Azaltma için Uç Kanallar Yöntemi

Karanlık Kanal Önceli (DCP: Dark Channel Prior), doğal, yüz, metin ve düşük aydınlatmalı görüntüler dahil sahnelerde kör bulanıklık azaltma konusunda destek sunmaktadır. Ancak, sınırlamaları mevcuttur ve giriş görüntüsüne parlak pikseller hakimken bulanıklık çekirdeği tahminini destekleme olasılığı düşmektedir. Net görüntülerdeki parlak piksellerin, bulanıklaştırma işleminden sonra parlak olma ihtimalinin ortadan kalktığı gözlemlenmektedir. Bu gözlem, parlak kanal önceli (BCP: Bright Channel Prior) olarak tanımlanmaktadır. Uç Kanallar Önceli (Yan ve diğerleri, 2017) (ECP: Extreme Channels Prior) yönteminde ise hem karanlık hem de parlaklık öncellerinden faydalanmaktadır. ECP (Yan ve diğerleri, 2017) yöntemi ortak bir öncel olarak değerlendirilmektedir. Genel olarak, mevcut yöntemlerin çoğu, doğal görüntülerin istatistiklerinin ön bilgisinden yararlanmaktadır. Ağır kuyruklu gradyan dağılımları önceli (Fergus, 2006; Levin ve diğerleri, 2009), normalleştirilmiş seyreklik önceli (Krishnan ve diğerleri, 2011), l_0 -düzenli gradyan (L. Xu, Zheng ve Jia, 2013), tekrarlayan yama (Michaeli ve Irani, 2014) ve yoğunluk ve gradyan önceli (Pan ve diğerleri, 2016a) gibi yöntemler bunlara örnek olarak verilebilmektedir. Yakın zamanda, karanlık kanala dayalı öncel (Pan ve diğerleri, 2016b) ile bir kör görüntü bulanıklık azaltma yöntemi ortaya konulmuştur. Bu algoritma, çekirdek tahmini için gizli görüntülerin karanlık kanalının seyrekliğini kullanmakta ve diğer yaklaşımlara kıyasla daha iyi sonuçlar üretmektedir. Bununla birlikte, daha büyük derecelerde parlak pikseller veya parazit, bu koşullar altında karanlık kanal önceli geçerli olmadığı için bu algoritmanın performansını önemli ölçüde etkileyebilmektedir. ECP metodu (Pan ve diğerleri, 2016b) çalışmasından esinlenerek, parlak kanal piksellerinin, bulanıklaştırma sonrasında artık parlak olmadıkları (Şekil 5.6) ifade etmektedir. Bu özellik ara net tahminden görüntüsünden, bulanıklık çekirdeği tahmini için kullanılabilir. ECP yönteminin temel katkıları aşağıdaki gibidir:

1. Hem DCP hem de BCP'nin avantajlarını kullanarak Extreme Channels Prior (ECP) adında yeni bir doğal görüntü önceli önerilmektedir.

2. ECP ortaya konulurken, 5.000 doğal görüntünün bulanıklaştırma işleminden sonra parlak kanal piksellerinin yoğunluğunun azaldığı (bulanık olmayanlar ile karşılaştırıldığında), ilgili çalışmada kanıtlanmıştır.
3. Seyrekliği değiştirmek ve verimli bir optimizasyon şeması geliştirmek için bir l_0 -düzenleme terimi olarak çekirdek tahminine yardımcı olmak için parlak kanal öncelinden yararlanılmaktadır.



Şekil 5.6. Bulanıklaşma ile parlak kanal piksel yoğunluğu arasındaki ilişki

Bulanıklaşma ile parlak kanal piksel yoğunluğu düşmektedir. Parlak kanalı artırmak görüntünün restorasyonuna katkı sağlamaktadır. Karanlık kanala dayalı (Pan ve diğerleri, 2016b) çalışmada, bulanıklaştırma işlemi sırasında karanlık kanal piksellerinin nasıl değiştiği göz önüne alınarak genetik bir yaklaşım önerilmektedir. Karanlık kanal önceli ilk olarak (He, Sun ve Tang, 2010) çalışmasında tanıtılmıştır. Doğal sahne parçalarının çoğunda, renk kanallarından en az birinin sıfıra çok yakın yoğunluğa sahip bazı piksellere sahip olduğu gözlemlene dayanmaktadır. Devamında doğal görüntülerin karanlık kanalının sıfır olmadığı bunun yerine seyrek olduğu görüşünü kullanarak bulanıklık çekirdeği tahmini için bu seyreklik kullanılmıştır. Bu çalışma, çeşitli kıyaslama veri kümelerinde güçlü performansa sahip olsa da görüntüde karanlık piksel yoksa, karanlık kanal önceli, ara gizli görüntü tahminine yardımcı olmayabilir. ECP metodunda, öncelikle BCP ortaya konulmuş, daha sonra BCP ile DCP birleştirilerek ortaya yeni bir istatistiksel öncel ortaya çıkarılmıştır. Parlak kanal önceli olan BCP aşağıdaki gibi tanımlanmaktadır:

$$B(I)(x) = \max_{(y) \in \Omega(x)} \left(\max_{c \in \{r, g, b\}} I^c(y) \right) \quad (5.23)$$

Burada x piksel konumu, I^c , I 'nin bir renk kanalıdır ve $\Omega(x)$, x merkezli yerel bir parçayı belirtmektedir. İlgili denklemde belirtildiği üzere, parlak kanal, $\max_{c \in \{r, g, b\}}$ ve $\max_{(y) \in \Omega(x)}$ olmak üzere iki maksimum operatörünün çıktısıdır. I gri ölçekli (gray-scale) bir görüntüyse, yalnızca ikinci operatör sonuç verecektir. Bu kavramı kullanarak, ışığın olmadığı veya gölgenin hakim olduğu durumlar dışında, $B(I)$ 'nin yoğunluğunun yüksek ve bire yakın olması gerektiği sonucuna varılır. Bu sonuca BCP adı verilmektedir. Parlak kanaldaki yüksek yoğunluklar temel olarak üç faktöre bağlıdır:

1. Işık. Örneğin güneş ışığı, diğer mevcut ışık kaynaklarından gelen ışık ve gündüzleri güneş tarafından aydınlatılan gökyüzü bölgeleri,
2. Beyaz renkteki yüzeyler. Örneğin, beyaz duvarlar, zeminler, vb.
3. Parlak renkli nesneler veya yüzeyler. Örneğin yeşil yaprak, sarı çiçek ve mavi su gibi renk kanallarından en az birinde yüksek yansımaya sahip nesneler de yüksek yoğunluklara yol açacaktır.

Bulanıklık azaltmada BCP'yi kullanabilmek için, bulanıklaştırma işleminden sonra parlak kanalların nasıl değiştiği analiz edilmiştir. Tüm bulanıklık azaltma problemlerinde, gürültünün olabildiğince küçük olduğu kabul edilmektedir.

$$\begin{aligned}
 b(x) &= \sum_{Z \in \emptyset(x)} I\left(x - z + \frac{p}{2}\right) k(z) \\
 &\leq \sum_{Z \in \emptyset(x)} \max_{(y) \in \Omega(x)} I(y) k(z) \\
 &= \max_{(y) \in \Omega(x)} I(y) \sum_{Z \in \emptyset(x)} k(z) \\
 &= \max_{(y) \in \Omega(x)} I(y)
 \end{aligned} \tag{5.24}$$

Yukarıda yer alan denklem, bulanıklaştırma işleminden sonra piksel x 'teki yoğunluğun, x merkezli görüntü parçasının piksellerin maksimum yoğunluğundan daha fazla olmadığını göstermektedir. Özellikle, x yerel parçadaki en parlak piksel ise, 4.denklem $b(x) \leq I(x)$ olur. Bu durum parlak kanalın tamamına yayılırsa;

$$B(b)(x) = B(I)(x) \tag{5.25}$$

olarak bulunmaktadır. Buradan, bulanık bir görüntünün parlak kanalının, karşılık gelen net görüntüden daha düşük piksel yoğunluklarına sahip olduğu görülmüş olmaktadır. Bu nedenle, çekirdek tahminine yardımcı olan bir özellik, bulanıklaştırma işleminin gizli görüntüde bire eşit olan piksel değerlerini azalttığını ortaya koymaktadır. Biçimsel olarak,

$$\|1-B(b)(x)\|_0 \geq \|1-B(l)(x)\|_0 \quad (5.26)$$

şeklinde ifade edilebilmektedir.

Önerilen Algoritma

Bulanıklık azaltma problemi, bir maksimum a posteriori (MAP) çerçevesinde (L. Xu ve diğerleri, 2013) formüle edilmiştir.

$$\{\hat{l}, \hat{k}\} = \arg \min_{l, k} l(l \otimes k, b) + \gamma p(k) + \lambda p(l) \quad (5.27)$$

burada $p(k)$ ve $p(l)$, sırasıyla bulanıklık çekirdeğinin ve gizli görüntünün öncelleridir.

Önerilen ECP

Önerilen BCP kullanılarak, ters BCP'nin l_0 düzenlemesi (regularization) aşağıdaki gibi ortaya konulmuştur:

$$P(l) = \|1-B(l)\|_0 \quad (5.28)$$

Ardından hem önerilen BCP'nin hem de daha önceden kullanılmakta olan DCP'nin avantajlarından yararlanarak görüntü bulanıklığı azaltmak için yeni bir Uç Kanallar Önceli (ECP) önerilmiştir.

$$P(l) = \|1-B(l)\|_0 + \|D(l)\|_0 \quad (5.29)$$

Amaç Fonksiyonu

Önerilen ECP tabanlı regülasyon terimini (Pan ve diğerleri, 2016b) 'deki son tek görüntü bulanıklaştırma çerçevesiyle birleştirildiğinde, görüntü bulanığı azaltma için önerilen amaç fonksiyonu aşağıdaki gibi olmaktadır:

$$\{\hat{l}, \hat{k}\} = \arg \min_{l, k} \|l \otimes k - b\|_2^2 + \gamma \|k\|_2^2 + \mu \|\nabla l\|_0 + \lambda \|D(l)\|_0 + \eta \|1 - B(l)\|_0 \quad (5.30)$$

Bu denklemde; γ , μ , λ , η karşılık gelen ağırlık parametreleridir. Denklemdaki ilk terim, restore edilen görüntü l 'nin konvolüsyonunun, bulanık görüntü b ile tutarlı olması gerektiğini kısıtlamak için veri doğruluğu terimidir. $l \otimes k$ ve b arasındaki fark genellikle l_2 normu (Cho ve Lee, 2009; Zhou, 2014) veya l_1 normu (L. Xu ve Jia, 2010; Ren ve diğerleri, 2016) kullanılarak sınırlandırılmaktadır. ECP çalışmasında veri doğruluğu işlevi için l_2 -normu kullanılmaktadır. İkinci terim, bulanık çekirdek k 'nin kararlı olmasını gerektiren bir kısıtlamadır. ECP yöntemi bu terim için de Hızlı Fourier Dönüşümü (FFT) (Cho ve Lee, 2009; L. Xu ve Jia, 2010) ile çözülebilen l_2 normunu seçmektedir. Üçüncü terim, keskin görüntü gradyanlarını koruma ve ince olanları kaldırmayı sağlarken, dördüncü düzensileştirme terimi, karanlık kanalların seyrekliğini sağlamaktadır.

Optimizasyon

Amaç fonksiyonunun çözümünü elde etmek zorlayıcı olduğu için, bunun yerine ikinci dereceden bölme algoritmasına dayalı alternatif bir minimizasyon algoritması kullanılmaktadır (L. Xu, Lu, Xu ve Jia, 2011). Ara gizli görüntü l ve bulanıklık çekirdeği k , alternatif olarak bunlardan birini sabitleyerek tahmin edilir, yani optimizasyon problemi iki alt probleme çevrilmiş olmaktadır:

$$\hat{l} = \arg \min_l \|l \otimes k - b\|_2^2 + \mu \|\nabla l\|_0 + \lambda \|D(l)\|_0 + \eta \|1 - B(l)\|_0 \quad (5.31)$$

$$\hat{k} = \arg \min_k \|l \otimes k - b\|_2^2 + \gamma \|k\|_2^2 \quad (5.32)$$

Keskin Görüntünün Elde Edilmesi (Görüntü Restorasyonu)

l_0 düzenleştirme teriminin hesaplama açısından zor olduğunu göz önünde bulundurarak, yarı ikinci dereceden bölme tekniğine dayalı yeni bir yöntem önerilmiştir. Sırasıyla $D(l)$, $1 - B(l)$ ve ∇l 'ye karşılık gelen; $g = (g_h, g_v)^T$ olmak üzere sırasıyla yeni yardımcı değişkenler p , q ve g tanımlanmıştır, amaç fonksiyonunu aşağıdaki gibi düzenlenmiştir:

$$\{\hat{l}, \hat{g}, \hat{p}, \hat{q}\} = \arg \min_{l, g, p, q} \|l \otimes k - b\|_2^2 + \alpha \|\nabla l - g\|_2^2 + \beta \|D(l) - p\|_2^2 + \omega \|1 - B(l) - q\|_2^2 + \mu \|g\|_0 + \lambda \|p\|_0 + \eta \|q\|_0 \quad (5.33)$$

Burada α , λ ve η pozitif ceza parametreleridir. Yukarıdaki optimizasyon problemi, diğer değişkenler sabitlenirken alternatif olarak l , p , q ve g 'yi ayrı ayrı minimize ederek çözülebilmektedir. Gizli görüntü l 'yi çözmek için amaç fonksiyonu aşağıdaki gibidir.

$$\hat{l} = \arg \min_l \|l \otimes k - b\|_2^2 + \alpha \|\nabla l - g\|_2^2 + \beta \|D(l) - p\|_2^2 + \omega \|1 - B(l) - q\|_2^2 \quad (5.34)$$

Optimizasyonun kolaylığı için, tutarlılık için $1 - B(l)$ ifadesini $D(1 - l)$ 'ye dönüştürülmüştür. $D(l)$ işleminin lineer olmaması dikkat çekici olup, bu ifade eşdeğer bir doğrusal operatör M ile değiştirilmektedir. M temelde, görüntüdeki pikseli karanlık kanalına eşleyen bir eşleştirme matrisidir. Şu şekilde tanımlanmaktadır:

$$M(x, y) = \begin{cases} 1, & y = \arg \min_{y \in \Omega(x)} l(y) \\ 0, & \text{aksi halde} \end{cases} \quad (5.35)$$

Bulanıklık Çekirdeği Tahmini

Bu alt problemde, gradyan görüntülerde hızlı bulanıklaştırma yöntemi (Cho ve Lee, 2009) ve veri doğruluğu teriminin L_2 normu kullanılarak alınan görüntü l ile bulanıklık $\hat{k}$ çekirdeği optimize edilmektedir:

$$\hat{k} = \arg \min_k \|\nabla l \otimes k - \nabla b\|_2^2 + \gamma \|k\|_2^2 \quad (5.36)$$

Bu denklem aslında en küçük kareler minimizasyon problemini ifade etmektedir. Çözüm FFT ile aşağıdaki şekilde hesaplanabilmektedir:

$$\hat{\mathbf{k}} = \mathbf{F}^{-1} \left(\frac{\overline{\mathbf{F}(\nabla \mathbf{l})} \mathbf{F}(\nabla \mathbf{b})}{\overline{\mathbf{F}(\nabla \mathbf{l})} \mathbf{F}(\nabla \mathbf{l}) + \gamma} \right) \quad (5.37)$$

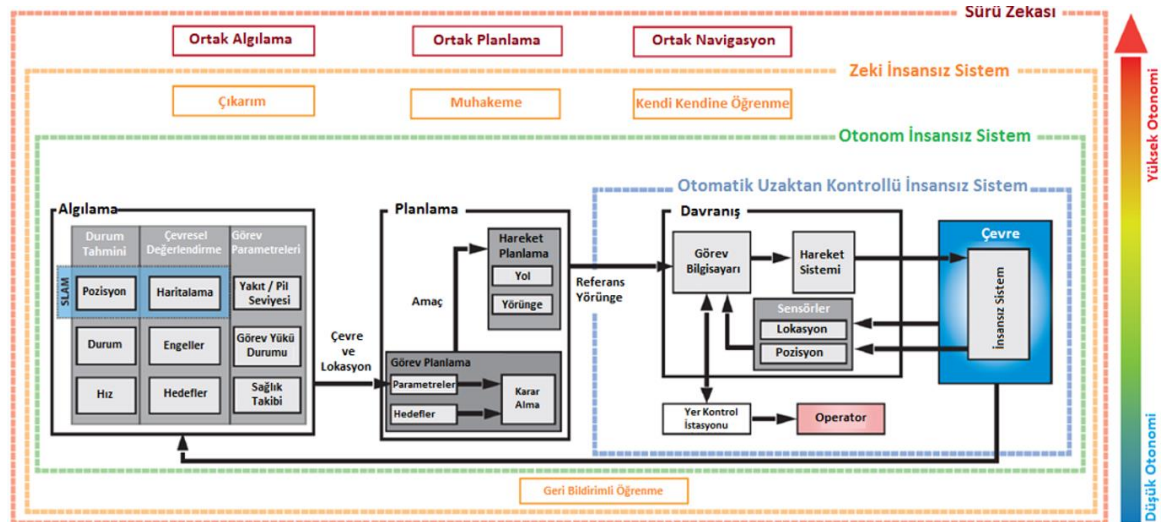
6. ÖNERİLEN METOT VE DENEYSEL ÇALIŞMALAR

Günümüzde robotlar ve insansız sistemlere, sağlık, tarım, madencilik, sürücüsüz araçlar, gezegen keşifleri ve nükleer tesis güvenliği gibi yaygın bir alanda pek çok alanda büyük görevler düşmektedir (Ben-Ari ve Mondada, 2018). Uzaktan kontrollü ya da tamamen otonom olması fark etmeksizin robotik ve insansız sistemlerin kendi konumunu ve çevresini bilmesi yönetilecek görevin başarılı bir şekilde sonuçlandırılması için oldukça önem arz etmektedir. Dış mekânlarda operasyon gerçekleştirecek insansız sistemler çoğunlukla konum bilgisini elde etmek için GNSS uydularından yayınlanan pozisyon ve zaman sinyallerini kullanmaktadır. Fakat GNSS sinyallerinin kolayca karıştırılabilir olması, insansız sistemlerin uyduların görüş açısından (line-of-sight) çıkmış olması, etrafta uydu sinyallerinin yanlış değerlendirilmesine neden olacak şekilde yansıtıcı / yön değiştirici pek fazla engellerin bulunması (multipath problemi) ve mağara / tünel gibi iç mekânlarda bu sinyallerin alınamaması gibi nedenlerle GNSS sinyallerinden lokasyon kestirimine alternatif yöntemler akademik hayatta ve endüstride önemli çalışma konuları arasında yer almaktadır (Şimşek, Satir ve Bilge, 2021).

Eş zamanlı konumlama ve haritalama (SLAM), GNSS sinyallerinin ulaşmadığı ya da tam olarak kullanılamadığı ortamlarda insansız platformların konumlama yapabilmesi için 30 yılı aşkın süredir robotik ve insansız sistemler alanında aktif ve gelişen bir araştırma alanı durumundadır (Zaffar ve diğerleri, 2018). Eş zamanlı konumlama ve haritalama (SLAM) sayesinde, çevreye ait karakteristik ve yapısal özellikler kullanılarak çevrenin haritasının çıkarılması ve eş zamanlı olarak oluşturulan harita üzerinde robotik ve insansız sistemlerin konumlarının belirlenmesini sağlamaktadır. Ayrıca oluşturulan bu harita, kullanıcıya tekrar tekrar kullanılabilmesi amacıyla da sunulabilmektedir (Makhubela, Zuva ve Agunbiade, 2018). SLAM metotları kullanılarak oluşturulan bilgiler (harita ve izlenen yörünge) insansız sistemi uzaktan kontrol eden operatör için anlam taşıyacağı gibi otonomi seviyesi yüksek yarı/tam otonom insansız sistemlerin verimli ve istenen şekilde operasyon gerçekleştirebilmesi için hayati öneme sahiptir. SLAM gibi pozisyon tahmini sağlayan algılama teknolojileri, operatör yardımıyla harekât gerçekleştiren insansız sistemlerin otonomi seviyelerini yükseltmek amacıyla da kullanılan sistemlerdir. SLAM teknikleri yardımıyla, bir harita oluşturulduğunda, ortaya çıkan harita konum tahminine ek olarak, insansız platformun operasyonuna yardımcı olan engelden kaçınma, obje tanıma ve görev

planlama gibi amaçlar için de kullanılabilir (Bailey ve Durrant-Whyte, 2006). Haritanın bu şekilde kullanılabilmesi için, haritada yer alan yapıların / nesnelerin uzaklıklarının, derinlik bilgisinin ve obje ve platform büyüklüklerinin bilinmesi de önem taşımaktadır.

İnsansız sistemlerin, çalışma mantığını genel olarak özetleyen algılama-planlama-davranış döngüsü aşağıda yer alan Şekil 6.1’de gösterilmiştir (Elbanhawi, 2017). Bu şekilden görüleceği üzere, uzaktan kontrollü sistemlere (operatör tarafından kontrol edilen sistemler) algılama ve planlama teknolojileri kazandırıldığında, ortaya çıkan sistem, otonom insansız sistem olarak adlandırılmaktadır (Elbanhawi, 2017). SLAM teknolojilerinin, pozisyon bulma ve haritalama tekniklerini içerir şekilde algılama teknolojileri arasında yer aldığı görülebilir. Otonom sistemlerde, planlanan göreve uygun davranışların gerçekleştirilmesi için algılama teknolojilerinin oldukça hassas şekilde çalışması gerekmektedir.



Şekil 6.1. Algılama-planlama-davranış döngüsü

SLAM konusunda yapılan ilk çalışmaların uzaklık sensörü (akustik ve LIDAR) temelli yaklaşımlar olduğu görülmektedir. Bu sensörler derinlik verisini yüksek doğrulukla (<3 cm hassasiyet) sunmasına rağmen, içinde bulunan çevrenin özellikleri hakkında detay bilgi vermemektedir. LIDAR sensörler, akustik sensörlere göre çevre hakkında daha fazla bilgi sağlamaktadır. Fakat LIDAR sensörler, görsel sensörler (kameralar) ve akustik sensörler ile karşılaştırıldığında hem daha ağır hem de daha fazla güç harcayan (50-200 W) sensörlerdendir. Devam eden çalışmalarda ise görsel sensörler (kameralar) kullanılarak

ortam özelliklerinin daha detaylı olduğu fakat derinlik verisinin yeterince güvenilir olmadığı yaklaşımlar ortaya çıkmıştır (Post, 2020). Bu sorunu çözmek için ise hem uzaklık hem de görsel verilerin bir arada kullanıldığı RGB-D ve stereo kameraların tercih edildiği uygulamalar da vardır (Cadena ve diğerleri, 2016). Fakat bu sensör ve kameraların ağırlıkları ve boyutları monoküler RGB kameralara göre oldukça fazladır. Dolayısıyla mikro-nano ölçekteki insansız araçlarda bu kameraların kullanımı ağırlık ve güç gereksinimleri açısından mümkün olmamaktadır. Bununla birlikte monoküler kameraların kullanıldığı uygulamalarda derinlik bilgisinin elde edilmesi daha zor olup, hesaplama maliyetleri de daha yüksektir. Fakat günümüz teknolojisinde özellikle nano-mikro insansız platformlar için, en uygun sensörlerin monoküler kameralar olduğu söylenebilir. SLAM metotlarında kullanılabilecek sensör çeşitleri ve bu sensörlere ait özellikler aşağıdaki tabloda özetlenmiştir (Zaffar ve diğerleri, 2018).

Çizelge 6.1. Sensör çeşitlerinin karşılaştırılması

Sensör Tipi	Harcanan Güç (W)	Derinlik Aralığı (m)	Ücret (\$)	Boyut
Akustik	0.01-1	2-5	10-500	Çok Kompakt
LIDAR	50-200	50-300	5k-100k	Büyük
Monoküler Kamera	0.001-10	Mümkün Değil	100-5k	Çok Kompakt
RGB-D	2-5	3-5	150-400	Kompakt
Stereo Kamera	2-15	5-20	500-5k	Kompakt
Dinamik Görsel Sensör	0.15-1	Mümkün Değil	3k-5k	Kompakt
Tüm yönlü Kamera	1-20	Mümkün Değil	100-1k	Çok Kompakt

Tabloda görüleceği üzere hem boyut olarak küçük hem de harcanan güç açısından avantajlı olan aynı zamanda çevrenin özellikleri hakkında da detaylı bilgi verebilen sensör olarak akıllı telefonlarda da tercih edilmekte olan monoküler kameralar göze çarpmaktadır. Monoküler kameralar kullanıldığında SLAM tekniği görsel eş zamanlı konumlama ve haritalama ismini (vSLAM) almaktadır. Bu teknikte anlık olarak harita oluştururken, eş zamanlı olarak bu harita üzerinde gösterilmek üzere kamera / platform pozunu da tahmin edilmektedir (Mur-Artal ve Tardós, 2017). vSLAM, diğer sensörlerin kullanıldığı SLAM metotları ile karşılaştırıldığında teknik olarak daha zor uygulanabilen bir metottur (Taketomi ve diğerleri, 2017). vSLAM uygulamaları öznitelik temelli ve doğrudan olmak üzere iki ayrı kategoride değerlendirilebilir. Öznitelik temelli yaklaşımlar, resimde yer almakta olan öznitelik noktalarını (ORB, SIFT, SURF gibi) kullanarak takip ve

haritalandırma işlevini gerçekleştirirken, doğrudan metotlar ise anlık olarak gelen resim karesinin bir sonraki karede yer alan gri seviyedeki değişimi karşılaştırılarak aynı işlevi yerine getirmektedir (Taketomi ve diğerleri, 2017).

vSLAM tekniğinde 3 ana aşama bulunmaktadır. Bunlar başlatma (initialization), takip (tracking) ve haritalama (mapping) aşamalarıdır. Başlatma aşamasında belirli bir koordinat sistemi tanımlanarak 3 boyutlu ortam ilk yakalanan öznitelik noktaları ile oluşturulmuş olur. Bu ortamdaki öznitelik noktaları ya da yapısal özellikler çıkarılarak insansız sistemin bu 3 boyutlu ortamdaki yeri başlangıç noktası olarak belirlenmiş olur. Takip aşamasında ise öznitelik temelli metotlarda öznitelik noktaları ya da doğrudan metotlar için pikseller / anahtar noktalar takip edilerek başlangıç noktasından itibaren ne kadar uzaklaşıldığı takip edilmeye çalışılır. Haritalandırma aşamasında ise başlangıç aşamasında oluşturulan 3 boyuttaki koordinatlara daha önceden bilinmeyen, yeni karşılaşılan yerler eklenerek haritanın kümülatif olarak oluşturulması sağlanır. Fakat vSLAM algoritmalarının gerçek zamanlı olarak çalıştırılabilmesi için gerekli olan teknolojiler ise yeniden konumlama ve global harita optimizasyonu teknolojileridir.

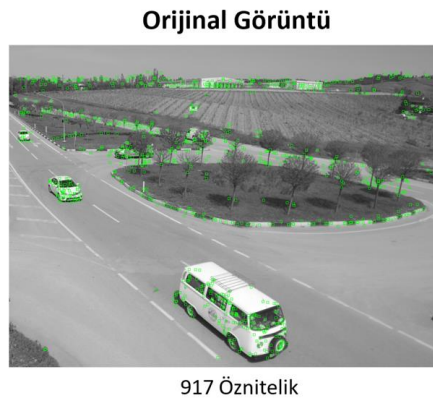
Takip işlemi, kameranın hızlı hareketleri ya da bazı düzensizliklerden (disturbance) dolayı doğru şekilde yapılamadığında yeniden konumlama yapmak gerekmektedir. Aksi halde sistem takipteki ilk kayıptan (tracking loss) itibaren çalışmasını doğru şekilde sürdürememektedir. Yeniden konumlanmanın (relocalization) yer almadığı vSLAM ya da görsel odometre algoritmaları genellikle akademik çalışmalar olmakla birlikte, bu tarz çalışmaların gerçek dünyada kullanılamamaktadır.

Haritalar genellikle kamera hareketinin doğru tahmin edilememesinden kaynaklı birikmiş tahmin hataları (scale drift gibi sebeplerle) içermektedir. Bu hataları en aza indirmek için global harita optimizasyonu yapılmaktadır. Bu aşamada tüm haritanın bütünü ele alınarak harita yeniden oluşturulmaktadır. Bazı kamera hareketlerinin ardından yeniden başlangıç noktası civarına gelindiğinin farkına varıldığında (veri tabanında tutulan öznitelik tanımlayıcıları / piksel şablonları karşılaştırılarak) daha önce yapılmış olan birikmiş hatanın tespiti ve ortadan kaldırılması mümkün olmaktadır. Elde edilen bu hatanın giderilmesi için döngü kapatma (loop closure) algoritmaları kullanılmaktadır. Döngünün tespiti için elde edilen yeni resim kareleri eskileri ile karşılaştırılmaktadır. Eğer birbirine çok yakın özellikler içeren resim kareleri mevcutsa, buradan daha önce geçildiğine dair bir

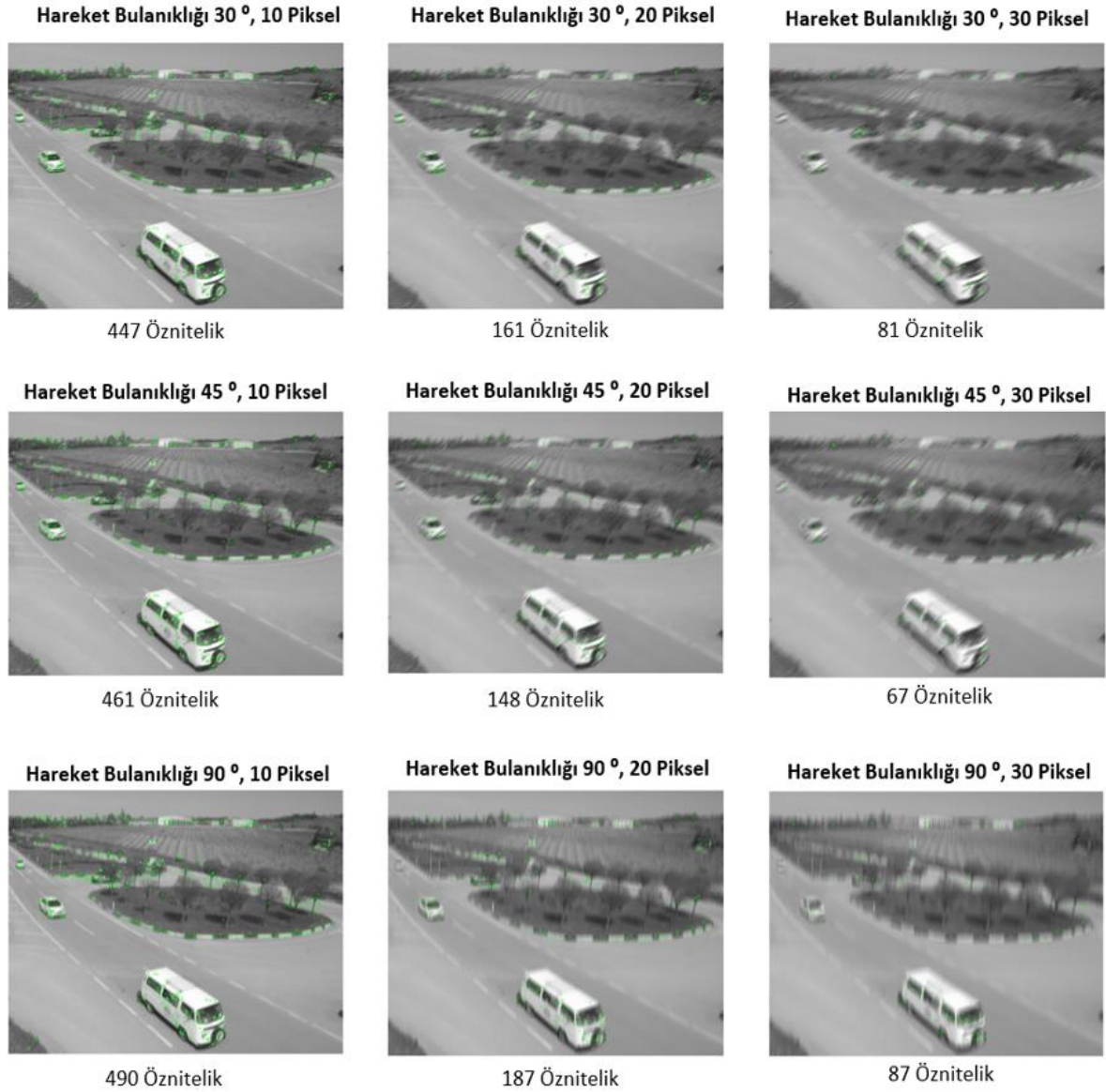
sonuca varılabilir. Bu durumda yörünge ve harita bu yeni bilgiye göre güncellenmektedir. Hareket bulanıklığı bahsi geçen bozulmalara neden olan en önemli etkenlerden bir tanesidir.

6.1. Sentetik Hareket Bulanıklığının Öznitelik Tespitine Etkisi

Hareket bulanıklığı arttıkça, görüntü karelerinde özniteliklerin tespitinde zorlanmalar ve bulunan özniteliklerin sayısında düşüş meydana gelmektedir. Bu durumu gözlemlemek amacıyla, dron kullanılarak çekilmiş görüntülerde özniteliklerin hareket bulanıklığı altında tespiti gerçekleştirilmiştir. Görüntünün çekiminde DJI Matrice 100 markalı üzerinde kameranın stabilizasyonunu sağlayan gimbal bulunduran bir dron platformu kullanılmıştır. Orijinal resim karesi, dronda yer alan ve bozucu hareketleri engelleyen gimbal mekanizması sayesinde herhangi bir bulanıklık içermemektedir. Bu görüntüde 917 adet öznitelik tespiti gerçekleştirilmiştir. Bu orijinal görüntülere, sahada karşılaşılabilecek tipte 30, 45 ve 90 derecelik; 10, 20 ve 30 piksel büyüklüklerinde hareket bulanıklıkları sentetik olarak eklenmiştir. Eklenen sentetik bulanıklık ile tespit edilen özniteliklerin sayısında 67 anahtar noktaya kadar azalmalar gözlemlenmiştir. Bu durumda eklenen hareket bulanıklığı 30 piksel büyüklükte ve her iki ekseninde de (45 derece) bileşenlere sahip olduğunda ve özniteliklerin sadece %7,3'ü yakalanabilmiştir. Bu durum öznitelik takibine dayalı olarak işlem gerçekleştiren algoritmaların başarısının hareket bulanıklığından ciddi oranda etkilendiğini göstermektedir. Aşağıda yer alan Şekil 6.2'de orijinal görüntüde tespit edilen öznitelikler ve toplam öznitelik sayısı verilmiştir. Şekil 6.3'te ise çeşitli açı ve şiddette (piksel cinsinden) lineer hareket bulanıklığı altındaki görüntülerde, tespit edilen özniteliklerin azaldığı gözlemlenmiştir.



Şekil 6.2. Orijinal görüntüde öznitelik çıkarımı



Şekil 6.3. Hareket bulanıklığının öznitelik çıkarımına etkisi

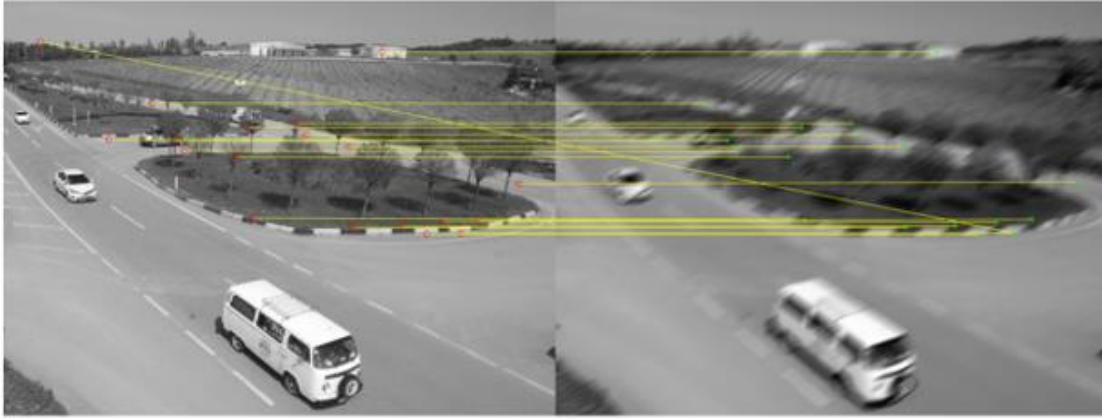
Şekil 6.3'ten de görüleceği üzere piksel bazında bulanıklık büyüklüğünün özniteliklerin çıkarımına negatif doğrusal bir etkisi mevcuttur. Bulanıklık büyüklüğü arttıkça, öznitelik çıkarımı azalmıştır. Bulanıklığın büyüklüğü 30 piksel mertebesine ulaştığında artık görüntü karesindeki bozulma insan gözüyle de anlaşılabilir seviyelere gelmektedir.

6.2. Sentetik Hareket Bulanıklığının Öznitelik Eşleştirmesine Etkisi

Hareket bulanıklığı öznitelik çıkarımını etkilediği gibi özniteliklerin eşleştirmesini de etkilemektedir. Bulanıklık, tespit edilen köşe, kenar ve benzeri özniteliklerin çevre pikseller ile keskin şekilde ayrılmasını engellemektedir. Bulanıklık nedeniyle öznitelikler

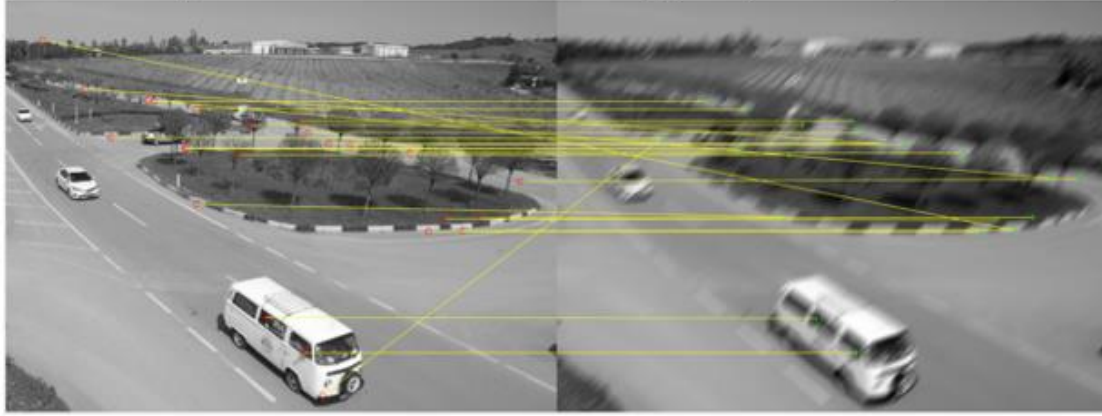
(kenar, köşe ve benzeri) çevre piksellere de dağılmakta ve piksel yoğunluğu arasındaki fark azalmaktadır. Bu durumda orijinal görüntüde yer alan öznitelikler ile bulanıklık görüntülerdekiler birbirinden oldukça ayrılmaktadır. Bu nedenle hareket bulanıklığı altında tespit edilen aynı pikseldeki öznitelikler, bulanıklık nedeniyle farklı öznitelik olarak değerlendirilebilirler. Çünkü daha önce anlatıldığı gibi öznitelik tanımlayıcılar her anahtar noktayı 128 ile 256 bitten oluşan ve çevre pikseller ile ilişkileri de kodlayan ikili öznitelik vektörleri ile ifade etmektedir. Bu nedenle hem anahtar noktada hem de etrafında yer alan piksellerdeki hareket nedenli bozulma öznitelik vektörlerinin değişmesine neden olmaktadır. Aynı görüntü karesinde ya da ardışık görüntülerdeki öznitelikler hareket bulanıklığı nedeniyle birbirlerinden o kadar farklılaşırlar ki öznitelik vektörleri birbirleri ile uyuşmamaya başlamaktadır. Bu nedenle aynı öznitelik farklı öznitelik gibi değerlendirilir ve öznitelikler arasında eşleştirme gerçekleştirilemez. vSLAM algoritmaları resim karelerinde aynı öznitelik eşleştirerek hareket tahmini gerçekleştirmektedirler. Eşleştirilen özniteliklerin sayısı ne kadar çok olursa, hareket tahmini o kadar hassas gerçekleştirilebilir. Aşağıda yer alan Şekil 6.4’de bulanıklık büyüklüğü 30 piksel olan, 30, 45 ve 90 derecelik hareket bulanıklığı içeren görüntüler ile orijinal görüntü arasındaki öznitelik eşleştirmesi gerçekleştirilmiştir. Şekilden görüleceği üzere bulanıklık nedeniyle eşleşme sayıları oldukça azalmış olup bazı öznitelikler ise farklı öznitelikler ile yanlış olarak eşleştirilmiştir.

Orijinal Görüntü Karesi ile Bulanık Görüntü Arası Eşleştirme (30 °, 30 Piksel)



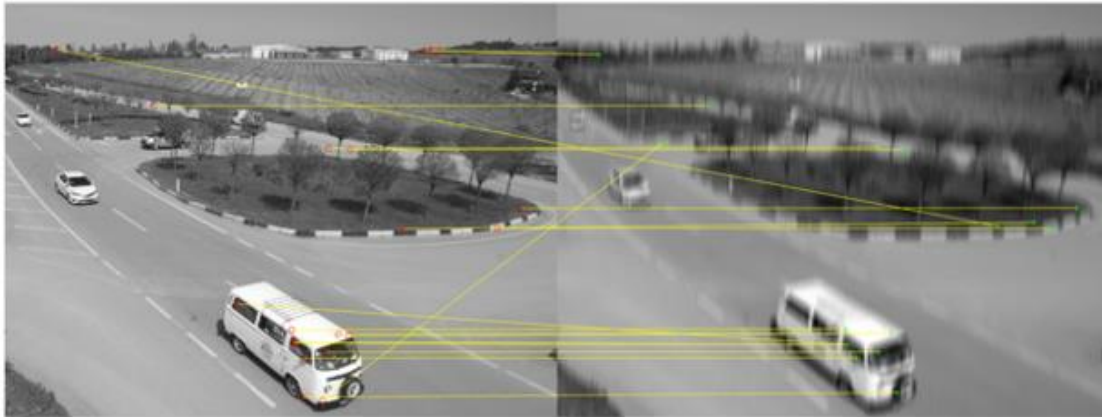
19 Öznitelik

Orijinal Görüntü Karesi ile Bulanık Görüntü Arası Eşleştirme (45 °, 30 Piksel)



24 Öznitelik

Orijinal Görüntü Karesi ile Bulanık Görüntü Arası Eşleştirme (90 °, 30 Piksel)



20 Öznitelik

Şekil 6.4. Hareket bulanıklığının öznitelik eşleştirmeye etkisi

Yukarıda sonuçları verilen eşleştirmeler sentetik olarak oluşturulmuş hareket bulanıklığının öznitelik tespiti ve eşleştirmesi üzerine olumsuz etkilerini göstermektedir. Hareket bulanıklığı altında tespit edilen 67 özniteliğin sadece 24 tanesinin eşlenebildiği

görülmektedir. 30 derece ve 30 piksellik hareket bulanıklığı altında ise tespit edilen 81 özniteliğin sadece 19 tanesi orijinali ile eşlenebilmektedir. Bu durum, hareket bulanıklığını nedeniyle az sayıda tespit edilen özniteliğin sadece yaklaşık %20'sinin eşleştirebileceğini göstermektedir. Buradan orijinal net resimde yer alan 917 adet özniteliğin bulanıklık altında sadece 19 adetinin vSLAM uygulamalarında takip aşamasında kullanılabileceği anlaşılmaktadır. Özet olarak orijinal keskin resimde yer alan özniteliklerin sadece %2 gibi düşük oranı takip amacıyla kullanılabilir durumdadır. Bu orana yanlış eşleştirmeler de dahil edildiğinde hareket bulanıklığının vSLAM uygulamalarını ne derece olumsuz etkilediği görülmektedir.

Sentetik bulanıklaştırmanın yaygın bir kullanımı da bir görüntüyü yumuşatmak ve daha temiz bir görsel elde etmek için bulanıklaştırmaktır. Gaussian ya da Median filtre kullanılarak bu tipte bir bulanıklaştırma yapmak ve hareketin ya da nesnelerin ortaya çıkarılmasını sağlamak da mümkündür. Doğal hareket bulanıklığı ise daha karmaşık bileşenler içermektedir. Doğal hareket bulanıklığı durumunda özniteliklerin ve doğrudan yöntemlerin kullandığı gri seviye farklılıklarının nasıl etkilendiği gerçekleştirilen deneysel çalışmalar ile ilerleyen başlıklarda detaylı olarak incelenmiştir.

6.3. Mikro / Nano Ölçekteki İnsansız Araçlarda Konumlama

Yakın gelecekte, sayıca çok, ucuz ve sürü halinde görev yapabilen mikro/nano insansız araçların ve robotların kullanımının artması planlanmaktadır. 2035 yılı ve ötesinde artık bu platformların savunma alanı başta olmak üzere pek çok alanda sahada görev alacak seviyede (teknolojik hazırlık seviyesi 9 ve üstü) olması beklenmektedir (Department of Defence [DoD], 2013). GNSS sinyallerinin kullanımına engel olan durumlardan biri de platform boyutlarının küçültülmesidir. Çünkü bilindiği üzere GNSS sinyalleri metre seviyesi konumlama hassasiyetlere sahip sinyallerdir (Aqel ve diğerleri, 2016). GNSS sinyalleri kullanılarak konumu tespit edilen bir nesne/ platform, aslında birkaç metre sapma ile bulunmuş olur. Boyutlar küçüldükçe, metre seviyesindeki konumlama hassasiyeti platformun kontrolü ve konumlanması açısından yetersiz kalmaktadır. Örneğin, 100 gram ağırlığında 10 santimetre skalasında boyutlara sahip bir quadcopter platformunun 10 metre hassasiyet ile konumunun bulunması, platform boyutundan yüzlerce kat fazla bir hata ile platformun konumlanması anlamına gelmektedir. Bu seviyedeki bir hata, platformun operasyonunu doğru şekilde sürdürmesine imkân

sağlayamayabilir. Hatta, böyle bir durumda platformun konumu tam olarak bilinemediği için çarpışmaya maruz kalması ihtimali de artmaktadır. Bu nedenle, mikro ve nano seviyedeki insansız platformların konumlanması ve kontrolünün sağlanması daha büyük platformlara göre daha zordur. Özellikle bu platformların ağırlık kısıtlarından dolayı gelişmiş kameralar taşıyamaması, güç kaynaklarının sınırlı olması da bu platformlarda yapılabilecek işlemleri kısıtlamaktadır. Ayrıca bu platformlar, tekrar kullanımı da mümkün olmayacak tek kullanımlık olarak sarf edilebilir bir şekilde, sürü halinde kullanılması planlanan platformlardır. Bazı durumlarda sahaya gönderilen bir insansız aracın görevini tamamlandıktan sonra dönmesi istenmeyebilmektedir. Bu nedenle kullanılacak malzemelerin ve sensörlerin oldukça maliyet uygun olması tercih edilmelidir. Aksi takdirde, yüzlerce/binlerce hatta bazı senaryolarda milyonlarca mikro/nano insansız araçların kullanımı mümkün olmamaktadır.

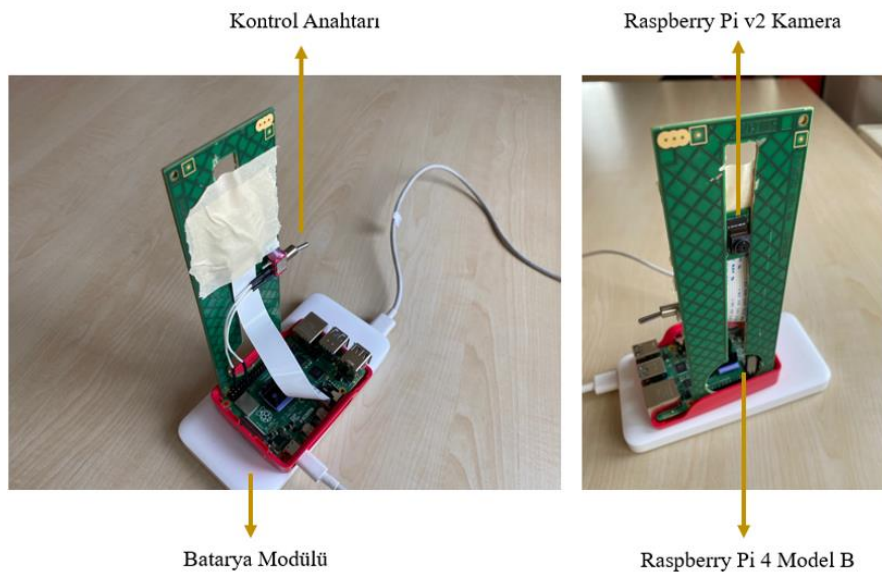
Nano / mikro insansız araçlarda vSLAM algoritmalarının kullanılabilmesi için tercih edilebilecek en uygun seçeneklerden biri, daha kompakt, düşük maliyetli, kolayca kalibre edilebilen ve hafif monoküler RGB kameralardır (Y. Chen ve diğerleri, 2018). Bu tipte kameralar, genellikle herhangi bir gimbal ya da kontrol sağlayan ek bir mekanizma olmadan, doğrudan platform üzerine montelenmiş olarak kullanılmaktadır. Bu tipte bir kullanımda karşılaşılabilecek en temel problemler daha önce de bahsedilmiş olan görüntü ölçek belirsizliği ve hareket kaynaklı bulanıklıktır (Yu ve diğerleri, 2011).

Daha önce gerçekleştirilen çalışmalarda bu problemlerin çözümü stereo (Howard, 2008; Azartash ve diğerleri, 2014) ve RGB-D (Fabian ve Clayton, 2014; A. Huang, Bachrach ve Henry 2017; Fang ve Zhang, 2015) kamera kullanımı veya görsel ataletsel SLAM (viSLAM) metotlarıdır. Fakat bu tip kameraların kullanımı hem boyutlar hem ağırlık hem de güç tüketimi açısından dezavantajlar barındırmaktadır. Ek olarak mikro /nano insansız hava araçları (İHA) taşıdıkları bataryanın küçüklüğü ve sağlanan gücün kısıtlı olması nedeniyle yalnızca yaklaşık 10 dakika kadar havada kalabilmektedirler. Bu nedenle viSLAM gibi platform üzerinde işlem gücü gerektiren metotlar bu tipteki platformlar için uygun olmayabilmektedir. Bunun yerine platformdan alınan görüntünün, bir kontrol paneline aktarılması ve bu kontrol panelinde yer alan işlemcinin kullanılması daha fazla operasyon zamanı sağlayabilmektedir. Ayrıca platform üzerinde yer alan ataletsel ölçüm biriminin kalibrasyonu için gereken süre 3 dakikaya kadar çıkabilmektedir. Bu durum viSLAM yöntemlerinin kullanımının bir diğer dezavantajı olarak değerlendirilmektedir.

Bu teze konu olan çalışmada özellikle mikro/nano İHA'larda ve robotlarda uygun maliyetli, düşük güç tüketimli ve hafif monoküler kameraların kullanıldığı durumlara vSLAM algoritmalarının hareket bulanıklığına olan hassasiyetleri detaylı olarak analiz edilmiş ve literatürde şimdiye kadar yer almamış, özgün bir vSLAM çerçevesi önerisinde bulunulmuştur. Bu çerçeve hem doğrudan hem de öznitelik temelli yaklaşımlara uygun şekilde tasarlanmış olup, çerçevenin çalışma performansı oluşturulan özel veri kümelerinde test edilmiştir.

Kamera Seçimi ve Kalibrasyonu

Belirlenen vSLAM algoritmalarının testi için nano mikro insansız sistemlerde ve robot platformlarında sıkça kullanılan iğne deliği kamera tipi seçilmiştir. Bu kameraların ortak özelliği hem boyut ve ağırlık hem de maliyetten tasarruf sağlamalarıdır. Yapılan çalışmada iğne deliği kameralardan en yaygın kullanılanlardan biri olan Raspberry Pi v2.1 (Resim 6.1) kamerası kullanılmıştır. Veri kümesi oluşturulurken ilgili kamera ile 640x480 boyutlarında saniyede 20 resim karesi (fps:frame per second) hızında kayıt alınmıştır. Kontrol anahtarı kullanılarak kayıt başlatmak ve tamamlamak mümkün olabilmektedir. Literatürde yer alan veri kümelerinin birçoğunda çözünürlüğü yüksek ve anlık titreşimlerden arındırılmış görüntüler yer almaktadır. Düşük çözünürlükte ve gimball kullanılmadan oluşturulmuş veri kümesi üstünde vSLAM algoritmalarının görev icra edebilmesi nano mikro boyuttaki insansız sistemler ve robotlar için önemli bir yetenektir.



Resim 6.1. Raspberry Pi kartı ve üzerinde Raspberry Pi v2.1 model kamera

vSLAM algoritmalarının görüntüden öznitelik / doku çıkarımı yapabilmesi açısından kamera kalibrasyonu elzemdir. Kullanılan kameranın merceğinin özellikleri vSLAM algoritmalarının performansını önemli ölçüde etkilemektedir. Veri kümesi oluşturan her enstitü / üniversite / araştırma merkezi kendi kullanmış olduğu kameranın kalibrasyon parametrelerini paylaşmaktadır. Bu parametreler vSLAM algoritmalarının resim karesinde hangi bölgelerde işlem yapacağını belirlemektedir. Bu parametreler sayesinde daha 2 boyutlu resim uzasyından 3 boyutlu gerçek dünyaya projeksiyon yapmak mümkün olabilmektedir. Kamera kalibrasyonu için en sık kullanılan yöntemlerden biri olan siyah-beyaz damalı yüzey / satranç tahtası kullanılarak Raspberry Pi v2.1 kameranın kalibrasyon matrisi ve bozulma parametreleri aşağıdaki gibi bulunmuştur:

$$\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 518.695 & 0 & 332.380 \\ 0 & 518.757 & 229.837 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.1)$$

$$[k_1 \quad k_2 \quad p_1 \quad p_2 \quad k_3] = [0.24 \quad -0.358 \quad -0.009 \quad 0.005 \quad 0.029] \quad (6.2)$$

Oluşturulan bu kalibrasyon matrisi, ilgili kamera için özeldir. Bu işlemi bir kere gerçekleştirmek yeterli olup, her yeniden başlama durumunda kalibrasyon tekrarına gerek olmamaktadır. Bu durum sadece kamera kullanılan vSLAM metotlarının avantajlarından biridir.

Veri Kümesi ve Test Senaryoları

Enstitüler, araştırma merkezleri ve üniversiteler tarafından daha önceden oluşturulmuş veri kümelerinde (EUROC (Burri ve diğerleri, 2016), TUM Mono (Engel ve diğerleri, 2016) ve KITTI (Geiger ve diğerleri, 2013) gibi) genellikle yüksek kalitede, pahalı ve herhangi bir ağırlık limiti olmayan kameraların kullanıldığı gözlemlenmiştir. Bu ve benzeri veri kümelerinde yer alan platforma yaptırılan hareketler genellikle ani ve keskin hareketler içermemektedir. Bahsi geçen veri kümelerine ve bu çalışma kapsamında oluşturulan veri kümesine ait özelliklerin özeti Çizelge 6.2’de verilmiştir. Çizelge 6.2’den görüleceği üzere hem ağırlık hem maliyet açısından oldukça avantajlı bir kamera seçimi yapılmıştır.

Çizelge 6.2. Veri kümelerinin sensör özelliklerinin karşılaştırılması

Özellikler	KITTI (Geiger ve diğerleri, 2013)	EUROC (Burri ve diğerleri, 2016)	TUM Mono (Engel ve diğerleri, 2016)	Oluşturulan Veri Kümesi
Kamera	FL2-14S3C-C (2 adet)	MT9V034 (2 adet)	IDS UI- 3241LE-M-GL	Raspberry Pi v2.1
Kamera Fiyatı	562.40\$ (Toplam:1124,8\$)	66,25\$ (Toplam:132,5\$)	357\$	37,5\$
Kamera Ağırlığı	58 gram (Toplam:116 gr)	10 gr (Toplam:20 gr)	12 gr	3 gr
Platform	Otomobil	İHA	Elde Taşınabilir Platform	Elde Taşınabilir Platform
Güç Tüketimi	2,5W (Toplam:5W)	1,65W (Toplam 3,3W)	1,5W	1W

Aşağıda yer alan Resim 6.2’de ilgili veri kümelerine ait kamera görselleri verilmiştir. Bu kameralar yukarıdaki tabloda belirtildiği üzere bazı veri kümelerinde birden çok adetli kullanılmıştır. EUROC, TUM Mono ve bu çalışma kapsamında kullanılan kameralar birbirleriyle uyumlu olup aynı taşıyıcı karta monte edilebilir durumdadır.

KITTI (Geiger, 2013)
FL2-14S3C-CEUROC (Burri, 2016)
MT9V034TUM Mono (Engel, 2016)
IDS UI-3241LE-M-GLOluşturulan Veri
Kümesi

Resim 6.2. Veri kümeleri oluşturulurken kullanılan kameralar

EUROC ve TUM Mono veri kümeleri incelendiğinde platformun sapma (yaw) hareketini oldukça nadir şekilde ve belirli bir kontrol ile yaptığı gözlemlenmektedir. Bu durumun, bu veri kümelerinin çoğunlukla, vSLAM yöntemlerinin konsept doğrulamalarının yapılmasından dolayı olduğu düşünülmektedir.

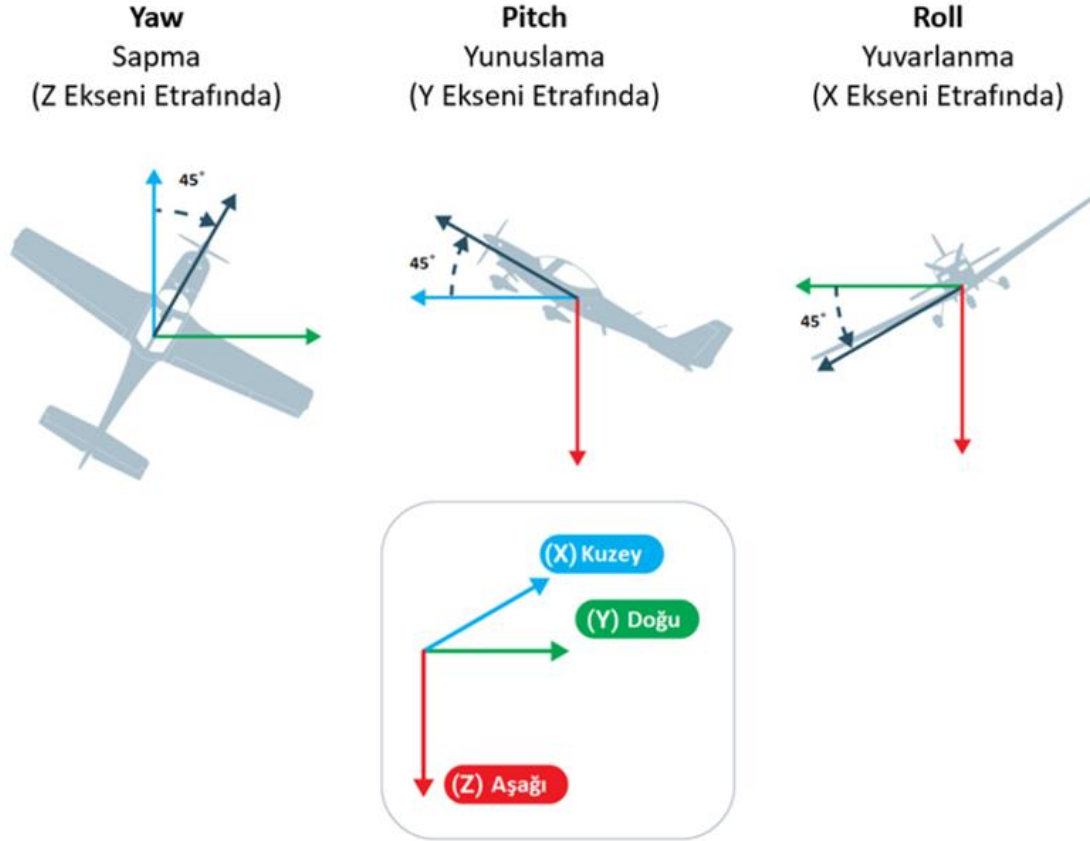
Platformun hızlı sapma yaptığı durumda özellikle bir dönemeç mevcutsa, pek çok yeni öznitelik ile karşılaşılacak ve eski öznitelikler hareketin neden olduğu bulanıklık nedeniyle gözlemlenemeyecektir. Ayrıca platformun zorlu hareketlere maruz kaldığı durumda yapabileceği ani hareketler de hareket bulanıklığına sebep olabilir. Özellikle bu etkilerin gözlemlenmesi amacıyla yeni bir veri kümesinin oluşturulması gerekli görülmüştür.

Literatürde ön plana çıkmış olan öznitelik temelli ve doğrudan vSLAM algoritmalarının, daha düşük kalitede, daha ucuz ve hafif kameralar kullanıldığında performanslarını gözlemlemek amacıyla zorlayıcı koşullara sahip veri kümeleri oluşturulmuştur. Bu veri kümeleri gerçek dünyada karşılaşılabilecek problemlere yönelik daha zorlu koşulları da içermektedir. Veri kümeleri mikro /nano boyutlardaki İHA'ların daha atik ve çevik oldukları değerlendirilerek oluşturulmuştur.

Zorlu Koşullara Sahip Veri Setinin Oluşturulması

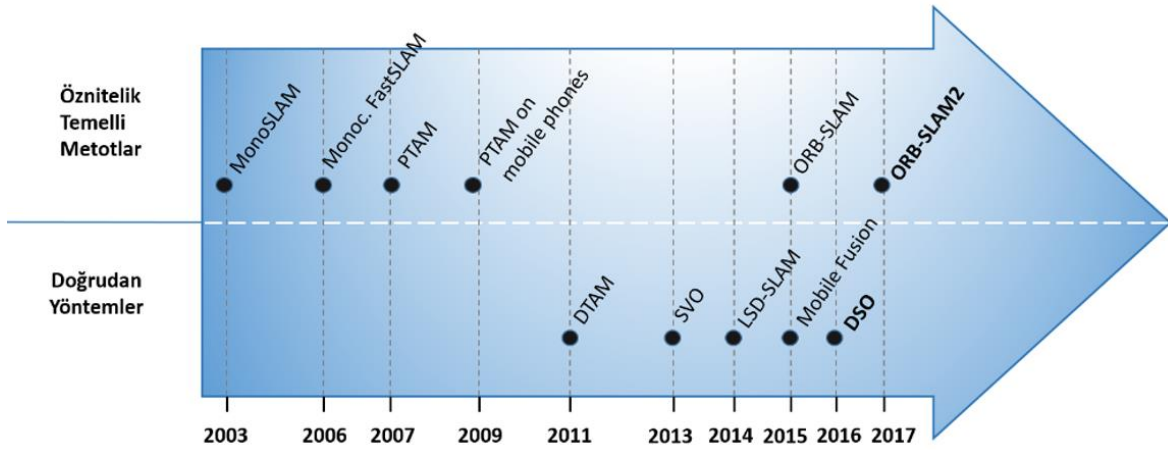
İlk veri kümesi seti iç mekânda 140 santimetre (cm) yükseklikte uçmakta olan mikro dron hareketleri taklit edilerek oluşturulmuştur. İlgili veri kümesi vSLAM algoritmalarının ani titreşimlerden ve İHA'ların uçuş esnasında gerçekleştirdiği yunuslama (pitch), yuvarlanma (roll), sapma (yaw) hareketlerinden kaynaklı görüntü kaymalarına karşı duyarlılıkları gözlemlemek için oluşturulmuştur. İlgili veri kümesinde yavaş (ortalama 0,8 m/s) ve hızlı (ortalama 1,5 m/s) olarak belirlediğimiz iki hız seviyesinde yunuslama ($\pm 45^\circ$), yalpalama ($\pm 45^\circ$) ve sapma ($\pm 45^\circ$) açılarının her birinin ayrı ayrı yer aldığı toplam üçer adet veri kümesi ve üç hareketin aynı anda yapıldığı (bütünleşik) ayrı bir veri kümesi oluşturulmuştur. Ayrıca yine her iki hız seviyesinde veri kümemizde bir merkez etrafında geniş bir dairesel dönüş yaptığımız (dönüş 1) ve 90° sapma (yaw) yaparak döndüğümüz (dönüş 2) ikişer adet senaryoya yer verilmiştir. Veri kümeleri iç mekânda koridor ortamında bahsedilen video kamera ile çekilmiştir. Koridor ortamları dokular bakımından zayıf ortamlar olduğu için vSLAM algoritmaları için zorlu ortamlardır. Literatürde koridor ortamları için ayrıca çözümler arayan çalışmalar da mevcuttur (G. Zhang ve Suh, 2012). Ayrıca belirlenen algoritmaların pozisyon tahminlerinin değerlendirilmesi için dikdörtgen bir katta döngüyü tamamladığımız 2 adet senaryomuz (döngü 1 ve döngü 2) veri kümemizde yer almaktadır. Döngü içeren senaryolarda gerçekleştirilen hareketler daha yumuşak olup, algoritmaların döngü kapama performanslarının gözlemlenmesi için oluşturulmuşlardır. Özetlenecek olursa; veri kümesinde görüntü duyarlılıklarını test etmek

amacıyla yavaş seviyede 6 adet, hızlı seviyede 6 adet ve pozisyon tahminleri için de 2 adet olmak üzere toplamda 34438 adet resim karesinden oluşan 14 ayrı veri kümesi yer almaktadır. Şekil 6.5'te sapma, yunuslama ve yuvarlanma hareketleri eksenlere göre gösterilmiştir.



Şekil 6.5. Veri kümesi oluşturulurken gerçekleştirilen zorlu hareketler

Bu veri kümelerinde özellikle, hareket bulanıklığının hangi hızlarda ve hangi durumlarda doğrudan ve öznitelik temelli metotları etkilediği gözlemlenmeye çalışılmıştır. Bu amaçla hem öznitelik temelli hem de doğrudan metotlarda en olgun seviyeye ulaşılmış olan ORB-SLAM2 (Mur-Artal ve Tardós, 2017), DSO (Engel ve diğerleri, 2017) ve LDSO (Gao, 2018) algoritmalarının performansı gözlemlenmek üzere seçilmiştir. Bu yöntemler, Şekil 6.6'te gösterildiği gibi kendi klasmanlarında en üst seviyelerde yer almaktadır. LDSO algoritması, DSO yöntemine döngü kapatma modülü eklenmesiyle oluşturulduğu için bu ismi almıştır.



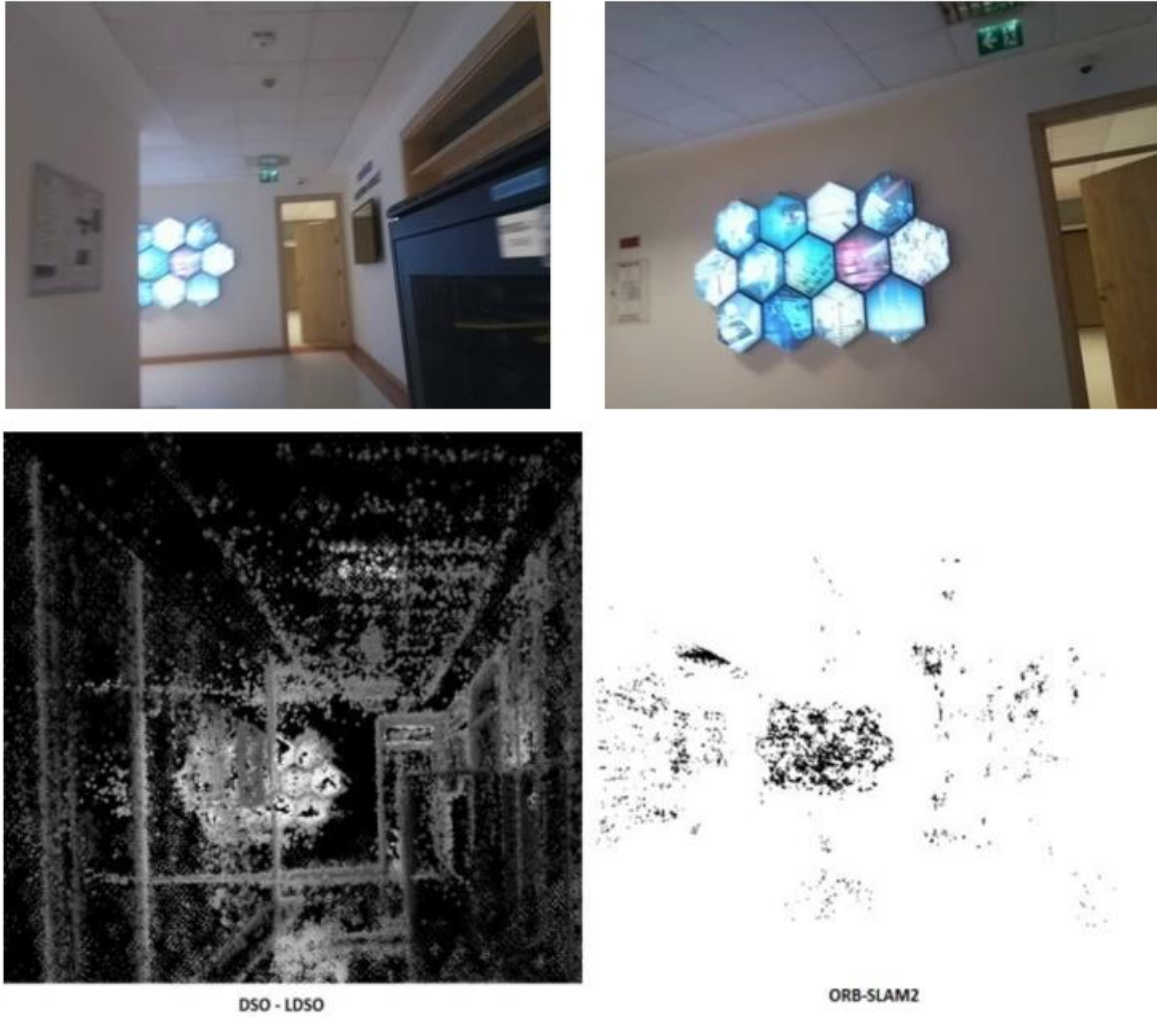
Şekil 6.6. vSLAM metotlarının kronolojik sıralaması

Seçilen algoritmalar, 4 çekirdekli Xeon W3520 işlemci ve 12 GB'lık RAM bulunan donanımda Linux 18.04 işletim sisteminde test edilmiştir. Bu donanımın ortalama seviyedeki bir kontrol istasyonuna benzer olduğu değerlendirilmektedir.

Zorlu Koşullara Sahip Veri Setine ait Deneysel Sonuçlar

Harita Anlaşılabilirlik Seviyesi

DSO ve LDSO aynı haritalama yöntemini (gri ölçek yoğunluğuna göre haritalama) kullandıkları için bu iki algoritma ilgili veri setinde aynı tip seyrek haritaları oluşturmaktadır. LDSO algoritmasının haritalama ve piksel takibi yaptığı modüller DSO algoritması ile aynıdır. Bu nedenle oluşan haritaların görselleri birbirleri ile aynı olmaktadır. ORB-SLAM2 algoritması ise bu iki algoritmayla karşılaştırıldığında çok daha az yoğunlukta bir nokta bulutu oluşturmaktadır. İlgili algoritmaların oluşturduğu haritalardan aynı koridora ait örnekler Şekil 6.7'de verilmiştir. Görüleceği gibi DSO ve LDSO algoritmaları anlaşılabilirlik seviyesi daha yüksek haritalar üretmektedir. Harita nokta bulut yoğunluğu diğer bir deyişle resim karelerinde bulunacak öznitelik/piksel sayısı artırılabilir. Fakat bu sayı artırıldığında gerçek zamanlı performansın oldukça düştüğü bilinmektedir. Bu nedenle özellik sayısının uygulamaya göre optimum bir noktada tutulması önemlidir. Yapılan testte her iki algoritma için anahtar nokta sayısı 2000 adette tutulmuştur.

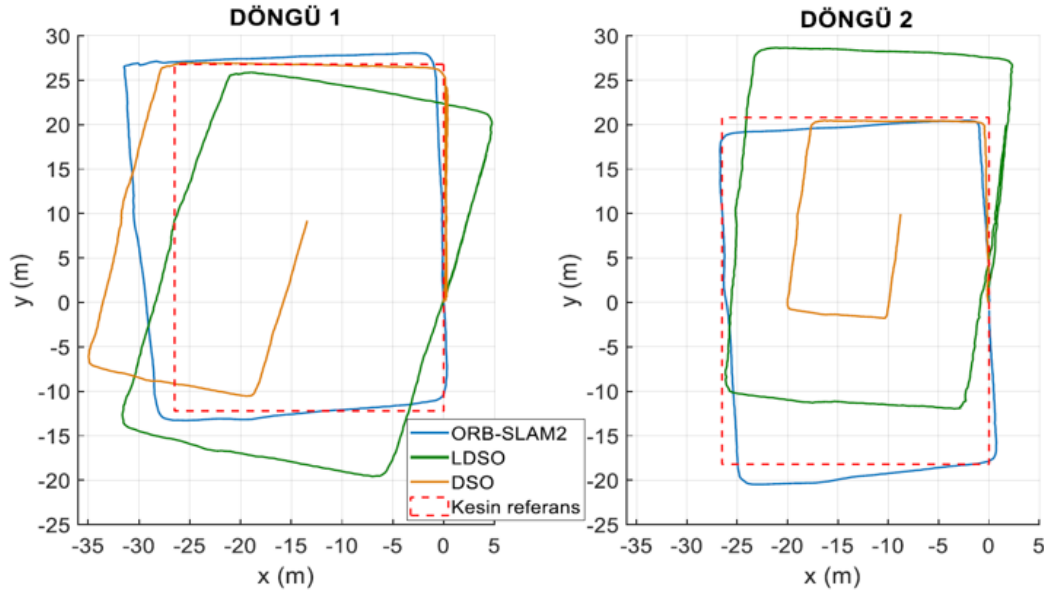


Şekil 6.7. DSO-LDSO ve ORB-SLAM2 nokta bulutu haritaları

Pozisyon Tahmini ve Lokalizasyon Performansı

DSO, LDSO ve ORB-SLAM2 algoritmalarının performansları uzun tarafı 39 metre kısa tarafı 26,5 metre olan dikdörtgen şeklindeki bir katta bir tam tur dönülerek hazırlanmış olan 2 adet veri setinde (Döngü 1, Döngü 2) karşılaştırılmıştır. İlgili veri kümesinde koridorun tam orta noktası belirlenerek bu orta nokta üzerinden sapmadan kayıt alınmıştır. Şekil 6.8’den görüleceği üzere kesin referansa en yakın sonuç ORB-SLAM2 algoritması tarafından tahmin edilmiştir. Algoritmaların testinde kullanılan Döngü 1, Döngü 2 veri kümelerinde kayıt alınmaya başlanan noktada tüm algoritmaların aynı başlangıç noktasına sahip olduğu fakat hareket sürdürüldükçe yörünge tahmininde sapmaların olduğu görülmektedir. Bununla birlikte ORB-SLAM2 algoritması ve LDSO algoritması burada bir döngü tamamlandığını bulabilmiş ve başlangıç noktasının üzerinden ikinci defa geçildiğini tespit etmiştir. Bu iki algoritmada döngü kapama (loop closure) tekniğinin doğru çalıştığı

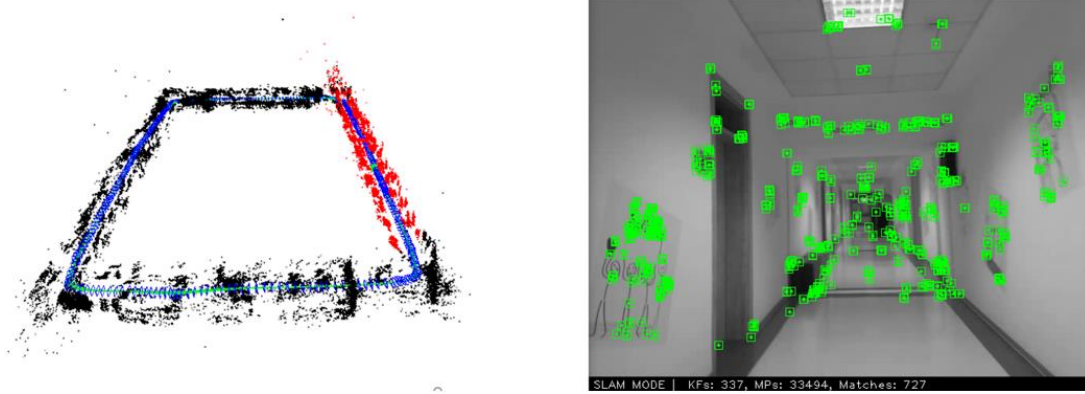
görülmektedir. DSO algoritması ise döngü kapama modülüne sahip olmadığı için beklendiği üzere döngüyü kapatamamaktadır fakat Döngü 1’de ilk ve ikinci dönüşleri, Döngü 2’de ise ilk dönüşü LDSO’dan daha başarılı şekilde tahmin etmiştir. Burada LDSO’nun döngü kapama işlemi yaparken DSO’daki bu başarılı dönüş tahminlerini kaydırdığı tespit edilmiştir.



Şekil 6.8. DSO, LDSO ve ORB-SLAM2 pozisyon ve lokalizasyon sonuçları

Sonuçlar incelendiğinde DSO algoritmasının ölçeklemede belirsizlik yaşadığı, gerçek uzunluğu aynı olan koridor parçalarını sanki daha kısaymış gibi tahmin ederek dönüş yaptığı görülmektedir. Döngü kapama özelliği bulunan LDSO ve ORB-SLAM2 algoritmalarında ise hareket devam ederken yörünge tahmininde biriken kümülatif hataların, döngü kapama aşamasında ortadan kaldırıldığı ve ölçeklemedeki belirsizliğin yok edildiği gözlemlenmiştir.

İkili öznitelik vektörlerinin yer aldığı veri kümesinde daha önceden yer alan özniteliklere hareket esnasında tekrar denk gelindiğinde, eşleştirme gerçekleşir ve ilgili anahtar karenin olduğu yerden tekrar geçildiği anlaşılır. O zamana kadar olan tüm kümülatif hata anahtar karenin olduğu yere tekrar gelindiğinde sıfırlanır. Bu işleme döngü kapama işlemidir. Oluşturulan veri kümesinde döngünün kapandığı ana ait ekran görüntüsü Resim 6.3’de verilmiştir.



Resim 6.3. Döngü 1 veri kümesine ait döngü kapama işlemi

Veri Kümesindeki Zorlayıcı Koşullar Altında Performans

vSLAM algoritmalarında öznitelik / piksel takibi ani dönüşler ve hareket / ışık kaynaklı bozulmalar nedeniyle koptuğunda eğer yeniden konumlama yapılamıyorsa takip ve haritalama işlemine son verilir ve insansız sistem nerede olduğunu olarak bilemez hale gelir. Çünkü bir önceki resim karesi ile arından gelen resim karesi arasındaki bağ tamamen ortadan kalkmış olur. Bu durumda konumlama ve haritalama işlemlerine devam edilemez. Çünkü daha önce bulunan yerler ile bu yeni yer arasında herhangi bir bağlantı kurulamamaktadır. İlgili algoritmaların veri kümemizde yer alan zorlu koşullardaki performansları sırasıyla aşağıdaki Çizelge 6.3 ve Çizelge 6.4 verilmiştir.

Çizelge 6.3. Yavaş (0,8 m/s) hızda vSLAM algoritmalarının performansları

vSLAM Algoritmaları	Yavaş Seviyedeki Koşullar (0.8 m/s)					
	Roll ($\pm 45^\circ$)	Pitch ($\pm 45^\circ$)	Yaw ($\pm 45^\circ$)	Bütünleşik	Dönüş 1	Dönüş 2
ORB-SLAM2	✓	✓	✓	✓	✓	✗
DSO	✓	✗	✓	✗	✓	✓
LDSO	✓	✗	✓	✗	✓	✓

Çizelge 6.4. Hızlı (1,5 m/s) hızda vSLAM algoritmalarının performansları

vSLAM Algoritmaları	Yavaş Seviyedeki Koşullar (1.5 m/s)					
	Roll ($\pm 45^\circ$)	Pitch ($\pm 45^\circ$)	Yaw ($\pm 45^\circ$)	Bütünleşik	Dönüş 1	Dönüş 2
ORB-SLAM2	✓	✗	✓	✗	✓	✗
DSO	✓	✗	✗	✗	✓	✗
LDSO	✓	✗	✗	✗	✓	✗

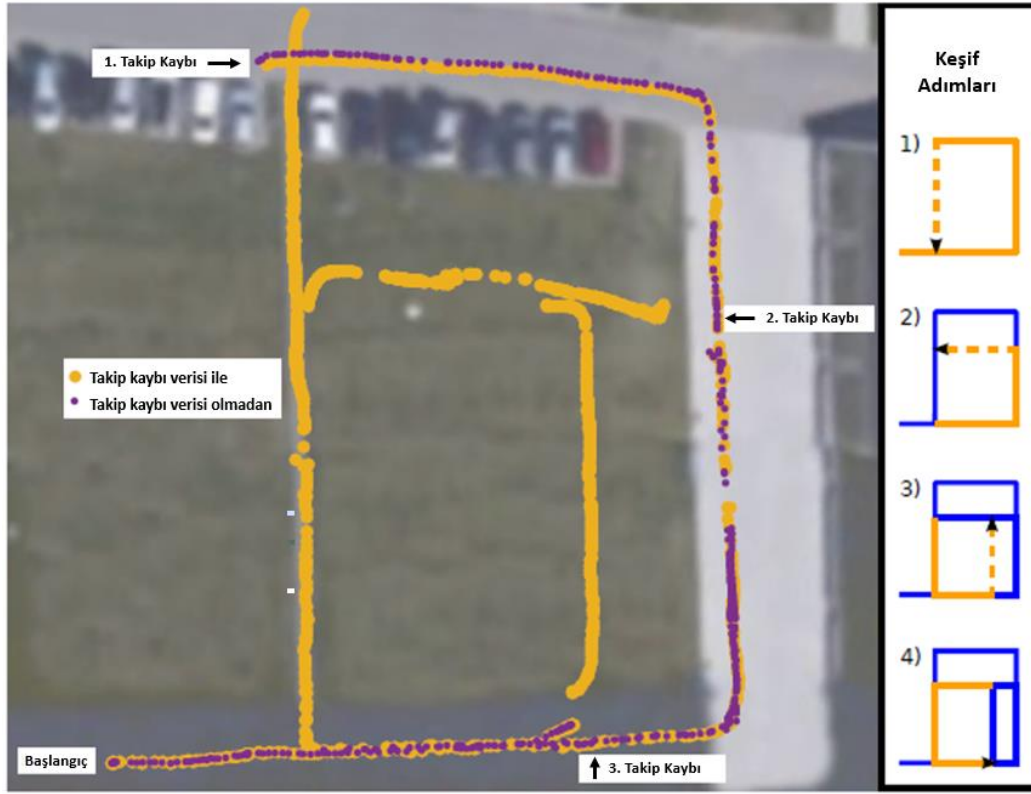
Çizelge 6.3 ve Çizelge 6.4 den görüleceği üzere ORB-SLAM2 algoritmasının zorlu koşullarda daha başarılı şekilde operasyon sürdürmeye uygun olduğu bu sonuçlardan çıkarılabilir. İlgili algoritma zorlu koşullarda kısa süreli takip kaybı yaşasa bile resim karelerindeki öznitelik noktalarını tanıyarak kendi konumu yeniden bulabilmektedir. Yani daha önce tespit etmiş olduğu ORB öznitelikleri, zorlu koşullarda kamera kadrajından çıksa bile tekrar bu özniteliklerin görüntülendiği durumda tekrar eşleştirme gerçekleştirilmekte ve ilgili anahtar resim karesi ile bağlantı kurulmaktadır. DSO ve LDSO algoritmalarında ise takip kaybı olduktan sonra yeniden konumlanmanın sağlanamadığı gözlemlenmiştir. Bununla birlikte ORB-SLAM2 algoritması platformun kendi ekseni etrafında 90° sapma (yaw) hareketi yaparak döndüğü ‘Dönüş 2’ veri kümesinde takip kaybı yaşarken diğer iki algoritma yavaş hızda bu dönüşü takip kaybı olmadan gerçekleştirebilmiştir. Bu durumun burada dönüşün yapıldığı duvarın dokusuz olmasının yani ORB özniteliklerinin çıkarılabileceği herhangi bir dokunun bulunmaması etkili olmuştur. Hızlı dönüşlerde ise hiçbir algoritma ‘Dönüş 2’ veri kümesinde başarılı olamamıştır. Dönüş 2’de tüm algoritmaların başarısız olma sebebinin dönüş hareketi esnasında meydana gelen hareket bulanıklığı olduğu değerlendirilmiştir.

Gerçekleştirilen bu deneyde, hızlı dönüş (sapma - yaw) ve yüksek dereceli ($\pm 45^\circ$) yunuslama (pitch) hareketine maruz kalınan durumlarda hem doğrudan hem de öznitelik temelli yöntemlerin takip kaybına uğradığı ve dönüş hareketi sonrasında koridorun yeni bölümü ile önceki bölümleri arasında bir bağlantı kuramadıkları gözlemlenmiştir. $\pm 45^\circ$ ’lik yunuslama hareketi operasyon şartlarında çok fazla karşılaşılabilecek bir durum değildir. Çünkü bu açıda gerçekleştirilen harekette İHA platformunun 45° ile yere doğru çakılma / pike hareketi ve ardından 45° derece ile yukarı doğru yükselme hareketi gerçekleştirmesi simüle edilmiştir. Bu tip bir hareket yalnızca algoritmaların tepkisini ölçmek için yapılmış

olup, normal operasyon sırasında gözlemlenebilecek bir durum olarak değerlendirilmemektedir. Çünkü özellikle dron platformlarında kamikaze dalış dışında bu tip hareket kullanılmamaktadır. Fakat hızlı dönüş hareketleri, agresif manevraları yapabilen nano / mikro İHA'larda oldukça sıkça karşılaşılan bir harekettir. Dolayısıyla bu hareketin neden olduğu takip kaybı önlenmeden, mikro / nano İHA'larda vSLAM metotları kullanılarak haritalama ve konumlama yapılması mümkün görülmemektedir. Bu nedenle ani dönüş hareketlerinin incelenmesi amacıyla çalışmalar genişletilmiştir.

6.4. vSLAM Tekniklerinde Takip Kaybına Yönelik Çözümler

Şimdiye kadar gerçekleştirilen akademik çalışmalarda, monoküler kamera tabanlı vSLAM yöntemlerinde meydana gelen hareket bulanıklığından kaynaklanan hataların azaltılması veya takip kaybının önlenmesi konusunda sınırlı sayıda çalışma yapılmıştır. Takip kaybına yönelik gerçekleştirilen çalışmalardan en önemlilerinden birinde (S. Schubert, Lange, Neubert ve Protzel, 2016) izleme kaybının ortaya çıkmasından sonra veri kaybının (ters tekrar) önlenmesi ve gerçek zamanlı uygulamalar için uygun olmasa da dallanan iş parçacığı yapısının (paralel takip) uygulanması ile çeşitli çözümler sunulmaktadır. Ters tekrar yönteminde takip kaybı meydana geldikten sonraki kısım tekrar üzerinde çalışılmak üzere kayda alınır. Alınan bu kayıt ile, kaybın gerçekleştiği yerin devamında gelen alanlara ait yerlere dair veri kaybı önlenmiş olur. Hareket tümüyle bittiğinde hareket kaybının olduğu kısmın haritası da çıkarılarak ana yörüngeye çevrimdışı olarak eklenmeye çalışılır. Bu yöntemde ilgili kısmın video kaydının gerçekleştirilmesi sadece o bölgenin görsel olarak kayda alınmasını sağlamaktadır. Bununla birlikte çevrimdışı yapılacak haritalamada manuel bazı çıkarımların yapılması da gerekmektedir. Ayrıca takip kaybının yaşandığı bölgelerin çok olması video kayıt boyutlarının beklenenin üstünde gerçekleşmesine neden olabilmektedir. Önerilen diğer yöntem olan paralel takip metodunda ise, bir takip kaybı meydana geldiğinde, yeni bir iş parçacığı üretilir. Bu iş parçacığı sanki vSLAM algoritması baştan başlatılmış gibi takip kaybının devamında çalışmaya başlamaktadır. Birden çok art arda takip kaybı olduğunda iş parçacıklarının sayısı da artmaktadır. En sonunda tüm iş parçacıklarının sonuçları birleştirilerek yörünge tam olarak çıkarılması çalışılmıştır. Her yeni iş parçasında 'döngü kapama' ve 'yerel haritalama' özellikleri devre dışı bırakılmaktadır. Bu yöntemde doğrudan yekpare bir harita elde edilmemekle birlikte, tüm iş parçacıklarının çıktıları ile yekpare bir yörünge çıkarılmaya çalışılmaktadır (S. Schubert ve diğerleri, 2016).



Resim 6.4. Paralel iş parçacığı ile örnek bir yörüngenin çıkarımı

Bu yöntemlerden geri tekrar yöntemi gerçek zamanlı olarak kullanılamamaktadır. Paralel iş parçacığı temelli yöntemde ise, gerçekleştirilen dönüşlerin 90° 'lik bir dönüş olduğu bilinerek, tüm yörünge çıkarımı (Resim 6.4) yapılmıştır. Fakat burada gerçekleştirilen dönüş hareketi örneğin 70° 'lik bir dönüş olup ve önceden bilinmeseydi, tüm yörüngenin doğru şekilde elde edilmesi mümkün olmaması muhtemeldir. Bu yöntemle ilgili bir diğer durum ise, tekrar tekrar takip kaybının meydana gelmesi durumudur. Takip kaybı sonrasında yeni bir iş parçacığı başlatıldığında çoklu takip kaybı ile de karşılaşılma ihtimali vardır. Bir takip kaybının ardından gelen yeni takip kaybı için de yeni iş parçacığı başlatılarak iş parçacıklarında dallanmaya neden olunabilmektedir. Bu durum CPU'nun iş yükünü ve hafıza kullanımını özellikle platform üzerindeki kullanımlarda ciddi derecede verimsizleştirecektir.

Benzer şekilde başka bir çalışmada (Pretto, 2007), insansı robotlar için, yürüme, dönme ve çömelme hareketlerinden kaynaklanan yoğunluğu şok ve titreşim nedenli hareket bulanıklığı etkilerine karşı sağlam olan nokta yayılma fonksiyonu tahmini kullanılarak bir öznitelik eşleştirici yöntemi sunulmuştur. Bu çalışmada ele alınan titreşimler daha çok adım atan bacaklı robotlara uygun olup, yoğunlukla titreşim ve darbe hareketlerini

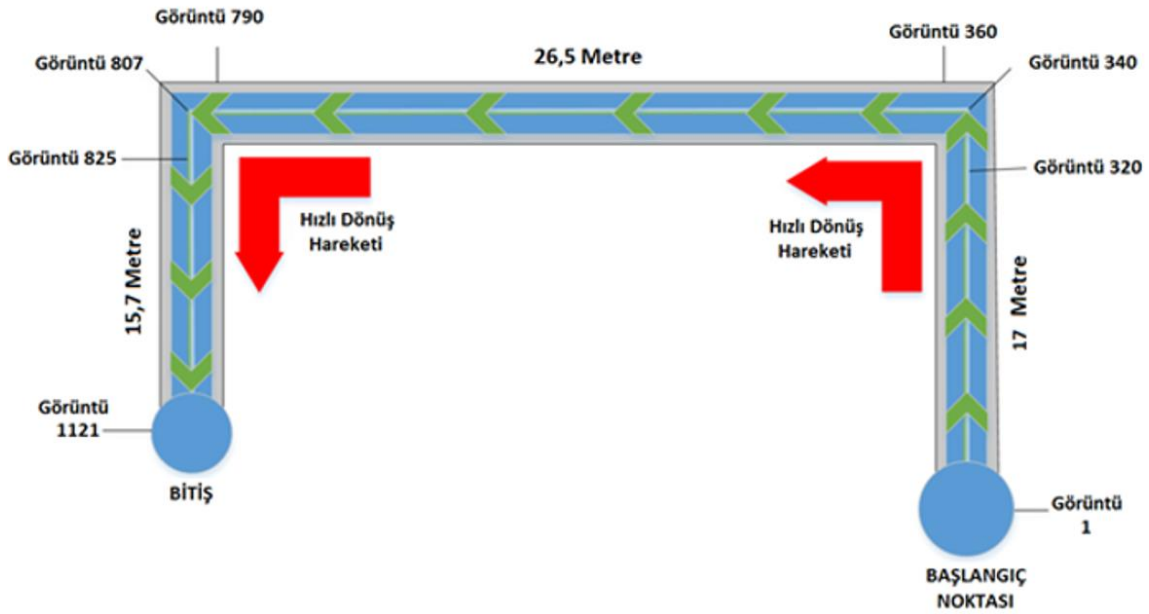
içermektedir. Bacaklı robot platformların hareket profili insansız hava araçlarının özellikle dönüş hareketleri düşünüldüğünde doğrudan uyarlanabilir bulunmamaktadır.

Yukarıda bahsedilen çalışmalara ek olarak, ilgili harita zenginliğini ve izleme performansını geliştirmek için kenarlar, çizgiler vb. gibi ek özniteliklerin tespit edilmesi ve kullanılmasına yönelik çeşitli araştırmalar da mevcuttur (Klein ve Murray, 2008; Pumarola, 2017). Hareket bulanıklığının yönü ve çizgilerin/kenarların yönü tutarlıysa, bu tür yaklaşımlar, izleme performansını, yani aracın yörünge tahminini iyileştirmek için uygun hale gelir. Ancak, haritalamada, yani 2B görüntü özelliklerinin 3B alana yansıtılmasında aynı performans elde edilemez: harita üzerinde hareket bulanıklığının neden olduğu kayan çizgiler nedeniyle haritalama performansı düşer. İzleme kaybını önlerken harita tutarlılığını sağlamak da oldukça önemlidir.

6.5. Hızlı Dönüş Hareketi İçeren Veri Kümesinin Oluşturulması

Daha önce gerçekleştirilen deneylerde elde edilen sonuçlar değerlendirildiğinde, nano/mikro İHA'ların en fazla karşılaşacağı takip kaybının agresif manevraları içeren ani dönüş hareketlerinde oldukları gözlemlenmiştir. Daha derin bir analizin gerçekleştirilmesi için, koridor ortamında ortalama 1,2 m/s hızda, dron hareketleri benzetimi ile bir deney gerçekleştirilmiştir. Deney ortamı, üç koridordan (17–26.5 ve 15,7 metrelik) ve iki keskin köşeden oluşmaktadır. Deney ortamının şematik bir görünümü Şekil 6.9'da çizilmiştir.

Şekil 6.9'de belirtildiği gibi, ileri hareket yeşil renkle, hızlı dönme hareketleri ise kırmızı renkle gösterilmiştir. Ayrıca dönme hareketlerine karşılık gelen başlangıç, bitiş noktaları ve görüntü numaraları da aynı şekilde yer almaktadır. Deney alanı sadece hızlı dönme hareketlerinden kaynaklanan hareket bulanıklığını gözlemlemek için kurulmuştur. İleriye doğru harekette takip kaybı, genellikle dokusuz ortamlarda, özellikle öznitelik tabanlı algoritmalarda meydana gelebilmektedir. Ancak deney alanımızda incelenen ORB-SLAM2 (Mur-Artal ve Tardós, 2017) ve DSO (Engel ve diğerleri, 2017) algoritmalarının 1,2 m/s ortalama hızda ileri hareketlerde hareket bulanıklığına karşı dirençli olduğu gözlemlenmiştir. Bu algoritmaların ileri hareketinde herhangi bir hareket kaybı meydana gelmemiştir.



Şekil 6.9. Deney ortamına ait çizim

İlgili donanım platformu ile 22 derece/s veya 0.384 rad/s hızında zengin özellikli köşelerin yakınında hızlı dönme hareketleri gerçekleştirilmiştir. Platform üzerinde daha önceki deneyde de kullanılan ve nano/mikro insansız sistemlerde ve robot platformlarında sıklıkla tercih edilen iğne deliği kamerası Raspberry Pi v2.1 kullanılmıştır. Daha önce de belirtildiği üzere yeni bir veri kümesi oluşturulmasının nedeni, akademide kullanıma sunulmuş veri kümelerinde, 90°'lik hızlı dönüşleri içermemesi ve kullanılan kameraların çoğunlukla kaliteli kameralar olup, nano / mikro seviyesi İHA'larda kullanılamayacak büyüklük, güç tüketimi ve maliyet özelliklerine sahip olmasıdır. Ayrıca düşük maliyetli kameraların, gelecekte tek kullanımlık, tekrar kullanılamaz araçlarda en çok tercih edilen kameralar olacağı değerlendirilmektedir (Department of Defence [DoD], 2013). Şekil 6.9'den görüleceği üzere gerçekleştirilen ani dönüş hareketinden sonra yeni bir koridora giriş gerçekleştirilmektedir. Bu durumda dönüş esnasında takip kaybı olması durumunda vSLAM algoritmasına yeni koridorda tanıdık gelen herhangi bir doku olmamaktadır. Platform kendini tamamen bambaşka bir yere taşınmış gibi değerlendirecektir.

Deneyde 140 cm yükseklikten ilgili kamera kullanılarak görüntüler alınmıştır. Platform, köşelere ulaşana kadar ortalama 1,2 m/s hızda ileri yönde hareket gerçekleştirmiş, köşelerde 90 derece keskin ve 0.384 rad/s hızında hızlı sapma (yaw) hareketi gerçekleştirilmiştir. Veri seti, ileriye bakan bir kamera ile ilgili koridor yörüngesinde 1121

görüntü karesi elde edilerek oluşturulmuştur. İlgili veri kümesinden bir kesit Resim 6.5'te verilmiştir.



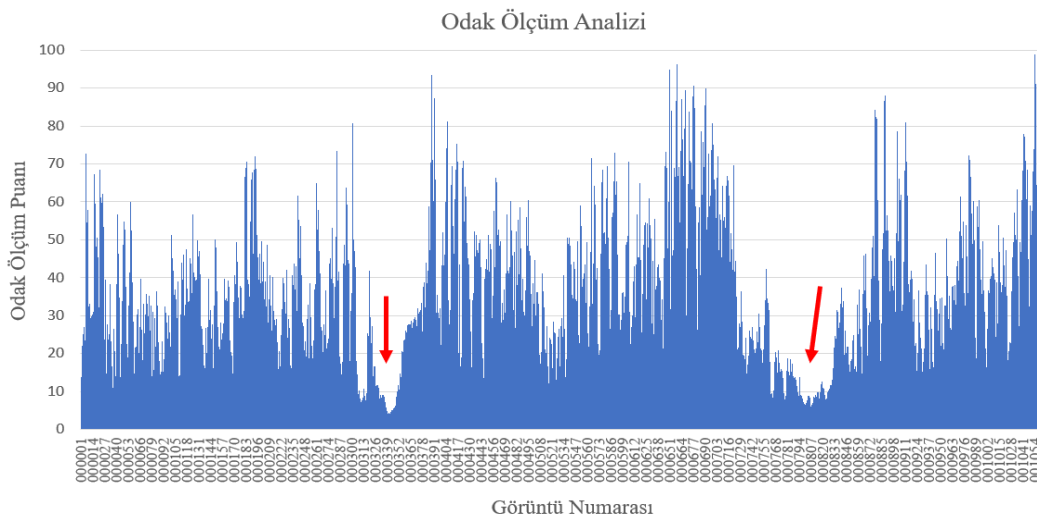
Resim 6.5. Veri kümesinde yer alan görüntülerin bir kısmını gösteren resim

Veri setinden hedeflenen özniteliklerin çıkarılması için görüntü işleme hızı oldukça önemlidir. Farklı işlem hızları için hareket bulanıklığı seviyesi değişmemekte, ancak dedektör ve tanımlayıcıya (descriptor) tahsis edilen işlem süresi değişmektedir. Örneğin, işlem hızı 20 ile 10 fps arasında değiştirildiğinde, bir resim karesi için ayrılan işlem süresi iki katına çıkar ve böylece ilgili resim karesinde işlemi tamamlamak için daha fazla zaman verilebilmektedir. Bununla birlikte, karşılık gelen pozlandırma süresi her işlem hızı için aynı tutulmuş ve bulanıklık etkisi herhangi bir değişiklik olması istenmemiştir. Oluşturulan veri setindeki görüntüler, ORB-SLAM2 (Mur-Artal ve Tardós, 2017) ve DSO (Engel ve diğerleri, 2017) algoritmalarında 20, 10 ve 5 fps işlem hızında işlenmiştir. Her iki algoritmada da seçilen işlem hızlarında hızlı dönme hareketlerinde izleme kaybı meydana gelmiştir. Algoritmalar dönüş hareketlerinden sonra takibe devam edememişler ve çalışmalarını durdurmuşlardır. Koridor ortamında kahve makinesi, dolap, kapı vb. çeşitli nesneler ayırt edici özelliklerle yerleştirilmiştir. Bu sayede ileri yöndeki harekette öznitelik çıkarma ve üçgenleme işlemleri kolaylıkla gerçekleştirilmiştir. Gerçekleştirilen deneyde,

0.384 rad/s hızındaki platform dönüş hareketinin neden olduğu hareket bulanıklığının ORB-SLAM2 (Mur-Artal ve Tardós, 2017) ve DSO (Engel ve diğerleri, 2017) algoritmalarının belirtilen işlem hızlarında haritalama ve konumlama görevlerini yerine getirmesini engellediği görülmüştür.

6.5.1. Veri kümesinin detaylı analizi

Meydana gelen takip kayıplarının nedeni, odak ölçüm operatörü kullanılarak tespit edilmiştir. Hareket bulanıklığını detaylı şekilde analiz edebilmek için kullanılabilecek çok sayıda odak ölçüm operatörü vardır (Pertuz, Puig ve Garcia, 2013). Çalışmamızdaki veri seti, gerçek zamanlı uygulamalar için en uygun yöntemlerden biri olan ve daha önce mikroskop kameralarında keskinleştirme çalışmalarında kullanılan LAP4 yöntemi ile analiz edilmiştir. LAP4 yöntemi Laplace temelli bir operatördür. Laplace operatörleri, Laplace'ın ikinci türevi aracılığıyla görüntülerde yer alan kenarların sayısını ölçmeyi amaçlar. Laplace tabanlı operatörler, ikinci türev hesaplamaları nedeniyle gürültüye karşı çok hassastır (Junera, 2009). Laplacian Varyansı analizi sayesinde, hızlı dönme hareketi durumlarına karşılık gelen odak ölçüm puanının (OÖP, Focus Measure Score - FMS) diğer resim karelerine oranla nispeten düşük olduğu gözlemlenmiştir. Tüm veri kümesi, LAP4 operatörü ile analiz edildiğinde Şekil 6.10'da görüldüğü şekilde her görüntüye karşılık gelen odak ölçüm puanı elde edilmiştir.



Şekil 6.10. Veri kümesinde odak ölçüm analizi

Şekil 6.10. 'da kırmızı oklarla gösterilen ve odak ölçüm puanının nispeten düşük olduğu bölgelerde takip kaybının meydana geldiği tespit edilmiştir. Bu görüntü karelerinin hareket bulanıklığına daha fazla maruz kaldığı, göz ile tespit yöntemi ile de veri kümesi üzerinden gözlemlenmiştir. Ölçüm puanının yüksek olduğu görüntüler de yine beklendiği şekilde, daha az bulanıklık içeren ve resim netliğinin yüksek olduğu görüntülerdir.

Laplacian Varyansı (LAP4) Operatörü

Laplacian Varyansı (LAP4) odak ölçüm operatörü, ilk olarak mikroskoplarda otomatik odaklanma çalışmalarında ortaya çıkarılmıştır (Pertuz ve diğerleri, 2013). Bu operatör, görüntünün Laplace'ının varyans değerini kullanır. Çoğunlukla odaktan şekil çıkarımı (Shape From Focus - SFF) uygulamalarında kullanılmaktadır. Odaktan şekil çıkarmak için, her pikselin odak seviyesinin tespit edilmesi gerekmektedir. Odaktan şekil çıkarımında LAP4 operatörü Eş 6.1'de verilmiştir:

$$\sigma_{i,j} = \sum_{(i,j) \in \Omega(x,y)} (\Delta I(i,j) - \overline{\Delta I})^2 \quad (6.3)$$

Burada $\overline{\Delta I}$ ifadesi, $\Omega(x,y)$ bölgesindeki görüntünün Laplace'ının ortalamasıdır. Metod oldukça sade bir işlem içermekte ve metodun uygulanması yüksek işlem gücü gerektirmemektedir. Resim üzerindeki uygulamasında görüntünün gri tona çevrilmiş hali ile 3x3 boyutlarında Laplace çekirdeği konvolüsyon işlemine tabi tutulur. Çıkan sonucun varyansı alınarak odak ölçüm puanı hesaplanır. Laplace çekirdeği 3x3 boyutlarında aşağıda verildiği gibidir:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (6.4)$$

Aşağıda odak ölçüm puanı hesaplamasının, Python programlama diliyle yazılan kodu Resim 6.6'da verilmiştir. İlgili kod çevrimdışı olarak veri kümesinde yer alan tüm görüntülerin odak ölçüm puanını hesaplayarak bir metin (txt) dosyasına kaydetmektedir.

```

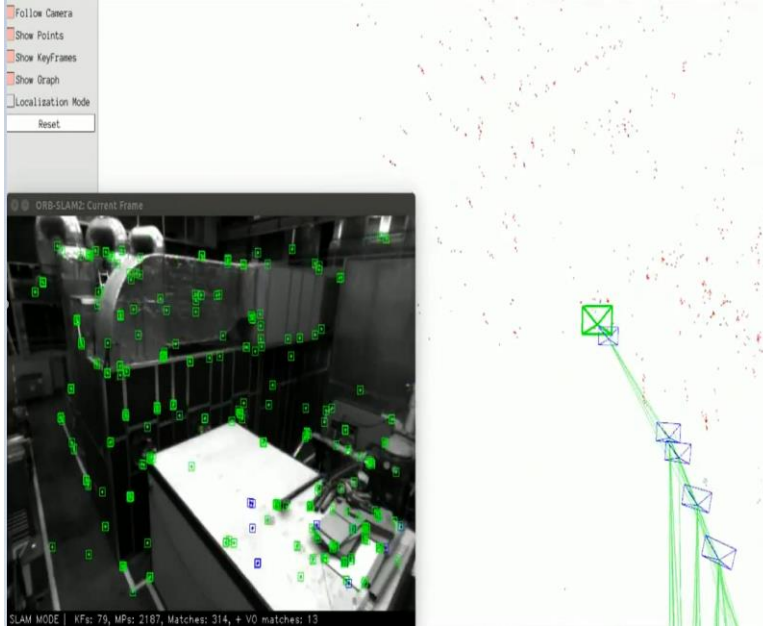
1  from imutils import paths
2  import argparse
3  import cv2
4  def variance_of_laplacian(image):
5      # önce görüntünün Laplace'ı hesaplanır
6      # ardından sonucun varyansı bulunur
7      return cv2.Laplacian(image, cv2.CV_64F).var()
8  ap = argparse.ArgumentParser()
9  ap.add_argument("-i", "--images", required=True,
10                 help="Görüntülerin bulunduğu klasörü yolu")
11  args = vars(ap.parse_args())
12  #OpenCV ile Bulanıklık Tespiti
13  #Sonuçların kaydedileceği text dosyası açılır
14  f= open("/home/bugras/Downloads/blur_detect/Odak_olcum_puani.txt","w+")
15  #Tüm görüntü girdileri sırasıyla alınır
16  for imagePath in paths.list_images(args["images"]):
17      # Görüntü yüklendikten sonra gri tona çevrilir
18      # Laplace Varyansı ile Odak Ölçüm Puanı (fms) bulunur
19      image = cv2.imread(imagePath)
20      gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
21      fms = variance_of_laplacian(gray)
22      z=imagePath+" "+str(fms)
23      f.write("Sonuc %s\r\n" %z))
24  f.close()

```

Resim 6.6. LAP4 yöntemine ait Python programlama dilindeki kodları

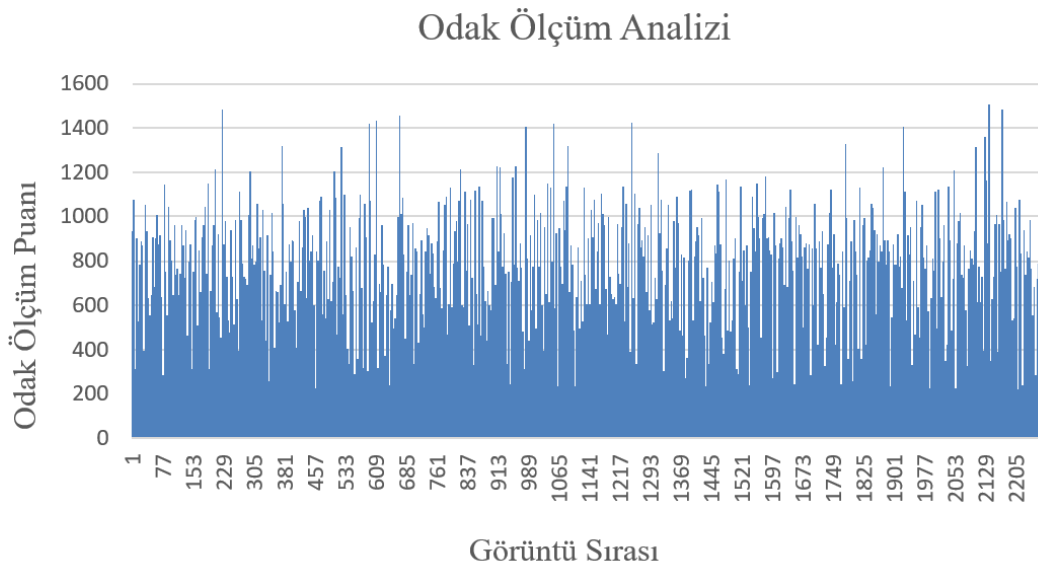
6.5.2. Mevcut veri kümelerinin analizi

TUM Mono, KITTI ve EUROOC veri kümelerinde en hızlı hareketlerin gerçekleştirildiği veri sekansları incelenmiştir. KITTI veri kümesi otomobil platformu üzerinde yerleştirilmiş stereo kamera ile oluşturulduğundan İHA’larda oluşabilecek hareket bulanığına benzerlik göstermemektedir. EUROOC veri kümesinde en hızlı hareketlerin yapıldığı ve herhangi bir ek sensör yerleşiminin bulunmadığı “MH05_difficult” isimli veri sekansı incelenmiştir. Bu veri kümesinde doğrudan dron kullanılarak veri sekansı oluşturulmuştur fakat dron ile agresif dönüş hareketi gerçekleştirilmemiştir. Yapılan dönme hareketi belirli bir açıda ve hareket kaynaklı bulanıklığa neden olmayacak şekildedir. Resim 6.7’de EUROOC veri kümesindeki “MH05_difficult” veri sekansına ait bir görsel paylaşılmıştır.



Resim 6.7. EUROC veri kümesi MH05_difficult isimli veri sekansından görüntü

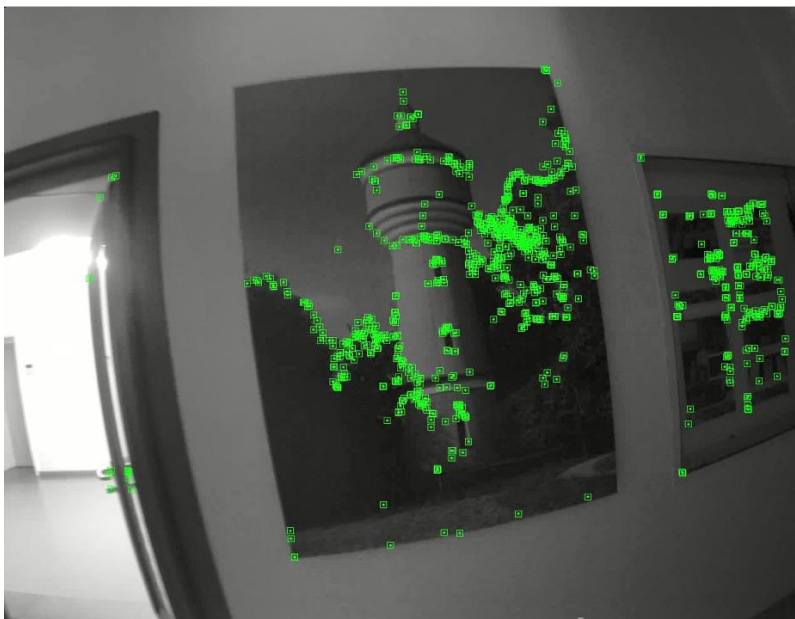
Bu MH05_difficult isimli sekansa ait odak ölçüm analizi gerçekleştirilmiştir. Her bir görüntüye karşılık gelen odak ölçüm puanı incelenmiştir. Elde edilen görüntülerin bulanıklık seviyesinin oldukça düşük olduğu ve bulanıklığa neden olacak hareketlerden kaçınıldığı görülmüştür. Bulanıklık seviyesinin yüksek olduğu (OÖP'nin düşük olduğu) görüntüler sadece belirli görüntüler olup hareket devamlılığı olmadığı için öznitelik takip kaybına neden olmamıştır. İlgili sekansa ait grafik Şekil 6.11'de verilmiştir.



Şekil 6.11. EUROC-MH05_difficult sekansı odak ölçüm analizi

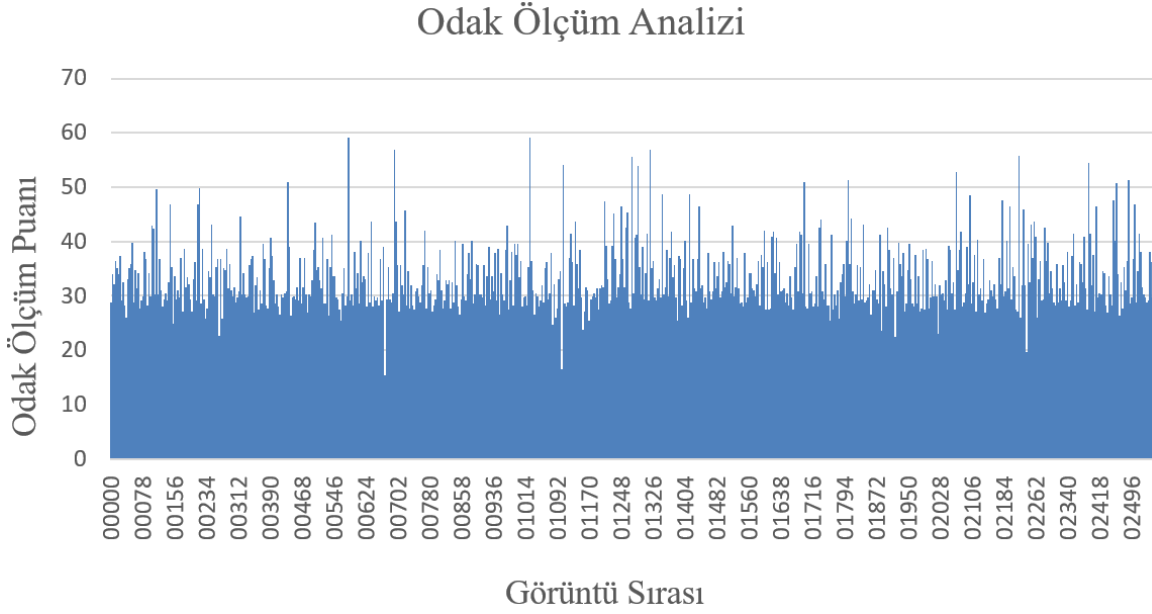
EUROC veri kümesi, yaklaşık 1600 gram ağırlığında bir dron ve üzerine yerleştirilmiş stereo kamera sistemi ile oluşturulmuştur. Dolayısıyla platformun kaldıracabileceği ağırlık miktarı fazladır. Bu nedenle daha kaliteli ve pahalı bir kamera seti kullanılabilmıştır.

Mikro / nano ölçekteki insansız platformlarda da kullanılabilir olduğu değerlendirilen veri kümesi ise TUM mono veri kümesidir. Bu veri kümesi oluşturulurken kullanılan kamera bu tez çalışması için oluşturulan veri kümesinde de kullanılan monoküler RGB kameradır. Bununla birlikte, daha ağır, güç tüketimi daha yüksek ve daha pahalı bir kamera kullanılmıştır. TUM mono veri kümesindeki en hızlı hareketlerin gerçekleştirildiği sekansların sekans 14 ve sekans 35 oldukları gözlemlenmiştir. TUM Mono veri kümesinde çok fazla karanlık- aydınlık geçişi yapıldığı dolayısıyla özellikle karanlık ortamdan aydınlık ortamlara geçilirken, ani parlamaların olduğu gözlemlenmiştir. Bu veri kümesinde özellikle ORB-SLAM2 (Mur-Artal ve Tardós, 2017) algoritmasında parlama kaynaklı takip kaybı gözlemlenmiştir. Bununla birlikte, görüntülerdeki bulanıklığın yine anlık olduğu ve devamlı bir hareketten (ileri yönde hızlı hareket ya da dönüş hareketi gibi) kaynaklanmadığı görülmüştür. TUM Mono veri kümesinden 35.sekansa ait bir sahnedeki ekran görüntüsü aşağıda yer alan Resim 6.8’de gösterilmiştir. Bu sekans incelendiğinde gerçekleştirilen hareketlerin gerçek saha koşullarından farklı olduğu, duvarlara yerleştirilen resimler ve panoların elde taşınan platform ile takip edildiği ve bu yolla takip kaybının önlenildiği gözlemlenmiştir.

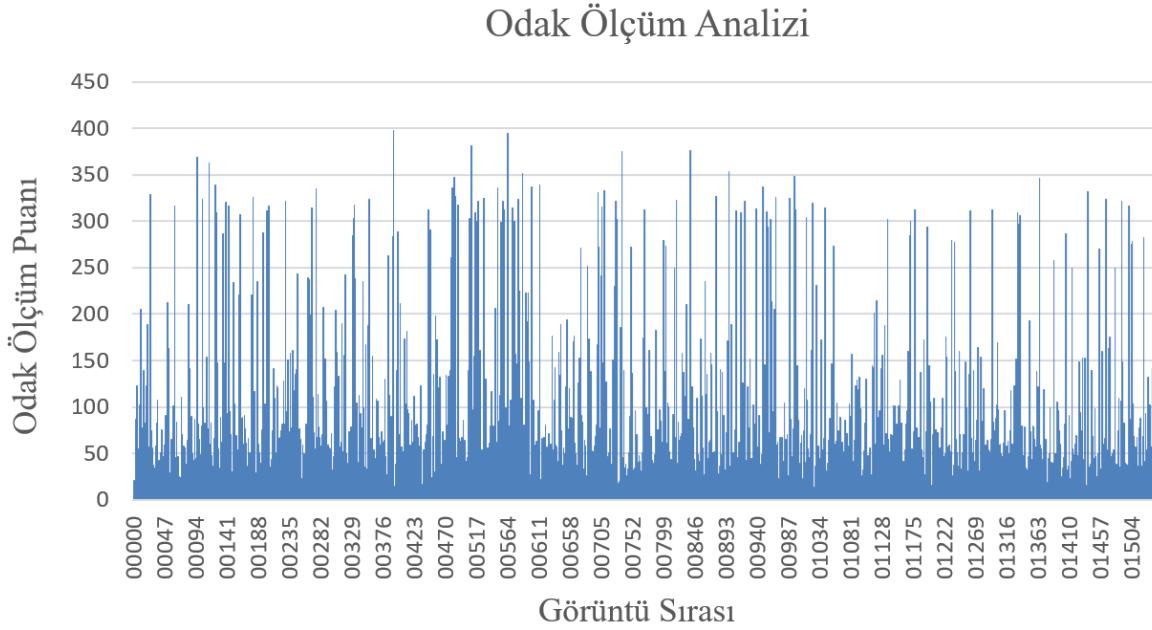


Resim 6.8. TUM Mono veri kümesinde yer alan 35.veri sekansından görüntü

TUM Mono veri kümesinde, bir diğer hızlı hareket içeren sekans ise 14.sekansdır. 35. ve 14. sekanslara ait odak ölçüm analizi sonuçları Şekil 6.12 ve Şekil 6.13'te gösterilmiştir.



Şekil 6.12. TUM Mono veri kümesi 35.sekansa ait odak ölçüm analizi



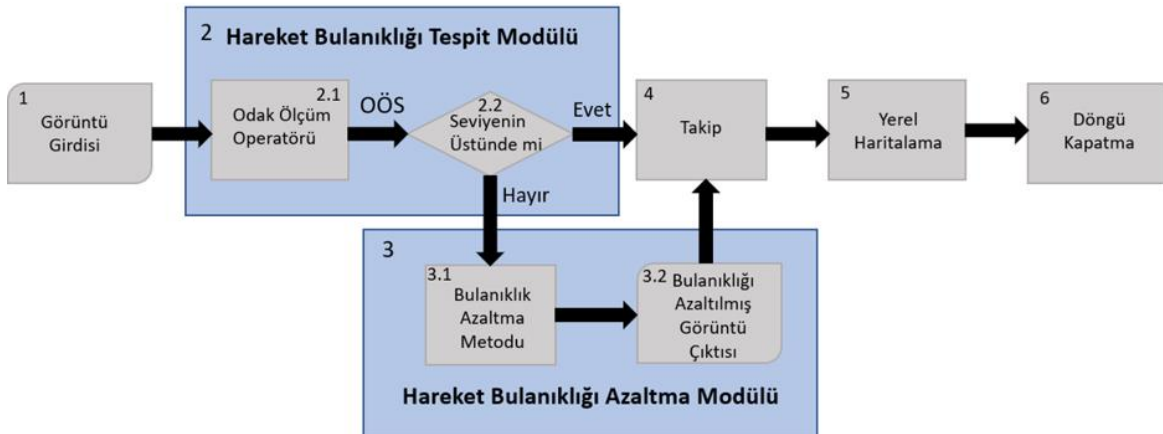
Şekil 6.13. TUM Mono veri kümesi 14.sekansa ait odak ölçüm analizi

Veriler incelendiğinde 35.sekansa ait görüntülerdeki bulanıklık seviyesinin ortalama olarak yüksek olduğu görülebilir. Fakat odak ölçüm puanı sonuçlarında bu çalışma kapsamında

hazırlanan veri kümesi daha düşük seviyelere inen, daha bulanık görüntüler içermektedir. Ayrıca gerçek saha uygulamalarına daha uygun senaryolar üretilmiştir.

6.6. Önerilen Yeni vSLAM Çerçevesi

Elde edilen sonuçlar, mikro / nano İHA’larda kullanılabilecek düşük maliyetli, hafif, kompakt kameraların gerçek dünya uygulamalarında hareket bulanıklığı nedeniyle çalışmalarını uygun şekilde gerçekleştiremediğini göstermektedir. Ayrıca bahsi geçen hızlı dönüş hareketlerinin olduğu görüntü karelerine ait odak ölçüm puanının diğer görüntü karelerinden daha düşük olduğu LAP4 operatörü kullanılarak gözlemlenmiştir. Hızlı hareketlerde meydana gelen hareket bulanıklığına dayanıklı olan daha önce vSLAM metotlarında kullanılmamış yeni vSLAM çerçevesi aşağıda yer alan Şekil 6.14’te yer almaktadır. Bu çerçeve hem doğrudan hem de öznitelik temelli vSLAM metotlarına uygun şekilde tasarlanmıştır. Önerilen yöntem ile hareket bulanıklıklarına daha dayanıklı, takip kaybı oranı daha düşük bir vSLAM çerçevesi oluşturmak amaçlanmıştır.



Şekil 6.14. Önerilen vSLAM çerçevesi

Önerilen vSLAM çerçevesinde, Şekil 6.14’te numaralandırılmış şekilde yer alan modüller numaralarına göre aşağıdaki gibi detaylandırılabilir:

1. Görüntü Girdisi
2. Hareket Bulanıklığı Tespit Modülü
 - 2.1. Odak Ölçüm Operatörü
 - 2.2. Karşılaştırma Operatörü

3. Hareket Bulanıklığı Azaltma Modülü
 - 3.1. Bulanıklık Azaltma Metodu
 - 3.2. Bulanıklığı Azaltılmış Görüntü Çıktısı
4. Takip Modülü
5. Yerel Haritalama Modülü
6. Döngü Kapatma Modülü

Önerilen hareket bulanıklığına dayanıklı vSLAM çerçevesinde, kameradan alınan görüntü girdisi (1) daha önceki vSLAM çerçevelerinde yer almayan, henüz takip kaybı oluşmadan sistemin hareket bulanıklığının seviyesinin tespit edilmesini sağlayan “hareket bulanıklığı tespit” modülüne (2) iletilmektedir. Bu modülde odak ölçüm operatörü (2.1) (Focus Measure Operatör- FMO) kullanılarak resmin bulanıklık seviyesi tespit edilir. Bu işlem sonucunda odak ölçüm puanı (OÖP) elde edilmektedir. Elde edilen bu değer karşılaştırma operatörüne (2.2) iletilmektedir. Eğer bulanıklık seviyesi daha önceden belirlenen eşik değeri seviyesinden yüksekse, yani OÖP belirlenen seviyenin altında kalıyorsa ilgili resim karesi takip işlemi için kullanılmadan önce “hareket bulanıklığı azaltma” (3) modülüne iletilmektedir. Bu modülde seçilen bulanıklık azaltma metodu (3.1) ile hareket bulanıklığı azaltılarak, bulanıklığı azaltılmış görüntü çıktısı (3.2) elde edilmektedir. Elde edilen bulanıklığı azaltılmış görüntü çıktısı (3.2) takip modülüne (4) iletilmektedir.

Eğer kameradan elde edilen görüntülerdeki bulanıklık seviyesi belirlenen seviyeden yüksek değilse, doğrudan takip modülüne (4) iletilerek işleme devam edilmektedir. Bu şekilde takip kaybı meydana gelmeden, vSLAM algoritmalarının çalışması sağlanabilmektedir.

“Hareket bulanıklığı tespit” (2) ve “hareket bulanıklığı azaltma” (3) modüllerinde çok farklı algoritmalar seçilerek kullanılabilir. Çerçeve seçilecek algoritmaların kullanımına ve doğrudan ve öznitelik temelli vSLAM algoritmaları ile entegrasyonuna imkân sağlamaktadır.

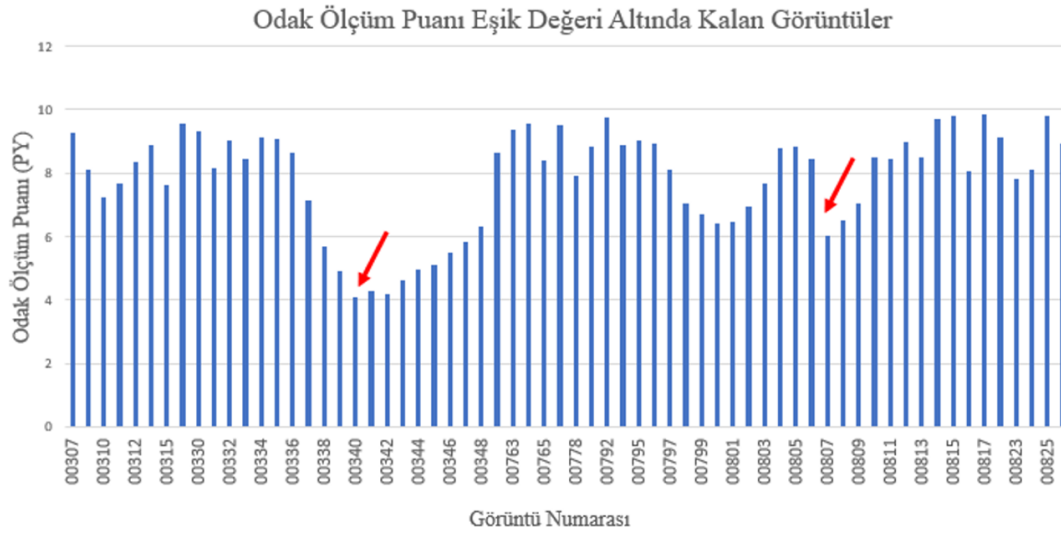
Bu çerçeve sayesinde vSLAM algoritmalarında, herhangi bir ek sensör ihtiyacı olmadan, özellikle agresif hareket eden ve ağırlık / batarya kullanım kısıtları olan insansız sistemlerde ve robotlarda yazılımsal olarak takip kaybının önüne geçilebilmektedir.

Daha detaylı bir anlatım ile, görüntü girdisi (1) alınarak önce “hareket bulanıklığı tespit” (2) modülüne iletilmektedir. Hareket bulanıklığı tespit modülü “odak ölçüm” (2.1) ve bir “karşılaştırma” (2.2) operatöründen oluşmaktadır. Görüntü girdisi, öncelikle odak ölçüm operatörüne iletilerek, resmin odak ölçümü (bu teze konu olan çalışmada seçilen LAP4 gibi) gerçekleştirilmektedir. Burada elde edilen sonuç resmin ne kadar odaklı olduğunu, yani keskinlik seviyesini ifade etmektedir. Bu operatörün çıktısı olarak elde edilen odak ölçüm puanı (OÖP), “karşılaştırma” (2.2) operatörüne gönderilmektedir. OÖP belirlenmiş (kameraya göre değişmekle birlikte vSLAM algoritmalarında piksel yoğunluğu cinsinden 10 birim yeterli olmaktadır) bir eşik değeri seviyesi ile karşılaştırılmaktadır. Karşılaştırma sonucunda eğer OÖP belirlenen seviyenin üstündeyse, resim karesinin bulanıklık seviyesi düşük olarak kabul edilir. Eğer OÖP belirlenmiş olan seviyeden düşük ise, resim karesindeki bulanıklık miktarı yüksek olarak kabul edilmektedir. Bulanıklık seviyesi yüksek olan resim karesi “hareket bulanıklığı azaltma” modülüne (3) iletilmektedir.

“Hareket bulanıklığı azaltma” modülünde (3) istenilen hareket bulanıklığı azaltma metodu (3.1) (örn: PMP, ECP gibi) seçilerek kullanılabilir. Çerçevenin yapısı kullanıcının kullandığı kameraya en uygun yöntemi seçmesine imkân tanımaktadır. Burada kullanılan metot ile resim karesinin bulanıklığı azaltılarak, restore edilmiş, netlik seviyesi yüksek resim elde edilmektedir. İlgili resim takip modülüne (4) iletilerek özniteliklerin / piksellerin takibi işlemine devam edilmektedir. Devam eden aşamalar daha önce kullanılmakta olan vSLAM çerçeveleriyle aynıdır. Takip aşamasında, sahnedeki özniteliklerin / piksellerin değişim oranı elde edilerek anahtar resim kareleri (keyframes) ataması gerçekleştirilmektedir. Anahtar resim kareleri ve öznitelikler/pikseller “yerel haritalama” modülüne (5) iletilmektedir. Bu modülde 2 boyutlu resim evreninden, 3 boyutlu gerçek dünya projeksiyonu gerçekleştirilerek 3 boyutlu nokta bulutu haritası oluşturulmaktadır. Burada elde edilen anahtar resim kareleri, “döngü kapatma” modülüne (6) iletilerek, daha önce benzer anahtar resim karelerinin sistemde yer alıp almadığı kontrol edilmektedir. Eğer daha önce benzer anahtar resim kareleri sistemde yer alıyorsa, aynı bölgeden daha önce geçilmiş demektir. “Döngü kapatma” modülünde (6) aynı yerden daha önce geçilmiş ise bu bölgeler arası bağlantı kurularak harita yeniden oluşturulmaktadır. Daha önceden geçilmiş bölgenin, yeni bir yer gibi haritada gösterilmemesi için anahtar resim karesi eşleştirmesinin gerçekleştirilmesi önemlidir.

6.6.1. Hızlı dönüş hareketi içeren veri kümesinde önerilen çerçevenin testleri

Veri kümesi oluşturulurken kullanılan Raspberry Pi v2.1 kamerası için odak ölçüm puanı değerlendirildiğinde, 10 puan (piksel yoğunluğu cinsinden) üzerindeki bölümde herhangi bir takip kaybının olmadığı görülmüştür. Takip kaybının yaşandığı durumlarda odak ölçüm puanı değerinin 4-6 seviyesine kadar düştüğü gözlemlenmiştir. Önerilen çerçevede OÖP = 10 olacak şekilde seviyesi eşik değeri olarak seçilmiştir. Odak ölçüm puanı 10 değerinin altında kalan görüntü kareleri, tüm veri kümesinden ayırt edilerek bu görüntüler ile ayrı bir alt veri kümesi oluşturulmuştur. Şekil 6.15'te OÖP değeri eşik değerinin altında kalan, hareket bulanıklığı yüksek görüntüler ve bu görüntülere ait OÖP değerleri gösterilmiştir.

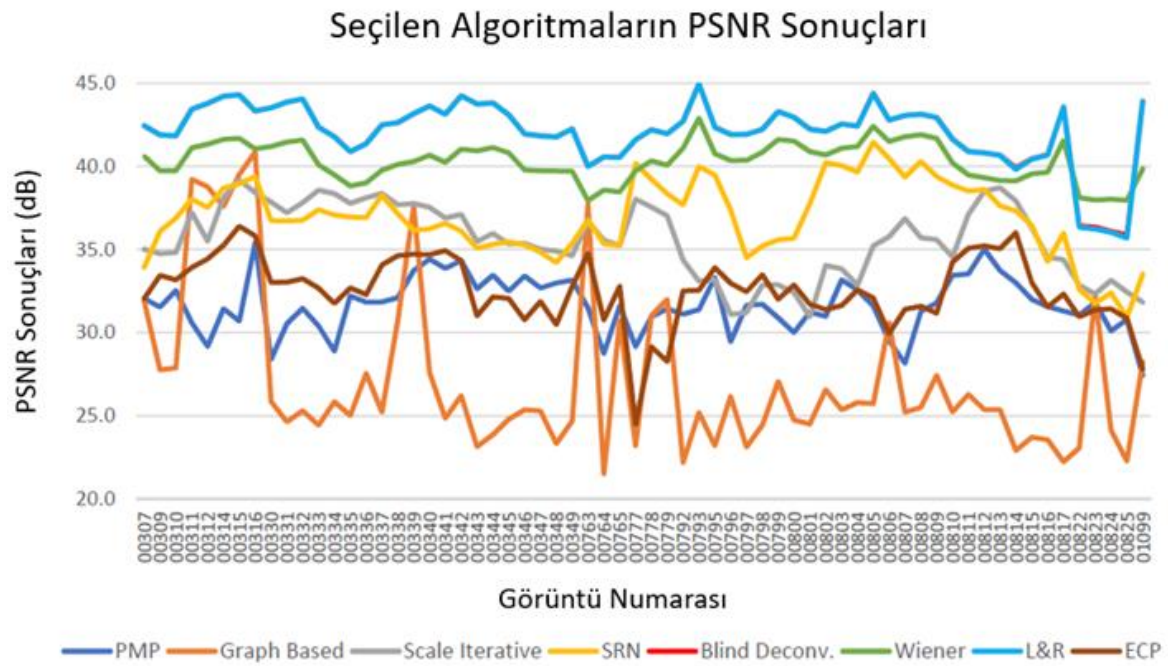


Şekil 6.15. Odak ölçüm puanı eşik değerinin altında kalan görüntüler

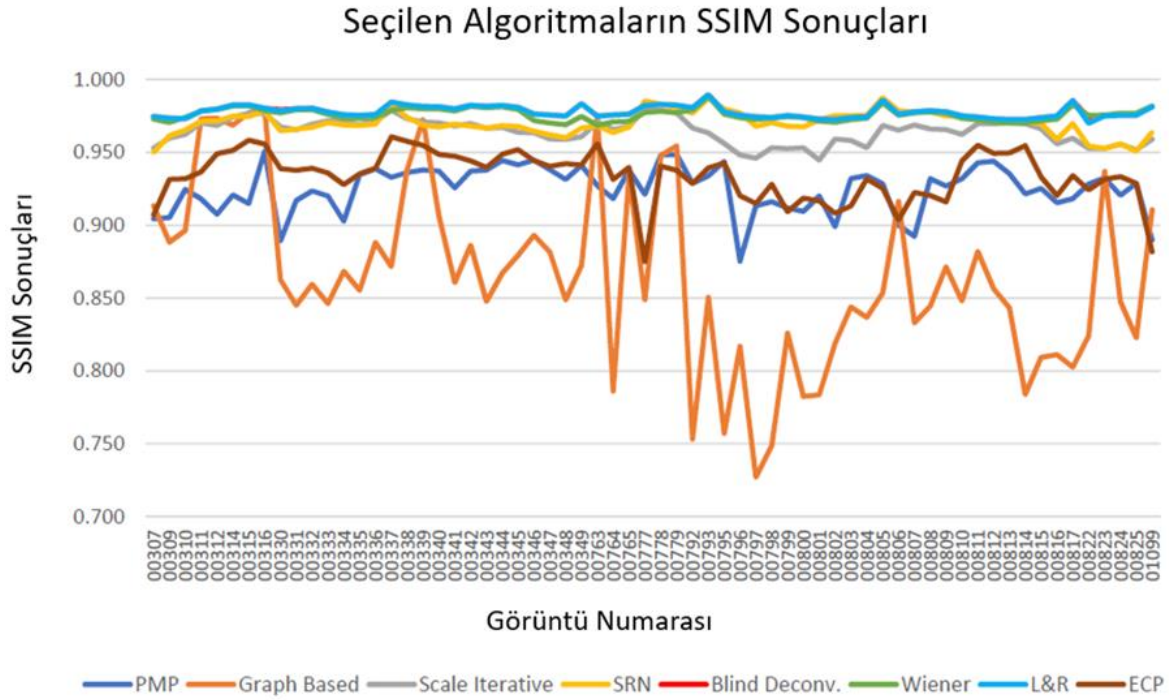
İlgili şekilde OÖP değerinin en düşük olduğu bölgeler kırmızı ok ile işaretlenmiştir. İlgili eşik değerinin altında 64 tane görüntü karesinin kaldığı gözlemlenmiştir. Önerilen vSLAM çerçevesine göre eşik değeri altında kalan resimler, hareket bulanıklığı azaltma modülüne iletilerek, görüntülerde yer alan bulanıkların azaltılması/ kaldırılması hedeflenmektedir. Bulanıklık azaltma metodu seçiminde herhangi bir kriter mevcut değildir. Bu kısım daha önce de belirtildiği üzere istenilen yöntemin kullanılabilmesi için esnek olarak tasarlanmıştır.

Daha önceki bölümlerde de değinildiği üzere hareket bulanıklığını azaltmak için pek çok yöntem mevcuttur. Bununla birlikte önerilen vSLAM çerçevesinin testlerinde kullanmak

için, klasik (Lucy ve Richardson, Wiener, BD), öncel tabanlı (ECP, PMP, GB) ve yapay zekâ temelli (SRN, SIUN) yöntemler ayrı ayrı test edilmiştir. Bu yöntemler sayesinde eşik değerinin altında kalan görüntü karelerini restore edilerek önerilen çerçeve kapsamında test edilmiştir. Elde edilen restore edilmiş görüntülere Laplacian Varyansı (LAP4) odak ölçüm operatörü uygulanarak bu yöntemlerin gerçekleştirdiği değişimler analiz edilmiştir. Ayrıca restore edilen görüntülerdeki değişiklikler farklı yaygın kullanılan metriklerle ölçülmüştür. Hareket bulanıklığına maruz kalan veri seti görüntüleri arasındaki korelasyonu ölçmek için kullanılan metrikler, Pik Sinyal Gürültü Oranı (PSNR) ve Yapısal Benzerlik (SSIM) metrikleridir. Klasik yöntemlerden Wiener filtresi, Lucy ve Richardson ve Kör ters konvolüsyon (BD) yöntemlerinde en uygun nokta yayılım fonksiyonunun bulunması için 1 pikselden 10 piksele kadar 360 derece tam daire taraması yapılmış ve en uygun nokta yayılım fonksiyonları kullanılmıştır. Her bir yöntemin PSNR ve SSIM değerleri Şekil 6.16 ve Şekil 6.17’de yer alan grafiklerde verilmiştir.



Şekil 6.16. Seçilen algoritmaların PSNR sonuçları



Şekil 6.17. Seçilen algoritmaların SSIM sonuçları

En çok tercih edilen görüntü kalitesi metrikleri olan PSNR ve SSIM metrikleri, seçilen yöntemlerin karşılaştırılmasında kullanılmıştır. Terminolojik olarak, PSNR, restore edilen görüntünün maksimum gücü ile bulanık görüntü girişinin maksimum gücü arasındaki orandır. SSIM ise, bulanık giriş görüntüsü ile restore edilen versiyon arasındaki benzerliği ölçmek için kullanılan başka bir göstergedir. Hesaplanan SSIM değerinin 1 değerine yakın olması, restore edilen görüntünün yapısal bilgilerinin orijinal bulanık olana çok benzediğini gösterir. SSIM değerinin her zaman 1'e yakın olması istenir. Bu değer 1'e yakın olmadığı durumlarda restore edilen resmin, orijinal resimden yapısal olarak farklılaştığı görülmektedir. PSNR ve SSIM analizlerinde, kör ters konvolüsyon (BD) ve Lucy ve Richardson (L&R) algoritmalarının benzer performans gösterdikleri gözlemlenmiştir. Deneysel ölçümlere göre, restore edilen görüntülerdeki iyileştirme bu iki algortmada en iyi performansa sahiptir. Bununla birlikte SSIM sonuçlarında GB yönteminin çok fazla salınım yaptığı ve restore edilen görüntülerin yapısal olarak bozukluğa uğradığı gözlemlenmiştir.

Bu metriklere ek olarak, görüntü bulanıklığı tespitinde kullanılan odak ölçüm puanının restore edilmiş görüntülerdeki değişimi gözlemlenmiştir. Restore edilen görüntülerin odak ölçüm puanlarının orijinal değeri ile karşılaştırılması Şekil 6.18'de ve önceki odak ölçüm

puanına göre elde edilen değişim Şekil 6.19'da gösterilmiştir. İlgili yöntemler kullanılarak odak ölçüm puanında artış sağlandığı tespit edilmiştir.

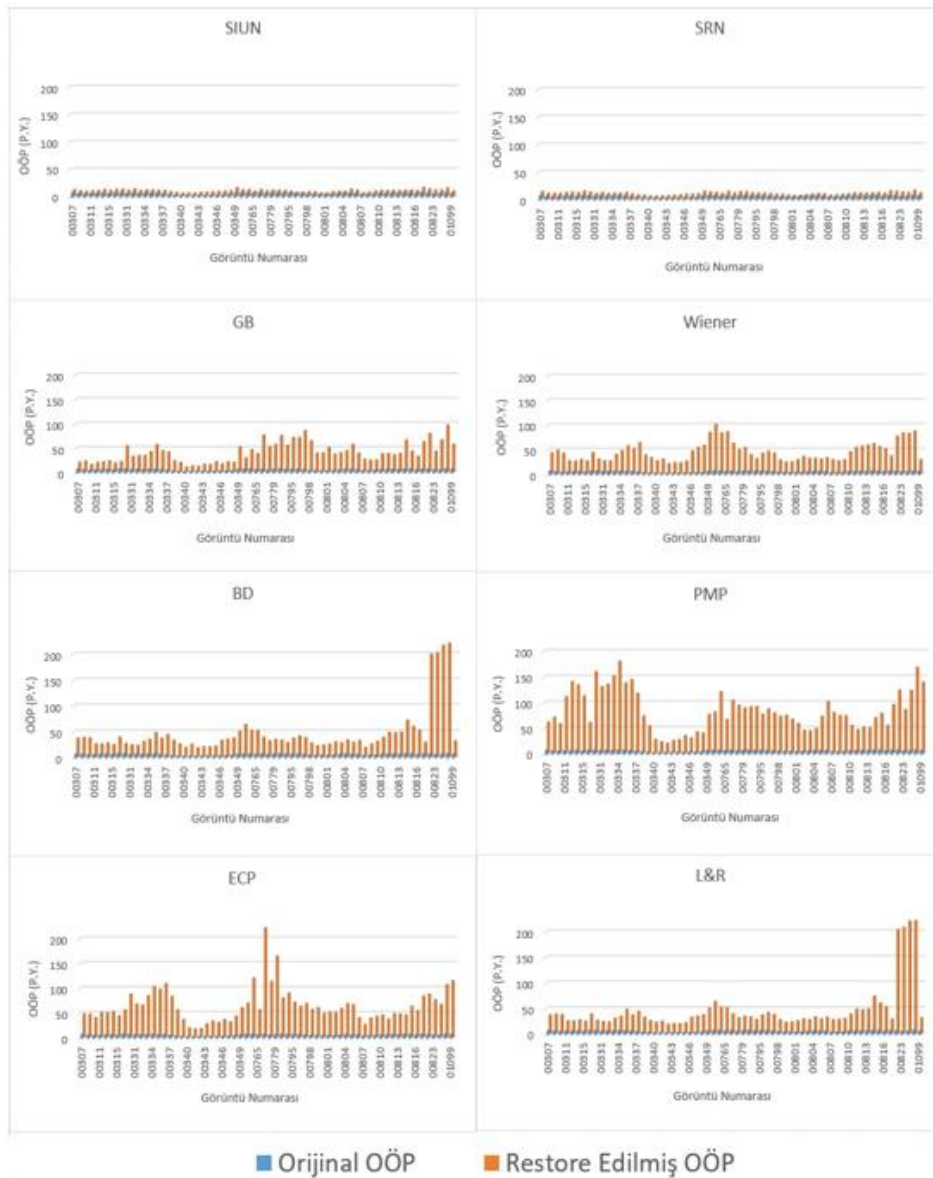
Odak Ölçüm Puanı Analizi

Her bir bulanıklık azaltma / ortadan kaldırma yöntemine ait odak ölçüm puanları Şekil 6.19'da verilmiştir. Odak ölçüm puanında meydana gelen ortalama değişimler açısından en başarılı algoritmalar sırasıyla PMP (Wen, 2020), ECP (Yan ve diğerleri, 2017), Wiener (Tukey, 1952), L&R edilmektedir (Lucy, 1974, Richardson, 1972), BD, Graph-Based (Bai, 2018), SRN (Tao ve diğerleri, 2018) ve SIUN (Ye, 2020) yöntemleridir. Odak ölçüm puanındaki kazanç, yapay zekâ temelli algoritmalar olan SIUN (Ye, 2020) ve SRN (Tao ve diğerleri, 2018) algoritmaları için en düşük seviyededir. Diğer algoritmalar, SIUN ve SRN tekniklerine kıyasla OÖP değerinde üç kat daha fazla artış sağlamışlardır. Çizelge 6.4'te bulanıklık azaltma yöntemlerinin ortalama OÖP değerleri verilmiştir. Bu değerler elde edilirken tüm görüntülere ait odak ölçüm puanlarındaki değişim toplanarak görüntü sayısı olan 64'e bölünmüştür. Dolayısıyla hem ileri yöndeki harekete hem de dönüş hareketlerine ait görüntülerdeki odak ölçüm puanlarındaki değişim hesaplamaya dâhil edilmiştir. Bu hesaplama ile OÖP eşik değeri olan 10'un altında yer alan bulanıklık seviyesi yüksek tüm resimlerdeki ortalama değişim görülmek istenmiştir. Doğrudan takip kaybının olduğu görüntülere ait analizlere ilerleyen bölümlerde ayrıca yer verilmiştir.

Çizelge 6.5. Bulanıklık azaltma metotlarına ait ortalama odak ölçüm puanları

Bulanıklık Azaltma/ Yok Etme Metodu	Ortalama Δ OÖP (Piksel Yoğunluğu)
PMP	74.53
ECP	51.23
Wiener	39.96
L&R	37.83
BD	33.82
GB	32.18
SRN	2.49
SIUN	2.18

Eşik değeri 10 puanın altında kalan tüm görüntüler ele alındığında sıralama Çizelge 6.5’de gösterildiği gibi olmaktadır. Bu noktada ECP algoritmasının ortama OÖP değişimine göre sıralamada 2.sırada yer almaktadır. Fakat ECP algoritması, ileri harekette meydana gelen bulanıkların azaltılmasında / yok edilmesinde yüksek başarı sağlarken, bulanıklığı daha fazla olan ani dönüş hareketlerinin yer aldığı kısımlarda daha az başarı sağlamıştır. Dolayısıyla bu metodun takip kaybına neden olmayan ileri hareketteki başarısı oldukça fazlayken, ani dönüşlerde aynı performans sağlanamamıştır. Dolayısıyla ortalama değişimdeki bu performansın takip kaybının azaltılmasındaki etkisi beklenen şekilde gerçekleşmemiştir. Bu sonuçlar vSLAM algoritmalarında gerçekleştirilen denemelere ait olup, ilerleyen başlıklarda detaylandırılmaktadır.



Şekil 6.18. Orijinal ve restore edilmiş görüntülerin Odak Ölçüm Puanları

Şekil 6.18’de veri kümesindeki bulanıklık içeren orijinal ve restore edilmiş resimlere ait odak ölçüm puanları gösterilmektedir. Orijinal görüntülere ait OÖP değeri mavi renkle, restore edilmiş resimlere ait OÖP ise turuncu renklerde verilmiştir. Görsellerde turuncu renklerin baskın olduğu görülmektedir. Bu durum, orijinal resimlerdeki OÖP’nin restore edilmiş resimlere oranla çok düşük kaldığını göstermektedir. Aşağıda yer alan Şekil 6.19’da ise, her yöntemin odak ölçüm puanında meydana getirdiği değişiklik ($\Delta\text{OÖP}$) gösterilmiştir.

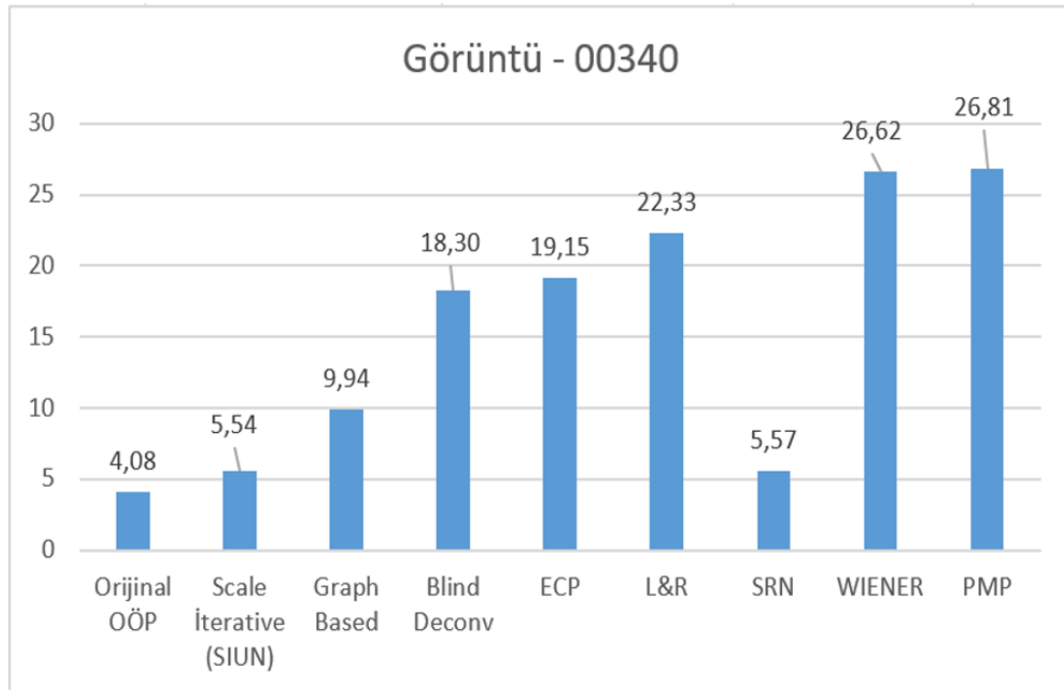


Şekil 6.19. Restore edilmiş görüntülerde OÖP değişimi

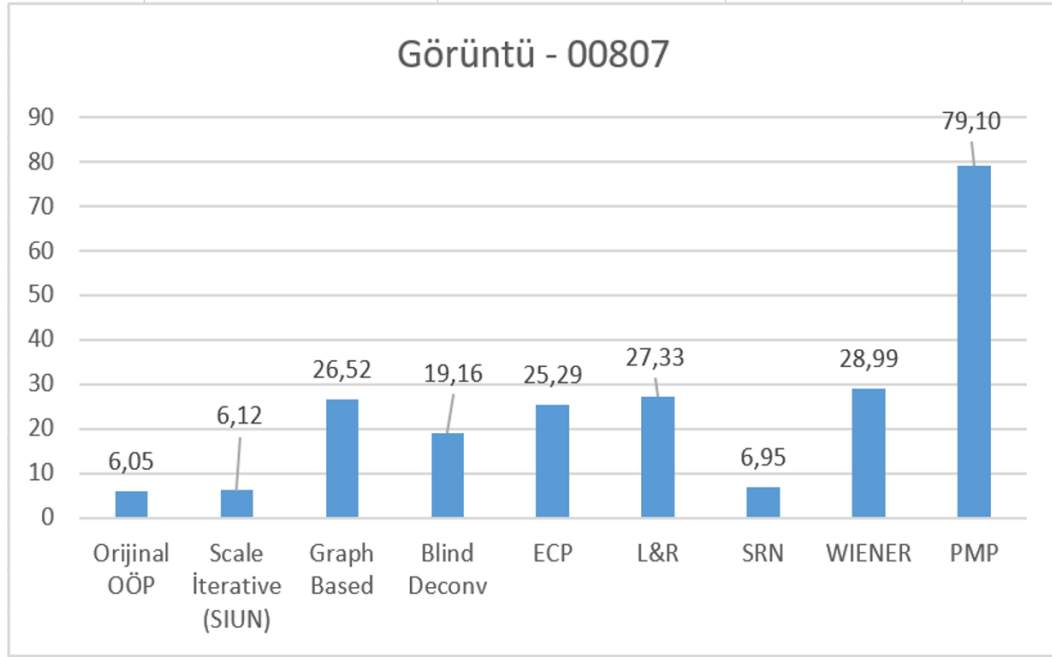
Odak ölçüm puanındaki değişim incelendiğinde, yapay zekâ temelli SIUN ve SRN algoritmalarının performansı en düşük seviyededir. Diğer algoritmalar, SIUN ve SRN tekniklerine kıyasla OÖP değerinde üç kat daha fazla artışa sağlamışlardır. Wiener ve L&R yöntemleri, diğer bulanıklaştırma yöntemleriyle karşılaştırıldığında (BD, GB, SIUN, SRN) nispeten daha yüksek ortalama $\Delta OÖP$ değerine sahiptir. Tablodan gerçekleştirilebilecek bir diğer çıkarım ise, Wiener ve L&R yöntemlerinin ECP yönteminden daha düşük $\Delta OÖP$ 'ye sahip olmasına rağmen, yukarıda gösterilen karşılaştırmadan çıkarılabilecek şekilde hızlı dönme hareketinin yer aldığı görüntülerde her iki yöntemin daha iyi bulanıklık azaltma performansına sahip olmasıdır.

6.6.2. Seçilen algoritmaların ani dönüş hareketindeki performansları

Tüm alt veri kümesindeki odak ölçüm puanında meydana gelen olumlu değişimler Şekil 6.18 ve Şekil 6.19'da incelenmiştir. Sonuçlar ECP gibi yöntemlerin ileri hareketlerdeki performansının, dönüş hareketlerinden daha fazla olduğu göstermektedir. Bu nedenle, seçilen algoritmaların ani dönüşlerin meydana geldiği köşe noktalarındaki performansları ayrıca incelenmiştir. Dönüşlerde bulanıklığın en fazla olduğu 00340 ve 00807 numaralı görüntüler için seçilen algoritmalarla göre odak ölçüm puanları piksel yoğunluğu cinsinden sırasıyla Şekil 6.20 ve Şekil 6.21'de verilmiştir.



Şekil 6.20. 00340 numaralı görüntüye ait odak ölçüm puanları



Şekil 6.21. 00807 numaralı görüntüye ait odak ölçüm puanları

Şekillerde yer alan grafikler incelendiğinde, sıralamada GB (Bai, 2018) metodunun iki görüntü arasında farklı sıralamaya sahip olduğu görülmektedir. GB (Bai, 2018) metodunun SSIM sonuçlarında, çok fazla salınım gösterdiğine daha önceki başlıklarda (Şekil 6.17) değinilmişti. Bu nedenle sıralamanın tam olarak belirlenmesi için buradaki elde edilen sonucun önerilen vSLAM çerçevesi altında vSLAM algoritmalarında denenerek teyit edilmesi gerekmiştir.

6.7. Hareket Bulanıklığına Dayanıklı vSLAM Çerçevesindeki Performanslar

Önerilen Hareket Bulanıklığına Dayanıklı vSLAM Çerçevesinin hareket bulanıklığı içeren görüntülerin takip kayıplarını azaltması / önlemesi hedeflenmiştir. Ayrıca ortaya konulan çerçeve tasarımının hem doğrudan hem de öznitelik temelli vSLAM algoritmalarına taban teşkil etmesi planlanmıştır. Hem takip kayıplarının ne ölçüde önlendiği hem de doğrudan ve öznitelik temelli algoritmalar ile uyumlulukları restore edilmiş görüntüler üzerinden sınanmıştır. Gerçekleştirilen çalışmada odak ölçüm puanı için belirlenen eşik değerin altında kalan 64 görüntü, seçilen bulanıklık azaltma / önleme algoritmaları kullanılarak restore edilip aşağıda yer alan Şekil 6.22 ve Şekil 6.23’de de gösterildiği üzere ayrı alt veri kümeleri oluşturulmuştur. Bu alt veri kümeleri, orijinal veri kümesindeki görüntüler ile her bir yöntem için değiştirilerek, her bir yöntem için tam veri kümesi oluşturulmuştur.

Lucy & Richardson**PMP****SIUN****SRN**

Şekil 6.22. Seçilen yöntemlerle oluşturulan alt kümeler

Wiener**BD****ECP****GB**

Şekil 6.23. Seçilen yöntemlerle oluşturulan alt kümeler-2

Doğrudan ve öznitelik temelli vSLAM algoritmalarının takip performansı, her hareket bulanıklığı azaltma / önleme algoritmasının çıktısı olan restore edilmiş görüntülere sahip veri kümelerinde gözlemlenmiştir. Restore edilmiş 8 farklı 1121 adet görüntü içeren tüm veri kümesi 5, 10 ve 20 fps işlem hızlarında işlenmiştir. Son teknoloji yöntemler olan öznitelik temelli ORB-SLAM2 (Mur-Artal ve Tardós, 2017) ve doğrudan DSO (Engel ve diğerleri, 2017) algoritmalarının performansları sırasıyla aşağıda yer alan Çizelge 6.6 ve Çizelge 6.7’de verilmiştir. vSLAM algoritmalarının toplamda kaç farklı işlem hızında başarılı olduğunu ortaya çıkarmak için takip puanı parametresi oluşturulmuştur. İşlem hızı sadece vSLAM algoritmalarının işleme hızı olup, saniyede kaç adet görüntünün

işleneceğini ifade etmektedir. Görüntülerdeki pozlandırma zamanı değişmediği için, görüntünün bulanıklık seviyesi değişmemektedir. Tablolarda yer alan onay sembolü “başarılı” sonuçlar için kullanılırken, çarpı sembolü sonucun “başarısız” olduğunu göstermektedir.

Çizelge 6.6. Önerilen vSLAM çerçevesinin ORB-SLAM2 algoritmasındaki başarısı

ORB_SLAM2	5 Fps	10 Fps	20 Fps	Takip Puanı
PMP	✓	✓	✓	3
Wiener	✓	✓	✓	3
L&R	✓	✓	X	2
ECP	✓	✓	X	2
BD	✓	X	X	1
SIUN	✓	X	X	1
SRN	✓	X	X	1
GB	X	X	X	0

Çizelge 6.7. Önerilen vSLAM çerçevesinin DSO algoritmasındaki başarısı

DSO	5 Fps	10 Fps	20 Fps	Takip Puanı
PMP	✓	✓	✓	3
Wiener	✓	✓	X	2
L&R	✓	✓	X	2
ECP	✓	X	X	1
BD	X	X	X	0
SIUN	X	X	X	0
SRN	X	X	X	0
GB	X	X	X	0

Her iki vSLAM algoritması ile elde edilen sonuçlar incelendiğinde, daha önce takip kaybına uğranılan durumların, önerilen yeni çerçeve ile aşılabildiği gözlemlenmiştir. Daha önce takip kaybına uğrayan algoritmalar, yeni çerçeve ile uygun durumlarda takip kaybına uğramadan işlemlerine devam etmiştir. Tablolardan çıkarılan bir diğer sonuç ise, GB algoritması ile restore edilen görüntülerin iki yöntem ile de başarısız olduğudur. Daha önce

bahsedildiği üzere Şekil 6.17’de GB yöntemi için hesaplanan SSIM verisi, en düşük değere sahip olup beklenmeyen ölçüde dalgalı davranış göstermektedir. Bu durum, GB algoritması kullanılarak restore edilen görüntülerin orijinal bulanık görüntülerle yapısal olarak eşleşmediğini göstermektedir. İlgili vSLAM metotlarının, bu yapısal değişiklikleri art arda gelen görüntüler gibi değerlendirmeyerek takip kaybına uğraması kuvvetle muhtemeldir. Bu nedenle gerçekleştirilen testler GB yönteminin, vSLAM algoritmalarında bulanıklaştırma işlemi için başarılı bir aday olmadığını göstermiştir.

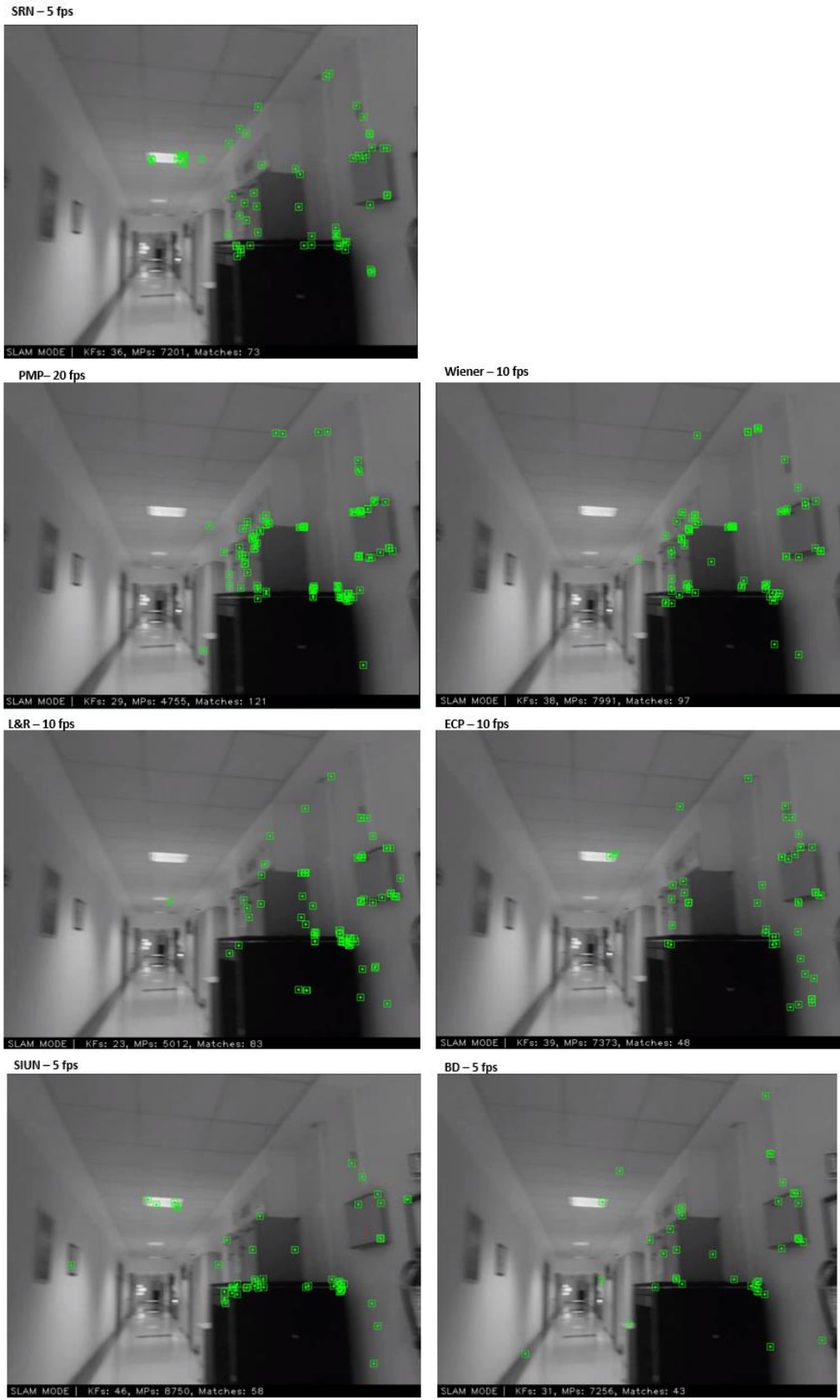
PMP metodu odak ölçüm puanında en yüksek değişimi sağladığı gibi, vSLAM algoritmalarındaki performansta da en fazla takip puanına sahip algoritma olmuştur. Önerilen hareket bulanıklığına dayanıklı vSLAM çerçevesinde kullanılmak üzere, PMP ve benzeri bir algoritmanın kullanımı daha verimli sonuçlar sağlamaktadır.

6.7.1. Restore edilen resimlerdeki öznitelik sayılarının karşılaştırılması

Önerilen çerçevede, daha önceden takip kaybına uğrayan öznitelik temelli ORB-SLAM2 algoritmasının ilgili hızlarda GB metodu dışındaki metotlarda başarılı şekilde çalıştığı gözlemlenmiştir. Her bir yöntem için takip kaybının önlendiği görüntülerdeki öznitelik sayıları tespit edilmiştir. Takip kaybının olduğu resim karesi Şekil 6.24’te gösterilmektedir. Bu şekilde hiçbir özniteliğin tespit edilemediği ve keskin dönüşün yer aldığı ve sonraki görüntüler arasında hareket bulanıklığı kaynaklı olarak eşlemenin sağlanamadığı görülmektedir. Görüntü daha önce de anlatıldığı üzere düşük çözünürlükte olup, koridor ortamı bulanık şekilde görülmektedir.



Şekil 6.24. Takip kaybının yaşandığı orijinal görüntü



Şekil 6.25. Önerilen yöntem ile ORB-SLAM2 algoritmasının tespit ettiği öz nitelikler

Şekil 6.25’de önerilen vSLAM çerçevesi kullanıldığında, takip kaybına uğramadan yakalanan öznitelikler gösterilmiştir. Daha önce takip kaybının olduğu, önerilen yeni vSLAM çerçevesi ile kayıpların engellendiği görüntülerdeki öznitelik sayıları bulanıklık azaltma metotlarına göre aşağıda yer alan Çizelge 6.8’de verilmiştir.

Çizelge 6.8. ORB-SLAM2 algoritması ile tespit edilen öznitelik sayıları

Bulanıklık Azaltma Yöntemi	Öznitelik Sayısı
PMP	121
Wiener	97
L&R	83
SRN	73
SIUN	58
ECP	48
BD	43

GB metodu herhangi bir işlem hızında başarılı olamadığı için, bu yönteme ait öznitelik sayısı verilmemiştir.

6.7.2. Eşik değeri seçiminin vSLAM çerçevesinin performansına etkileri

Bu teze konu olan çalışmada Odak Ölçüm Puanının eşik değeri aşağıda açıklanan formüle göre seçilmiştir. Veri kümesinde yer alan görüntülere Odak Ölçüm Puanı analizi uygulanarak en düşük 3 lokal minimum değeri bulunmaktadır. Daha sonra elde edilen bu 3 lokal minimum değeri karşılık gelen ağırlık ve lokal minimum altında kalan görüntülerin bulanıklık azaltma işlemi tercih değeri ile çarpılarak, toplanmaktadır. Daha sonra bu toplam 3’e bölünür. Elde edilen sonuç en yakın tam sayıya yuvarlanmaktadır. Tercih değeri ifadesi bulunurken, restore edilecek görüntü sayısı ne kadar az ise, o kadar yüksek bir değer elde edilmektedir. İlgili formüller Eş. 6.2’de ve Eş. 6.3’te gösterilmiştir.

$$\text{Lokal minimum için tercih değeri (TD)} = \frac{\text{Toplam görüntü sayısı}-\text{restore edilecek görüntü sayısı}}{\text{Toplam görüntü sayısı}} \quad (6.2)$$

$$\text{Eşik Değeri} = [(a_1 \times (1.\text{lokal minimum}) \times TD_1 + a_2 \times (2.\text{lokal minimum}) \times TD_2 + a_3 \times (3.\text{lokal minimum}) \times TD_3)] / 3 \quad (6.3)$$

Hem işlem yükünün istenen seviyede tutulması, hem de başarımın sağlanması için eşik değerinin optimum şekilde belirlenmesi önemlidir. Eşik değeri arttıkça, işlem yapılması gereken görüntü sayısının üstel olarak arttığı gözlemlenmiştir. Gerçekleştirilen test sonuçlarına göre optimum ağırlık değerleri sırasıyla (a_1 , a_2 ve a_3 için) 0,7, 1,8 ve 1,5 olarak elde edilmiştir. Aşağıda yer alan Çizelge 6.9’da veri kümesi OÖP’ındaki en düşük 3 lokal minimum, ilgili lokal minimumlara ait ağırlıklar ve tercih değerleri gösterilmiştir. Bu değerler kullanıldığında eşik değeri 10 olarak belirlenmektedir.

Çizelge 6.9. Veri kümesinde OÖP lokal minimumları, ağırlıklar ve tercih değerleri

Ağırlık Değerleri	En Düşük 3 Lokal Minimum	Tercih Değeri	Elde Edilen Eşik Değeri
0,7	4,08	0,99	10
1,5	12,035	0,91	
1,8	6,04	0,96	

6.7.3. Hareket bulanıklığına dayanıklı vSLAM çerçevesinin özellikleri

Yapılan testler sonucunda, önerilen hareket bulanıklığına dayanıklı vSLAM çerçevesiyle ilgili aşağıdaki tespitler yapılabilir:

Veri kümesindeki orijinal görüntüler, restore edilen görüntüler ile değiştirildiğinde bazı işlem hızlarında izleme kaybının önlenemediği gözlemlenmiştir. Öznitelik tabanlı yöntemlerde hareket bulanıklığı, öznitelik tespitini ve eşleştirmesini olumsuz etkilemektedir. Bulanık görüntü, belirtilen bulanıklık giderme yöntemleriyle restore edildiğinde, geri yüklenen görüntülerdeki ilgili özniteliklerin saptanmasında ve eşleştirilmesinde iyileştirmeler meydana gelmiştir.

Doğrudan yöntemlerde, hareket bulanıklığının pikseller arası yoğunluk farkını azalttığı ve takip performansını olumsuz etkilediği ve bulanıklık azaltma yöntemlerinin bu durumu ortadan kaldırabileceği gözlemlenmiştir.

vSLAM algoritmalarında elde edilen takip performansının odak ölçüm puanı ile doğrudan ilişkili olmasına rağmen PSNR ve SSIM değerleri ile doğru orantılı olmadığı deneysel olarak da doğrulanmıştır. Bununla birlikte, SSIM değerinin ani değişimlere uğradığı yöntemlerin yapısal olarak bozulmaya uğradığı görülmüştür. Yapısal bozulmaya uğrayan restore edilmiş görüntülerin eşleştirme performanslarının düştüğü ve bu tipte yöntemlerin önerilen vSLAM çerçevesine uygun olmadığı değerlendirilmiştir.

Bu teze konu olan çalışmanın vSLAM literatürüne iki önemli katkısı bulunmaktadır: İlk defa odak ölçüm operatörü vSLAM algoritmalarına dahil edilerek bulanıklık seviyesi tespiti gerçekleştirilmiştir. Bu sayede yüksek hareket bulanıklığı seviyelerine sahip kareler, takip kaybına neden olmadan algılanabilir. İkinci katkı ise yalnızca bulanıklık düzeyi yüksek kareler düzeltilerek algoritmanın hem haritalama hem de izleme aşamalarında çalışmaya devam etmesi sağlanmıştır.

Başlatma, takip, haritalama, bulanıklık algılama ve hareket bulanıklığı azaltma/ortadan kaldırma gibi özelliklerin mevcut çerçeveler ile karşılaştırmalı bir incelemesi aşağıdaki Çizelge 6.10'da gösterilmektedir. Önerilen çerçevenin hareket bulanıklığı altında daha dirençli olduğu ve başlatma aşaması altında uygulanabilir olduğu kolaylıkla görülmektedir.

Çizelge 6.10. Önerilen vSLAM çerçevesi ile mevcut çerçevelerin karşılaştırılması

Modüller	Öznitelik Temelli Yöntemler	Doğrudan Yöntemler	Önerilen Çerçeve
Başlatma	Uygulanmaktadır.	Uygulanmaktadır.	Uygulanmaktadır.
Takip	Hareket bulanıklığı altında takip kaybı	Hareket bulanıklığı altında takip kaybı	Hareket bulanıklığı altında daha dayanıklı takip
Haritalama	Haritalama hareket bulanıklığı nedeniyle sonlanabilir.	Haritalama hareket bulanıklığı nedeniyle sonlanabilir.	Hareket bulanıklığı altında daha dayanıklı haritalama

Çizelge 6.10. (devam) Önerilen vSLAM çerçevesi ile mevcut çerçevelerin karşılaştırılması

Modüller	Öznitelik Temelli Yöntemler	Doğrudan Yöntemler	Önerilen Çerçeve
Bulanıklık Tespiti	Uygulanmamaktadır.	Uygulanmamaktadır.	Uygulanmaktadır.
Bulanıklık Azaltma / Önleme	Uygulanmamaktadır.	Uygulanmamaktadır.	Uygulanmaktadır.

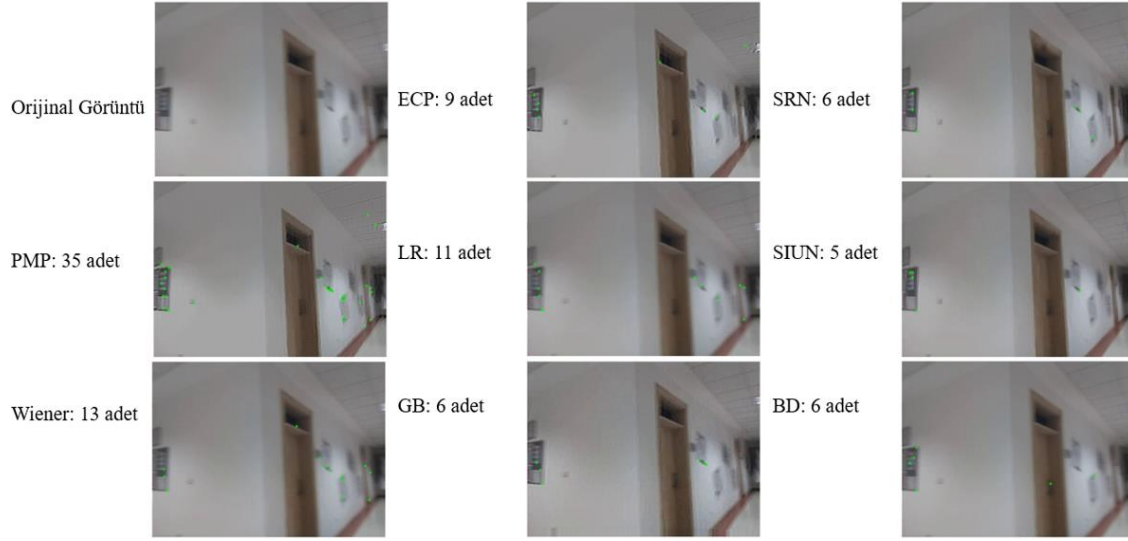
Çizelge 6.10’da ana modüller incelendiğinde, mevcut ve önerilen çerçevelerin farkları ortaya konulmuştur.

Önerilen çerçeve, mevcut çerçevelerde takip kaybına neden olan durumlarda vSLAM algoritmalarının normal operasyonuna devam etmesini sağlamaktadır. Bununla birlikte, eklenen modüllerin işlem gücü gereksinimlerini arttıracakları değerlendirilmekle birlikte, bulanıklık tespit modülü olarak LAP4 metodunun seçilmesi ve sadece belirlenen eşik değerinin altındaki görüntülere işlem uygulanması gerçek zamanlı çalışmaya daha yakın performans sağlamak amacıyla kurgulanmıştır.

6.8. Veri Kümesindeki Doğal Hareket Bulanıklığının Öznitelik Tespitine Etkisi

Daha önceki bölümlerde sentetik hareket bulanıklığının öznitelik çıkarımına ve eşleştirilmesine olan etkileri incelenmişti. Bu bölümde ise bu tez çalışması kapsamında oluşturulan ve doğal hareket bulanıklığı içeren veri kümesi üzerinde hem öznitelik çıkarımı hem de öznitelik eşleştirme performansları incelenecektir. Öznitelik çıkarımında 340 ve 825 numaralı görüntülerde yer alan ORB özellikleri (Şekil 6.26 ve Şekil 6.27) orijinal ve restore edilmiş görüntülerde gözlemlenmiştir. Ayrıca, ORB öznitelik eşleştirmesi 340-341, 807-808 görüntüleri (Şekil 6.28 ve Şekil 6.29) arasında gerçekleştirilmiştir. Ek olarak, 807 ve 808 numaralı görüntüler arasında Harris köşelerinin (Harris, 1988) ve SURF (Bay, 2006) özniteliklerinin eşleştirmesi her bulanıklık azaltma yöntemine göre yapılmıştır. vSLAM algoritmalarında daha önce de bahsi geçtiği üzere, öznitelikler bir veri tabanında toplanmaktadır. Ard arda gelen görüntü karelerinde eşleştirme yapılamasa bile, öznitelik daha önceki görüntü karelerinde tespit edilmiş ve veri tabanına eklenmiş ise tekrar eşleştirme yapılabilir. Bu başlık altında yapılan işlem, eşleştirme sayılarının ardışık görüntü karelerinde hangi seviyede olduğunu gözlemek için gerçekleştirilmiştir.

340 numaralı görüntüde yer alan ORB öznitelikleri



Şekil 6.26. 340 numaralı görüntüden ORB özniteliklerinin çıkarım sonuçları

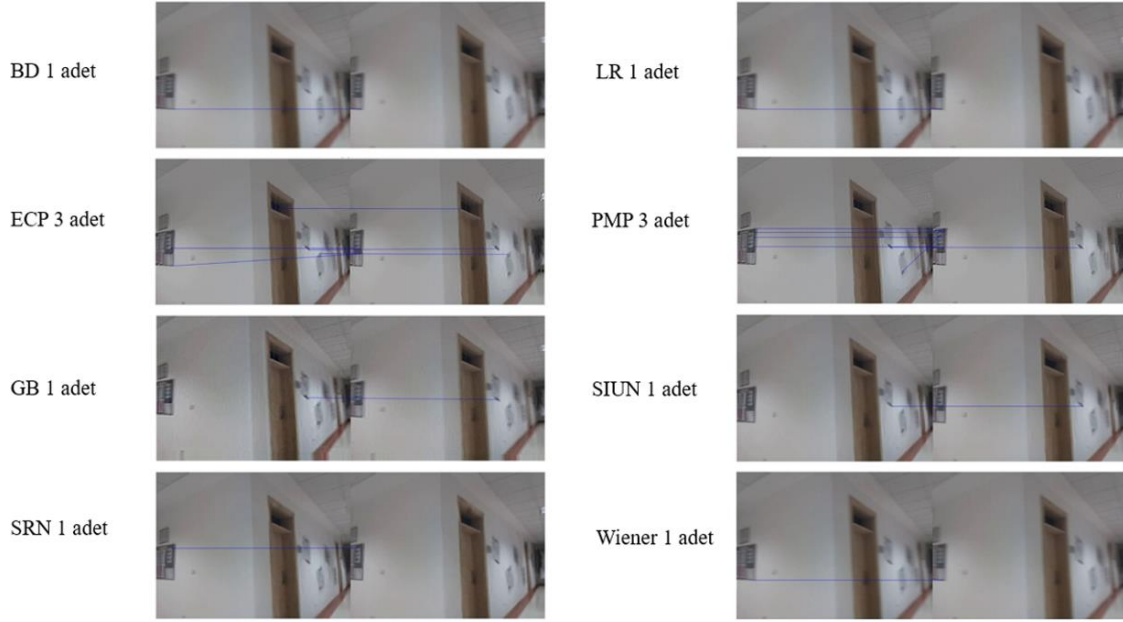
825 numaralı görüntüde yer alan ORB öznitelikleri



Şekil 6.27. 825 numaralı görüntüden ORB özniteliklerinin çıkarım sonuçları

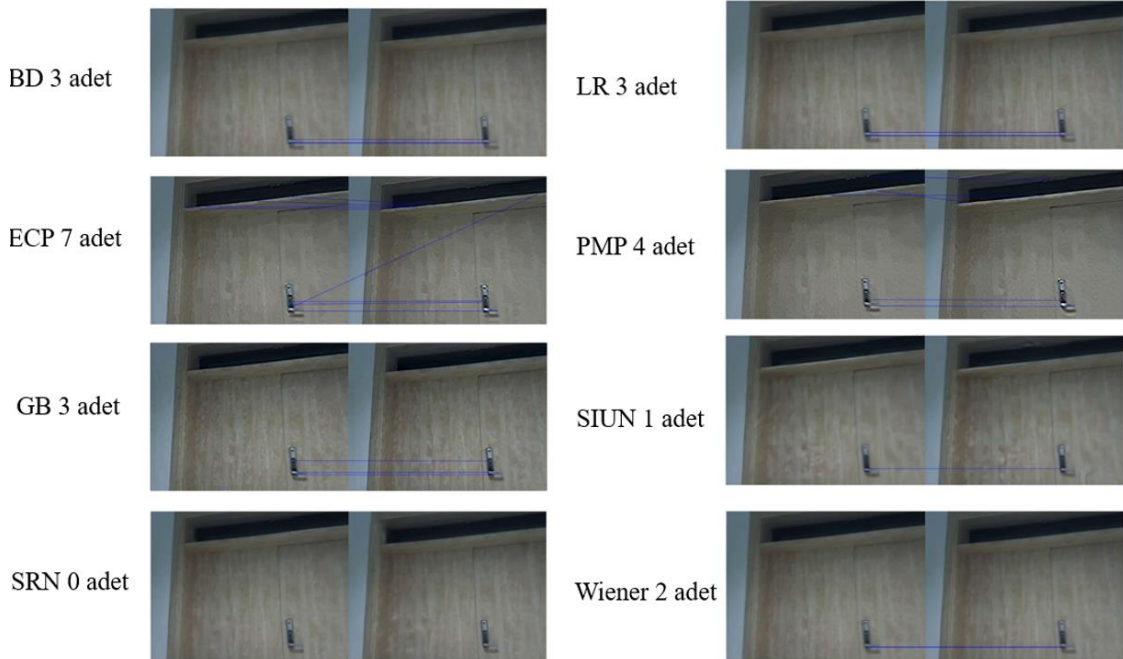
340-341 ve 807-808 ardışık görüntü kareleri için ORB özniteliklerinin eşleştirme sonuçları aşağıdaki Şekil 6.28 ile Şekil 6.29'da gösterilmiştir.

ORB Öznitelikleri Eşleştirme Sonuçları 340-341



Şekil 6.28. 340-341 numaralı görüntülerde ORB özniteliklerinin eşleştirmeleri

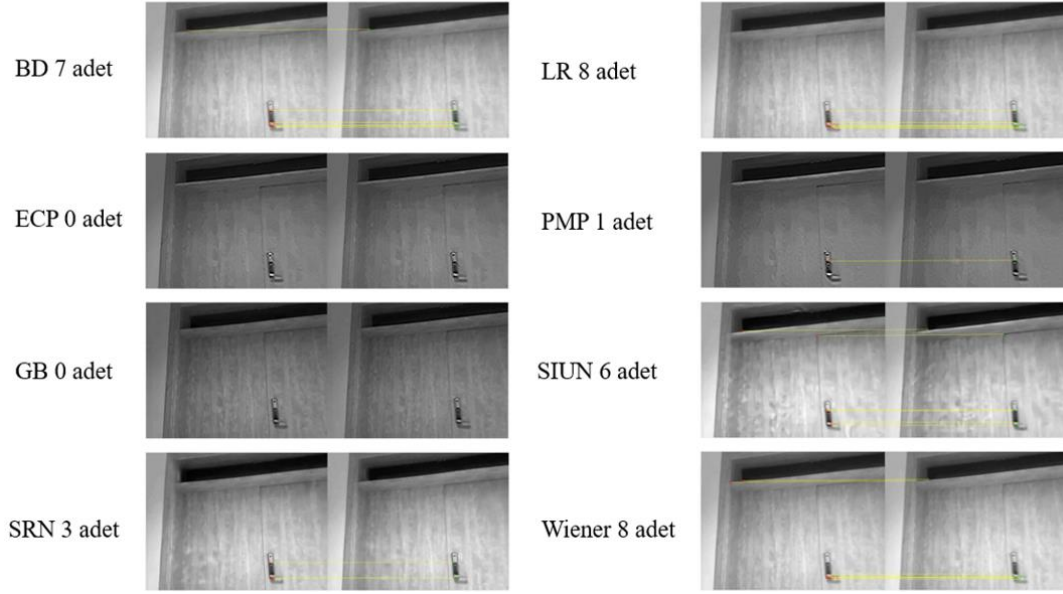
ORB Öznitelikleri Eşleştirme Sonuçları 807-808



Şekil 6.29. 807-808 numaralı görüntülerde ORB özniteliklerinin eşleştirmeleri

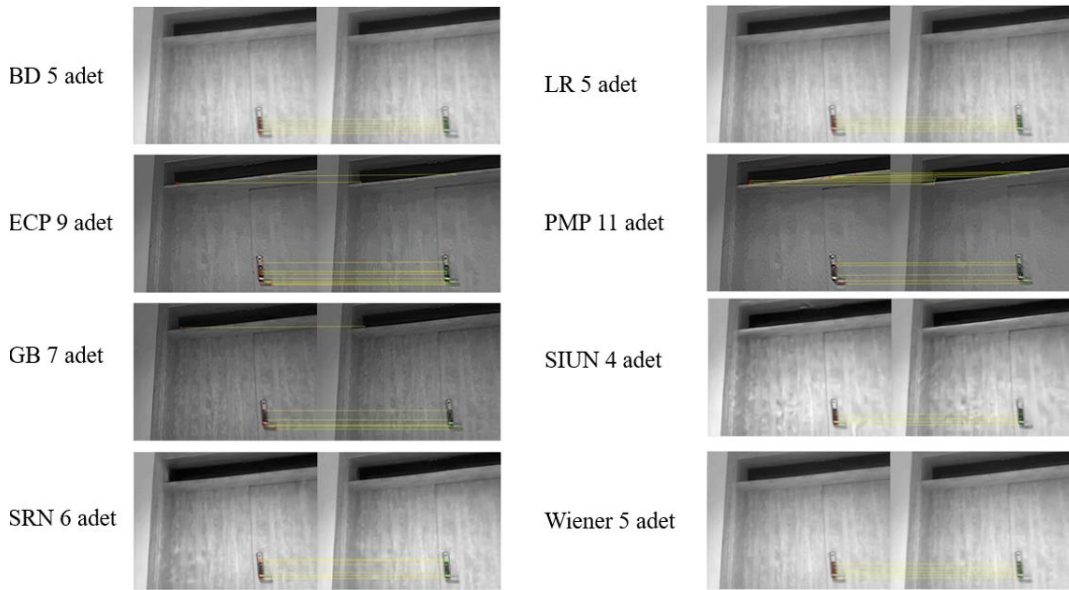
Harris köşe tespit algoritması (Harris, 1988) ve ve SURF (Bay, 2006) özniteliklerinin oluşturulan veri kümesindeki sonuçları aşağıdaki gibidir.

Harris Köşe Öznitelikleri Eşleştirme Sonuçları 807-808



Şekil 6.30. 807-808 numaralı görüntülerde Harris özniteliklerinin eşleştirmeleri

SURF Öznitelikleri Eşleştirme Sonuçları 807 - 808



Şekil 6.31. 807-808 numaralı görüntülerde SURF özniteliklerinin eşleştirmeleri

Şekil 6.30 ve Şekil 6.31’den görüleceği üzere, ardışık görüntü karelerinde bulanıklığın azaltıldığı restore edilmiş resimlerde eşleşmelerin bulanıklık azaltma yöntemlerinin başarımına göre arttığı gözlemlenmiştir.

6.9. AG-Mono: İHA'lar için Monoküler Görüntü Veri Kümesi

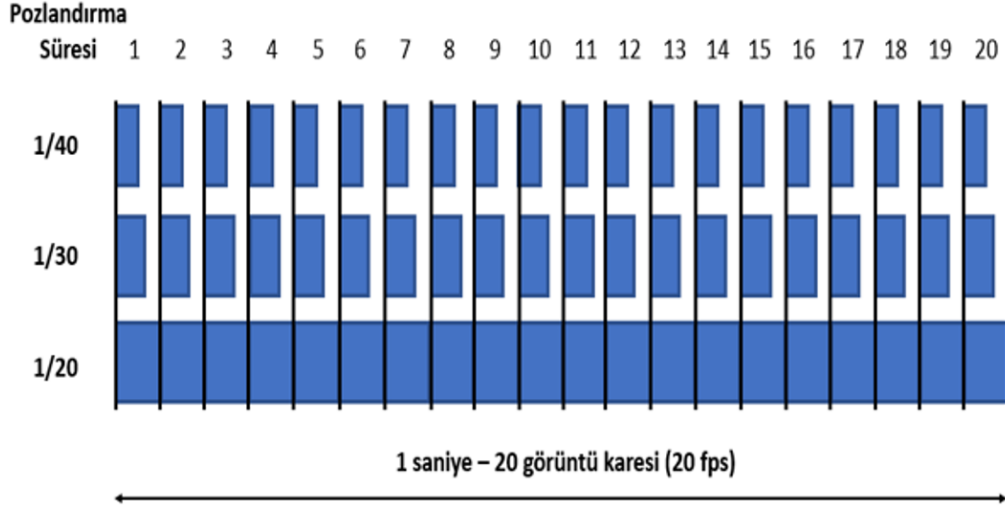
İHA'lar için görüntü tabanlı konumlama ve haritalama operasyonlarında, karşılaşılabilecek en zor operasyon senaryonlarından bir tanesi de 'U' şeklindeki dönüş hareketleridir. Bu dönüşlerde 'L' şeklindeki dönüşlerde de olduğu gibi dönüş gerçekleştirildikten sonra ilk defa karşılaşılan alanlar daha önce veritabanında bulunmayan öznitelikleri içermektedir. Dönüş 'U' şeklinde olduğunda bu durum iki kez tekrarlanmış olup, bu dönüşlerde öznitelikleri yakalamak daha zorlu bir hal almaktadır.

'U' şeklindeki dönüş hareketine yönelik önerilen çerçevenin etkinliğini gözlemlemek amacıyla AG-Mono ismiyle, akademik kullanıma açık görüntü veri kümesi oluşturulmuştur. AG-Mono veri kümesi, 16,5 x 4,5 metrelik koridor ortamında 22 derece/sn hızında 2 adet ard arda 'L' şeklindeki dönüşün toplamı olarak 'U' şeklindeki bir dönüşü içermektedir. Veri kümesinde ayrıca, pozlandırma süresinin etkisini de gözlemlemek amacıyla 3 farklı pozlandırma süresinde kayıtlar alınmıştır. Bu pozlandırma süresi ve görüntü özellikleri aşağıdaki Çizelge 6.11'de özetlenmiştir.

Çizelge 6.11. AG-Mono veri kümesi pozlandırma süreleri

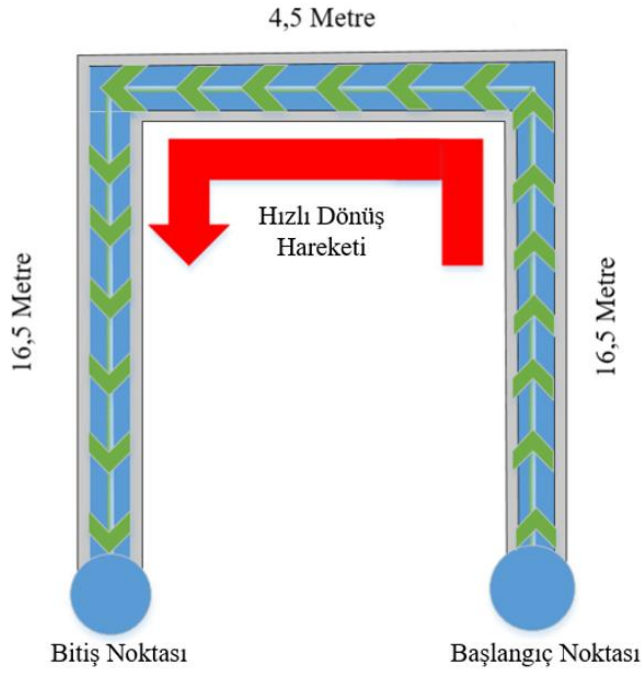
Görüntü Sekansı Adı	Görüntü Boyutu	Pozlandırma Süresi
AG-Mono-v1	640x 480	24 milisaniye $\approx 1/40$
AG-Mono-v2	640x 480	33 milisaniye $\approx 1/30$
AG-Mono-v3	640x 480	49 milisaniye $\approx 1/20$

Pozlandırma süresi, ışığın kamera optiğinden geçerek görüntünün oluşturulmasına kadar geçen süredir. Pozlandırma süresinin arttırılması hareket bulanıklığını riskini arttırırken, görüntünün oluşturulması için gerekli ışığın kamerada kullanılmasını sağlamaktadır. Bununla birlikte, daha az pozlandırma süresi bulanık görüntüyü önlerken, bu sürenin çok kısaltılması görüntü için gerekli ışığın alınamamasına ve karanlık görüntünün elde edilmesine neden olabilmektedir. Şekil 6.32'de farklı pozlandırma süreleri saniyede 20 görüntü kaydeden bir kamera için gösterilmiştir. Pozlandırma süresi gerçek zamanlı uygulamalarda şekilden de anlaşılacağı üzere 20fps için 1/20 saniyeden (50 ms) daha büyük olamamaktadır.



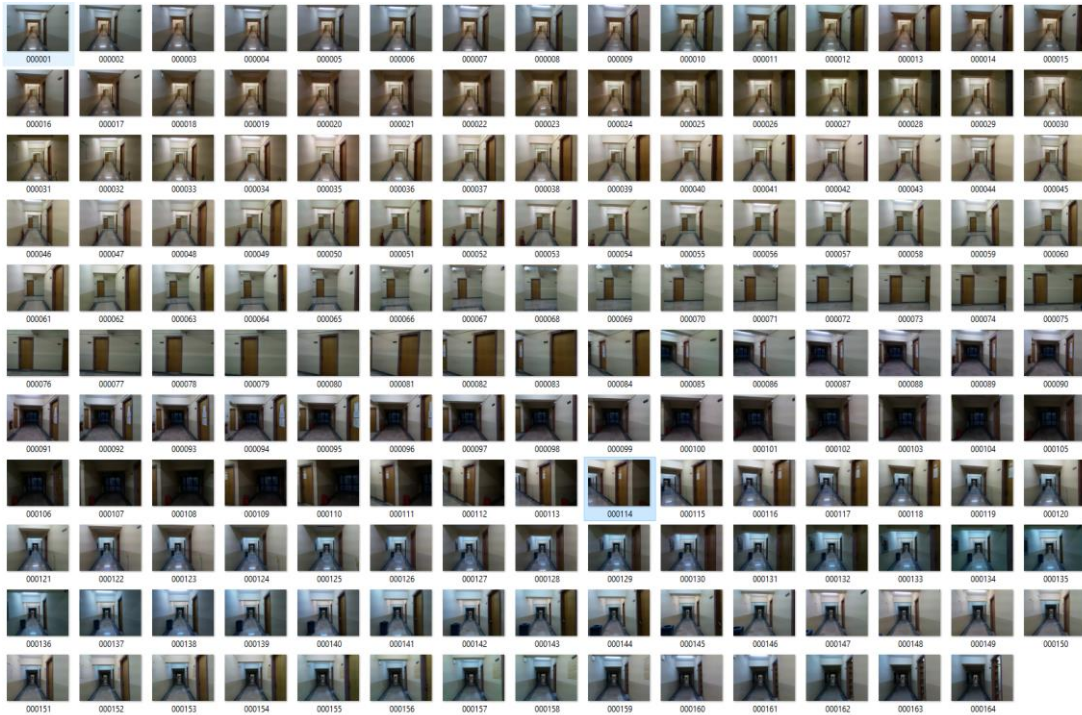
Şekil 6.32. Farklı pozlandırma sürelerinin 20fps görüntü akışındaki gösterimi

AG-Mono veri kümesinde gerçekleştirilen hareket, Şekil 6.33'te yer alan deney ortamına ait şematik çizimde gösterilmiştir. İlgili deney ortamında, 37,5 metrelik parkur yaklaşık 0,75m/s ileri hızda kat edilmiştir. Şekilde başlangıç ve bitiş noktaları mavi renk ile, ileri yöndeki hareketin yönü ise yeşil renkte yönü gösterilmektedir. Dönüş hareketinin gösteriminde ise kırmızı renk kullanılmıştır. Köşe noktalardaki dönüş hareketinde hız, daha önce oluşturulan veri kümesi ile aynı tutularak 22 derece/s (veya 0.384 rad/s) olarak gerçekleştirilmiştir. Şekilden de görüleceği üzere deney ortamında dönüş hareketine tamamlanır tamamlanmaz farklı bir koridor ortamına doğru yeniden dönüş hareketi gerçekleştirilmiştir. Veri kümesindeki görüntüler doğrudan kullanıldığında ORB-SLAM2 algoritmasında beklendiği üzere takip kaybı meydana gelmiştir.



Şekil 6.33. AG-Mono veri kümesine ait deney ortamının şematik gösterimi

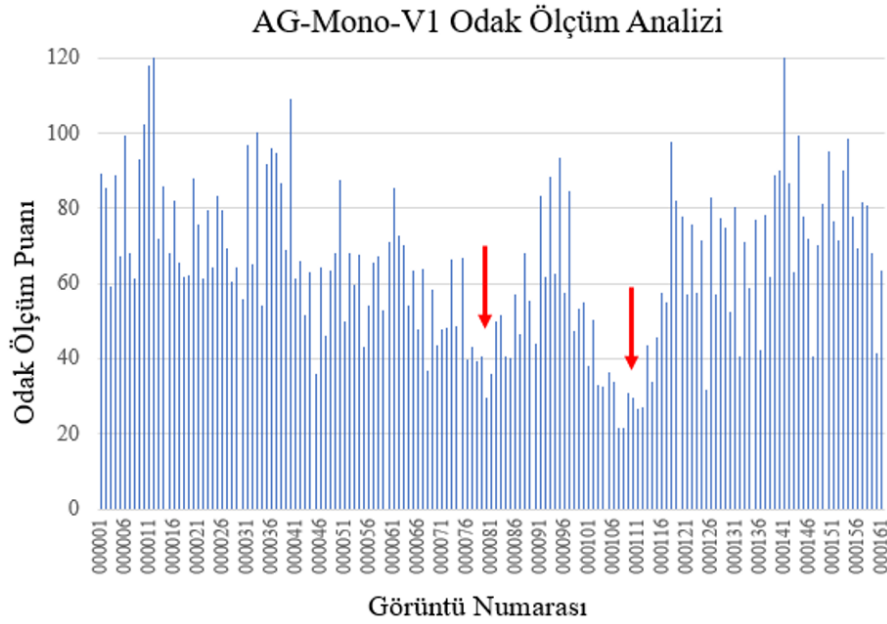
AG-Mono-v3 görüntü sekansına ait görüntü kümesi Resim 6.9’da gösterilmektedir. Değişen pozlandırma zamanlarına göre her bir görüntü sekansı için ayrı ayrı görüntü kümeleri mevcuttur.



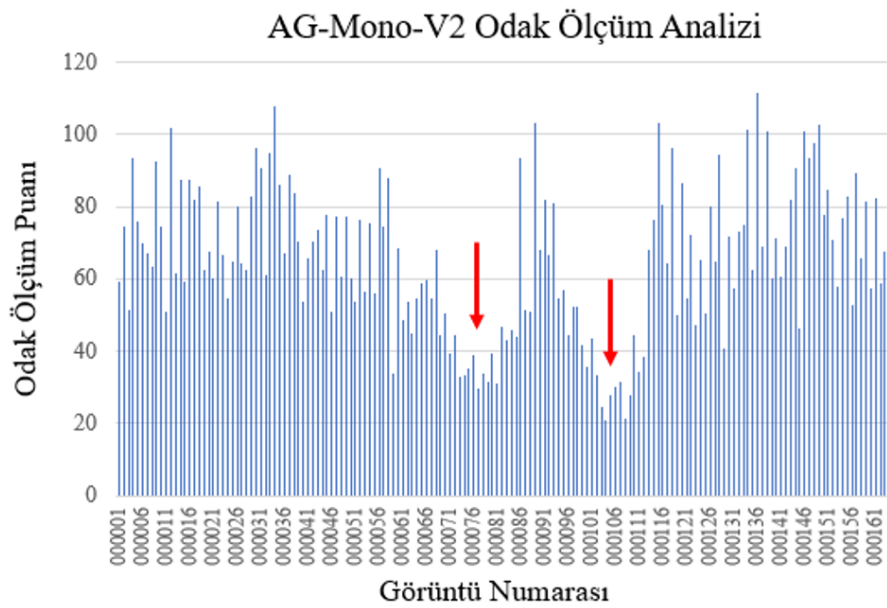
Resim 6.9. AG-Mono-V3 görüntü sekansına ait görüntü kümesi

6.9.1. AG-Mono veri kümesinin detaylı analizi

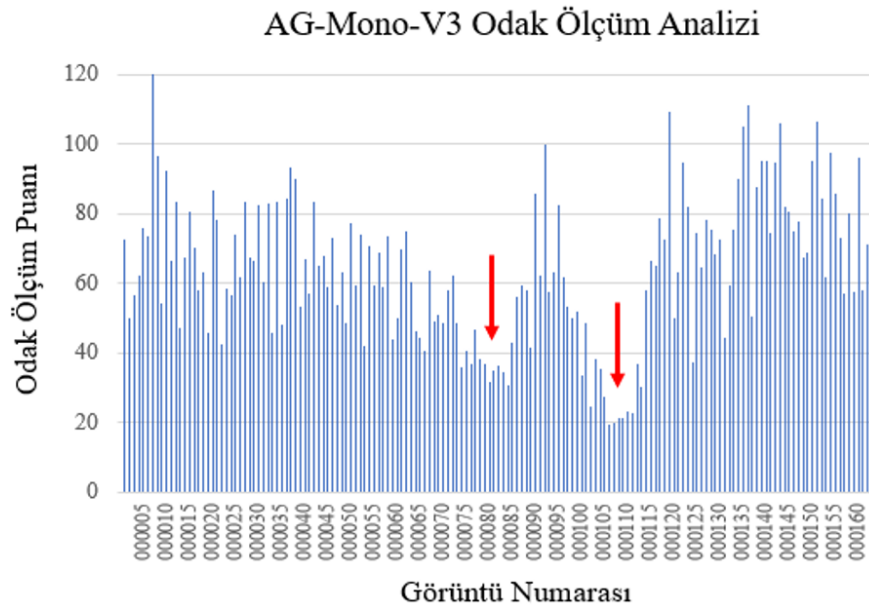
Farklı pozlandırma zamanlarına sahip görüntü sekansları için, OÖP analizi LAP4 yöntemi kullanılarak gerçekleştirilmiştir. Her bir yönteme ait OÖP sonuçları Şekil 6.34, Şekil 6.35 ve Şekil 6.36'da verilmiştir.



Şekil 6.34. AG-Mono-V1 görüntü sekansına ait OÖP sonuçları



Şekil 6.35. AG-Mono-V2 görüntü sekansına ait OÖP sonuçları



Şekil 6.36. AG-Mono-V3 görüntü sekansına ait OÖP sonuçları

Her bir görüntü sekansı için, odak ölçüm puanı analizi sonuçları Çizelge 6.12’de verilmiştir.

Çizelge 6.12. AG-Mono veri kümesi ortalama OÖP sonuçları

Görüntü Sekansı Adı	Pozlandırma Süresi	Ortalama OÖP
AG-Mono-v1	24 milisaniye $\cong 1/40$	65,09
AG-Mono-v2	33 milisaniye $\cong 1/30$	64,71
AG-Mono-v3	49 milisaniye $\cong 1/20$	63,79

Çizelge 6.12 incelendiğinde pozlandırma süresi daha küçük olan AG-Mono-v1 görüntü sekansının ortalama OÖP değerinin diğer sekanslardan yüksek olduğu dolayısıyla bu sekansa ait bulanıklığın diğer sekanslardan düşük olduğu ve sırasıyla AG-Mono-v2 ve AG-Mono-v3’ün bu sekansı takip ettiği görülmektedir. Buradan pozlandırma zamanındaki düşüşün, görüntülerdeki netliği arttırdığı görülmektedir. Pozlandırma süresi ne kadar uzun olursa, kamera ve önündeki nesnelerin hareket etmesi için süre o kadar uzamış olmaktadır. Bu durumda beklendiği üzere hareket bulanıklığının oluşması ihtimali uzun pozlandırma sürelerinde artmaktadır. AG-Mono veri kümesindeki her bir görüntü sekansına ait eşik değeri Eş.6.3’te verilen formüle göre hesaplanmıştır. AG-Mono veri kümesindeki OÖP

lokal minimumları, ağırlıklar ve tercih değerleri ile hesaplanan eşik değerleri Çizelge 6.13’de gösterilmiştir.

Çizelge 6.13. AG-Mono veri kümesi eşik değeri sonuçları

Görüntü Sekansı Adı	Ağırlık Değerleri	En Düşük 3 Lokal Minimum	Tercih Değeri	Elde Edilen Eşik Değeri
AG-Mono-v1	0,7	21,66	0,993	37
	1,5	31,88	0,925	
	1,8	29,55	0,962	
AG-Mono-v2	0,7	20,62	0,993	38
	1,5	40,62	0,852	
	1,8	29,62	0,920	
AG-Mono-v3	0,7	19,57	0,993	38
	1,5	37,32	0,932	
	1,8	30,77	0,871	

Elde edilen eşik değerlerinin altında kalan görüntüler daha önce başarılı olduğu kanıtlanan PMP yöntemi ile restore edilmiştir.

6.9.2. Önerilen çerçevenin AG-Mono veri kümesindeki performansı

Restore edilen görüntüler veri kümesindeki original bulanık görüntüler ile değiştirilmiş ve 20fps veri işleme hızında ORB-SLAM2 algoritmasında performans testleri gerçekleştirilmiştir. Gerçekleştirilen testlerde daha önce hareket bulanıklığı nedeniyle takip kaybına uğrayan bölgelerde algoritmanın takip kaybına uğramadığı gözlemlenmiştir.

AG-Mono-V1 görüntü sekansında takip kaybının yaşandığı hareket bulanıklığı yüksek olan 87. ve 110. görüntü karelerinde elde edilen sonuçlar Şekil 6.37’de gösterilmiştir. Şekilde daha önce öznitelik tespiti yapılamayan görüntü karelerinde sırasıyla 107 ve 144 adet öznitelik eşleştirmesi yapıldığı gözlemlenmiştir. Resim 6.10’da önerilen çerçeve ile ORB-SLAM2 yöntemi kullanılarak başarıyla tamamlanan ‘U’ şeklindeki dönüşü ait oluşturulan harita gösterilmiştir.

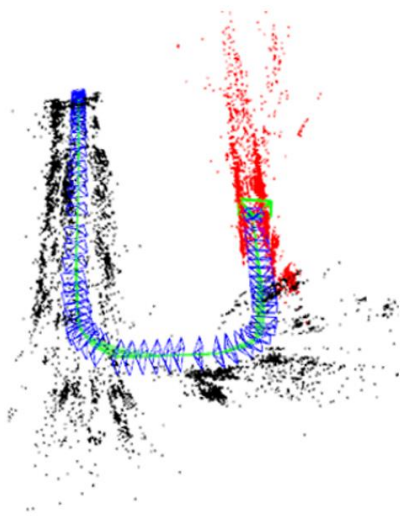


(a)



(b)

Şekil 6.37. AG-Mono-V1 görüntü sekansı 87. (a) ve 110. görüntü (b) kareleri



Resim 6.10. AG-Mono-V1 görüntü sekansına başarılı dönüş haritası

AG-Mono-V2 ve AG-Mono-V3 görüntü sekanslarında yer alan takip kaybına uğrayan restore edildikten sonra kullanılanbilen görüntü kareleri Şekil 6.38 ve Şekil 6.39’da gösterilmiştir.



(a)



(b)

Şekil 6.38. AG-Mono-V2 görüntü sekansı 83. (a) ve 108. görüntü (b) kareleri



(a)



(b)

Şekil 6.39. AG-Mono-V3 görüntü sekansı 83. (a) ve 108. görüntü (b) kareleri

7. SONUÇ VE ÖNERİLER

Günümüzde robotların ve insansız araçların birçok alanda kullanılması mümkün hale gelmiştir. Yakın gelecekte mikro/nano boyutlu İHA kullanımının artırılması planlanmaktadır. İnsansız araçların beklenen operasyonları gerçekleştirebilmesi için konumlama ve haritalama teknolojileri olmazsa olmazdır. GNSS sinyalleri mikro/nano İHA'lar tarafından çoğunlukla doğrudan kullanılmamaktadır. Çünkü GNSS'in metre düzeyindeki hata seviyeleri platformların boyutundan yüzlerce kat daha fazladır. Bu hata oranı mikro/nano ölçekteki platformların güvenli operasyon gerçekleştirmelerine uygun değildir. Ayrıca bu hassasiyet seviyesinde konumlandırma da bu platformlar için mümkün olmamaktadır. Ek olarak, pek çok insansız aracın ve robotların da iç mekanlarda çalışması beklenmektedir. Platformların ağırlık ve boyut kısıtları nedeniyle, hafif, kompakt, düşük maliyetli, kolayca kalibre olabilen ve düşük güç tüketimine sahip monoküler RGB kameralar bu platformlara en uygun kameralar arasında yer almaktadır. Ağırlık ve boyut sınırlamaları nedeniyle bir gimbal kullanmak da çoğunlukla (en azından mevcut durumda) mümkün değildir. Görsel SLAM (vSLAM) yöntemleri, monoküler RGB kameralı sistemlerin konumlarını belirlemek için tercih edilen yöntemlerdendir.

Bu tür platformlar, platform üzerinde yerleşik hesaplama için uygun pil kapasitesine sahip değildir. Bu nedenle, iletilen video ile kontrol istasyonunda görüntü işleme uygulamaları koşturulmaktadır. Böylece iletilen video hem operatörün platformu kontrol etmesi için hem de vSLAM uygulamalarında haritalama ve konumlama için kullanılabilir.

Bu teze konu olan çalışma özellikle mikro / nano seviyedeki insansız hava araçları ve robotların sadece düşük maliyetli ve küçük boyutlu monoküler kamera kullanılması ile görüntü temelli eş zamanlı konumlama ve haritalama performansını arttıracak yeni bir çerçeve sunmaktadır. Bu çerçeve ile mevcut çerçevelerde hareket bulanıklığından etkilenen vSLAM algoritmaları hareket bulanıklığına daha dayanıklı hale getirilmiştir.

Önerilen vSLAM çerçevesinde, hareket bulanıklığını tespit etmek için bir odak ölçüm operatörü olarak Laplacian Varyansı (LAP4) seçilmiştir. Tercih edilen LAP4 yönteminde, görüntünün tek bir kanalı Laplacian çekirdeği ile konvolüsyon işlemine tabi tutulur ve yanıtın varyansı hesaplanarak odak ölçüm puanı bulunur. Odak ölçüm puanı eşğin

üzerindeyse vSLAM süreci beklendiği gibi devam eder. Aksi takdirde, görüntü bulanıklığı azaltma yöntemleriyle görüntüler restore edilir. Bulanıklığın azaltılması / önlenmesi, tüm kareler için değil, yalnızca eşiğin altındaki kareler için uygulanır. Bu sayede işlem süresi uygun bir seviyede tutulur. Görüntülerin bulanıklaştırma teknikleriyle geri yüklendiği vSLAM algoritmalarında takip performansının arttığı ve takip kaybı oranının azaldığı gözlemlenmiştir.

İlerleyen dönemde manevra kabiliyeti yüksek binlerce mikro / nano İHA'nın çeşitli görevlerde tek kullanımlık olarak kullanılabileceği düşünüldüğünde bu çalışma ile daha maliyet etkin bir çözüm ortaya konulduğu değerlendirilmektedir. Mikro / nano İHA'ların henüz çok erken örnekleri piyasada mevcuttur ve ilerleyen dönemde bu tipte insansız platformlara olan ilginin artması yüksek olasılıktadır. Kullanımdaki artış ile hafif ve kompakt sensörler ile bu platformların görevlerini gerçekleştirmeleri için teknolojik arayışlar da artacaktır.

Önerilen hareket bulanıklığına dayanıklı vSLAM çerçevesi sayesinde, hareket bulanıklığı kaynaklı takip kaybının önlenmesi ile daha önce hızlı hareketlerde özellikle ani dönüş hareketlerinde takip kaybına uğrayan vSLAM algoritmalarının takip kaybına uğramadığı gözlemlenmiştir. Burada kullanılan bulanıklık azaltma algoritmalarının da önemli olduğu ve uygun algoritmanın seçimi konusunda da detaylı analizler çalışmada verilmiştir.

İHA'ların operasyonlarda saldırgan manevra yeteneklerinden kaynaklı oluşacak hızlı dönüş hareketleri hareket bulanıklığına neden olabilmektedir. Hareket bulanıklığı en son teknoloji vSLAM metotlarında takip kayıplarına neden olmaktadır. Bu nedenle bu teze konu olan çalışmanın konusu, mevcut algoritmaların takip kaybına uğradığı durumlara yeni bir çevre ile sonuç bulmak olmuştur. Önerilen çerçeve hem 'L' şeklindeki hem de 'U' şeklindeki dönüşlerde başarılı olmuştur. İlerleyen dönemde, mikro /nano İHA'ların kullanımının artmasıyla sadece kamera kullanarak konumlama ve haritlama yapan sistemlere yönelik çözüm arayışlarının artacağı düşünülmektedir.

Önerilen çerçeve daha önce takip kaybını önlemeye yönelik olarak yapılmış az sayıdaki çalışmalardan farklı olarak bulanıklık tespiti ve bulanıklık azaltma modülleriyle vSLAM algoritmalarına yeni modüller eklenmiştir. Bu modüller mevcut durumda takip

kayıplarının önlenmesini/en aza indirilmesini sağlamaktadır. Fakat gerçek zamanlı kullanımda işlem gücü gereksinimlerinin de arttıracağı düşünülmektedir.

Oluşturulan veri kümelerinde ortamda yer alan nesneler statik olup herhangi bir hareketli nesne yer almamaktadır. Hareket bulanıklığının kaynaklarından biri de ortamdaki nesnelerin pozlandırma süresi boyunca hareket etmesidir. Oluşturulan veri kümelerinde, hareketli nesnelerin de yer alması bulanıklık seviyesini daha karmaşık bir hale getirebilmektedir. Bu durum vSLAM algoritmalarının daha zorlayıcı bir etki ile karşılaşmasına yol açabilir. İlerleyen çalışmalarda hareketli nesnelerin olduğu veri kümelerinde vSLAM çerçevesinin başarısının test edilmesi faydalı olacaktır.

Olay (event) kameraları geliştiricileri, bu kameraları nano ölçekte ucuza ürettiklerinde, vSLAM uygulamaları için monoküler RGB kameralara bir alternatif olarak değerlendirilebilirler. Bununla birlikte, olay kameralarının ortamın ayrıntılarını sağlaması beklendiğinde ek bir yeniden yapılandırma aşamasına gerek duyulmaktadır. Diğer yandan detaylı çevresel veriler RGB kameralar ile kolaylıkla sağlanabilmektedir. Ayrıca, RGB kameraların çıktı görüntüleri, operatörlerin platformu kontrol etmesini kolaylaştıran insan görüşüne benzerlik taşımaktadır. RGB kameraların, maliyet açısından ciddi avantajları mevcuttur.

Odak ölçüm puanındaki kazanç, yapay zekâ temelli algoritmalar olan SIUN ve SRN algoritmaları için en düşük seviyede çıkmıştır. Bu teze konu olan çalışmada yapay zekâ temelli algoritmaların 640 x 480 çözünürlüğü için daha önceden eğitilmiş oldukları hazır modeller kullanılmıştır. Bu yöntemler için çok daha fazla veri ile model eğitimi özel olarak gerçekleştirilmiş olsaydı sonuçların elde edilenden daha başarılı olabileceği değerlendirilmektedir.

Yapılan çalışma, önerilen çerçevenin doğruluğunun ispat edilmesi anlamında önemli bir aşamayı geride bırakmıştır. Bu çerçevenin gerçek zamanlı olarak da uygulanması gelecek çalışmalarda gerçekleştirilecektir. Daha geniş bir bakış açısı ve vizyon ile ilerleyen dönemde platformların sürü olarak kullanımında tek yer istasyonundan konumlama ve haritalama ve birden çok kameranın etkin kullanımı konularında da çerçevede değişiklikler ve geliştirmeler yapılmalıdır. Özellikle birden çok kameradan akan görüntü verilerinin eş zamanlı olarak ya da belirli bir önceliklendirme ile tespitinin

gerçekleştirilmesi konusunda çalışmalarla yapılan çalışma sürü operasyonlarına uyumlu hale getirilebilir.

Sonuçlar incelendiğinde karşılaşılan bir diğer sonuç, hareket bulanıklığının azaltıldığı restore edilmiş görüntülerin kullanıldığı durumda doğrudan yaklaşımlarda performansın, öznitelik temelli yaklaşımlara göre daha az başarılı olduğu görülmektedir. Önerilen çerçevenin bazı bulanıklık azaltma yöntemlerinde doğrudan yöntemlere beklenen faydayı sunmadığı gözlemlenmiştir. Bu nedenle önerilen vSLAM çerçevesinin kullanımında, öznitelik temelli ya da doğrudan yöntemler arasından seçilen vSLAM algoritmasına en uygun bulanıklık azaltma yönteminin seçimi daha başarılı sonuçlar elde edilmesini sağlayacaktır. Bu durum, GB yöntemi ile bulanıklık azaltılmasında da gözlemlenmiştir. Kullanılan yöntem bulanıklığı azaltırken ya da insan algısına daha uygun şekilde görüntünün restorasyonunu yaparken resmin yapısal özelliklerini bozuyorsa, bu yöntemin özniteliklerin eşleştirilmesinde ve piksellerin takibinde de olumlu sonuç vermediği gözlemlenmiştir. Bu nedenle seçilecek bulanıklık azaltma yönteminin uygun yöntemler arasından seçilmesi gerekmektedir. Bu teze konu olan çalışmada PMP yönteminin seçilen kamera ve oluşturulan veri kümelerinde başarılı olduğu gözlemlenmiştir.

KAYNAKLAR

- Agrawal, A., Xu, Y., and Raskar, R. (2009). Invertible motion blur in video. In *ACM SIGGRAPH 2009 papers*, 28(3), 1-8.
- Agrawal, A., Xu, Y., and Raskar, R. (2012). *U.S. Patent No. 8,229,244*. Washington, DC: U.S. Patent and Trademark Office.
- Almeida, M. S. and Almeida, L. B. (2009). Blind and semi-blind deblurring of natural images. *IEEE Transactions on Image Processing*, 19(1), 36-52.
- Anwar, S., Huynh, C. P., and Porikli, F. (2018). Image deblurring with a class-specific prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(9), 2112-2130.
- Aqel, M. O., Marhaban, M. H., Saripan, M. I., and Ismail, N. B. (2016). Review of visual odometry: types, approaches, challenges, and applications. *SpringerPlus*, 5(1), 1-26.
- Atcioglu, B. (2016). Micro unmanned air vehicles formation process and models, *International Journal of Mechanical and Production Engineering (IJMPE)*, 4(6), 91-95.
- Austin, R. (2011). *Unmanned aircraft systems: UAVS design, development and deployment* (First Edition). John Wiley & Sons, 5-9.
- Azartash, H., Banai, N., and Nguyen, T. Q. (2014). An integrated stereo visual odometry for robotic navigation. *Robotics and Autonomous Systems*, 62(4), 414-421.
- Bai, Y., Cheung, G., Liu, X., and Gao, W. (2018). Graph-based blind image deblurring from a single photograph. *IEEE Transactions on Image Processing*, 28(3), 1404-1418.
- Bai, Y., Jia, H., Jiang, M., Liu, X., Xie, X., and Gao, W. (2019). Single-image blind deblurring using multi-scale latent structure prior. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(7), 2033-2045.
- Bailey, T. and Durrant-Whyte, H. (2006). Simultaneous localization and mapping (SLAM): Part II. *IEEE Robotics and Automation Magazine*, 13(3), 108-117.
- Bascle, B., Blake, A., and Zisserman, A. (1996, April). *Motion deblurring and super-resolution from an image sequence*. Paper presented at European Conference on Computer Vision, Berlin, 571-582.
- Bay, H., Tuytelaars, T., and Van Gool, L. (2006, May). *Surf: Speeded up robust features*. Paper presented at European Conference on Computer Vision, Berlin, 404-417.
- Ben-Ari, M. and Mondada, F. (2018). Robots and their applications. *Elements of robotics* (Third Edition). Springer, Cham, 1-20.

- Bengio, Y., Louradour, J., Collobert, R., and Weston, J. (2009, June). *Curriculum learning*. Paper presented at 26th Annual International Conference on Machine Learning, New York, 41-48.
- Berger, P., Hannak, G., and Matz, G. (2017). Graph signal recovery via primal-dual algorithms for total variation minimization. *IEEE Journal of Selected Topics in Signal Processing*, 11(6), 842-855.
- Boracchi, G., Foi, A., Katkovnik, V., and Egiazarian, K. (2008, March). Deblurring noisy radial-blurred images: spatially adaptive filtering approach. *Image Processing: Algorithms and Systems VI. International Society for Optics and Photonics*, 6812, 68121D.
- Brunelli, R. (2009). *Template matching techniques in computer vision: theory and practice* (Second Edition). John Wiley & Sons, 43-73
- Bunschoten, R. and Krose, B. (2003, September). *Visual odometry from an omnidirectional vision system*. Paper presented at 2003 IEEE International Conference on Robotics and Automation, Taipei, Taiwan, 577-583.
- Burgard, W., Hebert, M., and Bennewitz, M. (2016). World modeling. *Springer handbook of robotics* (Second Edition). Springer, Cham, 1135-1152.
- Burri, M., Nikolic, J., Gohl, P., Schneider, T., Rehder, J., Omari, S., Archtelik, W., and Siegwart, R. (2016). The EuRoC micro aerial vehicle datasets. *The International Journal of Robotics Research*, 35(10), 1157-1163.
- Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., and Leonard, J. J. (2016). Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Transactions on Robotics*, 32(6), 1309-1332.
- Cai, J. F., Ji, H., Liu, C., and Shen, Z. (2011). Framelet-based blind motion deblurring from a single image. *IEEE Transactions on Image Processing*, 21(2), 562-572.
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6), 679-698.
- Carrillo, L. R. G. López, A. E. D., Lozano, R., and Pégard, C. (2012). *Quad rotorcraft control: vision-based hovering and navigation* (First Edition). Springer Science & Business Media, 71-102.
- Chakrabarti, A. (2016, October). *A neural approach to blind motion deblurring*. Paper presented at European Conference on Computer Vision, Amsterdam, 221-235.
- Chan, T. F. and Wong, C. K. (1998). Total variation blind deconvolution. *IEEE Transactions on Image Processing*, 7(3), 370-375.

- Chen, L., Fang, F., Wang, T., and Zhang, G. (2019). *Blind image deblurring with local maximum gradient prior*. Paper presented at IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 1742-1750.
- Chen, Y., Zhou, Y., Lv, Q., and Deveerasetty, K. K. (2018, August). *A review of v-slam*. Paper presented at 2018 IEEE International Conference on Information and Automation (ICIA), Fujian, 603-608.
- Cho, H., Wang, J., and Lee, S. (2012, October). *Text image deblurring using text-specific properties*. Paper presented at European Conference on Computer Vision, Berlin, 524-537.
- Cho, S. and Lee, S. (2009). *Fast motion deblurring*. Paper presented at Special Interest Group on Computer Graphics and Interactive Techniques Asia, Yokohama, Japan, 1-8.
- Cho, T. S., Paris, S., Horn, B. K., and Freeman, W. T. (2011, June). *Blur kernel estimation using the radon transform*. Paper presented at Conference on Computer Vision and Pattern Recognition CVPR 2011, Colorado, 241-248.
- Couprie, C., Grady, L., Najman, L., Pesquet, J. C., and Talbot, H. (2013). Dual constrained TV-based regularization on graphs. *SIAM Journal on Imaging Sciences*, 6(3), 1246-1273.
- Cremers, D. (2017, May). *Direct methods for 3d reconstruction and visual slam*. Paper presented at 2017 15th IAPR International Conference on Machine Vision Applications (MVA), Japan 34-38.
- Crete, F., Dolmieri, T., Ladret, P., and Nicolas, M. (2007, February). The blur effect: perception and estimation with a new no-reference perceptual blur metric. *Human Vision and Electronic Imaging XII, International Society for Optics and Photonics*, (6492, 64920I).
- Cvišić, I. and Petrović, I. (2015, September). *Stereo odometry based on careful feature selection and tracking*. Paper presented at 2015 European Conference on Mobile Robots (ECMR), Lincoln, United Kingdom, 1-6.
- Cvišić, I., Ćesić, J., Marković, I., and Petrović, I. (2018). SOFT-SLAM: Computationally efficient stereo visual simultaneous localization and mapping for autonomous unmanned aerial vehicles. *Journal of Field Robotics*, 35(4), 578-595.
- Davison, A. J. (2003, October). Real-time simultaneous localisation and mapping with a single camera. Paper presented at IEEE International Conference on Computer Vision, Nice, 1403-1404.
- Davison, A. J., Reid, I. D., Molton, N. D., and Stasse, O. (2007). MonoSLAM: Real-time single camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6), 1052-1067.

- Delbracio, M. and Sapiro, G. (2015). *Burst deblurring: Removing camera shake through fourier burst accumulation* Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Boston, 2385-2393.
- Delmerico, J. and Scaramuzza, D. (2018, May). *A benchmark comparison of monocular visual-inertial odometry algorithms for flying robots*. Paper presented at 2018 IEEE International Conference on Robotics and Automation (ICRA), Australia, 2502-2509.
- Dey, N., Blanc-Féraud, L., Zimmer, C., Kam, Z., Olivo-Marin, J. C., and Zerubia, J. (2004, April). *A deconvolution method for confocal microscopy with total variation regularization*. Paper presented at 2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro, Arlington, USA, 1223-1226.
- Elbanhawi, M., Mohamed, A., Clothier, R., Palmer, J. L., Simic, M., and Watkins, S. (2017). Enabling technologies for autonomous MAV operations. *Progress in Aerospace Sciences*, 91, 27-52.
- El-Henawy, I. M., Amin, A. E., and Kareem Ahmed, H. A. (2014). A comparative study on image deblurring techniques. *International Journal of Advances in Computer Science and Technology (IJACST)*, 3(12), 01-08.
- Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 48(1), 71-99.
- Elmoataz, A., Lezoray, O., and Bougleux, S. (2008). Nonlocal discrete regularization on weighted graphs: a framework for image and manifold processing. *IEEE Transactions on Image Processing*, 17(7), 1047-1060.
- Engel, J., Koltun, V., and Cremers, D. (2017). Direct sparse odometry. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(3), 611-625.
- Engel, J., Schöps, T., and Cremers, D. (2014, September). *LSD-SLAM: Large-scale direct monocular SLAM*. Paper presented at European Conference on Computer Vision, Zurich, 834-849.
- Engel, J., Sturm, J., and Cremers, D. (2013). *Semi-dense visual odometry for a monocular camera*. Paper presented at IEEE International Conference on Computer Vision, Sydney, 1449-1456.
- Engel, J., Usenko, V., and Cremers, D. (2016). A photometrically calibrated benchmark for monocular visual odometry. *arXiv preprint arXiv:1607.02555*.
- Everaerts, J. (2006). New Platforms—Unconventional Platforms for Remote Sensing. *EuroSDR Project Presentations*. 24-25.
- Fabian, J. R. and Clayton, G. M. (2014, July). *Adaptive visual odometry using RGB-D cameras*. Paper presented at 2014 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Besançon, France, 1533-1538.

- Fang, Z. and Zhang, Y. (2015). Experimental evaluation of RGB-D visual odometry methods. *International Journal of Advanced Robotic Systems*, 12(3), 26.
- Fergus, R., Singh, B., Hertzmann, A., Roweis, S. T., and Freeman, W. T. (2006). *Removing camera shake from a single photograph*. Paper presented at Special Interest Group on Computer Graphics and Interactive Techniques 2006, Boston, 787-794.
- Forster, C., Zhang, Z., Gassner, M., Werlberger, M., and Scaramuzza, D. (2016). SVO: Semidirect visual odometry for monocular and multicamera systems. *IEEE Transactions on Robotics*, 33(2), 249-265.
- Fraundorfer, F. and Scaramuzza, D. (2012). Visual odometry: Part ii: Matching, robustness, optimization, and applications. *IEEE Robotics & Automation Magazine*, 19(2), 78-90.
- Gao, X., Wang, R., Demmel, N., and Cremers, D. (2018, October). *LDSO: Direct sparse odometry with loop closure*. Paper presented at 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, 2198-2204.
- Geiger, A., Lenz, P., Stiller, C., and Urtasun, R. (2013). Vision meets robotics: The kitti dataset. *The International Journal of Robotics Research*, 32(11), 1231-1237.
- Golban, C., Istvan, S., and Nedeveschi, S. (2012, June). *Stereo based visual odometry in difficult traffic scenes*. Paper presented at 2012 IEEE Intelligent Vehicles Symposium, Madrid, 736-741.
- Goldstein, A. and Fattal, R. (2012, October). *Blur-kernel estimation from spectral irregularities*. Paper presented at European Conference on Computer Vision, Berlin, 622-635.
- González, D., Patricio, M. A., Berlanga, A., and Molina, J. M. (2019). A super-resolution enhancement of UAV images based on a convolutional neural network for mobile devices. *Personal and Ubiquitous Computing*, 23(3), 1-12.
- Gonzalez, R., Rodriguez, F., Guzman, J. L., Pradalier, C., and Siegwart, R. (2013). Control of off-road mobile robots using visual odometry and slip compensation. *Advanced Robotics*, 27(11), 893-906.
- Gonzalez, R., Rodriguez, F., Guzman, J. L., Pradalier, C., and Siegwart, R. (2012). Combined visual odometry and visual compass for off-road mobile robots localization. *Robotica*, 30(6), 865-878.
- Grisetti, G., Stachniss, C., and Burgard, W. (2005, April). *Improving grid-based slam with rao-blackwellized particle filters by adaptive proposals and selective resampling*. Paper presented at 2005 IEEE International Conference on Robotics and Automation, Barcelona, 2432-2437.

- Handa, A., Whelan, T., McDonald, J., and Davison, A. J. (2014, May). *A benchmark for RGB-D visual odometry, 3D reconstruction and SLAM*. Paper presented at 2014 IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, 1524-1531.
- Hansen, P. C., Nagy, J. G., and O'leary, D. P. (2006). *Deblurring images: matrices, spectra, and filtering* (First Edition). Society for Industrial and Applied Mathematics, 24-28.
- Harris, C. and Stephens, M. (1988, August). A combined corner and edge detector. Paper presented at Alvey Vision Conference, Manchester, 15(50), 10-52.
- Hartley, R. and Zisserman, A. (2003). *Multiple View Geometry in Computer Vision* (Second Edition). Cambridge University Press, 2-3.
- He, K., Sun, J. and Tang, X. (2010). Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12), 2341-2353.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). *Deep residual learning for image recognition*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, 770-778.
- Hidane, M., L  zoray, O., and Elmoataz, A. (2013). Nonlinear multilayered representation of graph-signals. *Journal of Mathematical Imaging and Vision*, 45(2), 114-137.
- Hirsch, M., Schuler, C. J., Harmeling, S., and Sch  lkopf, B. (2011, November). *Fast removal of non-uniform camera shake*. Paper presented at 2011 International Conference on Computer Vision, Barcelona, 463-470.
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780.
- Howard, A. (2008, September). *Real-time stereo visual odometry for autonomous ground vehicles*. Paper presented at 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, Nice, 3946-3952.
- Hradi  , M., Kotera, J., Zemc  k, P., and   roubek, F. (2015, September). Convolutional neural networks for direct text deblurring. *The British Machine Vision Association*, 10(2), 1-13.
- Hu, W., Cheung, G., and Kazui, M. (2015). Graph-based dequantization of block-compressed piecewise smooth images. *IEEE Signal Processing Letters*, 23(2), 242-246.
- Hu, Z., Cho, S., Wang, J., and Yang, M. H. (2014a). *Deblurring low-light images with light streaks*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, 3382-3389.

- Hu, Z., Xu, L., and Yang, M. H. (2014b). *Joint depth estimation and camera shake removal from single blurry image*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, 2893-2900.
- Huang, A. S., Bachrach, A., Henry, P., Krainin, M., Maturana, D., Fox, D., and Roy, N. (2017). *Visual odometry and mapping for autonomous flight using an RGB-D camera*. Paper presented at 18th Symposium on Robotics Research, Puerto Varas, Chile, 235-252.
- Huang, G. (2019, May). *Visual-inertial navigation: A concise review*. Paper presented at 2019 International Conference on Robotics and Automation (ICRA), Montreal, 9572-9582.
- İnternet: Alan Kendrick J.D. Current State of the Art of Micro Air Vehicles. Erişim adresi: http://nerac.com/wpcontent/uploads/2014/09/MAV_FinalSep14.pdf.: Son Erişim Tarihi:22.12.2021.
- İnternet: Bouguet, J.Y. Camera calibration toolbox for Matlab. Erişim adresi: http://www.vision.caltech.edu/bouguetj/calib_doc/index.html. Son Erişim Tarihi:22.12.2021.
- İnternet: Department of Defence. Unmanned Systems Integrated Roadmap FY2013–2038. (Washington, DC: Department of Defense, 2013) Erişim adresi: <http://archive.defense.gov/pubs/DOD-USRM-2013.pdf>. Son Erişim Tarihi:22.12.2021.
- İnternet: Kudan. An Introduction to Simultaneous Localisation and Mapping. Erişim adresi: <https://www.kudan.io/post/an-introduction-to-simultaneous-localisation-and-mapping>. Erişim: Son Erişim Tarihi:22.12.2021.
- İnternet: Loughborough University, (2011). Corporate Identity Guidelines. Erişim adresi: <http://www.lboro.ac.uk/designandprint/downloads/corporate-identity/lboro-corporate-identityguidelines.pdf> pp 38. Son Erişim Tarihi:22.12.2021.
- İnternet: Medium. Introduction to Orb Oriented Fast and Rotated Brief. Erişim adresi: <https://medium.com/data-breach/introduction-to-orb-oriented-fast-and-rotated-brief-4220e8ec40cf>. Son Erişim Tarihi:22.12.2021.
- Jia, J. (2007, June). *Single image motion deblurring using transparency*. Paper presented at 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, 1-8.
- Jiang, Y., Xiong, G., Chen, H., and Lee, D. J. (2014). Incorporating a wheeled vehicle model in a new monocular visual odometry algorithm for dynamic outdoor environments. *Sensors*, 14(9), 16159-16180.
- Jiang, Y., Xu, Y., and Liu, Y. (2013). Performance evaluation of feature detection and matching in stereo visual odometry. *Neurocomputing*, 120, 380-390.

- Johnson, A. E., Goldberg, S. B., Cheng, Y., and Matthies, L. H. (2008, May). *Robust and efficient stereo feature tracking for visual odometry*. Paper presented at 2008 IEEE International Conference on Robotics and Automation, California, 39-46.
- Joshi, N., Szeliski, R., and Kriegman, D. J. (2008, June). *PSF estimation using sharp edge prediction*. Paper presented at 2008 IEEE Conference on Computer Vision and Pattern Recognition, Alaska, 1-8.
- Juneja, M. and Sandhu, P. S. (2009). Performance evaluation of edge detection techniques for images in spatial domain. *International Journal of Computer Theory and Engineering*, 1(5), 614.
- Jung, S. W., Kim, T. H., and Ko, S. J. (2009). A novel multiple image deblurring technique using fuzzy projection onto convex sets. *IEEE Signal Processing Letters*, 16(3), 192-195.
- Kheradmand, A. and Milanfar, P. (2014). A general framework for regularized, similarity-based image restoration. *IEEE Transactions on Image Processing*, 23(12), 5136-5151.
- Kicman, P. and Narkiewicz, J. (2013). *Concept of integrated INS/visual system for autonomous mobile robot operation*. Paper presented at Marine Navigation and Safety of Sea Transportation Symposium, Poland, 20-24.
- Klein, G. S. and Drummond, T. (2005, September). A Single-frame visual gyroscope. *The British Machine Vision Association*. 3(1), 1-8.
- Klein, G. S. and Drummond, T. W. (2004). Tightly integrated sensor fusion for robust visual tracking. *Image and Vision Computing*, 22(10), 769-776.
- Klein, G. and Murray, D. (2008, October). *Improving the agility of keyframe-based SLAM*. Paper presented at European Conference on Computer Vision, Berlin, 802-815.
- Klein, G. and Murray, D. (2009, October). *Parallel tracking and mapping on a camera phone*. Paper presented at 2009 8th IEEE International Symposium on Mixed and Augmented Reality, Orlando 83-86.
- Krishnan, D. and Fergus, R. (2009). Fast image deconvolution using hyper-Laplacian priors. *Advances in Neural Information Processing Systems*, 22(2009), 1033-1041.
- Krishnan, D., Tay, T., and Fergus, R. (2011, June). *Blind deconvolution using a normalized sparsity measure*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Colorado, 233-240.
- Kupyn, O., Budzan, V., Mykhailych, M., Mishkin, D., and Matas, J. (2018). *Deblurgan: Blind motion deblurring using conditional adversarial networks*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Utah, 8183-8192.

- Kümmerle, R., Grisetti, G., Strasdat, H., Konolige, K., and Burgard, W. (2011, May). *g2o: A general framework for graph optimization*. Paper presented at 2011 IEEE International Conference on Robotics and Automation, Shanghai, 3607-3613.
- Labbé, M. and Michaud, F. (2019). RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation. *Journal of Field Robotics*, 36(2), 416-446.
- Lelégard, L., Delaygue, E., Brédif, M., and Vallet, B. (2012). Detecting and correcting motion blur from images shot with channel-dependent exposure time. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3(2012), 341-346.
- Leutenegger, S., Lynen, S., Bosse, M., Siegwart, R., and Furgale, P. (2015). Keyframe-based visual-inertial odometry using nonlinear optimization. *The International Journal of Robotics Research*, 34(3), 314-334.
- Levin, A., Fergus, R., Durand, F., and Freeman, W. T. (2007). Image and depth from a conventional camera with a coded aperture. *ACM Transactions on Graphics (TOG)*, 26(3), 70.
- Levin, A., Weiss, Y., Durand, F., and Freeman, W. T. (2009, June). *Understanding and evaluating blind deconvolution algorithms*. Paper presented at 2009 IEEE Conference on Computer Vision and Pattern Recognition, Florida, 1964-1971.
- Levin, A., Weiss, Y., Durand, F., and Freeman, W. T. (2011). Understanding blind deconvolution algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12), 2354-2367.
- Li, F., Yu, J. and Chai, J. (2008, June). *A hybrid camera for motion deblurring and depth map super-resolution*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Alaska, 1-8.
- Li, L., Pan, J., Lai, W. S., Gao, C., Sang, N., and Yang, M. H. (2018). *Learning a discriminative prior for blind image deblurring*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Utah, 6616-6625.
- Lin, T. C., Hou, L., Liu, H., Li, Y., and Truong, T. K. (2018). Reconstruction of single image from multiple blurry measured images. *IEEE Transactions on Image Processing*, 27(6), 2762-2776.
- Liu, Z., Yeh, R. A., Tang, X., Liu, Y., and Agarwala, A. (2017). *Video frame synthesis using deep voxel flow*. Paper presented at IEEE International Conference on Computer Vision, Venice, 4463-4471.
- Lovegrove, S., Davison, A. J., and Ibanez-Guzmán, J. (2011, June). *Accurate visual odometry from a rear parking camera*. Paper presented at 2011 IEEE Intelligent Vehicles Symposium (IV), Baden, Germany, 788-793.

- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 91-110.
- Lucy, L. B. (1974). An iterative technique for the rectification of observed distributions. *The Astronomical Journal*, 79, 745.
- Maimone, M., Cheng, Y. and Matthies, L. (2007). Two years of visual odometry on the mars exploration rovers. *Journal of Field Robotics*, 24(3), 169-186.
- Makhubela, J. K., Zuva, T., and Agunbiade, O. Y. (2018, December). *A review on vision simultaneous localization and mapping (VSLAM)*. Paper presented at 2018 International Conference on Intelligent and Innovative Computing Applications (ICONIC), Mauritius, 1-5.
- Martinez, G. (2013, January). *Monocular visual odometry from frame to frame intensity differences for planetary exploration mobile robots*. Paper presented at 2013 IEEE Workshop on Robot Vision (WORV), Florida, 54-59.
- McManus, C., Furgale, P., and Barfoot, T. D. (2013). Towards lighting-invariant visual navigation: An appearance-based approach using scanning laser-rangefinders. *Robotics and Autonomous Systems*, 61(8), 836-852.
- Michaeli, T. and Irani, M. (2014, September). *Blind deblurring using internal patch recurrence*. Paper presented at European Conference on Computer Vision, Zurich, 783-798.
- Montemerlo, M. and Thrun, S. (2003, September). *Simultaneous localization and mapping with unknown data association using FastSLAM*. Paper presented at 2003 IEEE International Conference on Robotics and Automation, New Jersey, 1985-1991.
- Mouats, T., Aouf, N., Sappa, A. D., Aguilera, C., and Toledo, R. (2014). Multispectral stereo odometry. *IEEE Transactions on Intelligent Transportation Systems*, 16(3), 1210-1224.
- Mur-Artal, R., Montiel, J. M. M., and Tardos, J. D. (2015). ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, 31(5), 1147-1163.
- Mur-Artal, R. and Tardós, J.D. (2017). Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE Transactions on Robotics*, 33(5), 1255–1262.
- Nagatani, K., Ikeda, A., Ishigami, G., Yoshida, K., and Nagai, I. (2010). Development of a visual odometry system for a wheeled robot on loose soil using a telecentric camera. *Advanced Robotics*, 24(8-9), 1149-1167.
- Nah, S., Hyun Kim, T., and Mu Lee, K. (2017). *Deep multi-scale convolutional neural network for dynamic scene deblurring*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 3883-3891.

- Narvekar, N. D. and Karam, L. J. (2011). A no-reference image blur metric based on the cumulative probability of blur detection (CPBD). *IEEE Transactions on Image Processing*, 20(9), 2678-2683.
- Newcombe, R. A., Lovegrove, S. J., and Davison, A. J. (2011, November). *DTAM: Dense tracking and mapping in real-time*. Paper presented at 2011 International Conference on Computer Vision, Barcelona, 2320-2327.
- Nistér, D. (2004a). An efficient solution to the five-point relative pose problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6), 756-770.
- Nistér, D., Naroditsky, O., and Bergen, J. (2004b, June). *Visual odometry*. Paper presented at 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2004, Washington. 1-1.
- Nistér, D., Naroditsky, O., and Bergen, J. (2006). Visual odometry for ground vehicle applications. *Journal of Field Robotics*, 23(1), 3-20.
- Nourani-Vatani, N. and Borges, P. V. K. (2011). Correlation-based visual odometry for ground vehicles. *Journal of Field Robotics*, 28(5), 742-768.
- Oliva, A. and Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, 155(2006), 23-36.
- Ong, E., Lin, W., Lu, Z., Yang, X., Yao, S., Pan, F., Jiang, L., and Moschetti, F. (2003, July). *A no-reference quality metric for measuring image blur*. Paper presented at Seventh International Symposium on Signal Processing and Its Applications, Paris, 469-472.
- Pan, J., Hu, Z., Su, Z., and Yang, M. H. (2014). *Deblurring text images via L0-regularized intensity and gradient prior*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, 2901-2908.
- Pan, J., Hu, Z., Su, Z., and Yang, M. H. (2016a). l0-regularized intensity and gradient prior for deblurring text images and beyond. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(2), 342-355.
- Pan, J., Sun, D., Pfister, H., and Yang, M. H. (2016b). *Blind image deblurring using dark channel prior*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, 1628-1636.
- Pan, J., Sun, D., Pfister, H., and Yang, M. H. (2017). Deblurring images via dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(10), 2315-2328.
- Pang, J. and Cheung, G. (2017). Graph Laplacian regularization for image denoising: Analysis in the continuous domain. *IEEE Transactions on Image Processing*, 26(4), 1770-1785.

- Pathak, S., Moro, A., Fujii, H., Yamashita, A., and Asama, H. (2019, March). *Distortion-resistant spherical visual odometry for UAV-based bridge inspection*. Paper presented at International Workshop on Advanced Image Technology (IWAIT), Long Beach, 316-321.
- Paz, L. M., Piniés, P., Tardós, J. D., and Neira, J. (2008). Large-scale 6-DOF SLAM with stereo-in-hand. *IEEE Transactions on Robotics*, 24(5), 946-957.
- Pertuz, S., Puig, D., and Garcia, M. A. (2013). Analysis of focus measure operators for shape-from-focus. *Pattern Recognition*, 46(5), 1415-1432.
- Pretto, A., Menegatti, E., and Pagello, E. (2007, January). *Reliable features matching for humanoid robots*. Paper presented at 2007 7th IEEE International Conference on Humanoid Robots, Pittsburgh, USA, 532-538.
- Pumarola, A., Vakhitov, A., Agudo, A., Sanfeliu, A., and Moreno-Noguer, F. (2017, May). *PL-SLAM: Real-time monocular visual SLAM with points and lines*. Paper presented at 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapur, 4503-4508.
- Ragot, N., Khemmar, R., Pokala, A., Rossi, R., and Ertaud, J. Y. (2019, July). *Benchmark of visual slam algorithms: Orb-slam2 vs rtab-map*. Paper presented at 2019 Eighth International Conference on Emerging Security Technologies (EST), Essex, 1-6.
- Rahtu, E., Heikkilä, J., Ojansivu, V., and Ahonen, T. (2012). Local phase quantization for blur-insensitive image analysis. *Image and Vision Computing*, 30(8), 501-512.
- Raskar, R., Agrawal, A., and Tumblin, J. (2006). *Coded exposure photography: motion deblurring using fluttered shutter*. Paper presented at Special Interest Group on Computer Graphics and Interactive Techniques 2006 Papers, New York, 795-804.
- Rav-Acha, A. and Peleg, S. (2005). Two motion-blurred images are better than one. *Pattern Recognition Letters*, 26(3), 311-317.
- Ren, W., Cao, X., Pan, J., Guo, X., Zuo, W., and Yang, M. H. (2016). Image deblurring via enhanced low-rank prior. *IEEE Transactions on Image Processing*, 25(7), 3426-3437.
- Richardson, W. H. (1972). Bayesian-based iterative method of image restoration. *Journal Optical Society of America*, 62(1), 55-59.
- Ronneberger, O., Fischer, P., and Brox, T. (2015, October). *U-net: Convolutional networks for biomedical image segmentation*. Paper presented at International Conference on Medical Image Computing and Computer-assisted Intervention, Munich, 234-241.
- Royer, E., Lhuillier, M., Dhome, M., and Lavest, J. M. (2007). Monocular vision for mobile robot localization and autonomous navigation. *International Journal of Computer Vision*, 74(3), 237-260.

- Rudin, L. I., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4), 259-268.
- Scaramuzza, D. and Fraundorfer, F. (2011). Tutorial: Visual odometry. *IEEE Robotics & Automation Magazine*, 18(4), 80-92.
- Scaramuzza, D. and Siegwart, R. (2008). Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles. *IEEE Transactions on Robotics*, 24(5), 1015-1026.
- Scaramuzza, D., Fraundorfer, F., and Pollefeys, M. (2010). Closing the loop in appearance-guided omnidirectional visual odometry by using vocabulary trees. *Robotics and Autonomous Systems*, 58(6), 820-827.
- Schelten, K., Nowozin, S., Jancsary, J., Rother, C., and Roth, S. (2015, January). *Interleaved regression tree field cascades for blind image deconvolution*. Paper presented at 2015 IEEE Winter Conference on Applications of Computer Vision, Waikoloa Beach, USA, 494-501.
- Schubert, D., Goll, T., Demmel, N., Usenko, V., Stückler, J., and Cremers, D. (2018, October). *The TUM VI benchmark for evaluating visual-inertial odometry*. Paper presented at 2018 IEEE International Conference on Intelligent Robots and Systems (IROS), Madrid, 1680-1687.
- Schubert, S., Lange, S., Neubert, P., and Protzel, P. (2016) *Map Enhancement with Track-Loss Data in Visual SLAM*. Paper presented at International Conference on Intelligent Robots and Systems (IROS), Daejeon, Korea, 9-14.
- Schuler, C. J., Christopher Burger, H., Harmeling, S., and Scholkopf, B. (2013). *A machine learning approach for non-blind image deconvolution*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Portland, 1067-1074.
- Schuler, C. J., Hirsch, M., Harmeling, S., and Schölkopf, B. (2015). Learning to deblur. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(7), 1439-1451.
- Servières, M., Renaudin, V., Dupuis, A., and Antigny, N. (2021). Visual and Visual-Inertial SLAM: State of the Art, Classification, and Experimental Benchmarking. *Journal of Sensors*, (2021), 26-50.
- Shan, Q., Jia, J., and Agarwala, A. (2008). High-quality motion deblurring from a single image. *ACM Transactions on Graphics (TOG)*, 27(3), 1-10.
- Sheng, B., Li, P., Fang, X., Tan, P., and Wu, E. (2019). Depth-aware motion deblurring using loopy belief propagation. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(4), 955-969.

- Shuman, D. I., Narang, S. K., Frossard, P., Ortega, A., and Vandergheynst, P. (2013). The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Processing Magazine*, 30(3), 83-98.
- Siddiqui, R. and Khatibi, S. (2015). Robust visual odometry estimation of road vehicle from dominant surfaces for large-scale mapping. *Intelligent Transport Systems*, 9(3), 314-322.
- Sieberth, T., Wackrow, R., and Chandler, J. H. (2016). Automatic detection of blurred images in UAV image sets. *ISPRS Journal of Photogrammetry and Remote Sensing*, 122, 1-16.
- Singh, P. and Jain, L. (2013, December). *Noise reduction in ultrasound images using wavelet and spatial filtering techniques*. Paper presented at 2013 2nd International Conference on Information Management in the Knowledge Economy, Punjab, India, 57-63.
- Soltani, H., Taghirad, H. D., and Ravari, A. N. (2012, May). *Stereo-based visual navigation of mobile robots in unknown environments*. Paper presented at 20th Iranian Conference on Electrical Engineering, Tehran, Iran, 946-951.
- Sreehari, S., Venkatakrishnan, S. V., Wohlberg, B., Buzzard, G. T., Drummy, L. F., Simmons, J. P., and Bouman, C. A. (2016). Plug-and-play priors for bright field electron tomography and sparse interpolation. *IEEE Transactions on Computational Imaging*, 2(4), 408-423.
- Steinbrücker, F., Sturm, J., and Cremers, D. (2011, November). *Real-time visual odometry from dense RGB-D images*. Paper presented at 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), Barcelona, 719-722.
- Strasdat, H., Montiel, J. M., and Davison, A. J. (2012). Visual SLAM: why filter?. *Image and Vision Computing*, 30(2), 65-77.
- Sturm, J., Engelhard, N., Endres, F., Burgard, W., and Cremers, D. (2012, October). *A benchmark for the evaluation of RGB-D SLAM systems*. Paper presented at 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Algarve, Portugal, 573-580.
- Stühmer, J., Gumhold, S., and Cremers, D. (2010, September). *Real-time dense geometry from a handheld camera*. Paper presented at Joint Pattern Recognition Symposium, Berlin, 11-20.
- Su, S., Delbracio, M., Wang, J., Sapiro, G., Heidrich, W., and Wang, O. (2017). *Deep video deblurring for hand-held cameras*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 1279-1288.

- Sun, J., Cao, W., Xu, Z., and Ponce, J. (2015). *Learning a convolutional neural network for non-uniform motion blur removal*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Boston, 769-777.
- Sun, L., Cho, S., Wang, J., and Hays, J. (2013, April). *Edge-based blur kernel estimation using patch priors*. Paper presented at IEEE International Conference on Computational Photography (ICCP), Cambridge, 1-8.
- Susnjara, A., Perraudin, N., Kressner, D., and Vandergheynst, P. (2015). Accelerated filtering on graphs using lanczos method. *arXiv preprint arXiv:1509.04537*.
- Şimşek, B., Satir, S., and Bilge, H. Ş. (2021, June). *Performance comparison of direct and feature based vSLAM algorithms*. Paper presented at 2021 29th Signal Processing and Communications Applications Conference, İstanbul, 1-4.
- Taketomi, T., Uchiyama, H., and Ikeda, S. (2017). Visual SLAM algorithms: a survey from 2010 to 2016. *IPSP Transactions on Computer Vision and Applications*, 9(1), 1-11.
- Tao, X., Gao, H., Liao, R., Wang, J., and Jia, J. (2017). *Detail-revealing deep video super-resolution*. Paper presented at IEEE International Conference on Computer Vision, Venice, 4472-4480.
- Tao, X., Gao, H., Shen, X., Wang, J., and Jia, J. (2018). *Scale-recurrent network for deep image deblurring*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Sake Lake City, 8174-8182.
- Tardif, J. P., Pavlidis, Y., and Daniilidis, K. (2008, September). *Monocular visual odometry in urban environments using an omnidirectional camera*. Paper presented at 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, Nice, 2531-2538.
- Triggs, B., McLauchlan, P. F., Hartley, R. I., and Fitzgibbon, A. W. (1999, September). *Bundle adjustment a modern synthesis*. Paper presented at International Workshop on Vision Algorithms, Berlin, 298-372.
- Tukey, J. W. (1952). The extrapolation, interpolation and smoothing of stationary time series with engineering applications. *Journal Storage*, 24(1), 319-321.
- Üzer, F. (2010). *Camera motion blur and its effect on feature detectors*. Master's Thesis, Middle East Technical University, Ankara, 24-19.
- Van Hamme, D., Goeman, W., Veelaert, P., and Philips, W. (2015). Robust monocular visual odometry for road vehicles using uncertain perspective projection. *Journal on Image and Video Processing*, 2015(1), 1-21.
- Wang, Y., Fevig, R., and Schultz, R. R. (2008, October). *Super-resolution mosaicking of UAV surveillance video*. Paper presented at 2008 15th IEEE International Conference on Image Processing, Bordeaux, 345-348.

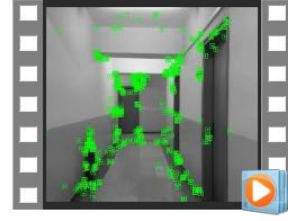
- Webster, C. B. and Reeves, S. J. (2007, August). *Radial deblurring with ffts*. Paper presented at 2007 IEEE International Conference on Image Processing, Texas, 1-101.
- Wen, F., Ying, R., Liu, P., and Qiu, R. C. (2019). Robust PCA using generalized nonconvex regularization. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(6), 1497-1510.
- Wen, F., Ying, R., Liu, Y., Liu, P., and Truong, T. K. (2020). A simple local minimal intensity prior and an improved algorithm for blind image deblurring. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(8), 2923-2937.
- Wenying, W., Jingxin, H., and Yi, Y. (2008, December). *Identification of blurred extent with rotation motion blurred image*. Paper presented at 2008 International Symposium on Information Science and Engineering, Shanghai, 669-672.
- Whyte, O., Sivic, J., and Zisserman, A. (2014). Deblurring shaken and partially saturated images. *International Journal of Computer Vision*, 110(2), 185-201.
- Wieschollek, P., Hirsch, M., Scholkopf, B., and Lensch, H. (2017). *Learning blind motion deblurring*. Paper presented at IEEE International Conference on Computer Vision, Venice, 231-240.
- Williams, B., Cummins, M., Neira, J., Newman, P., Reid, I., and Tardós, J. (2009). A comparison of loop closing techniques in monocular SLAM. *Robotics and Autonomous Systems*, 57(12), 1188-1197.
- Williams, B., Klein, G., and Reid, I. (2007, October). *Real-time SLAM relocation*. Paper presented at 2007 IEEE 11th International Conference on Computer Vision, Rio De Janeiro, 1-8.
- Xiao, L., Wang, J., Heidrich, W., and Hirsch, M. (2016, October). *Learning high-order filters for efficient blind deconvolution of document photographs*. Paper presented at European Conference on Computer Vision, Amsterdam, 734-749.
- Xu, L. and Jia, J. (2010, September). *Two-phase kernel estimation for robust motion deblurring*. Paper presented at European Conference on Computer Vision, Crete, 157-170.
- Xu, L., and Jia, J. (2012, April). *Depth-aware motion deblurring*. Paper presented at 2012 IEEE International Conference on Computational Photography (ICCP), Seattle, 1-8.
- Xu, L., Lu, C., Xu, Y., and Jia, J. (2011, December). *Image smoothing via L0 gradient minimization*. Paper presented at 2011 Special Interest Group on Computer Graphics and Interactive Techniques Asia Conference, Hong Kong, 1-12.
- Xu, L., Ren, J. S., Liu, C., and Jia, J. (2014). Deep convolutional neural network for image deconvolution. *Advances in Neural Information Processing Systems*, 27, 1790-1798.

- Xu, L., Zheng, S., and Jia, J. (2013). *Unnatural l0 sparse representation for natural image deblurring*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Portland, 1107-1114.
- Xu, N., Price, B., Cohen, S., and Huang, T. (2017). *Deep image matting*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 2970-2979.
- Yan, Y., Ren, W., Guo, Y., Wang, R., and Cao, X. (2017). *Image deblurring via extreme channels prior*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 4003-4011.
- Ye, M., Lyu, D. and Chen, G. (2020). Scale-iterative upscaling network for image deblurring. *IEEE Access*, 8 (2020), 18316-18325.
- Yoo, J., Hwang, S. S., Kim, S. D., Ki, M. S., and Cha, J. (2014). Scale-invariant template matching using histogram of dominant gradients. *Pattern Recognition*, 47(9), 3006-3018.
- Yu, Y., Pradalier, C., and Zong, G. (2011, July). *Appearance-based monocular visual odometry for ground vehicles*. Paper presented at 2011 IEEE International Conference on Advanced Intelligent Mechatronics (AIM), Budapest, 862-867.
- Zaffar, M., Ehsan, S., Stolkin, R., and Maier, K. M. (2018, August). *Sensors, slam and long-term autonomy: a review*. Paper presented at 2018 NASA/ESA Conference on Adaptive Hardware and Systems (AHS), Edinburgh, 285-290.
- Zhang, G. and Suh, I. H. (2012). A vertical and floor line-based monocular SLAM system for corridor environments. *International Journal of Control, Automation and Systems*, 10(3), 547-557.
- Zhang, H., Dai, Y., Li, H., and Koniusz, P. (2019). *Deep stacked hierarchical multi-patch network for image deblurring*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, 5978-5986.
- Zhang, J., Pan, J., Lai, W. S., Lau, R. W., and Yang, M. H. (2017a). *Learning fully convolutional networks for iterative non-blind deconvolution*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 3817-3825.
- Zhang, J., Pan, J., Ren, J., Song, Y., Bao, L., Lau, R. W., and Yang, M. H. (2018). *Dynamic scene deblurring using spatially variant recurrent neural networks*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, 2521-2529.
- Zhang, J., Singh, S., and Kantor, G. (2014). *Robust monocular visual odometry for a ground vehicle in undulating terrain*. Paper presented at Field and Service Robotics, Miyagi, Japan, 311-326.

- Zhang, J., Zhao, D., Xiong, R., Ma, S., and Gao, W. (2014). Image restoration using joint statistical modeling in a space-transform domain. *IEEE Transactions on Circuits and Systems for Video Technology*, 24(6), 915-928.
- Zhang, K., Zuo, W., Gu, S., and Zhang, L. (2017b). *Learning deep CNN denoiser prior for image restoration*. Paper presented at IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 3929-3938.
- Zhang, L., Zhou, L., and Huang, H. (2016). Bundled kernels for nonuniform blind video deblurring. *IEEE Transactions on Circuits and Systems for Video Technology*, 27(9), 1882-1894.
- Zhang, Z. (1999, September). *Flexible camera calibration by viewing a plane from unknown orientations*. Paper presented at 7th IEEE International Conference on Computer Vision, Corfu, Greece, 666-673.
- Zhao, Y., Xu, S., Bu, S., Jiang, H., and Han, P. (2019). *GSLAM: A general SLAM framework and benchmark*. Paper presented at IEEE International Conference on Computer Vision, Seoul, 1110-1120.
- Zheng, S., Xu, L., and Jia, J. (2013). *Forward motion deblurring*. Paper presented at 2013 IEEE International Conference on Computer Vision, Sydney, 1465-1472.
- Zhou, Y. and Komodakis, N. (2014, September). *A map-estimation framework for blind deblurring using high-level edge priors*. Paper presented at European Conference on Computer Vision, Zurich, 142-157.
- Zoran, D. and Weiss, Y. (2011, November). *From learning models of natural image patches to whole image restoration*. Paper presented at 2011 International Conference on Computer Vision, Barcelona, 479-486.
- Zou, D. and Tan, P. (2012). Coslam: Collaborative visual slam in dynamic environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(2), 354-366.

EKLER

EK-1.Video görünümleri

ag_mono-v1_basarisiz_
donusag_mono-v1_önerilen_
metotag_mono-v2_basarisiz_
donusag_mono-v2_önerilen_
metotag_mono-v3_basarisiz_
donusag_mono-v3_önerilen_
metot



GAZİ GELECEKTİR..