

T.C.
NİĞDE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

OTONOM ARAÇLARDA KULLANILAN GÖRSEL SEYİR
SİSTEMLERİ İÇİN YENİ BİR YAKLAŞIM

MURAT PEKER

Mayıs 2016

T.C.
NİĞDE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

OTONOM ARAÇLARDA KULLANILAN GÖRSEL SEYİR
SİSTEMLERİ İÇİN YENİ BİR YAKLAŞIM

MURAT PEKER

Doktora Tezi

Danışman

Yrd. Doç. Dr. Fuat KARAKAYA

Mayıs 2016

Murat PEKER tarafından **Fuat KARAKAYA** danışmanlığında hazırlanan “**OTONOM ARAÇLARDA KULLANILAN GÖRSEL SEYİR SİSTEMLERİ İÇİN YENİ BİR YAKLAŞIM**” adlı bu çalışma jürimiz tarafından Niğde Üniversitesi Fen Bilimleri Enstitüsü **Elektrik-Elektronik Mühendisliği** Ana Bilim Dalı’nda Doktora tezi olarak kabul edilmiştir.

Başkan : Prof. Dr. Tülay YILDIRIM, Yıldız Teknik Üniversitesi

Üye : Prof. Dr. Nurettin ACIR, Bursa Teknik Üniversitesi

Üye : Doç. Dr. Halis ALTUN, KTO Karatay Üniversitesi

Üye : Yrd. Doç. Dr. Bekir Sami TEZEKİCİ, Niğde Üniversitesi

Üye : Yrd. Doç. Dr. Fuat KARAKAYA, Niğde Üniversitesi

ONAY:

Bu tez, Fen Bilimleri Enstitüsü Yönetim Kurulunca belirlenmiş olan yukarıdaki jüri üyeleri tarafından/....../20.... tarihinde uygun görülmüş ve Enstitü Yönetim Kurulu’nun/....../20.... tarih ve sayılı kararıyla kabul edilmiştir.

...../...../20...

Doç. Dr. Murat BARUT
MÜDÜR

TEZ BİLDİRİMİ

Tez içindeki bütün bilgilerin bilimsel ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

Murat PEKER

ÖZET

OTONOM ARAÇLARDA KULLANILAN GÖRSEL SEYİR SİSTEMLERİ İÇİN YENİ BİR YAKLAŞIM

PEKER, Murat

Niğde Üniversitesi

Fen Bilimleri Enstitüsü

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Danışman

:Yrd. Doç. Dr. Fuat KARAKAYA

Nisan 2016, 123 sayfa

Bu tez çalışmasında, görsel algılayıcılar ile elde edilen görüntülerden, özgün algoritmalar yardımıyla çıkarılan yer izleri ve tanımlayıcılar kullanılarak, insansız araçlar için gürbüz görsel seyir sistemi gerçekleştirilmiştir. Görsel yer izi tabanlı konum değişimi hesaplama yöntemleri, bir görüntü akışı içerisindeki görüntülerde tekrarlanan yer izlerini tespit ederek konum değişimini bulmaya yarar. Tez kapsamında yeni bir yer izi tespit algoritması önerilmiştir. Önerilen yer izi tespit algoritması literatürdeki veri kümeleri üzerinde test edilmiş ve yaygın olarak kullanılan yer izi bulma algoritmasına (SIFT) göre daha fazla yer izi tespit ettiği benzetimlerle gösterilmiştir. İkinci olarak bu tez çalışmasında, yönelim açısı hesabı gerektiren yer izi tanımlayıcı algoritmalarında kullanılabilecek başarıyı yüksek ve hızlı bir yönelim açısı hesaplama algoritması önerilmiştir. Önerilen yönelim açısı hesaplama algoritmasının, çeşitli görüntü bozulmalarında literatürde mevcut diğer algoritmalara kıyasla daha az etkilendiği veri kümeleri üzerinde gösterilmiştir. Üçüncü olarak yönelim açısı hesabını ortadan kaldıran bir tanımlayıcı eşleştirme algoritması ve bu algoritmada kullanılabilecek bir tanımlayıcı deseni önerilmiştir. Son olarak bu çalışma kapsamında önerilen algoritmalar ve Kalman filtresi kullanılarak gerçekleştirilen görsel seyir sistemi iki farklı veri kümesi üzerinde test edilerek elde edilen sonuçlar incelenmiştir.

Anahtar Sözcükler: Yer izi tanımlayıcıları, görüntü eşleştirme, öznelik eşleştirme, görsel seyir sistemi.

SUMMARY

A NEW APPROACH TO VISUAL NAVIGATION SYSTEMS FOR AUTONOMOUS VEHICLES

PEKER, Murat

Nigde University

Graduate School of Natural and Applied Sciences

Department of Electrical and Electronics Engineering

Supervisor : Assistant Professor Dr. Fuat KARAKAYA

April 2016, 123 pages

This thesis presents a robust visual navigation system for autonomous unmanned vehicles. For this purpose novel algorithms are proposed to extract features and descriptors from images captured by visual sensors. Visual landmark-based localization methods detect repeating landmarks from images in a video stream to determine location change. In this thesis a new landmark detection algorithm is proposed. The proposed landmark detection algorithm was tested on commonly used benchmark problems and it was shown that the proposed algorithm could detect more landmarks than algorithm (SIFT) well known in the literature. As a second contribution of the thesis, a fast and effective new method is proposed to reduce the overhead cost of orientation estimation and it could be easily integrated to any image matching algorithms which require orientation estimation. It was shown that the proposed orientation estimation algorithm performs better immunity to some of the image distortion types than other algorithms available in the literature. Also, a new feature matching algorithm, aimed to eliminate the need for orientation estimation, and a descriptor pattern, used with the algorithm, is proposed. Finally, the visual navigation system, utilizing the proposed algorithms and Kalman filter, was tested on two different benchmark data sets and the results are presented.

Keywords: Landmark descriptors, image matching, feature matching, visual navigation system.

ÖN SÖZ

Bu tez çalışmasında, görsel algılayıcılar ile elde edilen görüntülerden, özgün algoritmalar yardımıyla çıkarılan yer izleri ve tanımlayıcılar kullanılarak, insansız araçlar için gürbüz görsel seyir sistemi gerçekleştirilmiştir. Bu çalışma kapsamında üç yeni algoritma önerilerek, literatürdeki mevcut görüntü eşleştirme yöntemlerinden daha başarılı sonuçlar elde edilmiştir. Ayrıca tez kapsamında önerilen algoritmalar ve Kalman filtresi kullanılarak oluşturulan görsel seyir sisteminin başarımı, literatürde gezinim algoritmalarının testinde kullanılan iki farklı veri kümesi üzerinde gösterilmiştir.

Doktora tez çalışmamın yürütülmesi esnasında, çalışmalarına yön veren, bilgi ve yardımlarını esirgemeyen ve bana her türlü desteği sağlayan danışman hocam, Sayın Yrd. Doç. Dr. Fuat KARAKAYA'ya en içten teşekkürlerimi sunarım.

Çalışmam sırasındaki katkıları ve yön verici yorumlarından dolayı değerli hocalarım Sayın Prof. Dr. Tülay YILDIRIM, Prof. Dr. Nurettin ACIR, Doç. Dr. Halis ALTUN ve Yrd. Doç. Dr. Bekir Sami TEZEKİCİ'ye teşekkürlerimi sunarım.

Bu tezi, bu günlere gelmemde maddi ve manevi her türlü desteğini esirgemeyen babam Erkan PEKER'e, annem Nuriye PEKER'e, kardeşime ve özellikle her zaman yanımda olan hayat arkadaşım Zennur PEKER'e ithaf ediyorum.

İÇİNDEKİLER

ÖZET	v
SUMMARY	vi
ÖN SÖZ	vii
İÇİNDEKİLER	viii
ÇİZELGELER DİZİNİ	x
ŞEKİLLER DİZİNİ	xi
SİMGE VE KISALTMALAR	xiv
BÖLÜM I GİRİŞ	1
1.1 Tezin Amaç ve Kapsamı	2
1.2 Eş Zamanlı Konum Belirleme	4
1.3 Tez Ana Hatları	8
BÖLÜM II EŞ ZAMANLI KONUMLAMA ALGILAYICILARI.....	9
2.1 Giriş	9
2.2 Ataletsel Gezinim Sistemleri (INS).....	10
2.3 Küresel Konum Belirleme Sistemi (GPS).....	10
2.4 Görsel Algılayıcı Sistemler.....	11
BÖLÜM III GÖRSEL YER İZİ TESPİTİ	13
3.1 Giriş	13
3.2 Yer İzi Tanımı	14
3.3 Yer İzi Belirleme Yöntemleri.....	15
3.3.1 Kenar ve Köşe Bulma Tabanlı Yöntemler.....	17
3.3.2 Bölge Bulma Tabanlı Yöntemler.....	22
BÖLÜM IV YER İZİ TANIMLAYICILAR	34
4.1 Giriş	34
4.2 Tanımlayıcılar	37
4.2.1 Yönelim Açısı Hesaplanan Yöntemler	37
4.2.2 Yönelim Açısı Hesaplanmayan Yöntemler	42
BÖLÜM V ÖNERİLEN YÖNTEMLER VE BENZETİM SONUÇLARI	46
5.1 Yer İzi Tespit Yöntemi	46
5.1.1 Temel piksel değer karşılaştırmaları (TPDK)	46

5.1.2 Benzetim Sonuçları	50
5.2 Dilimlere Ayrılmış Yoğunluk Histogramı	54
5.2.1 SIH Algoritmasında Parametre Seçimi	56
5.2.2 Benzetim İçin Kullanılan Veri Seti.....	57
5.2.3 Benzetim Sonuçları	59
5.3 Dilimlere Ayrılmış Yönelimlerin Histogramı	66
5.3.1 SOH Tanımlayıcılarının Elde Edilmesi.....	67
5.3.2 AMDF Kullanarak Tanımlayıcı Eşleştirilmesi.....	68
5.3.3 SOH Tanımlayıcısı için Parametre seçimi.....	70
5.3.4 Benzetim Sonuçları	72
BÖLÜM VI GÖRSEL EŞ ZAMANLI KONUMLAMA VE HARİTALAMA.....	79
6.1 Giriş	79
6.2 Görsel Konum Değişimi Kestirimi.....	79
6.2.1 Problem Tanımı	80
6.2.2 Kamera Konum Değişimi Hesaplama.....	83
6.2.3 Stereo Kamera Kullanımı ve Derinlik Bilgisi.....	85
6.2.4 Yanlış Yer İzi Eşleştirmeleri.....	87
6.3 Kalman Filtresi	90
6.3.1 Kalman Filtre Modeli	90
BÖLÜM VII GÖRSEL SEYİR SİSTEMİ BENZETİM SONUÇLARI.....	92
7.1 Giriş	92
7.2 Benzetimde Kullanılan Veri Kümeleri.....	92
7.3 Benzetim Sonuçları.....	96
BÖLÜM VIII SONUÇ VE ÖNERİLER.....	106
KAYNAKLAR	108
ÖZ GEÇMİŞ	122
TEZ ÇALIŞMASINDAN ÜRETİLEN ESERLER	123

ÇİZELGELER DİZİNİ

Çizelge 5.1. Yer izi tespit benzetimlerinde kullanılan <i>Oxford</i> veri kümesinde yer alan görüntü kümeleri.....	51
Çizelge 5.2. SIH algoritmasının farklı dilim açıları ve histogram boyutları için eşleştirme performans matrisi.....	57
Çizelge 5.3. Tek bir görüntü bölgesinin yönelim açısı hesabı için gerekli olan ortalama süre değerleri.....	66
Çizelge 5.4. Algoritmaların iki görüntü arasındaki eşleştirmeleri hesaplamadaki ortalama süre ve hafıza karşılaştırmaları	78
Çizelge 7.1. <i>Kitti-00</i> senaryosundan elde edilen x yönündeki ortalama hatalar.....	99
Çizelge 7.2. <i>Kitti-00</i> senaryosundan elde edilen y yönündeki ortalama hatalar.....	100
Çizelge 7.3. <i>Kitti-01</i> senaryosundan elde edilen x yönündeki ortalama hatalar.....	103
Çizelge 7.4. <i>Kitti-01</i> senaryosundan elde edilen y yönündeki ortalama hatalar.....	103
Çizelge 7.5. <i>Kitti-05</i> senaryosundan elde edilen x yönündeki ortalama hatalar.....	105
Çizelge 7.6. <i>Kitti-05</i> senaryosundan elde edilen y yönündeki ortalama hatalar.....	105

ŞEKİLLER DİZİNİ

Şekil 1.1. Eş zamanlı konum belirleme sistemi genel şeması	3
Şekil 1.2. Bu çalışmada incelenecek konunun genel sistem içindeki yeri.....	3
Şekil 1.3. Tez kapsamında geliştirilecek sistem şeması.....	4
Şekil 2.1. INS yapısı	10
Şekil 2.2. Üçgenleme metodu gösterimi	11
Şekil 3.1. Görsel tanımlamada kenar, köşe, birleşim noktaları (Biederman, 1987)	14
Şekil 3.2. Kenar veya köşelerin ayırt ediciliklerinin gösterilmesi.....	16
Şekil 3.3. Öz değerlere göre grafiksel kenar ve köşe gösterimi	19
Şekil 3.4. Harris köşe bulma algoritması kullanılarak bulunan ilgi noktaları	20
Şekil 3.5. Shi-Tomasi tarafından önerilen algoritmanın kenar ve köşe ayırma grafiği	21
Şekil 3.6. Shi-Tomasi köşe bulma algoritması kullanılarak bulunan ilgi noktaları	21
Şekil 3.7. FAST algoritmasında kullanılan karar ağacı (Rosten ve Drummond, 2006) .	22
Şekil 3.8. MSER bölgelerini belirlemek için kullanılan eşiklenmiş görüntü dizisi.....	23
Şekil 3.9. Görüntü bölgesinin sınırlarının belirlenmesi (Mikolajczyk vd., 2005)	24
Şekil 3.10. MSER bölgeleri a) geniş güzel yapı b) eliptik yapı (Mikolajczyk vd., 2005)	25
Şekil 3.11. Giriş sinyalinin köşe tepkisi.....	25
Şekil 3.12. Sinyal üzerinde bir bölgenin Gauss tepkisinden elde edilen çekirdek birleşimi	26
Şekil 3.13. Laplace kullanılarak konvolüsyon uygulanan sinyalden elde edilen sistem tepkisi ($\sigma = 1$)	27
Şekil 3.14. Farklı σ değerleri için normalize edilmemiş ve normalize edilmiş sistem tepkisi ve ölçek seçimi	28
Şekil 3.15. Gauss ölçek uzayı	29
Şekil 3.16. LoG ve DoG sistem tepkisi	29
Şekil 3.17. DoG boyut uzayı.....	30
Şekil 3.18. DoG hesaplama rutini (Lowe, 2004)	30
Şekil 3.19. Elde edilen DoG görüntüleri.....	30
Şekil 3.20. İlgi noktalarının tespit edilmesi (Lowe, 2004)	31
Şekil 4.1. GLOH tanımlayıcı gösterimi (Mikolajczyk ve Schmid, 2005)	35

Şekil 4.2. Görüntü bölgesinin bölümlenmesi ve histogram çıkarımı	39
Şekil 4.3. FREAK algoritmasında kullanılan örnekleme bölgeleri (Alahi vd., 2012)....	41
Şekil 4.4. Taylor yönelim hesaplama algoritmasında kullanılan a) çekirdek düzeni b) vektörlerin toplanması	42
Şekil 4.5. Yerel koordinat sistemi	43
Şekil 4.7. Destek bölgeleri ve normalize edilmiş görüntü bölgeleri (Fan vd., 2011)	44
Şekil 4.8. Piksel değer gruplandırması (Fan vd., 2011).....	44
Şekil 5.1. Bir boyutlu (1D) sinyalin yerel minimum ve maksimum noktaları	47
Şekil 5.2. İki boyutlu görüntünün üç boyutlu olarak gösterimi	47
Şekil 5.3. TPKD yöntemin akış şeması	48
Şekil 5.4. Minimum/maksimum noktalarının bulunduğu piksel yerleşimleri	49
Şekil 5.5. TPKD ile bulunan örnek öznitelik noktaları.....	49
Şekil 5.6. Eşleştirilen öznitelik grafikleri	52
Şekil 5.7. TPKD ve SIFT'in eşleştirme performanslarının Google Haritalardan elde edilen görüntüler üzerindeki test sonuçları.....	53
Şekil 5.8. Görüntü parçasından alınan dilim parametreleri	54
Şekil 5.9. SIH algoritmasına ait işlem adımları.....	56
Şekil 5.10. Benzetim için kullanılan <i>Oxford</i> veri kümesinde yer alan görüntü kümeleri	58
Şekil 5.11. Eliptik görüntü bölgesinin normalize edilmesi.....	58
Şekil 5.12. Kullanılan görüntü kümelerine ait performans eğrileri	65
Şekil 5.13. Görüntü parçasının dilimlenmesi.....	67
Şekil 5.14. SOH tanımlayıcısının oluşturulması	68
Şekil 5.15. AMDF ve L2-norm uzaklığı kullanılması durumunda elde edilen performans farkı.....	69
Şekil 5.16. SOH tanımlayıcısında açısız değişim durumunda oluşan kayma (a) Referans görüntü parçası (Görüntü 1) ve SOH tanımlayıcısı (Dilim sayısı 20 ve her dilimdeki histogram boyutu 1 olarak seçilmiştir.), (b) 60 derece döndürülmüş görüntü parçası (Görüntü 2) ve SOH tanımlayıcısı, (c) Görüntü 1 ve Görüntü 2 arasındaki kayma düzeltildiğinde oluşan durum.....	70
Şekil 5.17. Değişik parametreler için SOH tanımlayıcı performansları (a) Dilim sayısı sabit ve 18 iken histogram boyutunun performansa etkisi, (b) Histogram boyutu sabit ve 8 iken dilim sayısının performansa etkisi.	71
Şekil 5.18. Yöntemlere ait performans eğrileri.....	77
Şekil 6.1. Kamera hareketinden konum hesaplama.....	79

Şekil 6.2. Stereo kamera yapılandırması	86
Şekil 6.3. Veri kümesine doğru uydurma problemi a) uydurulan doğrunun veri kümesinde çok az veriyi işaret etmesi durumu b) uydurulan doğrunun veri kümesinde çoğu veri için doğru olması fakat en iyi çözüm olmaması durumu c) veri kümesini en doğru temsil eden doğrunun bulunduğu durum.....	88
Şekil 6.4. Yanlış eşleştirmeye neden olan yer izlerinin belirlenmesi a-b) giriş görüntüleri c-d) ilk eşleştirilen yer izleri e-f) RANSAC kullanılarak elde edilen doğru eşleştirme sağlayan yer izleri	89
Şekil 7.1. <i>New College</i> veri kümesinin güzergahı (Smith vd., 2009).....	92
Şekil 7.2. <i>New College</i> veri kümesinden örnek görüntü kareleri a) Sol kamera görüntüleri b) Sağ kamera görüntüleri (Smith vd., 2009).....	93
Şekil 7.3. <i>Kitti</i> Veri kümesinin toplandığı platform (Geiger vd., 2012)	93
Şekil 7.4. <i>Kitti</i> veri kümesinden örnek görüntü kareleri a) Sol kamera görüntüleri b) Sağ kamera görüntüleri (Geiger vd., 2012)	94
Şekil 7.5. <i>Kitti</i> veri kümesinde kullanılan farklı senaryolar için gerçek gezinim rotaları	95
Şekil 7.6. <i>New College</i> veri kümesi 1. bölgesinden elde edilen sonuçlar a) SIFT kullanılarak elde edilen sonuçlar b) önerilen yöntemler kullanılarak elde edilen sonuçlar	96
Şekil 7.7. <i>New College</i> veri kümesinden elde edilen konum bilgileri	97
Şekil 7.8. <i>Kitti-00</i> senaryosundan elde edilen konum bilgileri.....	98
Şekil 7.9. <i>Kitti-00</i> senaryosundan elde edilen hata grafikleri a) x yönündeki hatalar b) y yönündeki hatalar.....	99
Şekil 7.10. <i>Kitti-01</i> senaryosundan elde edilen konum bilgileri.....	100
Şekil 7.11. <i>Kitti-01</i> senaryosundan örnek görüntü kareleri (Geiger vd., 2012).....	101
Şekil 7.12. <i>Kitti-01</i> senaryosundan elde edilen hata grafikleri a) x yönündeki hatalar b) y yönündeki hatalar.....	102
Şekil 7.13. <i>Kitti-05</i> senaryosundan elde edilen konum bilgileri.....	103
Şekil 7.14. <i>Kitti-05</i> senaryosundan elde edilen hata grafikleri a) x yönündeki hatalar b) y yönündeki hatalar.....	104

SİMGE VE KISALTMALAR

Simgeler

σ

Açıklama

Standart sapma

Kısaltmalar

Açıklama

BRIEF

İkili Gürbüz Bağımsız Temel Öznitelikler

BRISK

İkili Gürbüz Değişmez Ölçeklenebilir Öznitelikler

DoB

Kutu Filtrelerinin Farkı

DoG

Gaussların Farkı

DoH

Hessianların Determinantı

EZKB

Eş Zamanlı Konum Belirleme

EZKH

Eş Zamanlı Konumlama ve Haritalama

FAST

Hızlandırılmış Bölüt Testlerinden elde edilen Öznitelikler

FPGA

Alan Programlanabilir Kapı Dizisi

FREAK

Hızlı Retina Öznitelikleri

GLOH

Eğim Konumu ve Yönelim Histogramı

GKF

Genişletilmiş Kalman Filtresi

GPS

Küresel Konum Belirleme Sistemi

GSS

Görsel Seyir Sistemi

HOG

Yönelimli Gradyanların Histogramı

IMU

Ataletsel Ölçüm Birimi

INS

Ataletsel Gezinim Sistemi

İHA

İnsansız Hava Aracı

LIOP

Yerel Piksel Değeri Gruplandırması

LoG

Gaussların Laplası

MROGH

Destek Bölgelerinin Yönelim Histogramı

MSER

Maksimum Kararlılıktaki Uç Bölgeler

ORB

Yönelimli FAST ve Döndürülmüş BRIEF

PCA

Temel Bileşenler Analizi

RANSAC

Rastgele Örnek Uyuşumu

RGB-D

Kırmızı Yeşil Mavi – Derinlik

SIFT

Ölçek Bağımsız Öznitelik Dönüşümü

SURF

Hızlandırılmış Gürbüz Öznitelikler



BÖLÜM I

GİRİŞ

Son dönemde insansız araçlar üzerindeki çalışmalar büyük önem kazanmıştır. Araç içinde yer alan sürücü veya pilot, insan fizyolojisine bağlı fiziksel sınırlamalardan dolayı, hem sağlık riskini hem de operasyon maliyetini arttırmaktadır. Son zamanlarda yapılan çalışmalarda, araçlardan insan faktörünün uzaklaştırılması ile hem risklerin en aza indirilmesine hem de araç tasarımlarındaki fiziksel sınırların aşılmasına olanak sağlanmıştır. Dolayısıyla askeri ve sivil alanlardaki arama/kurtarma, gözlem, haritalama gibi görevlerde insansız araçların kullanımı da gün geçtikçe artmaktadır.

İnsansız araçlar arasında üzerinde en çok çalışılan İnsansız Hava Araçlarıdır (İHA). Literatürde insansız hava araçlarına yönelik birçok çalışma mevcuttur. Hava araçlarına uzaktan kumandalı hava araçları, robot hava araçları, pilotsuz uçaklar gibi farklı isimlendirmeler de yapılmıştır. İnsansız hava araçları bir yer istasyonundan kontrol edilebildiği gibi uzaktan kontrolün mümkün olmadığı durumlarda da bağımsız bir şekilde uçuşuna devam edebilmelidir. Aracın harici kontrolcüden bağımsız bir şekilde uçuşuna devam edebilmesi için konum, hız, rota gibi birçok bilginin anlık olarak takip edilmesi gerekir. Anlık bilgi takibi için kullanılmış birçok algılayıcı mevcuttur. Küresel konum belirleme sistemi (GPS), ataletsel ölçüm birimi (IMU), ataletsel gezinim sistemi (INS), kamera, radar vb. bu algılayıcılardan bazılarıdır. Algılayıcılar aracın rotasında gitmesi ve rotasını tayin etmesi için doğrudan etkilidir. Dolayısıyla sistemin başarımı da bu algılayıcıların başarımına bağlıdır. Literatürde, algılayıcılar kullanılarak konumlama yapılması Eş Zamanlı Konumlama ve Haritalama (EZKH) veya Eş Zamanlı Konum Belirleme (EZKB) problemi olarak isimlendirilmiştir (Leonard ve Durrant-Whyte, 1991).

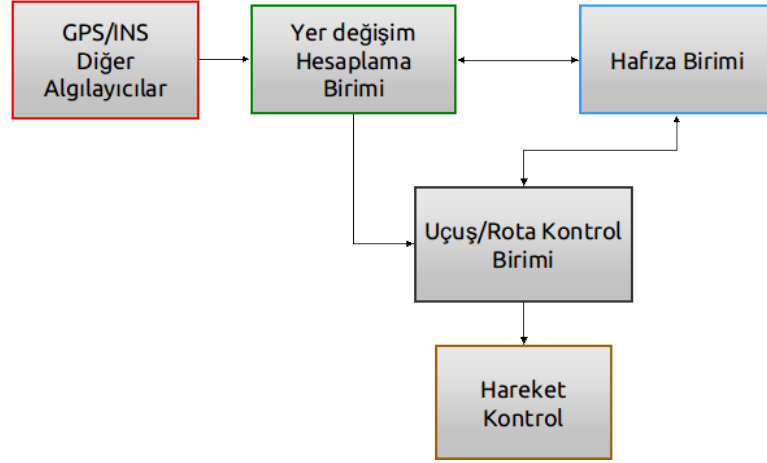
İnsansız araçlar için eş zamanlı konumlama ve haritalama problemi, başlangıç noktası bilinen bir aracın hareketine devam ederken, hareketin her noktasında algılayıcılardan veri toplaması ile başlar. EZKB ile hedeflenen, algılayıcılardan toplanan verilerin işlenerek, aracın içinde bulunduğu ortamın haritasının çıkarılması ve eş zamanlı olarak aracın haritada nerede konumlandığının takip edilmesidir.

Tez kapsamında, küresel konum belirleme sisteminin çalışmasının mümkün olmadığı durumlar (arazi koşulları, sinyallerin engellenmesi vb.) için, kamera görüntüleri kullanılarak konum değişimi hesaplanacaktır. Bu gibi sistemler görsel algılayıcı sistemler olarak adlandırılır. Tez çalışmasında, görsel algılayıcılar kullanılarak alınan görüntüler üzerinden yer izi noktalarının belirlenmesi ve yer izlerinin eşleştirilmesine yönelik gürbüz, kararlı ve başarıımı yüksek algoritmalar sunulacaktır.

1.1 Tezin Amaç ve Kapsamı

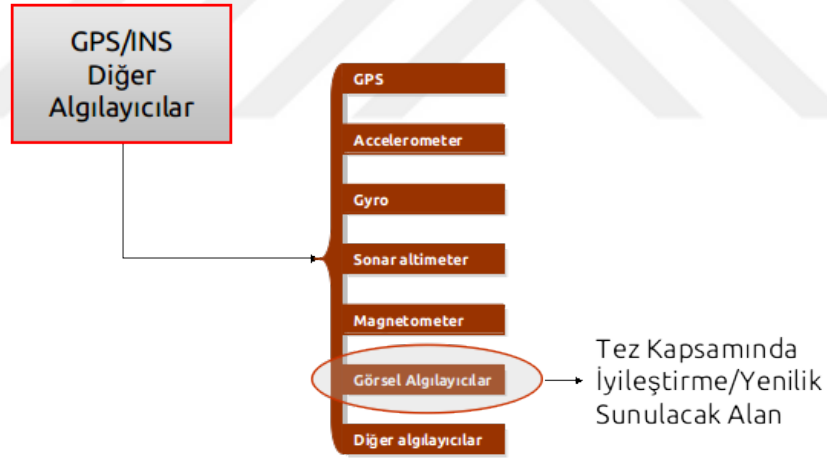
İnsansız araçlarda, yer izi çıkarımı için lazer ve radar tarayıcılar, sonar sistemleri gibi algılayıcıların maliyetlerinin yüksek olduğu ve yüksek irtifa içeren görüntülerde kullanılmasının etkin olmadığı literatürde yapılan çalışmalarda gösterilmiştir (Templeton vd., 2007). Bu nedenle maliyetin düşürülmesi ve yukarıda sayılan sistemlerin işlevsel olmadığı durumlarda görüntü işleme tabanlı yaklaşımlar aracın konumunun belirlenmesinde büyük önem taşımaktadır. Yer izi çıkarımında tarayıcıların yanı sıra ataletsel algılayıcılar (İvmeölçer, denge çarkı vb.) da kullanılmaktadır. Ataletsel algılayıcıların başarımları, GPS'in olmadığı durumlarda yeterli değildir. Bu nedenle insansız araçlar yüksek hızlarda seyrederken, GPS sinyalinde yaşanacak kesinti durumunda aracın rota dışına çıkmasının engellenmesi gerekmektedir. Ayrıca ataletsel algılayıcılardan elde edilen verilerin içerdiği gürültüden kaynaklanan artımlı hatalar konum kestiriminde büyük sapmalara sebep olmaktadır (Nemra ve Aouf, 2010).

Görüntü tabanlı yaklaşımlarda yer izi çıkarımı ve yer izlerinin belirleneceği noktaların bulunması için kullanılan özniteliklerin tespiti gibi konularda genellikle, ölçekten bağımsız öznitelik dönüşümü (SIFT), gürbüz özniteliklerin hızlandırılması (SURF), stereo görüntü işleme ve görsel akış (Optical-flow) çıkarımı gibi yöntemler kullanılmaktadır. Görüntü temelli sistemlerde, yer izlerinin çıkarılması ve daha önceden tanımlanmış bir yer izinin yeniden tanımlanması için kullanılacak olan öznitelik çıkarım algoritmalarının, araçların hızlı hareketleri nedeniyle görüntüde oluşacak olan büyük ışık değişimleri ve gürültü gibi etmenler ile başa çıkabilmeleri gerekmektedir. EZKB sistemi için genel sistem şeması Şekil 1.1'de verilmiştir.



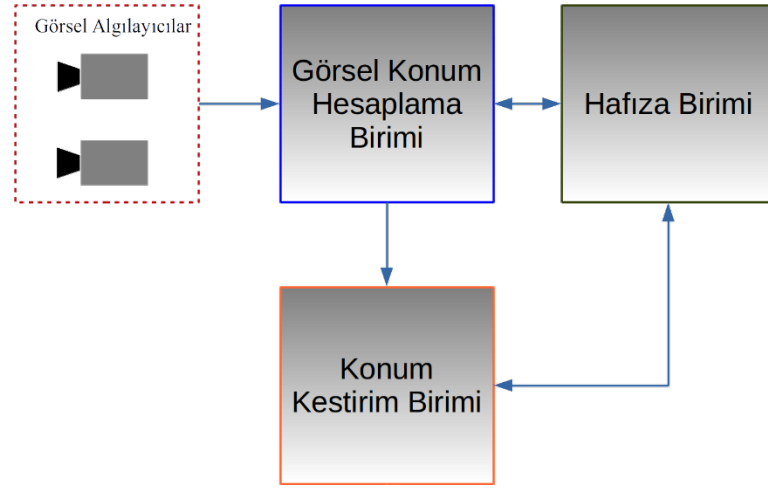
Şekil 1.1. Eş zamanlı konum belirleme sistemi genel şeması

Burada yer değiştirmenin hesaplanması için literatürde kullanılan algılayıcılar ve tez kapsamında ilgilenilen algılayıcılar Şekil 1.2’de gösterilmiştir. Tez kapsamında, sadece görsel algılayıcıların kullanıldığı EZKB sistemleri incelenmiştir. Önerilen yöntemler ile konum kestirim başarımı yüksek bir sistem gerçekleştirilmesi amaçlanmaktadır.



Şekil 1.2. Bu çalışmada incelenecek konunun genel sistem içindeki yeri

Tez kapsamında önerilen yöntemler kullanılarak Şekil 1.3’te gösterilen şemaya sahip bir sistem oluşturulması hedeflenmiştir.



Şekil 1.3. Tez kapsamında geliştirilecek sistem şeması

1.2 Eş Zamanlı Konum Belirleme

Son zamanlarda sivil ve askeri alanlarda otonom sistemlere olan ihtiyacın artması ile birlikte literatürde otonom (özerk) hareket edebilen araçlar konusundaki araştırmalar büyük bir ivme kazanmıştır. Bu çalışmalar; kapalı mekân (Bachrach vd., 2012; H. Chen vd., 2010; Hwang ve Song, 2011; Kuai vd., 2010; Tomic vd., 2012), dış mekan (Bailey, 2002; Kok vd., 2013; Paz vd., 2008; Schleicher vd., 2009; Semsch vd., 2009; Tisdale vd., 2009), su altı (Aulinas vd., 2010; Beall vd., 2010; Inoue vd., 2010; Kim ve Eustice, 2013) vb. bölgelerde keşif, gözlem, askeri görevler ve haritalama gibi çeşitli uygulama alanlarını içermektedir.

Araçların otonom hareket edebilmesine yönelik çalışmalar temel olarak eş zamanlı konum belirleme sistemlerini kapsamaktadır. Araçlar için eş zamanlı konum belirleme (EZKB) problemi, genel hareket planlama probleminin özel bir alanıdır (Goerzen vd., 2010). Tuna vd. (2012) çalışmasına göre odometri ve çeşitli harici algılayıcı verileri kullanılarak aracın içinde bulunduğu çevrenin sayısal ortamda yapılandırılması ve aracın bu ortamda konumlandırılmasını sağlayan sistemler Eş Zamanlı Konum Belirleme sistemleri olarak tanımlanmıştır. EZKB sistemleri için çözülmesi gereken önemli problemler şu şekilde sayılabilir: yapısı bilinmeyen ortamlarda konum ve rota belirlenmesi, hareket kontrolü (Besada-Portas vd., 2013).

EZKB sistemleri, algılayıcılar kullanılarak elde edilen bilgilerden yer değiştirme ve konum bilgilerinin elde edilmesini hedefler. Çevresel etkenler kullanılan algılayıcıların

(Küresel konum belirleme sistemleri - GPS, lazer tarama sistemleri, kameralar, atalet ölçüm birimleri vb.) çalışmalarını olumsuz etkileyebilmektedir. Örnek olarak yetersiz ışık şartlarında kameradan alınan bilgi yetersiz olabileceği gibi (Bachrach vd., 2012; Tomic vd., 2012), kapalı mekanlarda küresel konum belirleme sistemleri çoğunlukla kullanılmaz hale gelmektedir (Bachrach vd., 2012). Ayrıca bu sistemlerdeki problemler aracın hareket edebileceği serbest düzlem sayısının artması, kullanılan algılayıcıların hassasiyetleri, çevresel koşullar, hız ve ivmelenme gibi kısıtlamalar ile birlikte çözülmesi oldukça zor bir problem haline gelmektedir. Bu nedenle oluşan problem uzayı genişletmekte ve bu derece geniş bir problem uzayının çözümü için literatürde önerilmiş herhangi bir analitik algoritma bulunmamaktadır (Goerzen vd., 2010). Problemlerin çözümü için geleneksel kontrol sistemlerinin akıllı bileşenler kullanılarak geliştirilmesi gerekmektedir (Antsaklis vd., 1991).

EZKB kullanılarak kontrol edilmek istenen araçlar için, problem uzayının geniş olması ve çözümünün belirsiz olması nedeniyle araştırmacılar tarafından konu hakkında yaklaşımlar önerilmiştir. Bu yaklaşımlar Kalman filtresi (Grzonka vd., 2012), genişletilmiş Kalman filtresi (Huang vd., 2008; Huang ve Dissanayake, 2007; Kallapur ve Anavatti, 2006; Mao vd., 2007; Meng vd., 2010; Wang vd., 2012; Williams ve Crump, 2012), genişletilmiş bilgi filtresi (Bryson ve Sukkarieh, 2007; Liu ve Thrun, 2003; Li ve Aouf, 2012; Mao vd., 2007; Thrun vd., 2004; Wang ve Ma, 2011), kokusuz kalman filtresi (Sunderhauf vd., 2007; Jianli vd., 2011), parçacık filtresi (Lee ve Lee, 2009; Soleimani vd., 2010; Zang vd., 2006), Rao-Blackwellized parçacık filtresi (Elinas vd., 2006; Havangi vd., 2011; Wang vd., 2008) vb. olarak göze çarpmaktadır.

Smith vd. (1988; 1986) çalışmalarında, algılayıcı verileri ve yer izleri arasında geometrik bağıntı olduğunu belirtmiştir. Smith (1990) mobil bir robotun çevreden topladığı nesne bilgilerini kullanarak GKF tabanlı bir yaklaşım ile bilinmeyen bir ortamda gezinebileceğini göstermiştir. Bu çalışma ile günümüzde kullanılan GKF tabanlı EZKB çalışmalarının temeli oluşturulmuştur. GKF tabanlı EZKB sistemlerinde robotun pozisyonu ve tespit edilen nesnelerin (yer izleri) konumlarının birleşiminden oluşan bir durum vektörü kullanılmaktadır. Robot, üzerindeki algılayıcıları kullanarak çevresindeki nesneleri tespit edebilmektedir. Tespit edilen her bir nesne, daha önce tespit edilen diğer nesneler ile karşılaştırılmakta ve yeni nesne tespit edildiği zaman durum vektörünün güncellenmesi sağlanmaktadır.

Diğer bir çalışmada (Williams vd., 2002), GKF tabanlı yaklaşımdan kaynaklanan bir takım problemlere dikkat çekilmiştir. En büyük problemlerden bir tanesi ortamın taranmasına bağlı olarak artan durum parametrelerinin (yer izleri ve konum bilgileri) hesaplama yükünü çok hızlı bir şekilde arttırması ve buna bağlı olarak gerçek zamanlı çalışmanın mümkün olmaması ve kestirim başarımının olumsuz etkilenmesidir.

İnsansız hava araçları için tanımlanan dinamik hareket modelinin diğer ortam araçlarından daha karmaşık olması, oluşacak gürültü etkilerinin fazla olması ve araçların yüksek hızlarda seyretmesinden dolayı daha büyük sistematik problemler ortaya çıkmaktadır. Otonom hava araçları için genel olarak küresel konum belirleme sistemi ve ataletsel ölçüm birimlerini (IMU) ihtiva eden yapılar kullanılmaktadır. Küresel konum belirleme sisteminin verileri kullanılarak, ataletsel algılayıcılardan elde edilen veriler ile yapılan EZKB sistemindeki kestirimlerden kaynaklanan hatalar düzeltilebilmektedir (Rady vd., 2011). Yer izi çıkarımında, ataletsel algılayıcılardan kaynaklanan hataların en aza indirilmesi için ataletsel algılayıcılara ek olarak kamera, lazer, sonar, radar gibi algılayıcılarda kullanılmaktadır (Sergi vd., 2003; Tardós vd., 2002; Welle vd., 2010; Wu vd., 2013).

Araştırmacılar tarafından, otonom araçlar için çok sayıda görüntü temelli EZKB sistemi önerilmiştir. Bu gibi sistemlerde genel olarak kameradan alınan görüntü üzerinden yer izi bilgileri çıkarılmakta ve bu bilgilerin video kareleri arasındaki yer değiştirmeleri incelenmekte ve buna bağlı olarak aracın yer değiştirme bilgisi kestirilmeye çalışılmaktadır. Görüntü temelli sistemler, ataletsel gezinim sistemleri (INS) ile birlikte (Nemra ve Aouf, 2010; Xu vd., 2009) veya tek başlarına (Lui ve Jarvis, 2010) EZKB işlemi için kullanılabilir.

Bachrach vd. (2012) tarafından gerçekleştirilen çalışmada, RGB-D kamera ve IMU algılayıcıları kullanılarak mikro hava aracının karmaşık bir ortamda ve GPS verisinin engellendiği durumda, kontrolü gerçekleştirilmiştir. Çalışmada kullanılan RGB-D kamera, bünyesindeki lazer algılayıcı vasıtası ile derinlik bilgisine sahip renkli görüntü oluşturabilmektedir. RGB-D kameradan gelen görüntüler öncelikle bir ön işleme tabi tutulmaktadır. Ön işlem adımında giriş görüntüsü gri seviyeye çekilmekte, ardından sigma değeri 0.85 olan Gauss çekirdeği kullanılarak yumuşatma işlemi gerçekleştirilmektedir. Ayrıca ölçekten bağımsız yer izleri elde edebilmek için Gauss

piramitleri oluşturulmaktadır. Bu adımda Rosten ve Drummond (2006) tarafından önerilmiş olan hızlı öznitelik çıkarım algoritması kullanılarak her bir görüntü karesindeki önemli noktalar belirlenmektedir. Belirlenen bu noktalar derinlik bilgisi ile karşılaştırılmakta ve uyuşmanın olmadığı öznitelikler ihmal edilmektedir. Daha sonra bu adımlar başlangıç pozisyonunun kestirimi, özniteliklerin eşleştirilmesi ve hareket kestirimi ile sonuçlandırılmaktadır.

Tomic vd. (2012) tarafından gerçekleştirilen çalışmada IMU, lazer tarayıcı ve stereo kamera algılayıcıları birlikte kullanılmıştır. Böylece aracın hem iç hem de dış ortamlarda uçabilmesi amaçlanmıştır. Li vd. (2012b) tarafından yapılan çalışmada ise stereo kamera ve genişletilmiş bilgi filtresi kullanan EZKB sistemi önerilmiştir. Ayrıca Li vd. (2012a), boyuttan (Ölçekten) bağımsız öz-nitelik dönüşümü, gürbüz öz-niteliklerin hızlandırılması, rastgele örnek uyuşumu (RANSAC) gibi görüntü tabanlı yaklaşımları genişletilmiş bilgi filtresi, kokusuz (Unscented) Kalman filtresi, genişletilmiş Kalman filtresi gibi algoritmalar ile birlikte kullanarak gerçekleştirilen EZKB sisteminin başarımını irdelemiştir. Sunderhauf vd. (2007) ise tek kamera ve kokusuz Kalman filtresi kullanan bir sistem önermişlerdir.

Rady vd. (2011), GPS sinyalinin engellendiği durumlar için görüntü tabanlı konum belirleme sistemi önermişlerdir. Bu çalışmada önerilen EZKB sistemi iki fazdan oluşmaktadır. İlk fazda genel harita oluşturulmakta, ikinci fazda ise araç genel harita içerisinde konumlandırılmaktadır.

Angeli vd. (2006) sabit irtifada kabul edilen bir modelin eş zamanlı konumlandırılması için 2 boyutlu EZKB sistemi önermişlerdir. Ayrıca bu çalışmada, aracın daha önce geçtiği bir noktadan tekrar geçtiği tespit edilerek kestirim hatalarının azaltılması amaçlanmıştır.

İncelenen çalışmalarda görüldüğü üzere kestirim işlemi Kalman filtreleri, bilgi filtreleri, parçacık filtreleri, H-sonsuz filtresi tabanlı algoritmalar kullanılarak gerçekleştirilmektedir. Fakat bu algoritmaların kullandıkları parametre sayısına bağlı olarak ortaya çıkan işlem yükü problem oluşturmaktadır. Bu problemin çözümü için işlem yükü, alt haritalar metodu, parçalanmış durumların güncellenmesi, seyreltilmiş genişletilmiş bilgi filtresi gibi yaklaşımlar kullanılarak dağıtılmıştır. Görüntü işleme

tabanlı EZKB sistemlerinde, genel olarak çoklu kamera kullanımı gibi yaklaşımlar sayesinde daha doğru sonuç alınması amaçlanmaktadır. Fakat görüntü tabanlı EZKB sistemleri çok yoğun işlem yükü gerektirmektedir. Bu nedenle GPU, FPGA gibi paralel işlem yapabilme kabiliyetine sahip birimler kullanılarak gerçek zamanlı uygulamalar elde edilmesi amaçlanmıştır (Bleiweiss, 2010; Kok vd., 2013).

1.3 Tez Ana Hatları

Tezin sonraki bölümleri şu şekilde organize edilmiştir; 2. bölümde EZKB için insansız araçlarda kullanılan algılayıcı sistemleri, 3. bölümde görsel seyir sistemlerinde kullanılan yer izi tespit yöntemleri, 4. bölümde, literatürdeki, yer izi nokta veya bölgelerinden tanımlayıcı çıkarım yöntemleri, 5. bölümde ise tez kapsamında önerilen yöntemler ve benzetim sonuçları verilmiştir. 6. bölümde EZKB problemi için görsel seyir sistem modeli ve görsel seyir sistemlerinde kullanılan diğer görüntü işleme teknikleri incelenmiştir. 7. Bölümde ise oluşturulan görsel seyir sisteminin literatürdeki veri kümeleri ile test edilmesi sonucu elde edilen çıktılar irdelenmiştir.

BÖLÜM II

EŞ ZAMANLI KONUMLAMA ALGILAYICILARI

2.1 Giriş

Bu bölümde insansız araçlarda yaygın olarak kullanılan genel algılayıcılar ataletsel gezinim sistemi (Inertial Navigation System - INS), küresel konum belirleme sistemi (GPS) ve görsel algılayıcı sistemler başlıkları altında incelenmiştir.

Otonom araçlarda kullanılan gezinim algılayıcıları iki temel başlık altında incelenebilir. Bunlar ataletsel (Dead-reckoning) gezinim ve mutlak gezinim algılayıcıları olarak isimlendirilebilir. Ataletsel gezinim algılayıcılarında IMU adı verilen ataletsel ölçüm birimleri kullanılarak aracın ivme ve açısal hız bilgileri elde edilmektedir. Bu bilgiler kullanılarak aracın yer değiştirmesi hesaplanabilir. Her ne kadar bu algılayıcılar yüksek sayıda veri örnekleyebilse de, verilerin gürültü içeriği nedeniyle filtrelenmesi gerekmektedir. Elde edilen veriler filtrelenmesine rağmen gürültü içeriği, yer değiştirme bilgilerinde sürekli olarak bir kaymaya neden olmaktadır. Bu nedenle araştırmacılar mutlak algılayıcılar kullanarak bu kaymayı en aza indirmeye çalışmıştır.

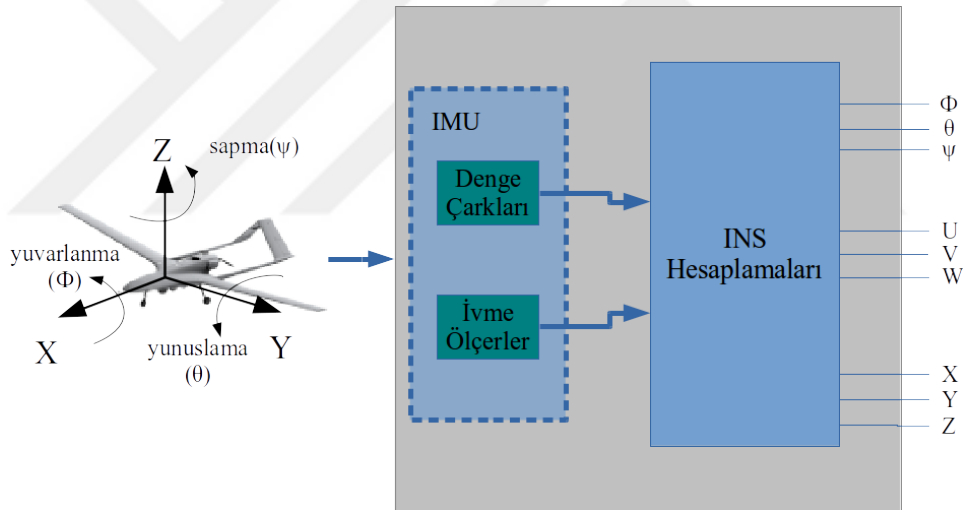
Mutlak algılayıcılara örnek olarak küresel konum belirleme sistemleri (GPS) gösterilebilir. IMU verilerinde oluşan kaymanın, belirli zaman aralıklarında toplanan, GPS verileri ile doğrultulması sonucu elde edilen karma veriler araştırmacılar tarafından oldukça ilgi görmektedir. Bu şekilde elde edilen konumlandırma bilgileri uzun soluklu ve küresel çapta uçuşlara olanak sağlamaktadır. Fakat GPS sinyalleri bazı durumlarda gerek iç, gerekse dış etkenlerden dolayı kesintiye uğrayabilmektedir. Bu durumlarda ise görsel algılayıcılar, GPS algılayıcılarına alternatif bir çözüm sunabilmektedir.

Son on yıl içerisinde, görsel algılayıcılar kullanılarak yer değiştirme hesaplanması çok önemli bir konu haline gelmiştir. Görsel algılayıcıların temel bileşeni olan kamera maliyetinin diğer algılayıcılara göre daha efektif olmasının yanı sıra kameralardan alınan görüntüler haritalama açısından çok önemli olan çevresel verileri görsel olarak sağlamaktadır.

2.2 Ataletsel Gezinim Sistemleri (INS)

Ataletsel gezinim sistemleri, ivmeölçerler (Accelerometer) ve denge çarkları (Gyroscope) gibi birden fazla algılayıcı barındırmaktadır ve bu algılayıcılardan elde edilen veriler kaynaştırılarak aracın anlık konum ve hız bilgisinde meydana gelen kaymalar en aza indirgenmektedir.

INS oluşturulabilmesi için algılayıcılardan toplanan verilerin kaynaştırılarak yer değiştirme problemine uygun hale getirilmesi gerekir. Şekil 2.1’de ivmeölçer ve denge çarkından oluşan basit bir INS yapısı verilmiştir. Burada (x, y, z) eksenlerine ait ivme verileri ve dönme oranları yüksek örnekleme hızında elde edilir. Bu veriler daha sonra işlenerek pozisyon verisine (X, Y, Z), hıza (U, V, W) ve durum (Φ , θ , ψ) bilgisine dönüştürülür.



Şekil 2.1. INS yapısı

2.3 Küresel Konum Belirleme Sistemi (GPS)

Küresel konum belirleme sistemi, konumları bilinen uydulardan yapılan yayınları toplanmakta ve bu veriler kullanılarak kendi konumunu hesaplamaktadır. GPS, konumunu belirleyebilmek için en az 4 uydu sinyaline ihtiyaç duyar. Bu uydulardan gelen sinyaller uyduların konumlarını da içerir. Bu sayede Şekil 2.2’de gösterilen üçgenleme metodu kullanılarak vericinin konumu belirlenmektedir.



Şekil 2.2. Üçgenleme metodu gösterimi

GPS sinyalleri iki farklı frekans kullanılarak yayılır. Bunlar 1575.42 MHz (L1) ve 1227.6 MHz (L2) olarak belirlenmiştir. L1 frekansı askeri kullanım için ayrılmıştır. Sivil kullanım için ise L2 frekansı kullanılır. GPS, alıcıya gelen sinyallerin zayıflamasını ve uydunun konum bilgisinin güvenilirliğini göz önüne alarak yer kestirimi yapar. Bazı durumlarda oluşan konum hataları 15 metreye kadar çıkabilir.

2.4 Görsel Algılayıcı Sistemler

Görsel algılayıcılar diğer algılayıcılara göre daha fazla bilgi içerir. Görüntünün alındığı konuma ait görüntüler sayesinde çevresel farkındalık yüksektir. Bu nedenle elde edilen görüntüler hem konum kestirimi hem de gezinim yapılan bölgenin güncel haritasının oluşturulmasında kullanılabilir. Görsel algılayıcı olarak kullanılan kameraların farklı yapıya sahip olanları da bulunmaktadır. Kızılötesi ve termal kameralar daha az bilgi içermeleri nedeniyle gezinim sistemlerinde fazla kullanılmamaktadır. Dolayısıyla bu tez kapsamında bu tür kameralardan alınan görüntüler incelenmemiştir. Standart kameralar (İğne deliği kamera), insan görmesine yakın görüntü sunmaları, düşük maliyetleri ve daha fazla ayırt edici öznetelik sağlamaları nedeniyle görsel algılayıcı sistemlerde tercih sebebidir.

Standart kameralardan elde edilen görüntüler detaylı bilgi sunsalar da görüntü boyutlarından dolayı hafıza gereksinimini arttırmırlar. Büyük boyutlu verilerin işlenmesi aşamasında kullanılan algoritmalar büyük önem taşımaktadır. Görüntüde yer alan her

pikselin işlenmesi yerine ardışıl görüntüler arasında bağlantı olan noktaların belirlenmesi temeli ile yer değiştirme hesabı yapan algoritmalar işlem yükünü azaltmaktadır.

Yer değiştirme hesabı için ardışıl görüntülerden tespit edilen yer izleri birbirleri ile karşılaştırılarak görüntüler arası benzer noktaların konumları bulunur. Çoklu kamera kullanılarak elde edilen görüntülerde derinlik bilgisi de elde edilebildiğinden tek kamera kullanılmasına göre yer değiştirme hesabının başarımı artmaktadır. Ayrıca derinlik bilgileri kullanılarak nokta bulutları oluşturulabilir ve gezinim yapılan bölgenin 3 boyutlu haritası çıkarılabilir.



BÖLÜM III

GÖRSEL YER İZİ TESPİTİ

3.1 Giriş

Görsel Seyir Sistemlerinde (GSS) amaç, aracın konum değişim bilgisinin kameralardan alınan bilgiler kullanılarak hesaplanmasıdır. Görsel sistemler tek başlarına bir seyir sistemi oluşturmak için kullanılabilecekleri gibi GPS/INS ile birlikte yer değiştirme bilgisinin iyileştirilmesi amacı ile kullanılabilir.

GPS sinyalinin olmadığı veya engellendiği durumlarda ataletsel algılayıcılardan toplanan bilgilerdeki gürültü içeriği nedeni ile hesaplama hataları oluşmaktadır. GSS kullanılan uygulamalarda bu hatalar iyileştirilerek konum kestirimi açısından kabul edilebilir sonuçlar elde edilebilmektedir.

Görsel seyir sistemlerinde görüntüde yer alan yer izi noktalarının veya bölgelerinin elde edilmesi için kullanılan görüntü işleme teknikleri büyük önem arz etmektedir. Çünkü elde edilen yer izleri sonraki adımlar için girdi teşkil etmektedir. Örneğin elde edilen yer izleri kullanılarak tanımlayıcılar oluşturulur ve bu tanımlayıcılar yer izlerinin görüntüler arasında eşleştirilmesini sağlar. Görüntü düzleminde eşleştirilen yer izlerinin 3 boyutlu uzaydaki yer değiştirme miktarı üçgenleme metodu kullanılarak elde edilebilir. Bu sayede görsel algılayıcılardan alınan bilgiler konum değiştirme bilgisine dönüştürülmektedir. GSS’de kullanılan bu yöntemler genel olarak 4 ana başlık altında toplanır ve bunlar şu şekilde sıralanabilir;

1. Yer izlerinin (ilgi noktalarının) tespit edilmesi,
2. Yer izi tanımlayıcıların belirlenmesi,
3. Yer izlerinin eşleştirilmesi,
4. Konum değişiminin hesaplanması ve gerçek dünya koordinat değişiminin elde edilmesi.

Görsel yer izleri ile aracın içinde bulunduğu çevre etkin bir şekilde ifade edilebilir. Ayrıca yer izi kullanılarak gerçekleştirilen algoritmalar, görüntüdeki bütün piksellerin

kullanılmasına kıyasla daha az hafıza gerektirir, bu sayede aracın içinde bulunduğu çevre hakkında zengin bilgi içeren bir veri kümesi oluşturulur. Bu sayede geniş bölgelerden toplanan veri kümelerinde geriye dönük sorgulama yapılarak uzun süreli gezinim gerçekleştirilebilir. Yer izi çıkarımı için kullanılan yöntemler bu bölümde incelenecektir.

3.2 Yer İzi Tanımı

Yer izleri aynı zamanda ilgi noktaları veya iyi öznitelikler olarak adlandırılabilir. Yer izleri görüntü içerisinde bir piksel olabileceği gibi piksel bölgeleri olarak da tanımlanabilir. İlgi noktaları, ilk olarak Biederman (1987) tarafından yapılan çalışma ile tanımlanmıştır. Bu çalışmada insanların nesneleri tanımlamak ve birbirlerinden ayırt edebilmek için kenar, köşe ve birleşim noktaları gibi bilgileri kullandığı tespit edilmiştir. Ayrıca, Şekil 3.1’de görülen ilk kolondaki nesnelerin, sadece bazı kenar köşe bilgileri kullanılarak dahi tanımlanabildiği gösterilmiştir.



Şekil 3.1. Görsel tanımlamada kenar, köşe, birleşim noktaları (Biederman, 1987)

Ayrıca bulunan ilgi noktalarının başka görüntü karelerinde de bulunabilmesi, bu noktaların en önemli özelliklerinden biridir. Başka bir deyişle bir görüntü karesinde tespit edilen ilgi noktalarının gürültü, ışık ve ölçek değişimi, geometrik bozulmalar vb. durumlara karşı tekrar elde edilebilir olması gerekir. İlgi noktalarının görüntü karesinde oluşacak genel bozulmalardan etkilenmemesi için yerellik özelliğinin olması görüntüde oluşacak şekilsel bozulmalara karşı sistemin eşleştirme performansını olumlu olarak

etkilemektedir. Bulunan ilgi noktalarının tekrar edilebilir ve ayırt edici özellikte olmaları önemlidir. Ayırt edici özellikte olmayan öznitelik noktaları sisteme gereksiz yük bindirebilir ve ilgi noktalarının eşleştirilmesi esnasında yanlış eşleştirmelere sebep olabilir.

Uygulamaların niteliğine bağlı olmakla birlikte bulunan öznitelik miktarı da önemli bir değişkendir. Bulunan öznitelikler ne kadar fazla olursa, eşleştirme başarımı da faydalı bir şekilde artmaktadır. Ayrıca ilgi noktalarının çıkarım hızı da çok önemli bir parametredir. Gerçek zamanlı çalışma gerekliliği bulunan uygulamalarda, özniteliklerin belirlenme hızı genel sistem hızını doğrudan etkilemektedir.

3.3 Yer İzi Belirleme Yöntemleri

Yer izi belirleme, görüntü içerisindeki bilginin anlaşılabilmesi açısından büyük öneme sahiptir. Görüntüde yer alan her pikseldeki bilginin işlenmesi için çok fazla işlem gücüne ihtiyaç duyulacağı gibi ayırt edici bilgiler içermeyen piksellerden dolayı eşleştirme başarımı da düşük olmaktadır. Bu nedenle görüntü içerisinde yer alan ayırt edici nokta veya bölgelerin bulunması konusu, araştırmacılar tarafından yoğun bir şekilde çalışılmaktadır. Başlıca yer izi belirleme yöntemleri üç ana başlık altında toplanabilir. Bunlar;

- i) Kenar Bulma: Roberts (Roberts, 1963), Prewitt (Prewitt, 1970), Sobel (Gonzalez ve Woods, 2002), Canny (Canny, 1986), Harris (Harris ve Stephens, 1988), SUSAN (Smith ve Brady, 1995),
- ii) Köşe Bulma: Harris, SUSAN, Shi-Tomasi (Shi ve Tomasi, 1994), FAST (Rosten ve Drummond, 2006), LoG (Lindeberg, 1998), DoG (Lowe, 2004), DoH (Lindeberg, 1998),
- iii) Bölge Bulma: LoG, DoG, DoH, DoB (Bay vd., 2008), MSER (Matas vd., 2004), PCBR (Deng vd., 2007).

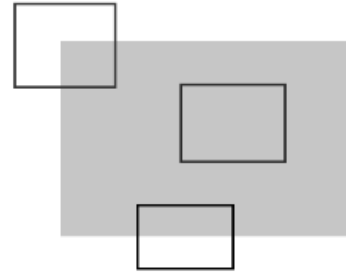
Yöntemlerden çoğu birden fazla başlık altında yer almaktadır. Bunun sebebi bu yöntemlerde bir pikselin kenar, köşe veya bir bölgeyi tanımlama ölçütü hesaplanmasıdır. Bu ölçüt kullanılarak istenilen özelliğe sahip olan pikseller seçilebilmektedir.

Kenar ve köşe bulma yöntemleri yapay görme uygulamalarının temel adımlarından biridir ve birçok nesne tanımlama ve bölütleme gibi uygulamalarda başarıyı artıran yöntemlerdendir. Kenar ve köşe bulma tabanlı yöntemler görüntüdeki değişim olan piksellerin belirlenmesi teorisi üzerine geliştirilmiştir ve bu amaçla görüntüdeki kenarların veya köşelerin bulunması hedeflenir.

Ayırt edici özellikleri önemli ölçüde düşük olduğundan kenar bulma algoritmalarının görüntü eşleştirme için kullanımı uygun görülmezken köşe bulma algoritmaları ise daha ayırt edici niteliklere sahiptir. Köşelerin ayırt ediciliğini incelemek için bir örnek vermek gerekirse; Şekil 3.2(a)'da A ve B ile gösterilen kısımlar köşe gibi ayırt edici bilgiler içermediğinden, görüntü içerisinde konumlandırılmaları oldukça zordur. C ve D ile verilen görseller ise nispeten daha kolay bir şekilde görüntü içerisinde konumlandırılabilir. Fakat bu görsellerinde kesin olarak hangi konumda olduğunu belirtmek mümkün değildir. E ve F de görülen parçaların görüntü içinde konumlandırılması, köşelerin ayırt edici özelliğinden dolayı daha kolaydır. Şekil 3.2(b)'de verilen görüntüde bu duruma dair bir örnek daha gösterilmiştir. Örneklerden de görüldüğü üzere görüntüden seçilecek özniteliklerin köşeler üzerinden çıkartılması daha anlamlıdır. Bölge bulma yöntemleri ise kenar veya köşe bulma yöntemlerinden, temsil ettikleri görüntü alanı bakımından ayrılırlar. Bu yöntemler görüntü içerisinde birbirinden farklılık gösteren görüntü bölgelerini tespit eder.



(a)



(b)

Şekil 3.2. Kenar veya köşelerin ayırt ediciliklerinin gösterilmesi

Yukarıda bahsedilen yöntemlerle tespit edilen ilgi noktaları, görüntüler arasında

oluşabilecek ölçek değişimlerinde yeniden bulunabilir olmalıdır. İlgili noktalarının ölçek değişimlerinden etkilenmemesi için araştırmacılar görüntü piramitleri ve ölçek uzaylarını önermiştir (Burt ve Adelson, 1983; Crowley ve Sanderson, 1987; Lindeberg, 1992). Bu çalışmalarda ölçek uzaylarını oluşturabilmek için giriş görüntüsü farklı boyutlara ölçeklenerek Gauss veya Laplace filtresi uygulanır ve böylece bir dizi görüntü elde edilir. İlgili noktalarının çıkarımında, oluşturulan boyut uzaylarından faydalanılmıştır. Ölçek değişimlerinden etkilenen ilgili noktası veya bölgesi çıkarım yöntemleri de ölçek uzayları kullanılarak ölçekten bağımsız hale getirilebilmektedir. Bölge bulma tabanlı yöntemlerden bazıları (LoG, DoG, DoH, DoB) bu şekilde elde edilmiştir.

3.3.1 Kenar ve Köşe Bulma Tabanlı Yöntemler

3.3.1.1 Harris

Moravec (1977; 1979) çalışmalarında, Moravec operatörünü Stanford aracının karmaşık bir ortamda gezinmesini hedefleyen bir çalışma sonucu önermiştir. Moravec çalışmasında, pikseller arası yoğunluk farkının yüksek olduğu noktaları ilgili noktaları olarak belirlemiştir. Esasen Moravec çalışmasında özellikle köşe bulmayı değil, görüntü akışı içinde tekrarlanan belirgin görüntü bölgelerini tanımlayabilmeyi amaçlamıştır. Moravec'in çalışmasında pikseller arasındaki yoğunluk farkına göre bir hata hesaplanmış ve bu hatayı en aza indirebilmek için hata fonksiyonu,

$$E_{x,y} = \sum_{u,v} w_{u,v} [I_{x+u,y+v} - I_{u,v}]^2 \quad (3.1)$$

ile tanımlanır. Burada w kaydırılan pencere fonksiyonunu ifade etmektedir. Pencere fonksiyonunun alabileceği değerler kümesi $\{(1,0),(1,1),(0,1),(-1,1),(-1,0),(-1,-1),(0,-1),(1,-1)\}$ 'dir. Ayrıca, $I_{u,v}$ pencerenin (u,v) noktasında merkezi bulunan piksel değerlerini temsil etmektedir. Moravec operatörü pencere fonksiyonunun alabileceği değerler sınırlı olduğundan 45° 'den küçük açısal değişimleri kapsamamaktadır. 1988 yılında Harris tarafından önerilen operatör ile Moravec'in algoritması geliştirilmiş ve küçük açısal değişimlerinde algılanabilmesi için Denklem (3.1), (u,v) merkez noktası etrafında Taylor açılımı kullanarak genişletilerek,

$$E_{x,y} = \sum_{u,v} w_{u,v} [xX + yY + O(x^2, y^2)]^2 \quad (3.2)$$

ifadesi elde edilir. Burada $O(x^2, y^2)$ terimi ihmal edilerek X ve Y türev operatörleri yaklaşım uygulanarak,

$$X = I * (-1, 0, 1) \cong \partial I / \partial x \quad (3.3)$$

$$Y = I * (-1, 0, 1)^T \cong \partial I / \partial y \quad (3.4)$$

ile tanımlanmıştır. Bu yaklaşımla beraber hata fonksiyonu,

$$E(x, y) = Ax^2 + 2Cxy + By^2 \quad (3.5)$$

ile ifade edilmiştir. Burada A , B ve C sırasıyla,

$$A = X^2 * w \quad (3.6)$$

$$B = Y^2 * w \quad (3.7)$$

$$C = (XY) * w \quad (3.8)$$

ile tanımlanır. Moravec'in çalışmasında w penceresi dikdörtgensel ve ikili olarak kullanıldığından elde edilen cevap gürültü içermektedir. Harris ise çalışmasında pencere fonksiyonunu Denklem (3.9)'da verildiği gibi Gauss olarak seçmiştir. Bu sayede daha yumuşak geçişler elde edilmektedir.

$$w_{u,v} = \exp - (u^2 + v^2) / 2\sigma^2 \quad (3.9)$$

Bu değişimden sonra E hata fonksiyonu,

$$E(x, y) = (x, y)M(x, y)^2 \quad (3.10)$$

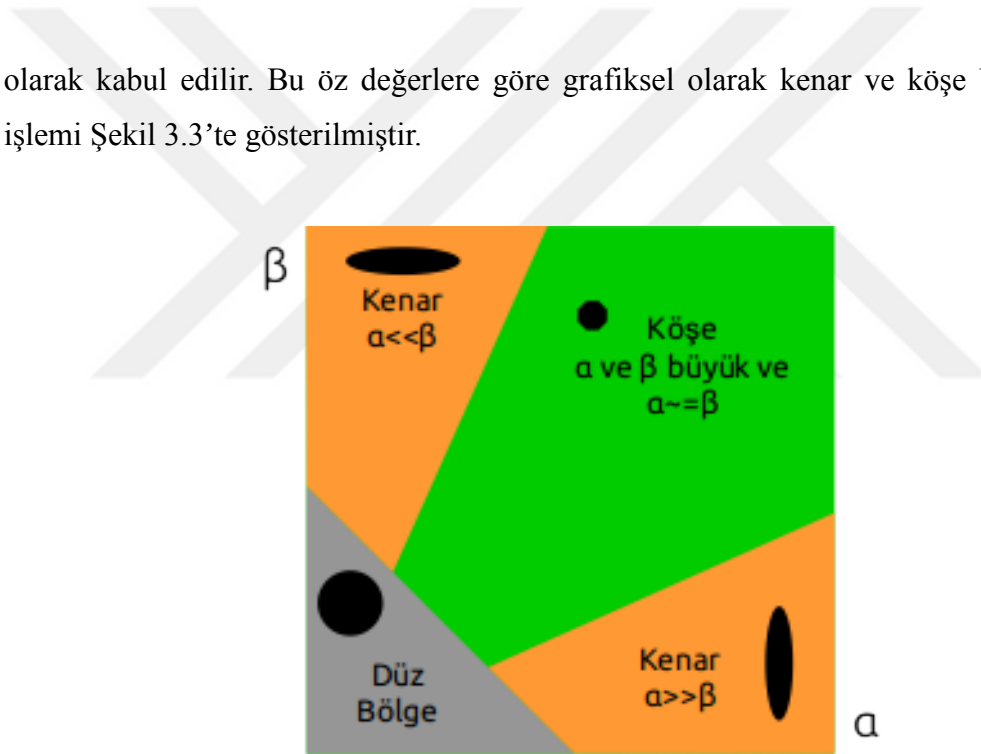
ile ifade edilir. Bu denklemde M , 2x2 boyutunda bir matristir ve

$$M = \begin{bmatrix} A & C \\ C & B \end{bmatrix} \quad (3.11)$$

şeklinde tanımlanmıştır. M matrisinin öz değerleri (α, β) hesaplandığında merkez piksel noktası için 3 farklı durum oluşur ve bu durumlara öz değerlerin önceden belirlenmiş bir sınır değerden küçük veya büyük olmasına göre karar verilir. Durumlara göre merkez piksel noktası;

1. Eğer öz değerlerden her ikisi de sınır değerden küçük ise merkez piksel noktası düz,
2. Eğer öz değerlerden bir tanesi sınır değerden büyük diğeri küçük ise merkez piksel noktası kenar,
3. Eğer öz değerlerden her ikisi de sınır değerden büyük ise merkez piksel noktası köşe,

olarak kabul edilir. Bu öz değerlere göre grafiksel olarak kenar ve köşe belirlenmesi işlemi Şekil 3.3'te gösterilmiştir.



Şekil 3.3. Öz değerlere göre grafiksel kenar ve köşe gösterimi

Fakat öz değerlerin hesaplanması işlem yükü bakımından ağır olduğundan Denklem (3.12) ve (3.13) kullanılarak R fonksiyonu,

$$Tr(M) = \alpha + \beta = A + B \quad (3.12)$$

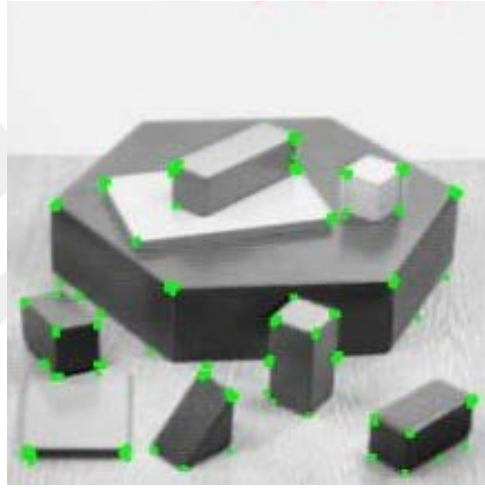
$$Det(M) = \alpha\beta = AB - C^2 \quad (3.13)$$

$$R = Det(M) - kTr^2 \quad (3.14)$$

ile hesaplanır. T kullanıcı tarafından seçilen bir eşik değeri olmak üzere, R 'nin aldığı değerler kullanılarak aşağıdaki kurallar ile merkez pikselin;

1. $|R| < T$ ise düz,
2. $R < 0$ ise kenar,
3. $R > T$ ise köşe,

olduğuna karar verilir. Bu sayede bir görüntü içerisinde sadece köşe noktaları seçilebilmektedir. Örnek olarak Şekil 3.4'te verilen görüntü üzerinde Harris köşe bulma algoritması kullanılarak belirlenmiş olan ilgi noktaları görülmektedir.



Şekil 3.4. Harris köşe bulma algoritması kullanılarak bulunan ilgi noktaları

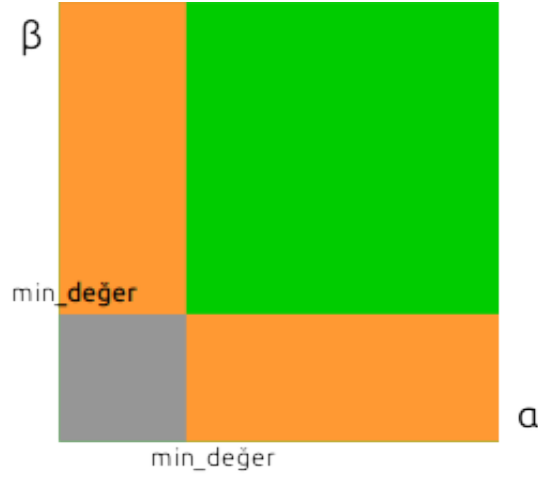
3.3.1.2 Shi-Tomasi

Harris'in çalışmasına küçük bir değişiklik uygulayarak Shi ve Tomasi (1994) tarafından bir yeni bir çalışma yayınlanmıştır. Bu çalışmada Harris köşe bulma algoritmasının karar fonksiyonu R 'nin hesaplanması için kullanılan yöntem doğrusallaştırılarak daha başarılı sonuçlar alınabildiği gösterilmiştir. Shi ve Tomasi tarafından önerilen R ,

$$R = \min(\alpha, \beta) \quad (3.15)$$

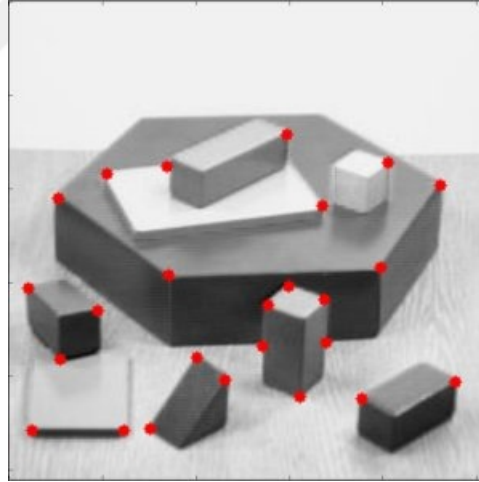
ile hesaplanır. Bu sayede elde edilen kenar ve köşe grafiği Şekil 3.5'te görüldüğü gibidir. Burada (3.15)'den elde edilen değer Şekil 3.5'te gösterilmiş olan

“min_değer”den büyük olması durumunda ilgili piksel köşe olarak kabul edilmektedir.



Şekil 3.5. Shi-Tomasi tarafından önerilen algoritmanın kenar ve köşe ayırma grafiği

Shi-Tomasi tarafından önerilen algoritma tarafından elde edilen ilgi noktaları Şekil 3.6’da verilmiştir.

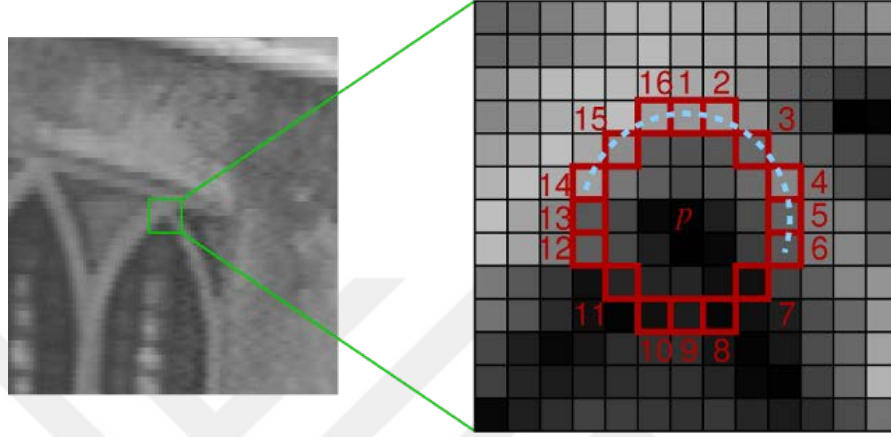


Şekil 3.6. Shi-Tomasi köşe bulma algoritması kullanılarak bulunan ilgi noktaları

3.3.1.3 FAST

FAST köşe bulma yöntemi, hızlandırılmış bölüt testlerinden elde edilen öznelik anlamına gelmektedir (Rosten ve Drummond, 2006). Köşe bulmada kullanılan en hızlı yöntemlerden biridir. Hızlı olmasının sebebi ise ön tanımlı karar ağaçları kullanarak pikselin köşe olup olmadığına karar vermesindendir. Karar ağaçları ikili piksel noktalarının olduğu dallardan oluşmaktadır. Karar ağacında yer alan bir dalda yer alan

iki piksel değeri arasında karşılaştırma yapılarak karar verilir. Karşılaştırmada kullanılan pikseller arasındaki farkın bir sınır değerinden büyük olması gerekir. FAST yönteminde kullanılan karşılaştırma çiftlerine örnek olarak Şekil 3.7 verilebilir. p merkez noktasını ve çevreleyen numaralandırılmış pikseller ise karar ağacındaki noktaları temsil etmektedir.



Şekil 3.7. FAST algoritmasında kullanılan karar ağacı (Rosten ve Drummond, 2006)

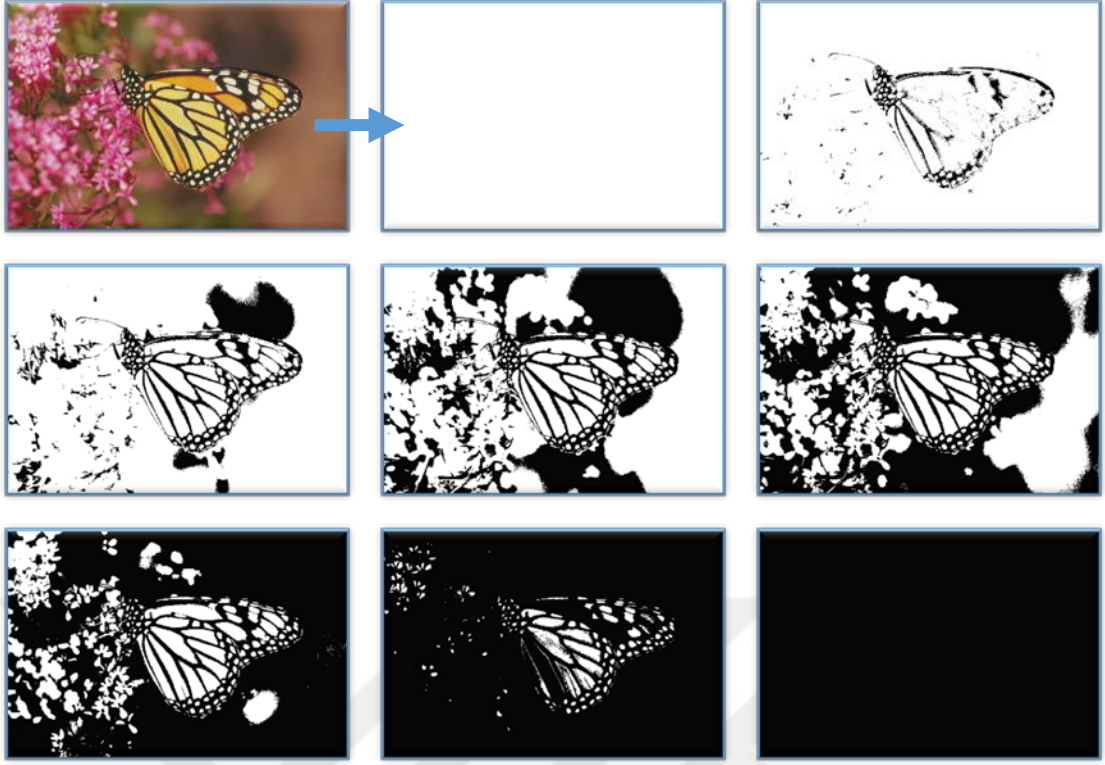
Kullanılan karar ağaçları daha büyük veya küçük boyutlarda seçilebilmektedir. Görüntü içeriğine göre boyut belirlenerek karşılaştırma için karar ağacı oluşturulur.

3.3.2 Bölge Bulma Tabanlı Yöntemler

Bu bölümde mevcut bölge bulma tabanlı yöntemler incelenecektir.

3.3.2.1 MSER

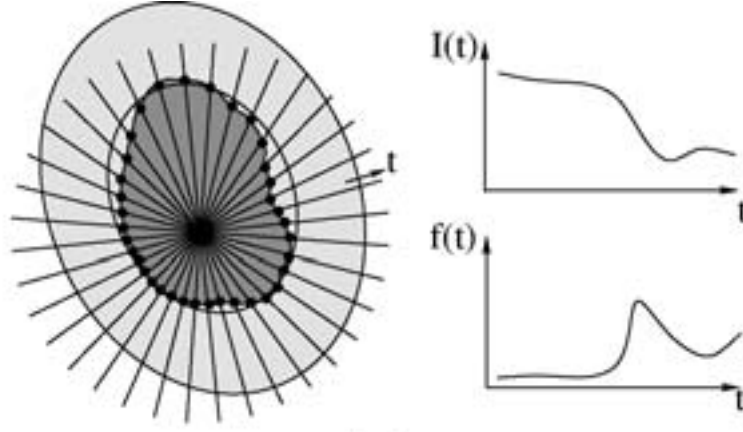
MSER yöntemi, azami kararlılıkta uçdeğer bölgeleri anlamına gelmektedir ve görüntüyü bağlantılı görüntü bileşenleri olarak ifade eder (Matas vd., 2004). MSER yönteminde görüntüden ilk olarak kararlı bölgeler çıkarılmalıdır. Bölge çıkarımı için giriş görüntüsüne farklı eşik değerleri uygulanarak siyah beyaz görüntüler elde edilir ve eşik seviyesinin altında kalan değerler siyah, üstünde ve eşit olan değerler ise beyaz olarak işaretlenir. Bu sayede beyazdan siyaha doğru geçiş yapan bir dizi görüntü oluşmaktadır (Şekil 3.8).



Şekil 3.8. MSER bölgelerini belirlemek için kullanılan eşiklenmiş görüntü dizisi

Bu görüntülerin en düşük eşik değerinden en büyük eşik değerine kadar sırayla verildiğini düşünelim. Eşik değerleri arttıkça görüntüde siyah noktalar belirmeye başlayacaktır. Bu siyah noktalar görüntü bölgeleri oluşturur ve eşik değeri arttıkça siyah bölgeler birleşmeye başlar. Beliren ve birleşen bu siyah noktalar görüntü bölgelerinin yerel en küçük değerlerini ifade etmektedir. Eşik değerleri kullanılarak oluşturulan her görüntü dizisinde bulunan yerel en küçük değer bölgeleri de bir dizi halinde hafızaya alınmaktadır. Bu en küçük değer bölgelerinden oluşan dizide en çok tekrarlayan bölgeler azami kararlılıktaki bölgeler olarak adlandırılır ve belirlenen bir değişim eşiği ile sürekli bulunabilirliği ölçülerek belirlenmektedir. Yerel en küçük bölgelerin bulunmasına benzer şekilde ters eşik sırasıyla gidilerek yerel en büyük noktalar bulunur.

Bulunan bu bölgeler sürekli bulunabilirlik eşiğinin yanı sıra, boyutlarına göre de elenir. Belirlenen alt ve üst eşik alan değerlerinden küçük ve büyük olan MSER bölgeleri elenir ve bu elemenden sonra geriye azami kararlılıkta uçdeğer bölgeleri kalır.

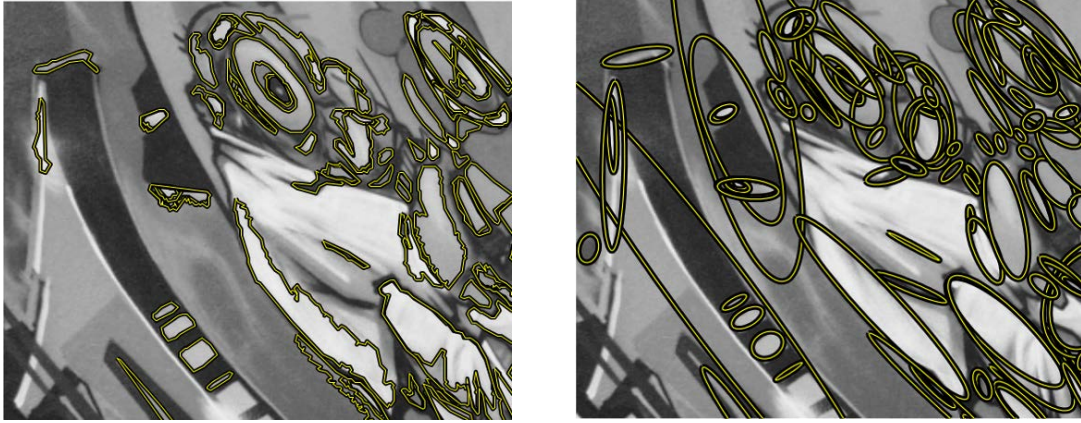


Şekil 3.9. Görüntü bölgesinin sınırlarının belirlenmesi (Mikolajczyk vd., 2005)

Elde edilen bölgeler rastgele bir yapıya sahip olacağından, bu bölgelerin görüntü eşleştirmede kullanılabilmesi için eliptik bir forma sokulması gerekmektedir. Eliptik parametreler Şekil 3.9'daki doğrusal hatlar boyunca,

$$f_I(t) = \frac{abs(I(t) - I_0)}{\max(\frac{\int_0^t abs(I(t) - I_0)}{t}, d)} \quad (3.16)$$

ile hesaplanmaktadır. t doğrusal hat üzerinde bir noktayı, $I(t)$ bu noktanın koordinatlarını, I_0 uç bölgenin merkez noktasını ifade eder. d sıfıra bölmeyi engellemek için eklenen küçük ve sabit bir sayıdır. $f_I(t)$ sınır geçiş bölgelerinde en büyük değerini alacaktır. Bu nokta görüntünün bakış açısı değişse bile değişmeyecektir. f_I ile elde edilen bu sınır noktaları birleştirilerek elde edilen şekil ile aynı ikinci dereceden moment değerlerini taşıyan bir elips ile değiştirilir. Bu işlemde elde edilen elips görüntü bölgesinin sınır çizgisinde olduğu gibi bakış açısından bağımsız olmaktadır. MSER ile elde edilen eliptik bölgeler Şekil 3.10'da gösterilmiştir.



(a)

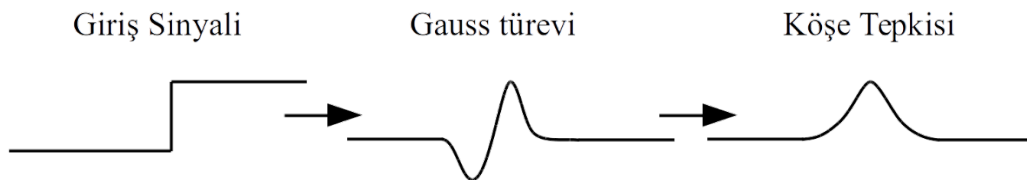
(b)

Şekil 3.10. MSER bölgeleri a) geliş güzel yapı b) eliptik yapı (Mikolajczyk vd., 2005)

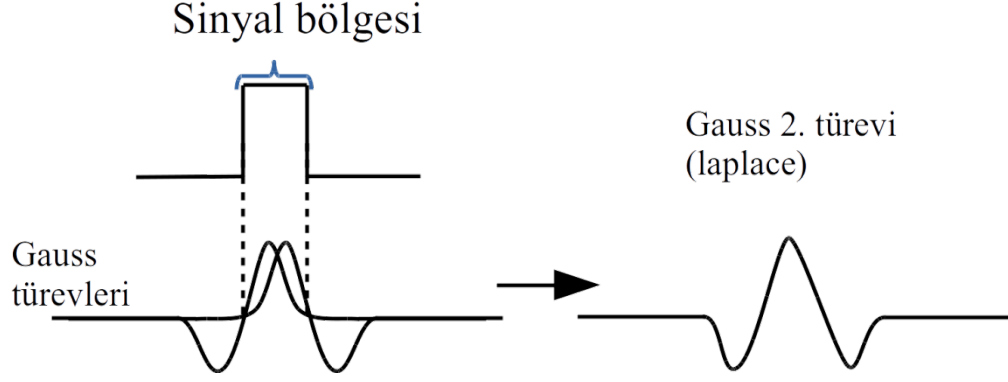
MSER bölgeleri bakış açısı değişimlerinden bağımsız olsa da ölçek değişimlerinden etkilenmektedir. Bunun için ölçek uzayında MSER çıkarımı için çalışma yapılmıştır (Forssen ve Lowe, 2007). LoG, DoG vb. ölçek uzayı kullanan yöntemler ise ölçek değişimlerinden daha az etkilenmektedir.

3.3.2.2 Ölçek Uzayı Kullanılarak Bölgelerin Bulunması

Ölçek uzayları, daha önceki bölümlerde bahsedildiği gibi ölçekten bağımsız ilgi noktalarının bulunabilmesi için kullanılan bir yöntemdir. Aynı zamanda ölçek uzayları kullanılarak görüntü bölgelerinin tespit edilmesi de mümkündür. İlgi noktalarının bulunması için kullanılan filtre çekirdek boyutlarının doğru seçilmesi ile elde edilecek olan ilgi noktası bir bölgeyi temsil edebilir. Doğru çekirdek boyutu belirlenmiş bir köşe bulma yöntemi bölge bulma algoritmasına dönüştürülebilir. Bu fikrin daha rahat anlaşılabilmesi için Şekil 3.11 ve Şekil 3.12 ile verilen sinyallerin Gauss türev çekirdeği ile konvolüsyonundan elde edilen köşe tepkileri incelenmelidir.



Şekil 3.11. Giriş sinyalinin köşe tepkisi



Şekil 3.12. Sinyal üzerinde bir bölgenin Gauss tepkisinden elde edilen çekirdek birleşimi

Şekil 3.12’de çok küçük bir sinyal bölgesinin tespit edilebilmesi için Gauss türevlerinin birleşiminin kullanılabileceği görülmektedir. Ayrıca bu birleşik sinyal asıl Gauss sinyalinin ikinci türevi ile ifade edilebilir. Gauss sinyalinin ikinci türevi kullanılarak görüntüye konvolüsyon uygulanır ise alınacak tepki sinyal bölgelerinde çok güçlü olacaktır. Fakat sinyal bölgesinin alanının büyümesi (ölçek büyümesi) durumunda kullanılan bu birleşik sinyalden iki ayrı tepki elde edilmektedir. Tam tersi düşünüldüğü durumda da sinyal bölgesi çok küçük (ölçek küçülmesi) ise alınacak tepki ile bölgenin tepkisi belirgin olmayacaktır. Dolayısı ile kullanılacak konvolüsyon çekirdeğinin boyutu çok önemlidir. İki boyutlu sinyal (görüntü) ele alındığında Denklem (3.17) ile gösterilen Gauss sinyalinin Laplace’ı (LoG) Denklem (3.18) ile hesaplanır.

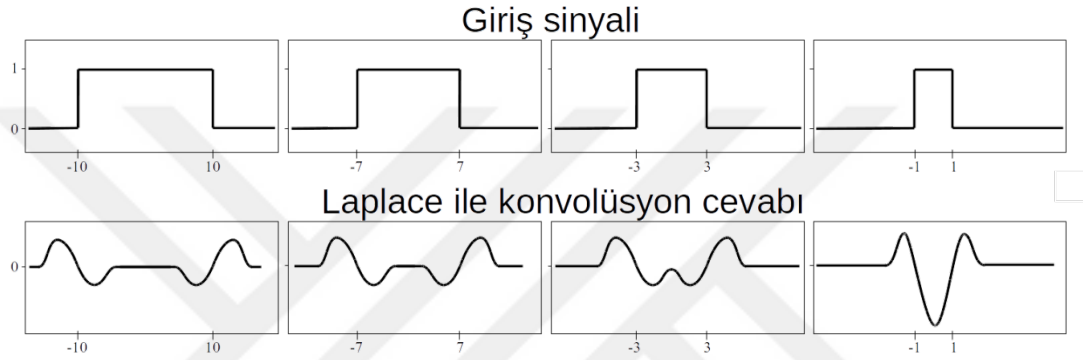
$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3.17)$$

$$L(x, y, \sigma) = \nabla^2 G(x, y, \sigma) = \frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2} \quad (3.18)$$

Burada görüntünün piksel koordinatları x ve y , görüntünün yumuşatma katsayısı σ ile gösterilir. G ise Gauss çekirdeği ile yumuşatılmış görüntüyü ifade etmektedir. σ ’nın seçimi hangi ölçekte görüntü bölgelerinin bulunacağını belirler. r yarıçapına sahip bir dairenin görüntü bölgesi olarak tanımlanabilmesi için Denklem (3.18)’in sıfırlarının dairenin yarıçapı genişliğinde olması gerekir. Laplace’ın sıfırları,

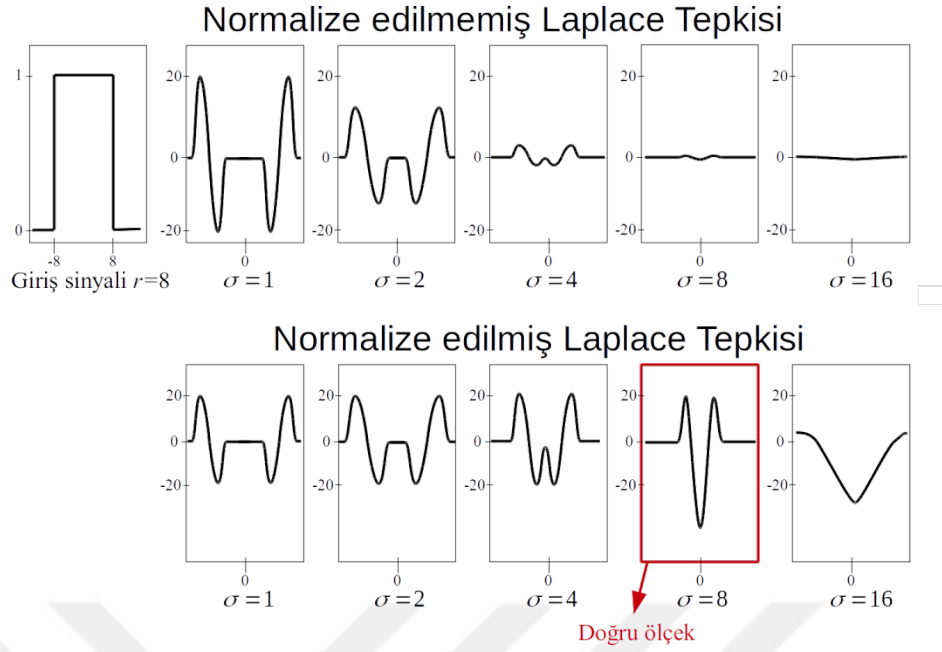
$$\left(1 - \frac{x^2 + y^2}{2\sigma^2}\right) = 0 \quad (3.19)$$

ile verilir. Bu eşitliğin çözümünden, en yüksek sonuç $\sigma = r/\sqrt{2}$ ile elde edilmektedir. r burada karakteristik ölçeği temsil etmektedir. Karakteristik ölçeğin kullanılan köşe bulma yöntemi değişmesi durumunda tekrar belirlenmesi gerekir. Yarıçapı 1 olan bir sinyal için sinyalin Laplace ile konvolüsyonundan Şekil 3.13'te gösterildiği gibi bir sonuç elde edilmektedir.



Şekil 3.13. Laplace kullanılarak konvolüsyon uygulanan sinyalden elde edilen sistem tepkisi ($\sigma = 1$)

Görüldüğü üzere sistemin verdiği tepkilerden en büyüğü ölçeğin tam uyduğu yarıçapın bir olduğu sinyaldir. Ölçek uzayları ile yapılmak istenen ise giriş görüntüsünün hangi ölçekte olduğunun belirlenmesidir ve farklı σ değerleri için giriş görüntüsünün verdiği tepkiler incelenerek gerçekleştirilebilir. Farklı σ değerleri kullanılması sistemin tepkisinde genliğin değişmesine sebep olmaktadır. Bu nedenle elde edilen sistem tepkilerinin normalize edilmesi gerekmektedir. Şekil 3.14'te farklı σ değerlerine karşılık sistem tepkisi ve normalize edilmiş tepkiler kullanılarak seçilen ölçek gösterilmiştir. Sistem tepkisi σ arttıkça, σ 'nın karesi oranında azalacaktır. Dolayısıyla sistem tepkisi σ^2 ile çarpılarak normalize edilmiştir.



Şekil 3.14. Farklı σ değerleri için normalize edilmemiş ve normalize edilmiş sistem tepkisi ve ölçek seçimi

İki boyutlu sinyaller (görüntüler) ele alındığında Denklem (3.19) normalize edilir ise,

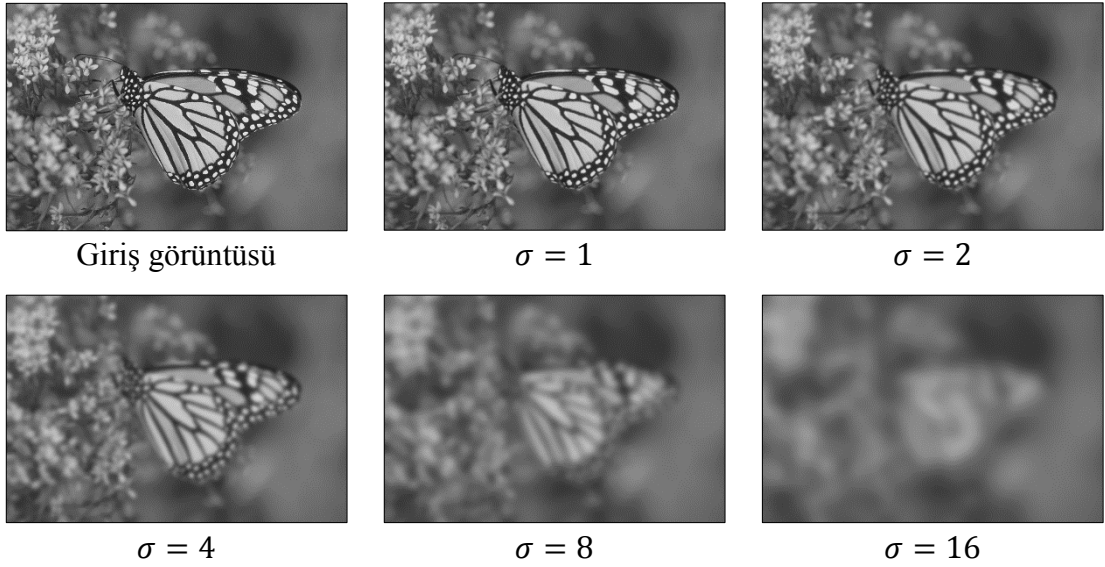
$$L_n(x, y, \sigma) = \nabla_n^2 G(x, y, \sigma) = \sigma^2 \left(\frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2} \right) \quad (3.20)$$

ifadesi elde edilir. Dolayısıyla ölçek uzayında her noktanın sistem tepkileri incelenerek doğru boyut ve konum elde edilebilir. Oluşturulan ölçek uzayı Şekil 3.15'te görüldüğü gibi olacaktır.

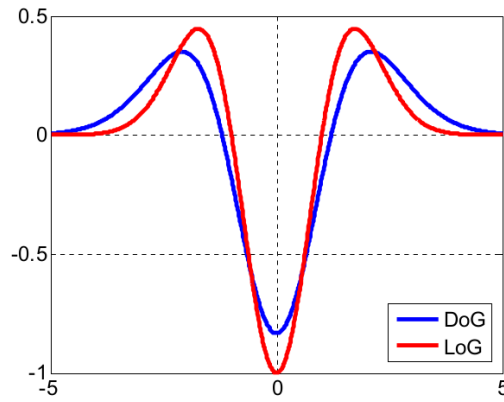
LoG'de yer alan türev işleminden dolayı hesaplama oldukça maliyetlidir. Bu sorunun çözümü için DoG algoritması, LoG algoritmasına bir yaklaşım olarak önerilmiştir (Lowe, 2004). Çalışmada σ değeri 1.6 kullanıldığı zaman DoG filtresinden alınan sistem tepkisinin LoG algoritmasına yakınsadığı gösterilmiştir. DoG,

$$\nabla_n^2 G(x, y, \sigma) = \frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma} \quad (3.21)$$

ile hesaplanır. DoG ile bir boyutlu sinyalden alınan sistem tepkileri incelendiği zaman oldukça yakın sonuçlar elde edildiği Şekil 3.16'da görülebilmektedir.



Şekil 3.15. Gauss ölçek uzayı



Şekil 3.16. LoG ve DoG sistem tepkisi

DoG algoritmasında farklı σ değerleri kullanılarak hesaplanan $G(x, y, \sigma)$ görüntüleri arasındaki farklar bulunur. Fark görüntüsü LoG ile elde edilen görüntüye benzer özellikler taşıdığından ölçek uzayı oluşturularak görüntü bölgeleri belirlenmektedir. Lowe çalışmasında DoG ile bölge çıkarımını, SIFT algoritmasının ölçekten bağımsız öznetelik çıkarımı aşamasında kullanmıştır. Lowe çalışmasında ölçekten bağımsız öznetelik çıkarımı için şu adımları uygulamaktadır;

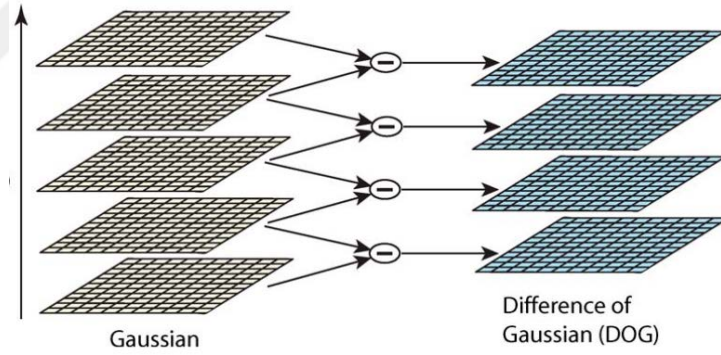
1. Ölçek uzayı oluşturulması: Giriş görüntüsünün boyutu 4 defa yarıya indirilerek farklı boyutlarda görüntüler elde edilmektedir.
2. Her bir ölçekte $\sigma = 1.6$ değerinden başlayarak artan Gauss yumuşatması

uygulanmış görüntüler elde edilir. Ölçeklenmiş orijinal görüntü ile birlikte toplam 5 adet yumuşatılmış görüntü oluşturulur. Oluşturulan boyut uzayı Şekil 3.17’de gösterildiği gibidir.



Şekil 3.17. DoG boyut uzayı

3. Aynı ölçekte yer alan yumuşatılmış görüntülerin farkları hesaplanır. Bu farklar her iki komşu görüntü için hesaplanır. Hesaplama rutini Şekil 3.18’de verilmiştir. Ardından oluşturulan DoG uzayı ise Şekil 3.19’da gösterilmiştir.



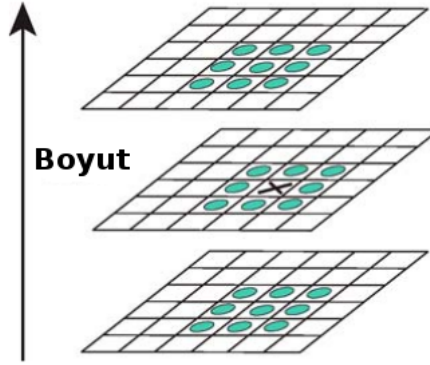
Şekil 3.18. DoG hesaplama rutini (Lowe, 2004)



Şekil 3.19. Elde edilen DoG görüntüleri

4. DoG görüntüleri kullanılarak her bir ölçek kendi içinde olacak şekilde en büyük değerli piksel noktaları bulunur. Bunun için her piksel değeri bir üst ve bir alt

boyuttaki komşu pikseller (26 piksel) ile karşılaştırılır ve hepsinden küçük veya hepsinden büyük ise ilgili piksel konumu öznitelik bölgesi olarak seçilir. Bu işlem Şekil 3.20’de gösterilmiştir.



Şekil 3.20. İlgi noktalarının tespit edilmesi (Lowe, 2004)

Bu sayede görüntüden öznitelik bölgeleri elde edilir. Fakat elde edilen bu ilgi noktaları görüntüye karşı oldukça hassastır. Görüntünün sınır bölgelerinde hesaplanan öznitelik bölgeleri düşük doğruluktur. Dolayısıyla bulunan özniteliklerin tam olarak yerleştirilmesi gerekir. Bu sorunun çözümü için Brown ve Lowe (2002) tarafından önerilmiş olan 3 boyutlu ikinci dereceden bir denklem ile öznitelik noktasının alt-konumu hesaplanmaktadır. Çalışmada, alt-konumu belirlenmiş öznitelik noktaları kullanılarak gerçekleştirilen sistemin eşleştirme performansının da arttığı gösterilmiştir. Bu yaklaşımda öznitelik noktası merkezde olacak şekilde DoG fonksiyonu üzerinde Taylor açılımı kullanılarak,

$$D(p) = D + \frac{\partial D^T}{\partial p} p + \frac{1}{2} p^T \frac{\partial^2 D}{\partial p^2} p \quad (3.22)$$

elde edilir. Bu denklemde D , DoG ile elde edilen öznitelik noktası etrafındaki değerlerdir ve $p = (x, y, \sigma)^T$ öznitelik noktasından olan uzaklıktır. p noktasının yerini bulabilmek için Denklem (3.22) sıfıra eşitlenir ve Denklem (3.23) elde edilir.

$$p = - \frac{\partial^2 D^{-1}}{\partial p^2} \frac{\partial D}{\partial p} \quad (3.23)$$

Brown ve Lowe çalışmalarında bu türevin yaklaşık olarak komşu uç değerlerinin farkına eşit olduğunu göstermişlerdir. Sonuç alt-konumu hesaplamak için 3x3 boyutundaki lineer sistem çözülerek işlem maliyeti düşürülür. Eğer elde edilen değer 0.5'ten büyük ise öznitelik noktası (p), yakın komşuluktaki başka bir öznitelik noktasına kayar ($p + \hat{p}$). Böylece ilgi noktasının tam yeri tespit edilir.

Ayrıca $D(\hat{x})$ fonksiyonun öznitelik noktasındaki değeri, kararsız ve düşük zıtlığa (Contrast) sahip noktaların elenmesi içinde kullanılabilir. Bu değer Denklem (3.23)'in Denklem (3.18)'de yerine koyulması ile,

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x} \quad (3.24)$$

olarak elde edilir. $D(\hat{x})$, fonksiyonun uç değerinin 0.3'ten düşük olması durumunda bu ilgi noktası elenmektedir. DoG algoritması aynı zamanda kenar bölgelerini de öznitelik noktası olarak tespit etmektedir. Kenar bölgelerinin elenebilmesi için 2x2 Hessian matrisi,

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (3.25)$$

kullanılır. Harris ve Stephens (1988) çalışmalarında, öz değerlerin (α, β) hesaplanmasına gerek kalmadan, sadece bu değerlerin oranlarını kullanarak kenar ve köşe ayrımı yapmışlardır. (α, β) arasındaki oran,

$$Tr(H) = \alpha + \beta = D_{xx} + D_{yy} \quad (3.26)$$

$$Det(H) = \alpha\beta = D_{xx}D_{yy} - (D_{xy})^2 \quad (3.27)$$

ile hesaplanır. Eğer Denklem (3.27) sonucu negatif ise öznitelik noktası elenir. Eğer Denklem (3.28) sağlanır ise bu ilgi noktası korunur ($r : (\alpha, \beta)$ arasındaki oran).

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r + 1)^2}{r} \quad (3.28)$$

Böylece ilgi noktalarının sayısı azaltılırken aynı zamanda eşleştirme başarımı yükseltilmiş olur. Lowe çalışmasında $\frac{(r+1)^2}{r}$ sınır değerini 10 olarak kullanmıştır.



BÖLÜM IV

YER İZİ TANIMLAYICILAR

4.1 Giriş

Yer izlerinin görüntüler arası eşleştirilmesi görsel seyir sistemlerinin başarımını etkileyen önemli noktalardan birisidir. Yer izlerinin yüksek doğrulukta eşleştirilebilmesi için bu noktalardan karşılaştırmada kullanılabilecek nitelikte bilgilerin çıkarılması gerekmektedir. Yer izlerini tanımlayabilmek için birçok yöntem önerilmiştir. Yer izi veya ilgi noktası tanımlayıcıları oldukça yoğun olarak kullanılmaktadır ve kullanım alanlarından bazıları şu şekilde sıralanabilir; 3 boyutlu yeniden oluşturma, hareketten yapı çıkarımı, nesne tanıma, sınıflandırma, görüntü birleştirme ve gezinim sistemleri için konum değişimi hesaplama vb. (Beall vd., 2010; Chang vd., 2007; Moreels ve Perona, 2006).

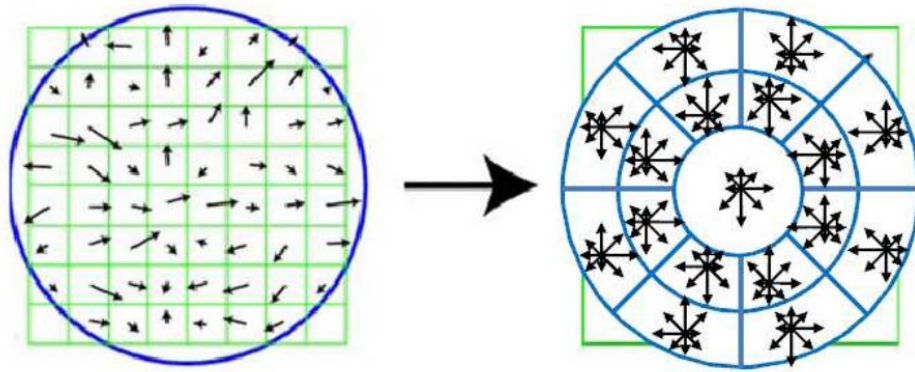
(Rosenthaler vd., 1992) ve (Heitger vd., 1992) çalışmaları görüntü içerisinden belirli bir bölgeye ait tanımlayıcı bilgisi çıkarımının literatürdeki ilk örneklerindendir. Bu çalışmalarda nöro-biyolojik mekanizmalardan esinlenilmiştir ve ilgi noktalarına karşılık gelen görüntü bölgelerinden Gabor benzeri filtreler kullanarak yerel enerji bilgisi hesaplanmıştır. Enerjinin birinci ve ikinci türevlerinin maksimumu bulunarak ilgi noktasının kalitesi belirlenmiştir.

Tanımlayıcılar ile ilgili 1992'den günümüze kadar birçok çalışma yapılmıştır. Lowe (2004) tarafından önerilen ölçek bağımsız öznitelik dönüşümü (SIFT) bu çalışmalar içerisinde en popüler olanıdır. Mikolajczyk (2005) çalışmasında karmaşık tanımlayıcılar ile bazı bölgesel tanımlayıcıları karşılaştırmıştır. Bu karşılaştırmaya göre SIFT ve SIFT'in değişik bir türü olan eğim konumu ve yönelim histogramı (GLOH) tanımlayıcılarının, şekil içeriği (Shape context), yönlendirmeli filtreler (Steerable filters), döner görüntüler (Spin images), fark değişmezleri (Differential invariants) ve moment değişmezleri (Moment invariants) gibi tanımlayıcılar karşısında daha başarılı olduğu gösterilmiştir. SIFT'in kabul gören yapısı kullanılarak birçok değişik tanımlayıcı geliştirilmiştir (Abdel-Hakim ve Farag, 2006; Bay vd., 2008; Ke ve Sukthankar, 2004; Morel ve Yu, 2009; Mortensen vd., 2005; Zhang vd., 2013).

SIFT tanımlayıcılarının boyutunu azaltabilmek için yapılan çalışmalar da mevcuttur (Ke ve Sukthankar, 2004). Bu çalışmanın amacı başarımı çok fazla etkilemeden temel bileşenler analizi (Jolliffe, 1986) (PCA) yöntemi kullanarak 128 adet değer içeren SIFT tanımlayıcılarının boyutunu değiştirebilmektir. Böylece sistemin hafıza gereksinimi de azaltılır.

SIFT benzeri SURF, GLOH ve HOG gibi birçok çalışma mevcuttur. Bunlar arasından SURF (Bay vd., 2008), SIFT algoritmasına yaklaşım uygulanan bir çalışmadır. Bu çalışmada SIFT algoritmasının hesaplama adımları integral görüntüleri kullanacak şekilde düzenlenmiştir. Böylece algoritmaya kabul edilebilir bir performans kaybı ile 10 kat hız kazandırılmıştır.

GLOH tanımlayıcıları (Mikolajczyk ve Schmid, 2005), SIFT tanımlayıcıların bir uzantısı olarak sunulmuştur. GLOH algoritmasında tanımlayıcıların açı gruplandırması ve parça ayrımı için logaritmik bir yapı kullanılmıştır. Bu şekilde elde edilen tanımlayıcıların SIFT tanımlayıcılarından daha yüksek başarımlar gösterdikleri belirtilmiştir. GLOH tanımlayıcı yapıları Şekil 4.1’de verilmiştir.



Şekil 4.1. GLOH tanımlayıcı gösterimi (Mikolajczyk ve Schmid, 2005)

Ayrıca Tola vd. (2008, 2010) yoğun görüntü eşleştirme uygulamaları için DAISY tanımlayıcılarını önermiştir. Morel ve Yu (2009) tarafından yapılan çalışmada ise SIFT tanımlayıcılarının bakış açısı değişiminin olduğu durumlarda performansını arttıran bir yöntem önerilmiştir. Ambai ve Yoshida (2011), performanstan ödün vererek, SIFT ve GLOH benzeri tanımlayıcılar için bakma tabloları kullanımı ve diğer yaklaşımlar ile yoğun ve gerçek zamanlı tanımlayıcılar (CARD) adında bir çalışma yapmıştır.

SIFT ve türevleri dışında farklı tanımlayıcı yapıları da bulunmaktadır. Berg ve Malik (2001), şablon eşleştirme uygulamaları için geometrik yumuşatma yöntemini önermiştir. Forssen ve Lowe (2007), azami kararlılıkta uç bölgeler (MSER) (Matas vd., 2004) algoritması tarafından bulunan bölgelerden tanımlayıcı çıkarımı ile alakalı bir yöntem geliştirmiştir. Chen vd. (2010), Weber Kanununu kullanarak geliştirilmiş bir tanımlayıcı önermiştir. Calonder vd. (2010), ikili gürbüz bağımsız temel öznitelikler (BRIEF) adlı çalışmalarında hızlı bir tanımlayıcı geliştirmiştir. Rublee vd. (2011), BRIEF tanımlayıcılarına yönelim hesabını ekleyerek, BRIEF tanımlayıcılarının açı bozulmalarından etkilenmemesini sağlamıştır. Alahi vd. (2012), insan gözünden esinlenerek farklı bir tanımlayıcı biçimi önermiştir (FREAK). Bu yöntemler gerçek zamanlı çalışabilmeleri için gerekli hıza sahipken başarımları SIFT gibi algoritmalarından geri kalmaktadır.

Bahsedilen yöntemler açısal bozulmalardan kaynaklanan eşleştirme hatalarının üstesinden gelebilmek için yönelim hesaplanmasına ihtiyaç duymaktadır. Bu yöntemlerin dışında yönelim açısı hesaplanmasına ihtiyaç duyulmayan çalışmalarda mevcuttur. Fan vd. (2011), destek bölgelerinden açıdan bağımsız öznitelikler çıkarımı ile MROGH tanımlayıcılarını önermiştir. Wang vd. (2011), Fan vd. çalışmasına benzer şekilde yerel piksel değerlerini kullanarak LIOP tanımlayıcılarını önermiştir. Fan vd. (2012), MROGH tanımlayıcılarının yanında piksel değerlerinden hesaplanan açığa bağlı olarak değişmeyen özellikleri kullanarak MRRID adında yeni bir tanımlayıcı önermiştir. Dubey vd. (2014), yerel sıralama biçimi kullanarak piksel değerlerinden açı bağımsız öznitelikler elde etmiştir. Xu vd. (2014) bölgesel değişmezleri hesaplayarak OSRI tanımlayıcılarını elde etmiştir. Zhou vd. (2014) ise örnekleme tabanlı bir tanımlayıcı (SLD) oluşturmuştur.

Tanımlayıcı performanslarının karşılaştırılması için literatürde önerilmiş yöntemler mevcuttur. Mikolajczyk vd. (2005) çalışmasında karmaşık tanımlayıcılar ile bazı bölgesel tanımlayıcıları karşılaştırmıştır. Ayrıca *yeniden tespit* ve $(1 - \text{hassasiyet})$ olmak üzere tanımlayıcılar için iki önemli ölçüt tanımlamıştır. Bulanıklık, sıkıştırma, ölçek, açısal değişim gibi bozunumlar içeren görüntü kümeleri önerilmiştir (*Oxford* veri kümesi). Daha sonra Moreels ve Perona (2006) yılında 3 boyutlu objeler üzerinde bazı yaygın tanımlayıcıların başarımlarını test etmişlerdir.

Yapılan çalışmalara bakıldığı zaman tanımlayıcıları iki bölümde incelemek mümkündür. Bunlar yönelim açısının hesaplandığı ve yönelim açısı hesaplanmasına ihtiyaç duymayan yöntemlerdir.

4.2 Tanımlayıcılar

4.2.1 Yönelim Açısı Hesaplanan Yöntemler

Literatürdeki yöntemlerin çoğunda tanımlayıcıların açı değişimlerinden etkilenmemesi için ilgi noktalarının tanımladığı görüntü bölgelerinden yönelim hesaplanması ve hesaplanan yönelim açısı kullanılarak görüntü bölgesinin açısının düzeltilmesi gerekmektedir. Açısal değişimlerin olduğu görüntülerde, açısı düzeltilmiş görüntü bölgesinden elde edilen tanımlayıcıların eşleştirme performansı da belirgin şekilde artmaktadır. Yönelim açısının hesaplanmasının tanımlayıcı eşleştirme performansını önemli ölçüde arttırdığı gösterilmiştir (Bellavia vd., 2014). Sıklıkla kullanılan SIFT, SURF, ORB, FREAK, BRISK gibi yöntemlerin kendi yönelim açısı hesaplama süreçleri vardır.

ORB algoritmasında, görüntü bölgesinin yönelimini bulmak için bu bölgenin ağırlık noktası bulunur. Ağırlık noktası ile merkez nokta arasındaki açı, bu görüntü bölgesinin yönelimini belirlemektedir. Bu hesap görüntü piksel değerlerinin kullanılması ile elde edildiğinden ekstra işlem maliyeti içermemektedir.

FREAK ve BRISK yöntemleri ise birbirlerine benzer şekilde yönelim hesabı yapmaktadır. Bu yöntemlerde önceden belirlenmiş örnekleme bölgeleri kullanılmaktadır. Bu örnekleme noktaları arasında dikey ve yatay yönde farklar hesaplanarak iki değer belirlenir. Bu değerler koordinatları (x, y) olan bir noktaya karşılık gelir. Bulunan nokta ile merkez nokta arasındaki açı değeri yönelim açısı olarak kullanılır.

4.2.1.1 SIFT

Lowe (2004) tarafından geliştirilen SIFT algoritması günümüzde halen çok popülerdir. Bu çalışma kapsamında önerilen tanımlayıcılar diğer araştırmacılar tarafından da

kullanılarak deęişik türevleri üretilmiştir (Abdel-Hakim ve Farag, 2006; Bay vd., 2008; Ke ve Sukthankar, 2004; Morel ve Yu, 2009; Mortensen vd., 2005; Zhang vd., 2013).

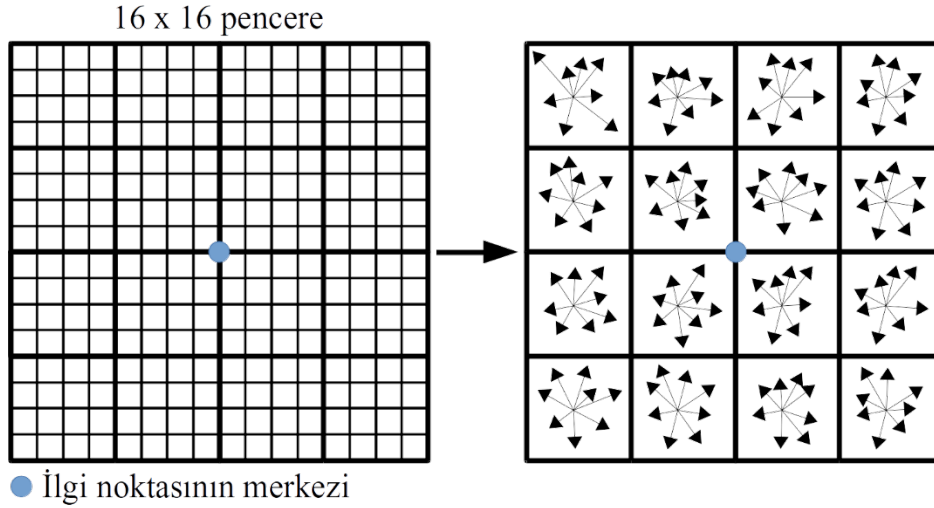
Bu algoritmada ilk olarak ölçekten bağımsız ilgi noktaları tespit edilir. Ölçekten bağımsız ilgi noktaları DoG kullanılarak hesaplanmaktadır. Belirlenen ilgi noktalarından elde edilen görüntü bölgelerini açıdan bağımsız hale getirebilmek için ilgi noktaları etrafında baskın yönelim açısı Denklem (4.1) ve yönelim büyüklükleri Denklem (4.2) ile hesaplanır.

$$\theta(x, y) = \arctan\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \quad (4.1)$$

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (4.2)$$

Burada L Gauss filtresi uygulanmış görüntüyü ifade etmektedir. θ ve m değerleri ilgi noktasının tanımladığı görüntü bölgesindeki pikseller için hesaplanır. Görüntü bölgesinin merkez noktası etrafında d yarıçaplı alandaki m değerleri için Gauss çekirdeęi kullanılarak ağırlıklandırılmış değerler elde edilir. Bu ağırlıklandırılmış m değerleri kullanılarak 10'ar derecelik açı gruplarından oluşan yönelim histogramı elde edilir. Elde edilen bu histogramın en büyük değerinin bulunduğu nokta baskın açı olarak kullanılır. Ayrıca bu algoritmada başarımın artırılması için ikincil açı değerleri de kullanılmaktadır. İkincil açılar en yüksek histogram değerinin %80'inden büyük açı değerlerinin, toplamda en fazla 3 açı değeri olacak şekilde alınması ile elde edilir. Lowe (2004) çalışmasında ikincil açı değerlerinin kullanımı ile başarımın arttığını göstermiştir.

Baskın açı değerlerinin hesaplanmasından sonra bulunan her bir açı değeri için farklı tanımlayıcılar oluşturulur. Her açı değeri için görüntü bölgesi açı miktarı kadar döndürülür. Yönlendirilmiş görüntü bölgesi 16x16 boyutunda parçalara bölünür. Bu 16x16 boyutundaki parçalardan 4x4 boyutlarında 16 farklı grup oluşturulur. 4x4 boyutundaki her bir gruptan 45'er derecelik açı gruplarından oluşan bir yönelim histogramı oluşturulur (Şekil 4.2).



Şekil 4.2. Görüntü bölgesinin bölünmesi ve histogram çıkarımı

Yönelim açısının hesaplanmasında olduğu gibi burada kullanılan m değerleri Gauss çekirdeği kullanılarak ağırlıklandırılmıştır. Bu sayede merkez noktasından uzak olan piksellerin etkisi daha düşük olmaktadır. Sonuçta toplamda 128 değerli bir vektör elde edilir. Bu vektörler 0 – 1 aralığına ölçeklendirilir. Lineer olmayan aydınlatma etkilerini azaltmak için vektör değerlerine 0,2’lik eşik uygulanır ve vektör tekrar 0 – 1 aralığına ölçeklendirilerek normalize edilmiş SIFT tanımlayıcıları elde edilir.

4.2.1.2 ORB

ORB algoritması (Rublee vd., 2011), FAST (Rosten ve Drummond, 2006) ve BRIEF (Calonder vd., 2012) algoritmalarının melez kullanımıyla elde edilmiş bir tanımlayıcıdır. ORB algoritması açılı değişimlerinden etkilenen BRIEF tanımlayıcılarını, yönelim bulma adımı ekleyerek geliştirmiştir. Algoritmada ilk olarak FAST ile ilgi noktaları bulunur ve ardından bulunan bu ilgi noktalarının tanımladığı görüntü bölgesinden yönelim açısı hesaplanır. Bulunan açı kullanılarak açısı düzeltilen görüntü bölgesinden, BRIEF tanımlayıcılarına benzer bir yöntemle tanımlayıcılar çıkarılır. ORB’nin işlem yükü SIFT’e göre daha azdır.

Çıkarılan bu ilgi noktalarından yönelim hesaplanması gerekmektedir. ORB algoritmasında kullanılan yönelim bulma algoritması basit olduğu halde güçlü bir yöntemdir. Herhangi bir ön işlem olmadan görüntü piksel değerleri kullanılarak yönelim açısı hesaplanır (Rosin, 1999). İki boyutlu bir görüntüde piksel değerlerinin ilişkin

moment değerlerini,

$$m_{pq} = \sum_{x,y} x^p y^q I(x,y) \quad (4.3)$$

ile hesaplanmıştır. $I(x,y)$ görüntü bölgesinin (x,y) noktalarındaki değerlerini ifade eder. Momentler kullanılarak görüntünün ağırlık merkezi polar koordinatlarda Denklem (4.4)'deki gibi gösterilir.

$$C\left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}\right) \quad (4.4)$$

C ağırlık merkezi ile görüntü bölgesinin merkezi arasındaki açı, yönelim açısını verir ve

$$\theta = \text{atan2}(m_{01}, m_{10}) \quad (4.5)$$

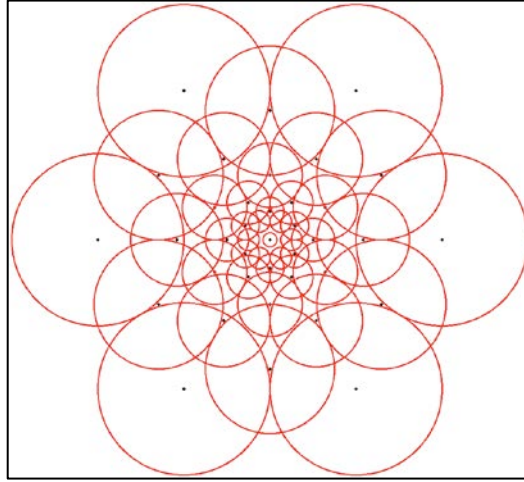
ile hesaplanır. Yönelim açısı hesaplanan görüntü bölgesinden tanımlayıcının çıkarımına geçilir. ORB tanımlayıcıları, literatürde ikili tanımlayıcılar olarak geçer. Görüntü bölgesindeki gürültüyü azaltmak için Gauss filtresi kullanılarak yumuşatılır. Bu sayede ORB yönteminde algoritmada görüntü bölgesindeki piksel değerleri doğrudan kullanılır.

BRIEF algoritmasında, n adet piksel noktası çifti Gauss dağılımı ile seçilir. Dolayısıyla merkez noktasına yakın piksel noktalarının seçilme olasılığı daha yüksektir. BRIEF ile ORB arasındaki bir farkta bu noktaların seçilmesindedir. BRIEF algoritmasında bu noktaların seçimi rastgele olmasına rağmen, ORB öğrenilmiş nokta çiftleri kullanmaktadır. Bu nokta çiftlerinde yer alan piksel değerleri karşılaştırılarak bir tanımlayıcı oluşturulur. Karşılaştırmada ilk nokta, ikinci noktadan büyük ise 1 küçük ise tanımlayıcıya 0 eklenir. Oluşturulan tanımlayıcıların boyutu da böylece daha küçük olmaktadır.

4.2.1.3 FREAK

FREAK (Alahi vd., 2012), karşılaştırmada kullanacağı simetrik görüntü bölgelerinin merkez noktalarını seçmek için ORB'ye benzer şekilde, makine öğrenmesi

kullanılmaktadır. Bu çalışmada karşılaştırılan simetrik bölgeler Şekil 4.3'te verilmiştir.



Şekil 4.3. FREAK algoritmasında kullanılan örnekleme bölgeleri (Alahi vd., 2012)

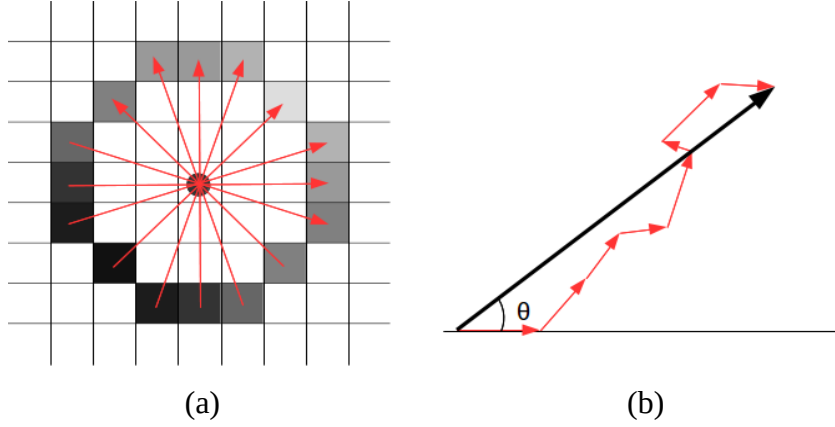
FREAK algoritmasında yönelim açısı, elde edilen 45 simetrik bölge kullanılarak,

$$O = \frac{1}{M} \sum_{P_o \in G} (I(P_o^{r1}) - I(P_o^{r2})) \frac{P_o^{r1} - P_o^{r2}}{\|P_o^{r1} - P_o^{r2}\|} \quad (4.6)$$

ile hesaplanır. Burada M bölge çiftlerinin sayısını, G kullanılan bölge çiftlerini ve P_o^i bölge çiftlerinin merkez konumlarını ifade etmektedir. Hesaplanan açı değeri kullanarak doğrultulan görüntü bölgesinden, ORB'de kullanılan karşılaştırma yöntemine benzer bir yöntem kullanılarak tanımlayıcılar elde edilir.

4.2.1.4 Taylor Yönelim Hesaplama Yöntemi

Taylor vd. (2009) tarafından yapılan çalışmada hızlı kenar bulma algoritması önerilmiştir. Bu çalışmada kullanılan kenar bulma çekirdeği kullanılarak simetrik pikseller arasında yatay ve dikeyde doğrultuda farklar hesaplanmıştır (Şekil 4.4).



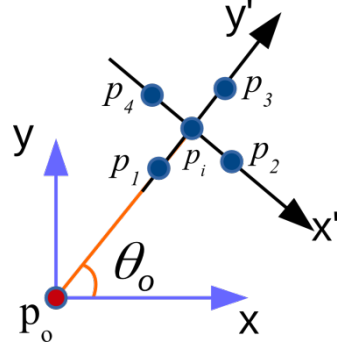
Şekil 4.4. Taylor yönelim hesaplama algoritmasında kullanılan a) çekirdek düzeni b) vektörlerin toplanması

Burada, görülen çekirdek düzeni üzerinde okla eşleştirilen her iki piksel arasındaki fark hesaplanır. Daha sonra hesaplanan bu değerler uç uca eklenerek vektörel toplam bulunur (Şekil 4.4(b)). Yönelim açısı (θ) elde edilen bu vektörden hesaplanmaktadır.

4.2.2 Yönelim Açısı Hesaplanmayan Yöntemler

Literatürde yönelim açısı hesaplanmadan, açıdan bağımsız öznelilikler ile elde edilen MROGH (Fan vd., 2011), OSRI (Xu vd., 2014), MRRID (Fan vd., 2012) gibi yöntemlerde mevcuttur. Bunlar yerel piksel değeri gruplandırması (Local Intensity Orders) ile elde edilen ayrı görüntü parçalarının histogramlarının farklı destek bölgelerinden (Support regions) çıkarılması ile hesaplanmaktadır. Bu yöntemler yönelim hesaplanması adımını içermemelerine rağmen hesaplama süresi olarak diğer yöntemlerden daha yavaştır.

Bahsedilen yönelim açısı hesabı gerektirmeyen tanımlayıcıların çıkarımında kullanılan yönelimli genliklerin, açısal değişimlerden etkilenmemesi gerekmektedir. Açısal değişimden bağımsız yönelimli genlikler Schmid vd. (2005) tarafından önerilmiş olan yönteme oldukça benzer bir şekilde gerçekleştirilmiştir. Schmid vd. (2005) yaptıkları çalışmada her bir pikselden hesaplanan genlikleri ve yönelimleri yerel koordinat sistemi kullanarak elde etmişlerdir. Bu koordinat sistemi Şekil 4.5'te gösterilmiştir.



Şekil 4.5. Yerel koordinat sistemi

Burada p_0 görüntü parçasının merkezini, p_i ise genlik ve yönelimin hesaplanacağı ilgili pikseli göstermektedir. Görüntü parçasında yer alan her bir piksel için hesaplanması gereken üç parametre vardır: pikselin görüntü parçasının merkezi ile yaptığı açı (θ_0), açı bağımsız genlik değeri (Rotated magnitude - Rm) ve açı bağımsız yönelim (Rotated gradient - Rg) değeri. Denklem (4.7) ile hesaplanan θ_0 değeri her bir pikselin, merkez piksele göre hangi açısal konumda yer aldığının bulunması için kullanılmaktadır.

$$\theta_0 = \tan^{-1} \left(\frac{p_{iy}}{p_{ix}} \right) \quad (4.7)$$

p_{iy} , p_{ix} değerleri ilgili pikselin, görüntü parçasının merkezine göre x ve y koordinatlarını temsil etmektedir. θ_0 değerinin bulunmasının ardından Rg ve Rm değerleri hesaplanır. Bu değerlerin hesaplanabilmesi için koordinat sistemi (x, y) den (x', y') 'a dönüştürülür. Koordinat sistemi dönüşümünden sonra p_i pikseli ve komşu pikselleri $\{p_1, p_2, p_3, p_4\}$, Şekil 4.5'te gösterilmiştir. Bu komşu piksellerin koordinatları θ_0 değerine bağlı olarak Denklem (4.8) ve (4.9) ile hesaplanır.

$$p_{jx} = p_{ix} - \cos \left(\theta_0 + (j - 1) \times \frac{\pi}{2} \right), \quad j = 1, 2, 3, 4 \quad (4.8)$$

$$p_{jy} = p_{iy} - \sin \left(\theta_0 + (j - 1) \times \frac{\pi}{2} \right), \quad j = 1, 2, 3, 4 \quad (4.9)$$

Bu hesaplamalardan anlaşılacağı üzere bulunan x ve y koordinatları alt-piksel noktalarına denk gelebilmektedir. Bu durumlarda komşu piksel değerleri doğrusal ara değer bulma uygulanarak belirlenir. $\{p_1, p_2, p_3, p_4\}$ değerleri kullanılarak Rm ve Rg değerleri,

$$dx = p_2 - p_4 \quad (4.10a)$$

$$dy = p_1 - p_3 \quad (4.10b)$$

$$Rg = \text{atan2}(dy, dx) \quad (4.10c)$$

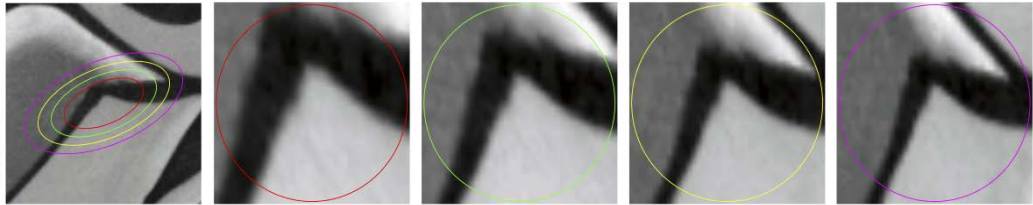
$$Rm = \sqrt{dy^2 + dx^2} \quad (4.10d)$$

ile hesaplanmaktadır. Bu değerler görüntüde oluşabilecek açı değişimlerinden bağımsızdır. Bu değerler kullanılarak yeni tanımlayıcılar elde edilebilir.

4.2.2.1 MROGH

Bu yöntemde ilgi noktalarından elde edilen görüntü bölgeleri diğer yöntemlerden yaklaşık olarak 1.2 kat büyüktür. Görüntü bölgesinde yer alan her piksel için Denklem (4.10c) ve (4.10d) ile gösterilen Rg ve Rm değerleri hesaplanmaktadır.

Daha sonra görüntü bölgesi, alt bölgelere ayrılarak destek bölgeleri elde edilir. MROGH, her destek bölgesinden yerel piksel değer gruplandırması ile elde edilen iç bölgelerin histogramlarını birleştirir. Destek bölgeleri Şekil 4.6'de gösterilmiştir.



Şekil 4.6. Destek bölgeleri ve normalize edilmiş görüntü bölgeleri (Fan vd., 2011)

Her bir destek bölgesinden Şekil 4.7'de gösterilen yerel piksel değer grupları elde edilir. Bu yerel piksel değer grup bölgelerinden Rg ve Rm (4.10c) ve (4.10d) değerleri kullanılarak, histogram değerleri elde edilir.



Şekil 4.7. Piksel değer gruplandırması (Fan vd., 2011)

Daha sonra bu histogramlar art arda eklenerek tanımlayıcı oluşturulur.

4.2.2.2 MRRID

MRRID algoritmasının, MROGH algoritmasına benzer şekilde bir tanımlayıcı çıkarım düzeni bulunmaktadır. Fakat çıkarılan histogramlar, Rg ve Rm değerlerini değil piksel değerlerini kullanır. Şekil 4.5'te görülen komşu piksel değerleri $\{p_1, p_2, p_3, p_4\}$ kullanılarak, CS-LBP (Heikkilä vd., 2009) çalışmasındaki yöntem ile hesaplanır. Sonrasında bu özniteliklerin histogramı oluşturulur. Destek bölgelerinden ve piksel grup bölgelerinden ile elde edilen bu histogramlar birleştirilerek MRRID tanımlayıcıları elde edilir.

BÖLÜM V

ÖNERİLEN YÖNTEMLER VE BENZETİM SONUÇLARI

5.1 Yer İzi Tespit Yöntemi

Bu bölümde yer izi (İlgi noktası, öznitelik noktası) tespiti için tez çalışmasında önerilen Temel piksel değer karşılaştırmaları yöntemi sunularak SIFT ile karşılaştırılacaktır.

Temel piksel değer karşılaştırmaları (TPDK) yöntemi hem gerçek zamanlı çalışmaya uygun hem de desen içermeyen görüntülerde bile yüksek başarımlı göstermektedir. Bu yöntem sunulmadan önce en hızlı yöntemlerden biri olarak bilinen FAST ve başarımlı konusunda referans olan SIFT yöntemleri incelenmiştir. FAST piksel tabanlı işlemler ve basit karşılaştırmalar ile köşelerin bulunmasını amaçlamaktadır. FAST yöntemi hızlı olsa da köşe bilgilerinin olmadığı durumlarda çok düşük performans sergilemektedir. Fakat bu çalışmada ihtiyaç duyulan algoritma, köşe bilgilerinin yetersiz olduğu durumlarda bile öznitelik çıkarabilmelidir.

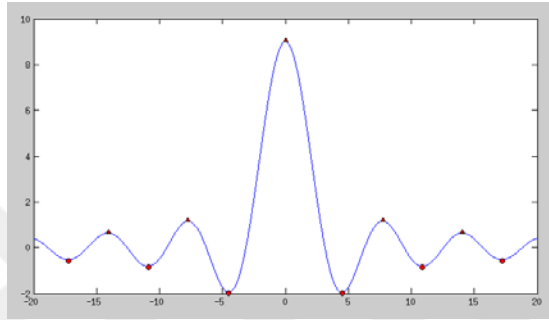
Görüntüden öznitelik çıkarımında genellikle SIFT algoritması referans olarak alınmaktadır. SIFT algoritması, DoG (Lowe, 2004) üzerindeki yerel uçdeğer noktalarını bulmaktadır. Daha sonra bulunan bu noktalardan, öznitelik çıkarılacak olan bölgeler, boyut uzayının yerel maksimum noktalarına bakılarak belirlenmektedir. Bu sayede bulunan öznitelikler hem boyuttan bağımsız hem de köşe ve bölge bilgisi içermektedir. Fakat SIFT algoritmasının zayıf olduğu nokta, desen içermeyen görüntülerde oldukça az sayıda öznitelik bulması veya hiç bulamamasıdır. Ayrıca bu algoritma yüksek işlem yükünden dolayı gerçek zamanlı çalışmaktan uzaktır. Bu tez kapsamında yapılan çalışmalarda ise desen bulunmayan bölgelerde dahi öznitelik çıkarımının artırılması hedeflenmektedir.

5.1.1 Temel piksel değer karşılaştırmaları (TPDK)

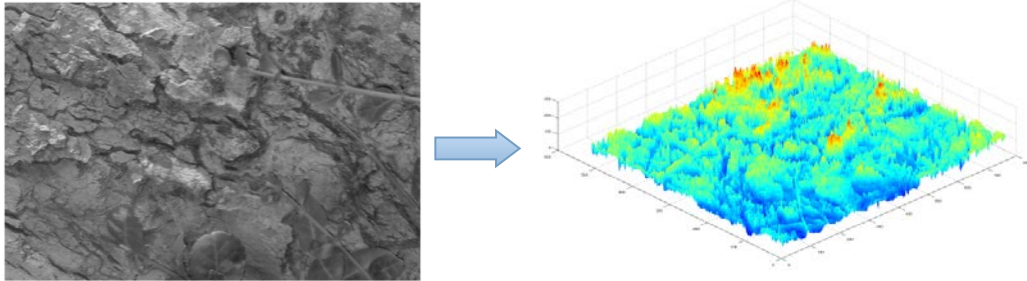
Bu çalışmada öznitelik bölgesi olarak, görüntü içerisindeki yerel minimum ve yerel maksimum noktalarının bulunması amaçlanmıştır. Bu noktalar görüntü içerisinde tekrarlanabilir noktalar olarak görülmektedir. Bir boyutlu sinyaller göz önüne

alındığında yerel maksimum ve minimum noktalarının boyutlandırma, öteleme gibi durumlarda konumlarının yeniden bulunabilirliğinin değişmediği gözlemlenmiştir. Bu nedenle öznitelik noktalarının önemli bir özelliği olan tekrar edilebilirlik açısından bu noktalar kullanılabilir olarak öne çıkmaktadır.

Bir boyutlu (1D) sinyal için yerel minimum ve maksimum noktaları Şekil 5.1’de gösterilmiştir. İki boyutlu bir görüntü ele alındığında oluşacak olan üç boyutlu gösterim ise Şekil 5.2’de verilmiştir.



Şekil 5.1. Bir boyutlu (1D) sinyalin yerel minimum ve maksimum noktaları

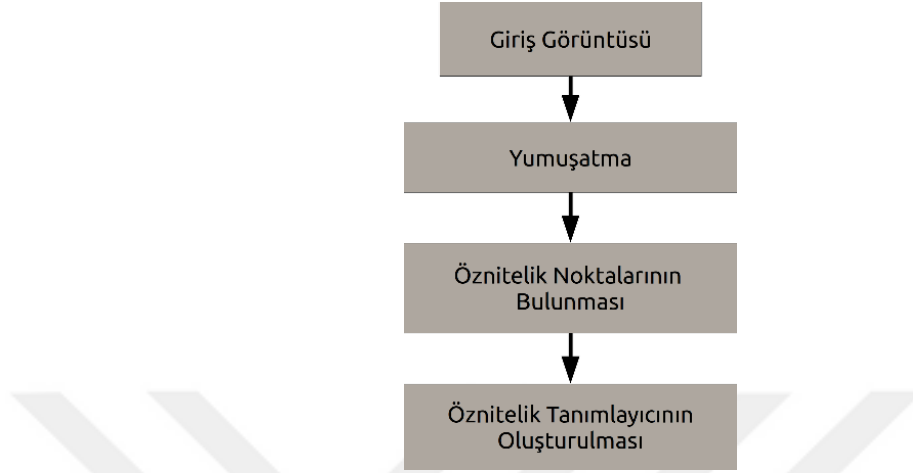


Şekil 5.2. İki boyutlu görüntünün üç boyutlu olarak gösterimi

Burada değişimin olduğu noktaların bulunması tanımlayıcı oluşturulması açısından önemlidir. Eğer tanımlayıcının hesaplanacağı bölgede ayırt edici özellikte değişimler bulunmaz ise resimler arası eşleştirme sürecinde hatalarla karşılaşılacaktır. Bu nedenle eşleştirme performansı düşmektedir. Eşleştirme performansı iki görüntü arasındaki doğru eşleştirme sayısını gösterir. Bu değer ne kadar büyük olursa hesaplanacak olan yer değiştirmede o kadar başarılı olacaktır.

Bahsedilen öznitelik noktalarını bulmak için görüntü üzerinde yatay ve dikey doğrultularda görüntünün yerel minimum/maksimum noktalarının belirlenmesi gerekir.

Fakat bu öznitelik noktalarının kararlı olabilmesi, görüntüde olabilecek gürültülerin bastırılması ile mümkündür. Bu doğrultuda görüntü üzerinde yapılacak olan işlemlere ait akış şeması Şekil 5.3'te gösterilmiştir.



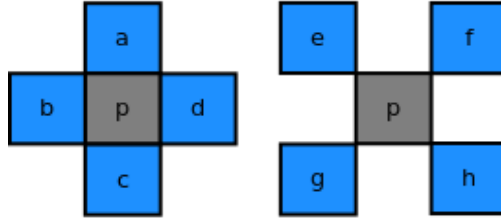
Şekil 5.3. TPDK yöntemin akış şeması

İlk olarak giriş görüntüsüne, gürültü kaynaklı hataları en aza indirebilmek için yumuşatma işlemi uygulanmalıdır. Bu sebeple giriş görüntüsüne Gauss filtresi uygulanarak görüntü içerisinde tepe ve vadi noktalarındaki geçişler yumuşatılmaktadır. Uygulanan Gauss yumuşatma filtresi (Denklem (3.17)) 3x3 boyutunda ve σ katsayısı ise 0.8 olarak seçilmiştir. Bu seçim farklı değerler kullanılarak elde edilmiştir. Giriş görüntüsüne, Denklem (5.1) ile verilen çekirdek kullanılarak konvolüsyon işlemi uygulanır.

$$k = \begin{bmatrix} 0.057 & 0.125 & 0.057 \\ 0.125 & 0.273 & 0.125 \\ 0.057 & 0.125 & 0.057 \end{bmatrix} \quad (5.1)$$

Görüntüye yumuşatma uygulanması, öznitelik çıkarımında gürültüden kaynaklanabilecek hataları azaltmaya yardımcı olur. Ardından öznitelik noktalarının tespit edilebilmesi için yatay ve dikey doğrultularda yerel minimum/maksimum noktalarının bulunması gerekir. Fakat bu noktaların sadece yatay ve dikey doğrultuda tespit edilmesi, görüntüde karşılaşılabilecek olan açısallı değişimlerde, özniteliklerin tekrar bulunmasını zorlaştıracaktır. Dolayısıyla oluşabilecek 45° ve katlarındaki açısallı değişiklikleri kapsayacak iki ayrı karşılaştırma çekirdeği önerilmiştir (Şekil 5.4). Çevresel piksel değerlerinin, merkez p piksel değerinden büyük veya küçük olduğu

noktalar öznitelik çıkarımı açısından önem arz etmektedir.



Şekil 5.4. Minimum/maksimum noktalarının bulunduğu piksel yerleşimleri

Şekil 5.4’te gösterilen çekirdekler kullanılarak aşağıdaki koşullara göre p noktası yerel uç nokta olarak belirlenmektedir. Burada;

1. $p < (a,b,c,d)$ veya $p < (e,f,g,h)$, ise yerel en küçük,
2. $p > (a,b,c,d)$ veya $p > (e,f,g,h)$, ise yerel en büyük.

Bu kurallar uygulanarak Şekil 5.5’te gösterilen öznitelik noktaları bulunmuştur.



Şekil 5.5. TPDK ile bulunan örnek öznitelik noktaları

Öznitelik noktalarının bulunmasından sonra bu noktaların tanımlayıcısının

oluşturulması gerekmektedir. Karşılaştırmalarda benzer test koşulları oluşturulabilmesi açısından SIFT tanımlayıcıları kullanarak algoritmanın başarımı test edilmiştir. SIFT tanımlayıcılarının açıdan bağımsız eşleştirme yapabilmesi için merkez yönelim açısının belirlenmesi gerekmektedir. Bu açı ilgili piksel çevresindeki bölgede baskın açı değeri belirlenerek bulunmaktadır. Dolayısıyla her yerel uç noktası için 8x8'lik bir çerçevede baskın açı değeri hesaplanır. Bu değer hesaplanması Bölüm 4.2.1'de anlatılmıştır.

Ayrıca SIFT tanımlayıcılarının çıkarılacağı bölgenin boyutunun belirlenmesi gereklidir. TPDK yönteminde öznitelik çıkarılacak bölgenin boyutu sabit olarak alınmıştır ve 24 piksel genişliğindedir. Burada bölge boyutu sabit olarak seçildiği için TPDK boyut değişimlerinden etkilenmektedir.

5.1.2 Benzetim Sonuçları

TPDK algoritmasının test edilebilmesi için Mikolajczyk (2005) tarafından önerilen *Oxford* veri kümesi kullanılmıştır. Veri kümesi boyut değişimi, bulanıklık, ışık değişimleri vb. birçok durum için artımlı bozulmalar içeren görüntü kümelerinden oluşmaktadır. TPDK algoritmasının performans benzetimlerinde kullanılan *Oxford* veri kümesine ait görüntü kümeleri Çizelge 5.1'de verilmiştir. Bunlardan; *Bikes* ve *Trees* görüntüleri artımlı yumuşatma, *Day&Night* ve *Leuven* görüntü kümeleri ışık değişimi, *Ceiling*, *Rome* ve *Semper* görüntü kümeleri açısal değişim içermektedir.

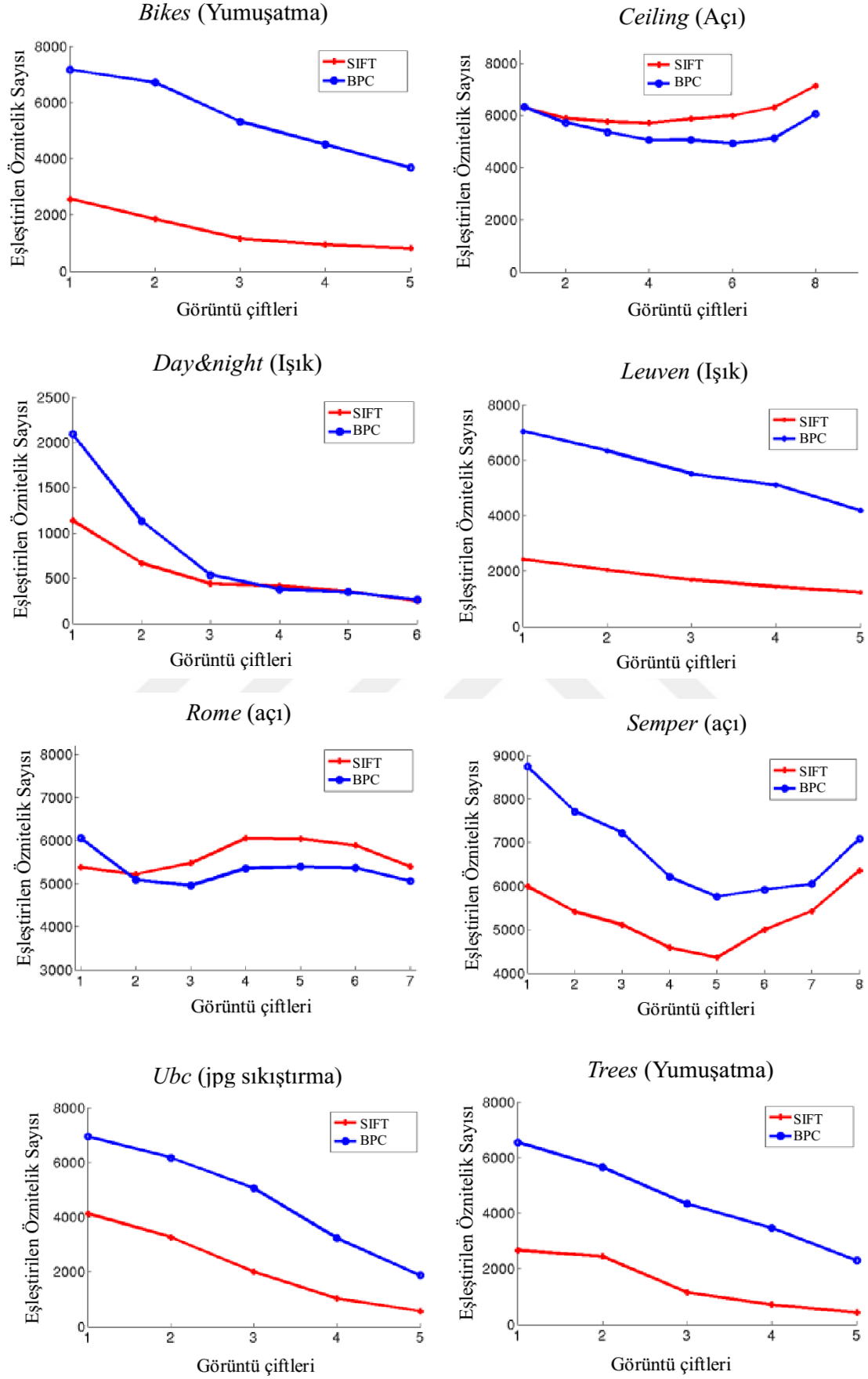
Görüntü kümelerinde yer alan görüntü sayıları değişiklik göstermektedir. Benzetim sonuçlarının verildiği Şekil 5.6'da grafik eksenleri *Eşleştirilen öznitelik sayısı* ve *Görüntü çiftleri*'dir. Görüntü kümelerindeki ilk görüntü ile aynı kümedeki diğer görüntüler arasında eşleştirilen yer izi sayıları incelenmiştir. *Görüntü çiftleri* ekseninde yer alan değerler eşleştirilen görüntü çiftinin kaçınıcı çift olduğunu belirtmektedir. Bu ekseninde yer alan 1 değeri 1. ve 2. görüntü arasındaki eşleştirmeyi, 2 değeri 1. ve 3. görüntü arasındaki eşleştirmeyi vb. ifade etmektedir.

Grafiklerden görüldüğü gibi *Bikes*, *Trees*, *Ubc*, *Leuven*, *Day&Night* ve *Semper* görüntü kümelerinde artan bozunumlara karşılık TPDK algoritması tarafından tespit edilen yer izlerinin sayısı SIFT algoritmasına göre 2-3 kat daha fazladır. *Rome* ve *Ceiling* görüntü kümelerinde ise TPDK algoritmasının eşleştirme sayısı SIFT'e nazaran daha düşüktür.

Bunun sebebi ise TPDK'nın açısız deęişimlere karşı hassas olmasıdır.

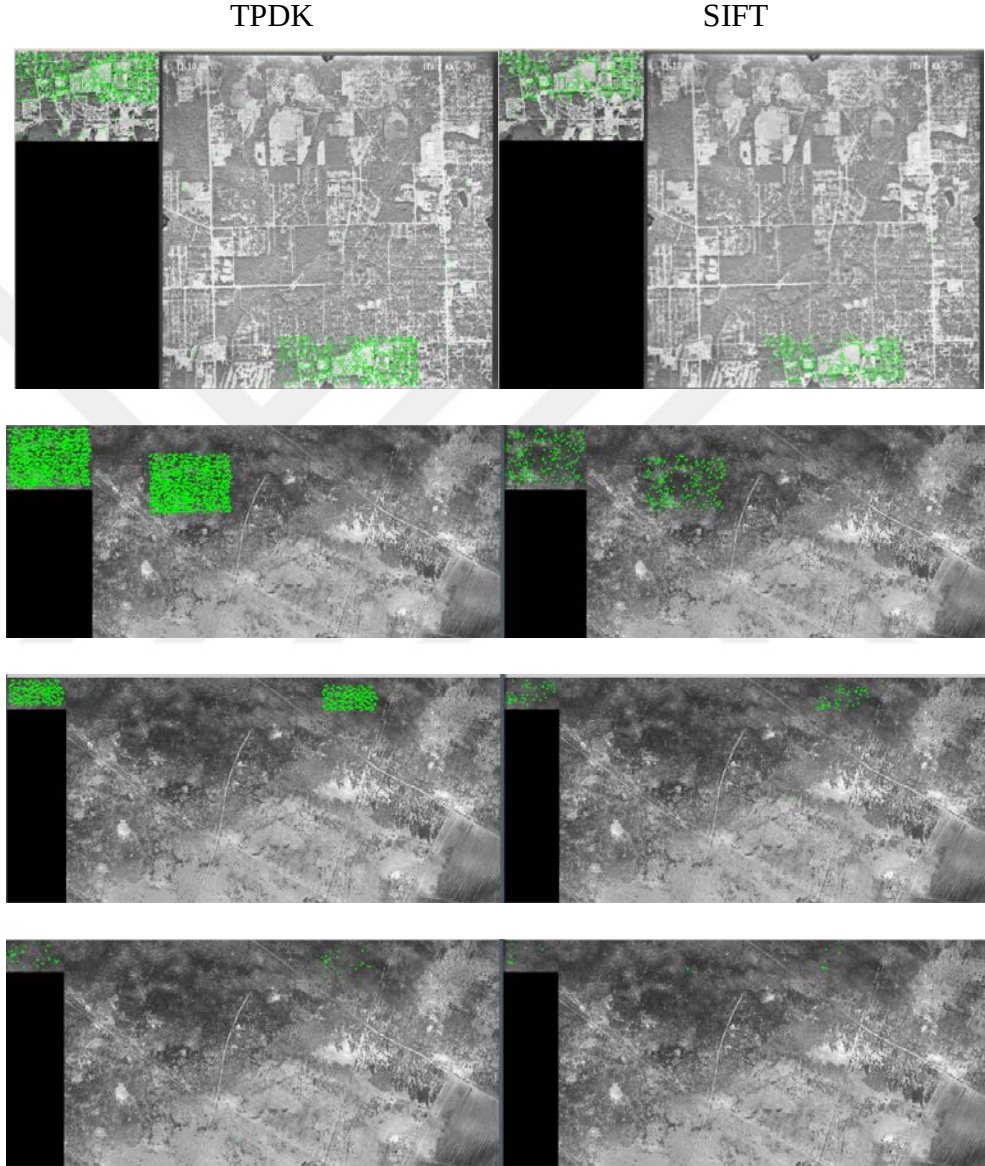
Çizelge 5.1. Yer izi tespit benzetimlerinde kullanılan *Oxford* veri kümesinde yer alan görüntü kümeleri

<i>Bikes</i> (Yumuşatma)	<i>Ceiling</i> (Açı)
	
<i>Day&night</i> (Işık)	<i>Leuven</i> (Işık)
	
<i>Rome</i> (açı)	<i>Semper</i> (açı)
	
<i>Trees</i> (Yumuşatma)	<i>Ubc</i> (jpg sıkıştırma)
	



Şekil 5.6. Eşleştirilen öznitelik grafikleri

Bu veri kümelerinden sonra, TPDK algoritmasının özellikle fazla desen içermeyen görüntülerde çalışması gözlemlenmiştir. Bunun için Google Haritalar® yardımıyla alınan görüntüler üzerinde TPDK ve SIFT kullanılarak çıkarılan tanımlayıcılar eşleştirilerek küçük bir parçanın görüntü içerisinde bulunması amaçlanmıştır. Elde edilen sonuçlar Şekil 5.7’de verilmiştir.



Şekil 5.7. TPDK ve SIFT’in eşleştirme performanslarının Google Haritalardan elde edilen görüntüler üzerindeki test sonuçları

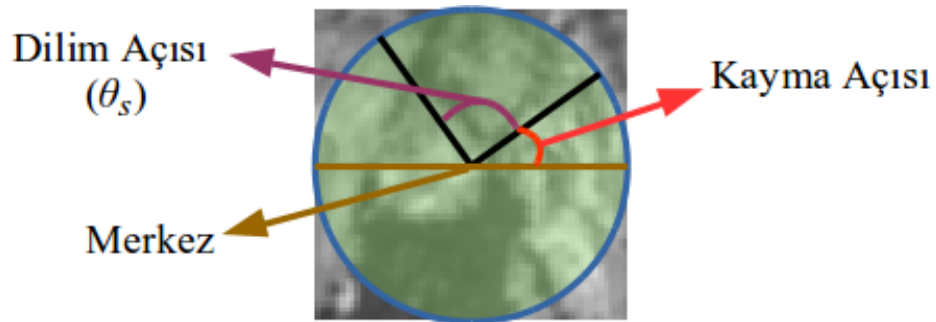
Bu görüntülerde Şekil 5.7(a)’da görülen eşleştirme sayısı TPDK ile 1468, SIFT ile 781 olarak elde edilmiştir. Nispeten daha az desen içeren diğer görüntülerde ise sırasıyla TPDK ile 1496, 364 ve 29 eşleştirme sağlarken, SIFT ile sırasıyla 229, 52 ve 9

eşleştirme elde edilmiştir.

Her ne kadar TPDK ile daha fazla ilgi noktası bulunabilse de bazı durumlarda bu öz nitelik sayısının azaltılması gerekmektedir. Bu durumlar için ise kenar olma ölçütü olarak Shi-Tomasi tarafından önerilen matris kullanılmaktadır. Bu matrisin kullanımı ile kenar olma derecesi yüksek olan ilgi noktaları tutularak gerekli sayıda ilgi noktası elde edilebilmektedir.

5.2 Dilimlere Ayrılmış Yoğunluk Histogramı

Dilimlere ayrılmış yoğunluk histogramı (Segmented Intensity Histogram - SIH) yöntemi ile öz nitelik noktalarından elde edilen görüntü parçalarının yönelim açılarının etkin bir şekilde bulunması amaçlanmaktadır. Bu bağlamda görüntü parçasının piksel değerleri toplamının yüksek olduğu görüntü dilimlerinin bulunması amaçlanmıştır. Görüntü parçasından alınacak olan dilimler Şekil 5.8’de gösterilen merkez açısı ve kayma açısı ile tanımlanabilir. Dilimler tarafından taranan bölgede kalan piksel değerleri toplanarak bir değer elde edilir. Bu bölgenin merkez açısı sabit kalmak koşulu ile, kayma açısı 1° adımlarla ve 0° - 360° aralığında değiştirilerek, bütün görüntü parçası taranır. Böylece görüntü parçasından bir histogram elde edilir. Bu histogramın tepe değerinin bulunduğu açı, yönelim açısı olarak tanımlanmıştır. Ayrıca SIH, sadece temel matematiksel işlemler ve görüntünün piksel değerlerini kullandığı için işlem yoğunluğu en alt seviyededir.



Şekil 5.8. Görüntü parçasından alınan dilim parametreleri

SIH ile histogram hesaplanabilmesi için görüntü parçasının her bir pikselinin hangi açı konumunda bulunduğunun belirlenmesi gerekir. Bir pikselin açı konumu Denklem (5.2) ile hesaplanmaktadır.

$$\theta = \text{atan2}(y - py, x - px) \quad (5.2)$$

Burada x ve y , işlem yapılacak olan pikselin görüntü parçasındaki konumunu belirtmektedir. px ve py değerleri ise görüntü parçasının orta noktasını belirtmektedir. 0° ile 360° aralığındaki her piksel değeri, ilgili histogram bölgesine artımlı olarak eklenir. Böylece 1° açı aralıklarında oluşturulan bu histogramın boyutu 360 olur. Histogram çıkarımını hızlandırmak için piksellerin açısal konumları önceden hesaplanmış bakma tablosundan (Look-up table) alınmaktadır.

Baskın açının bulunması için, dilim bölgesinde yer alan histogram değerlerinin toplanması gerekmektedir. Bu işlemi hızlandırmak için önceden hesaplanan histogram değerleri kullanılarak integral histogram oluşturulur. Bu sayede her bir görüntü bölgesinin toplam değerini hesaplamak için sadece bir çıkartma işlemi yapılması yeterlidir. İntegral histogramın hesaplanabilmesi için bu çalışmada,

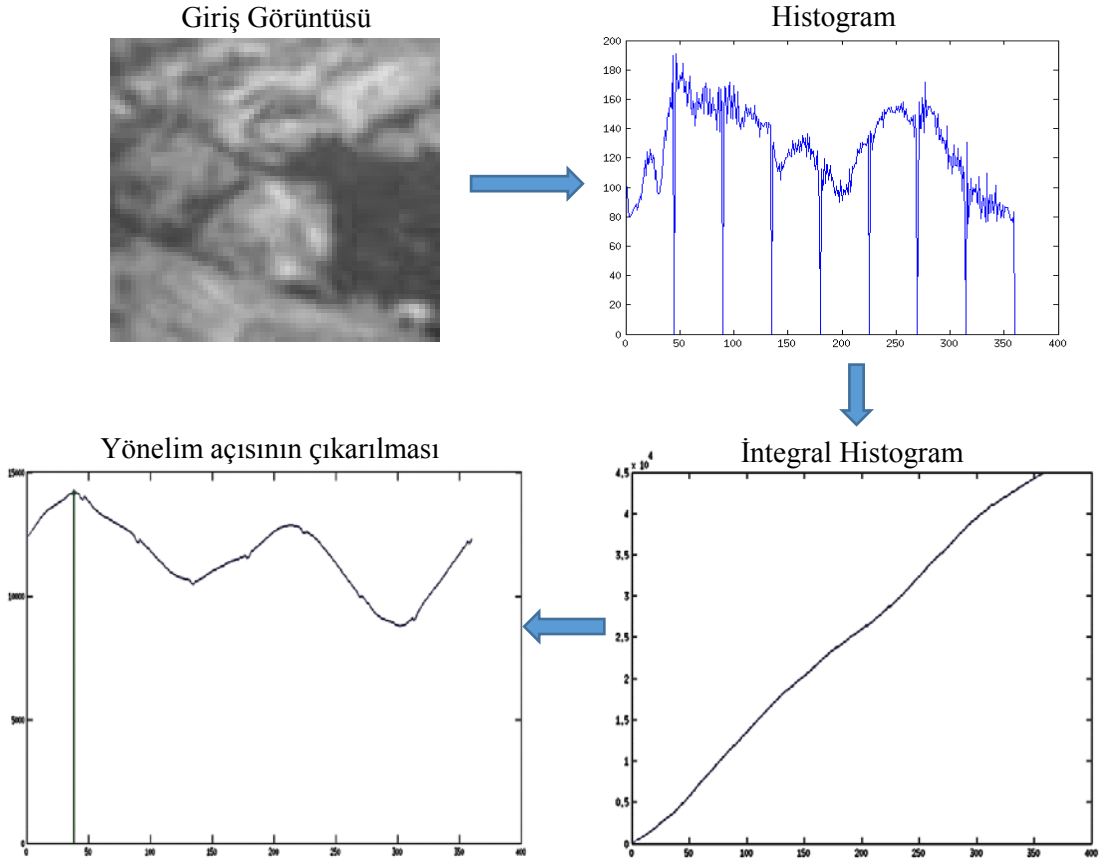
$$H(i) = \sum I(\theta == i) \quad (5.3)$$

$$\text{Yönelim Açısı} = \max \left(\sum_{i=0:360} H(i + ma) - H(i) \right) \quad (5.4)$$

olarak tanımlanmıştır. Burada ma sektör merkez açısı için kullanılmıştır. H hesaplanan görüntü sektörlerinden oluşan artımlı histogramı ifade etmektedir. 0° - 360° aralığındaki dilimlerin toplamaları ise

$$H(i + ma) - H(i) \quad (5.5)$$

ile hesaplanmıştır. Sonuç olarak elde edilen bu değerlerin tepe noktasının bulunduğu histogram noktası, işlenen görüntü parçasının yönelim açısı olarak alınmaktadır. SIH algoritmasında kullanılan işlem adımları Şekil 5.9'da verilmiştir.



Şekil 5.9. SIH algoritmasına ait işlem adımları

Bölüm 5.2.1’de *Oxford* veri kümesi üzerinde SIH ve diğer yönelim bulma algoritmalarının, SIFT tanımlayıcıları ile birlikte kullanılarak yer izi eşleştirme performansları incelenmiştir.

5.2.1 SIH Algoritmasında Parametre Seçimi

SIH algoritmasında dilim açısı ve histogram boyutu performansı etkileyen parametrelerdir. Bu parametrelerin küçük seçilmesi durumunda görüntü parçasında oluşabilecek gürültüler performansı olumsuz etkileyecektir. Çok büyük seçilmesi durumunda ise elde edilecek değerler arasındaki ayırt edicilik azalacağından yönelim açısının belirlenmesi zorlaşacaktır. Parametrelerin performans üzerine etkisini görebilmek amacıyla *Oxford* veri kümesi üzerinde dilim açısı ve histogram boyutunun farklı değerleri için performans sonuçları incelenmiştir. Performans değeri, görüntü kümelerindeki ilk görüntü ile aynı kümedeki diğer görüntüler arasında doğru eşleştirilen yer izlerinin, tespit edilen toplam yer izlerine oranını ifade etmektedir.

Çizelge 5.2. SIH algoritmasının farklı dilim açıları ve histogram boyutları için eşleştirme performans matrisi

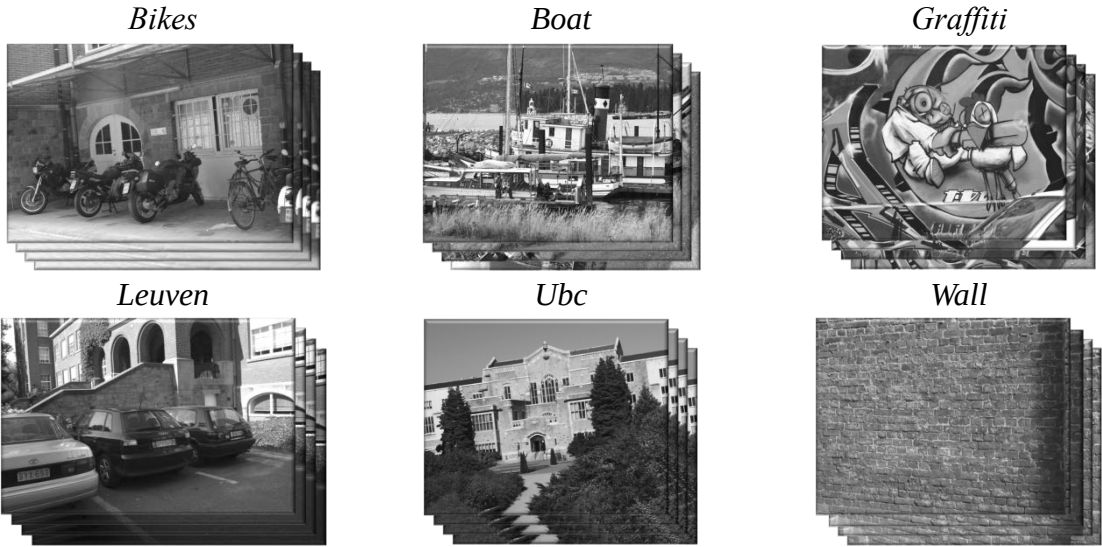
		Dilim Açısı									
		1	10	30	60	90	120	180	240	270	320
Histogram boyutu	120	62.9	60.6	63.8	63.2	63.4	0.0	0.0	0.0	0.0	0.0
	150	62.6	62.2	61.6	62.6	61.9	62.4	0.0	0.0	0.0	0.0
	180	62.9	61.6	60.2	62.4	63.4	62.2	0.0	0.0	0.0	0.0
	210	62.5	61.8	59.6	63.7	61.2	62.0	61.2	0.0	0.0	0.0
	270	59.9	62.9	56.2	59.4	60.8	61.1	62.6	62.0	0.0	0.0
	320	60.4	63.4	58.2	60.4	64.4	62.5	61.9	64.2	59.8	0.0
	360	59.8	61.9	58.3	60.9	64.5	59.2	63.6	62.3	64.2	61.6

Çizelge 5.2’de görüleceği üzere 90° merkez açısına sahip dilim açısı ve 360 histogram boyutu ile en yüksek eşleştirme oranı elde edilmiştir. Bu bölümde belirlenen en uygun parametre değerleri daha sonraki benzetimlerde kullanılmıştır.

5.2.2 Benzetim İçin Kullanılan Veri Seti

Bölüm 5.2’de ve 5.3’te tez kapsamında önerilen yöntemlerin test edilebilmesi için *Oxford* veri kümesi kullanılmıştır. Bu veri kümesi içerisinde boyut değişimi, bulanıklık, ışık değişimleri vb. gibi birçok farklı durum için görüntü kümeleri yer almaktadır. Bu veri kümesine ait görüntüler Şekil 5.10’da verilmiştir. Bunlardan; *Bikes* görüntü kümesi artımlı yumuşatma, *Leuven* görüntü kümesi ışık değişimi, *Boat* görüntü kümesi açısal değişim, *Graffiti* ve *Wall* görüntü kümeleri bakış açısı değişimi, *Ubc* görüntü kümesi ise JPEG bozulması içermektedir.

Yöntemlerin adil koşullar altında test edilebilmesi için tespit edilen öznitelik noktalarının bütün yöntemlerde aynı olması gerekmektedir. Bu nedenle Harris Affine Bölge bulma algoritması (Mikolajczyk vd., 2005) ile elde edilen öznitelik noktaları kullanılmıştır ve bu öznitelik noktaları eliptik bir alanı tanımlamaktadır. Eliptik bölge normalize edilerek, Şekil 5.12’de görüldüğü gibi dairesel yarıçapı 20.5 piksel ve boyutları 41x41 piksel olan görüntüler elde edilmektedir. Bu boyutlar literatürde tanımlayıcı performansını ölçmede sıklıkla kullanılmaktadır.

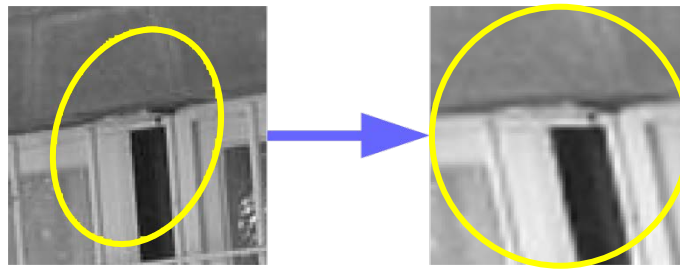


Şekil 5.10. Benzetim için kullanılan *Oxford* veri kümesinde yer alan görüntü kümeleri

Elde edilen eliptik bölge, $A \in R^{2 \times 2}$ simetrik matrisi ile gösterilir. Bu bölgede yer alan her X noktası için,

$$X^T A X \leq 1 \quad (5.6)$$

eşitliğinin sağlanması gerekmektedir. E birim matris, c yarıçap olmak üzere eğer $A = (1/c^2)E$ ise bölge dairesel, aksi durumda ise eliptik bir bölgedir. Burada yapılan işlem ile Şekil 5.11’de görülen eliptik formdaki bölge, dairesel forma dönüştürülmektedir.



Şekil 5.11. Eliptik görüntü bölgesinin normalize edilmesi

Normalize edilmiş bölgedeki noktalar X' ile ifade edilir ve

$$X'^T X' \leq r^2 \quad (5.7)$$

koşulunu sağlamalıdır. Burada r normalize edilmiş bölgenin yarıçapı olmak üzere bu çalışmada 20.5 piksel olarak seçilmiştir. Denklem (5.6) ve (5.7) birleştirilerek,

$$X = 1/r A^{-1/2} X' = T^{-1} X' \quad (5.8)$$

ifadesi elde edilmektedir. Böylece normalize edilmiş bölgedeki her bir X' noktasının, eliptik bölgedeki karşılığı olan X noktası belirlenir ve bu bölgedeki piksel değeri X' noktasına yazılır. Bu denklem $I(X) = I(X')$ ile gösterilebilir. Her bir X' noktası tam sayı olmayabilir. Bu durumda $I(X)$ değeri bilineer ara değerler ile hesaplanır.

Algoritmaların performansları Mikolajczyk (2005) tarafından önerilen yöntem kullanılarak ölçülmüştür. Burada *Geri çağırma* ve *1-kesinlik* olarak gösterilen performans ölçütleri,

$$\text{Geri çağırma} = \frac{\text{doğru pozitif eşleştirmeler}}{\text{yanlış pozitif eşleştirmeler}} \quad (5.9)$$

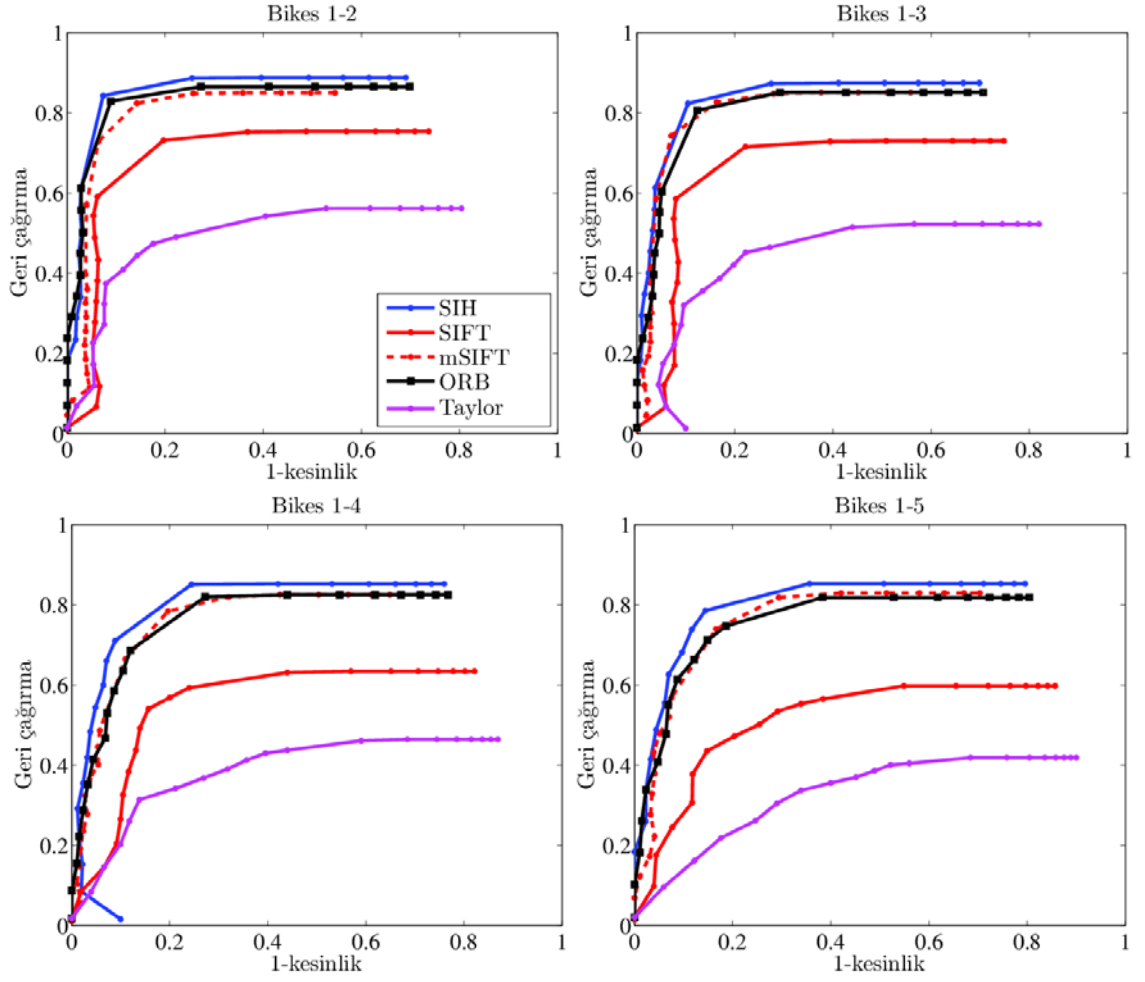
$$1 - \text{kesinlik} = \frac{\text{doğru pozitif eşleştirmeler}}{\text{toplam eşleştirilenler}} \quad (5.10)$$

ile hesaplanmaktadır. Basit anlamda *1-kesinlik* değeri arttıkça *Geri çağırma* değerinin de artması, tanımlayıcıların ayırt edici özelliklerinin yüksek olduğunu gösterir. Bu performans ölçümü daha sonraki bölümlerde tanımlayıcı performans ölçümlerinde de kullanılacaktır.

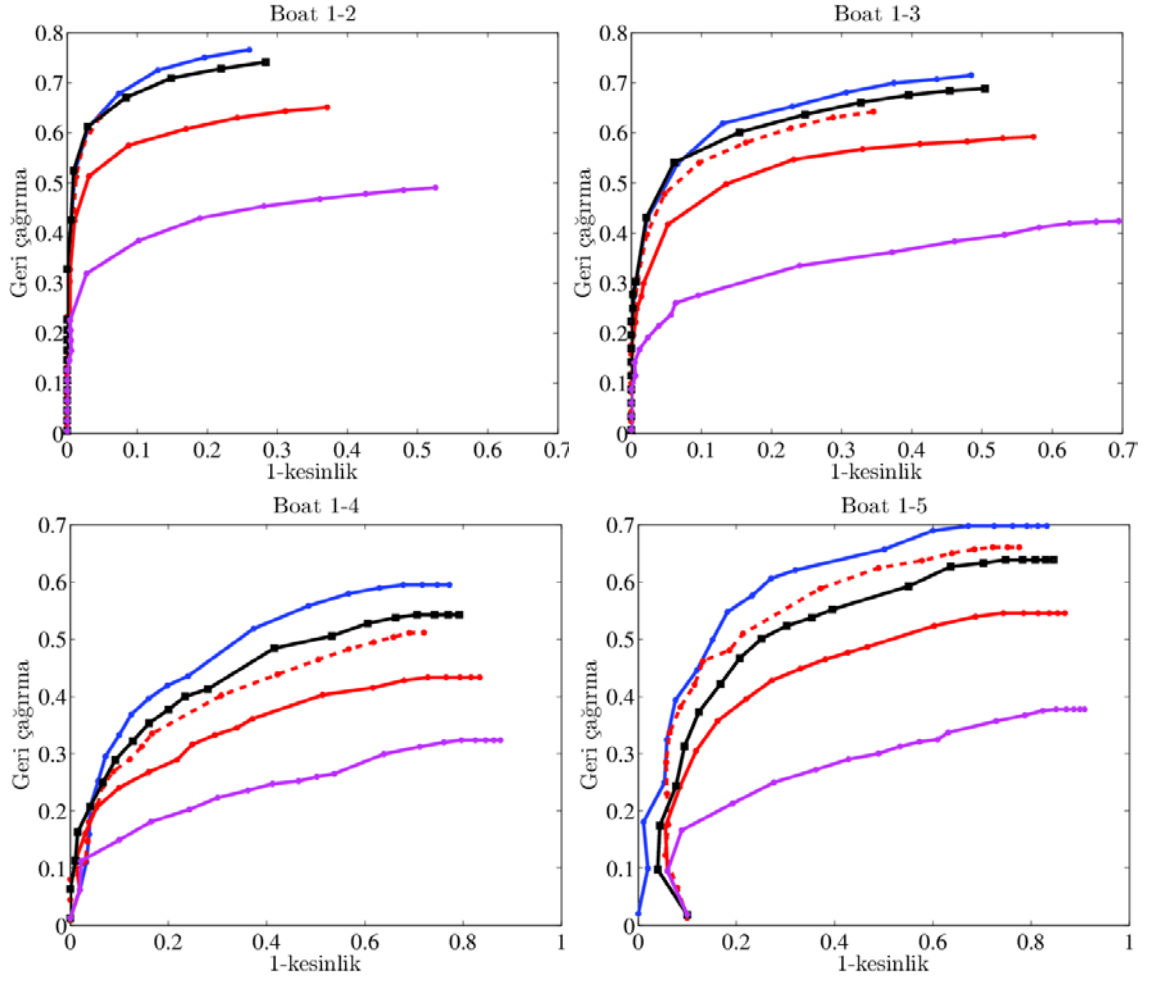
5.2.3 Benzetim Sonuçları

Karşılaştırma amacı ile *Oxford* veri kümesi üzerinde daha önce bahsedilen algoritmalar çalıştırılmış ve elde edilen sonuçlar Şekil 5.12'da verilmiştir. Burada SIFT etiketine sahip veriler, SIFT tanımlayıcılarının, SIFT yönelim bulma algoritması ile tek baskın açı çıkarılarak çalıştırılması sonucu elde edilen sonuçları göstermektedir. mSIFT ise SIFT yönelim bulma algoritmasında eşik değeri 0.8'den büyük olan en çok 3 yönelim değerinin kullanıldığı SIFT tanımlayıcılarının performansını göstermektedir. ORB ise ORB algoritmasında kullanılan yönelim bulma yöntemi kullanılarak elde edilen yönelim açılarının SIFT tanımlayıcılarında kullanılması ile elde edilmiş sonuçları

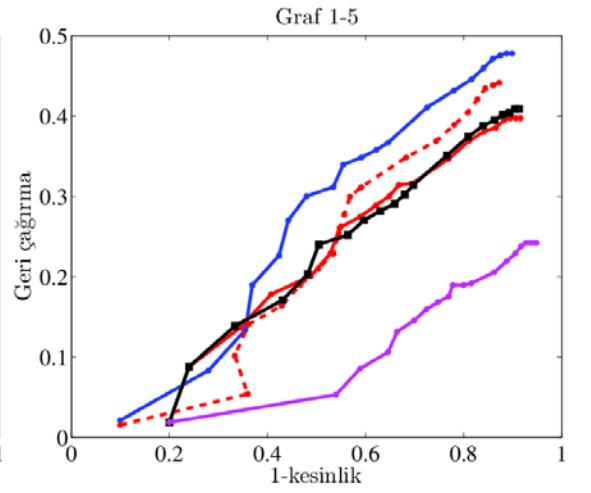
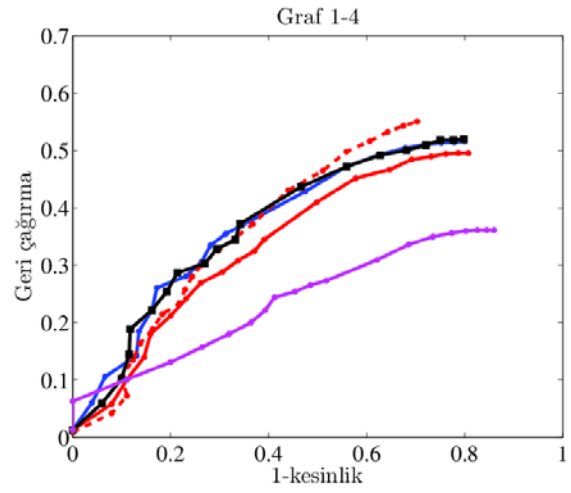
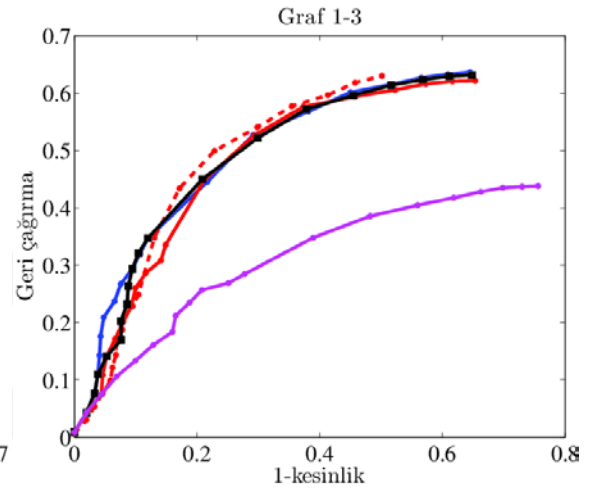
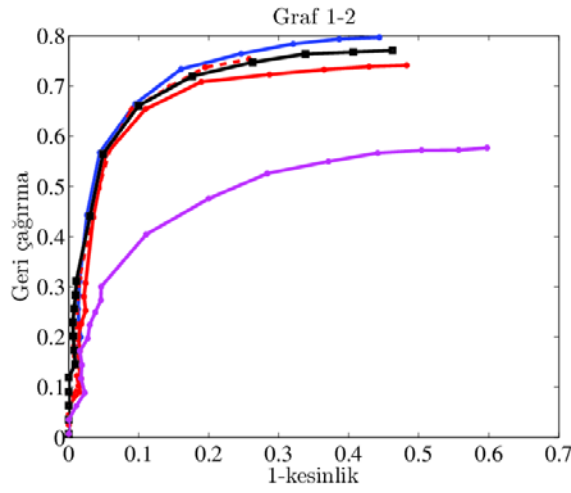
göstermektedir. Taylor ise (Taylor ve Drummond, 2011) tarafından önerilmiş olan yönelim bulma algoritmasının SIFT tanımlayıcıları ile kullanılmasından elde edilen sonuçları göstermektedir. Son olarak SIH yönelim bulma algoritması ile SIFT tanımlayıcılarının hesaplanması sonucunda elde edilen sonuçları göstermektedir.



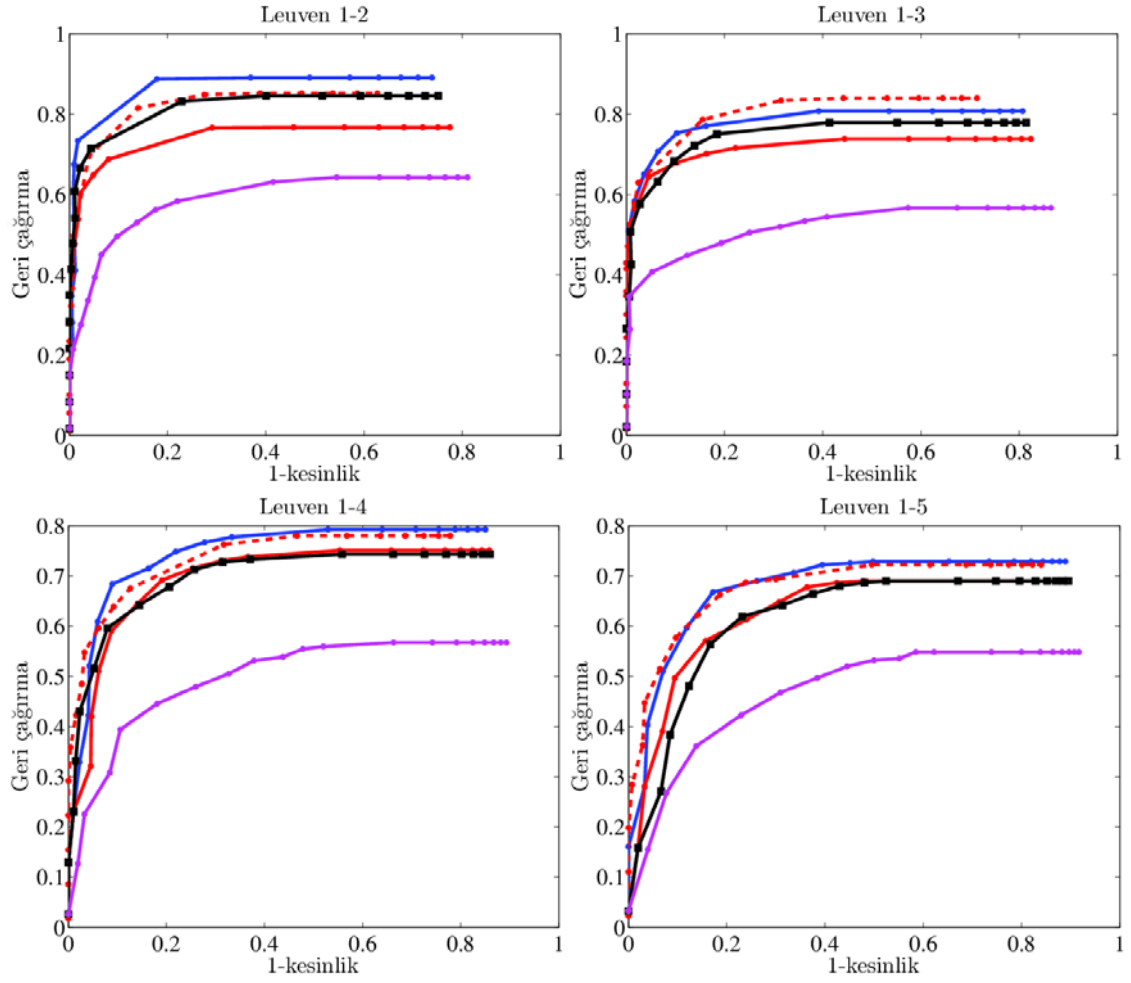
a) *Bikes* (Yumuşatma)



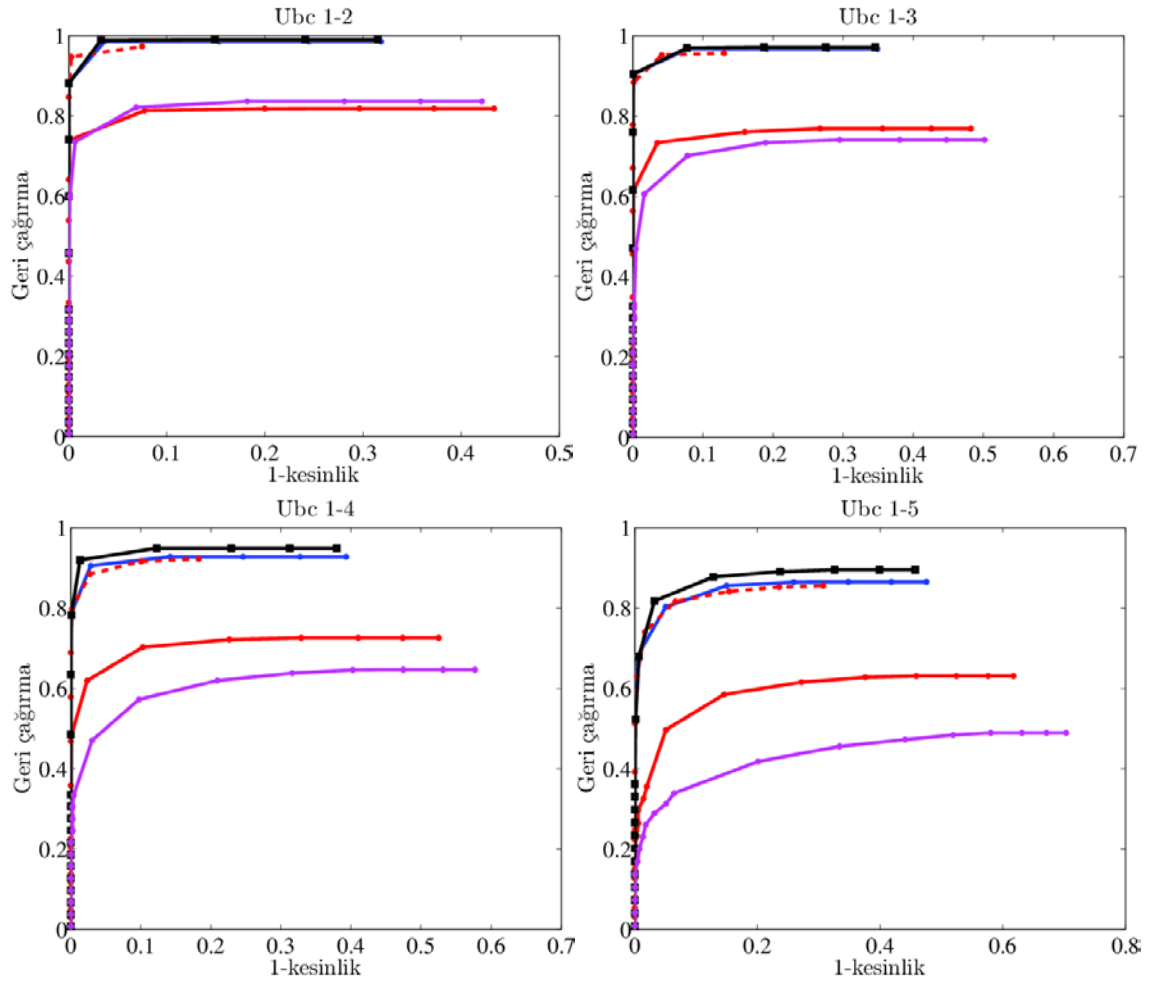
b) *Boat* (Açı ve Boyut)



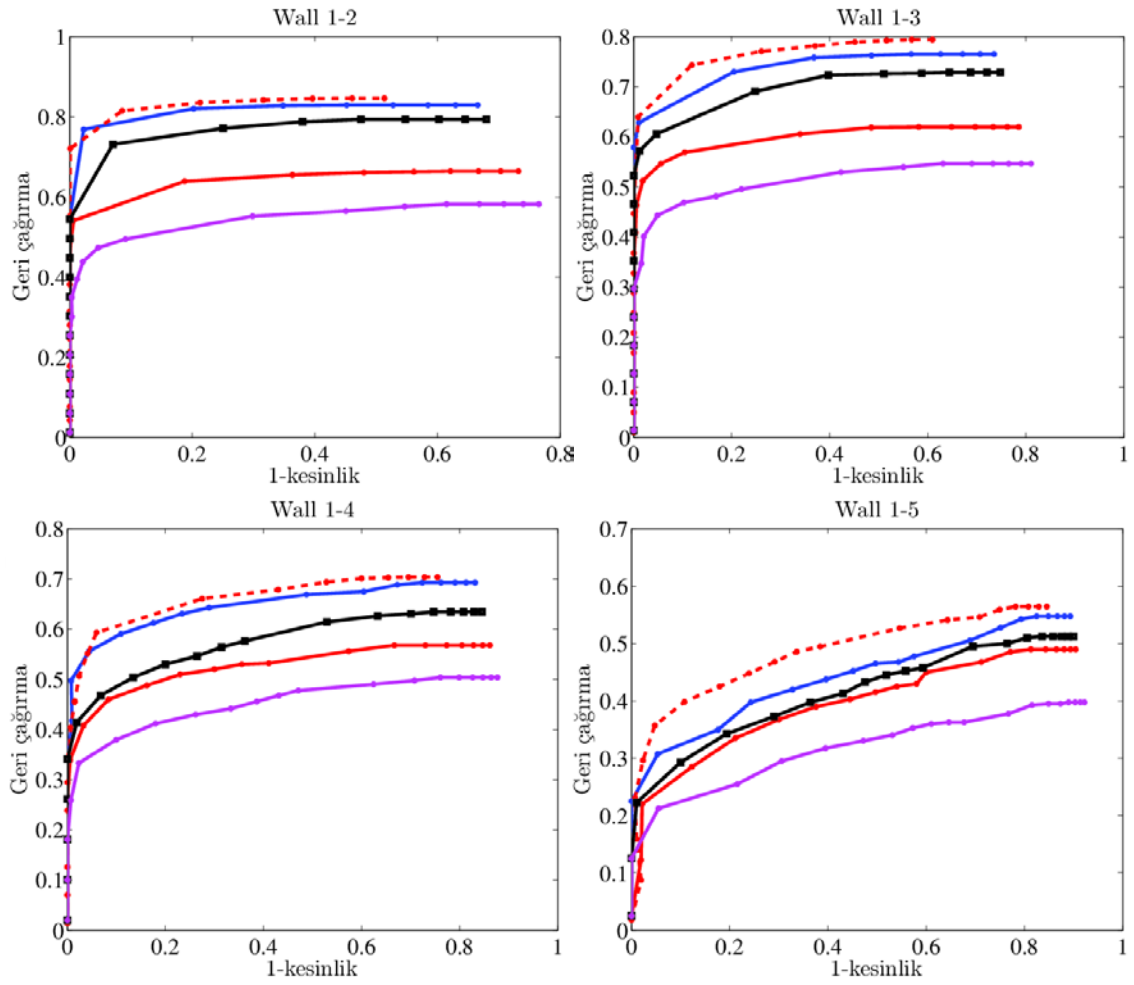
c) *Graffiti* (Açı ve Bakış açısı)



d) *Leuven* (Işık değişimi)



e) *Ubc* (JPEG bozulması)



f) Wall (Bakış açısı)

Şekil 5.12. Kullanılan görüntü kümelerine ait performans eğrileri

Her bir yöntemin çalışma süreleri, Intel Core i7 3.40 GHz işlemci, 8 GB bellek ve Ubuntu işletim sisteminin bulunduğu bir bilgisayarda c/c++ programlama dilleri ile yazılmış çalıştırılabilir dosyalar kullanılarak elde edilmiştir. 16000 görüntü bölgesi kullanılarak, tek bir görüntü bölgesinin yönelim açısı hesabı için gerekli olan ortalama süre değerleri elde edilmiş ve Çizelge 5.3'te verilmiştir. SIFT ile öznelik bölgesinin yönelim açılarının hesaplanması ortalama 1034 μ s sürerken, SIH 17 μ s de, ORB ise 14 μ s de aynı hesaplamaları yapabilmektedir. Fakat ORB algoritması açısal değişimlerin $\pm 20^\circ$ den büyük olduğu durumlarda diğer algoritmalara göre düşük performans göstermektedir.

Çizelge 5.3. Tek bir görüntü bölgesinin yönelim açısı hesabı için gerekli olan ortalama süre değerleri

Yöntem	Süre (µs.)
SIFT/mSIFT	1034
SIH	17
ORB	14
Taylor	13

SIFT ve mSIFT benzer algoritmalar ile yönelim açısı hesapladıkları için benzer sürede işlem yapmaktadır. ORB ve Taylor süre olarak SIH'den daha iyi olmasına rağmen performansları düşüktür. SIH başarımdan ödün vermeden SIFT tanımlayıcılarının çıkarım süresini kısalttığı elde edilen süre değerlerinden de anlaşılmaktadır.

5.3 Dilimlere Ayrılmış Yönelimlerin Histogramı

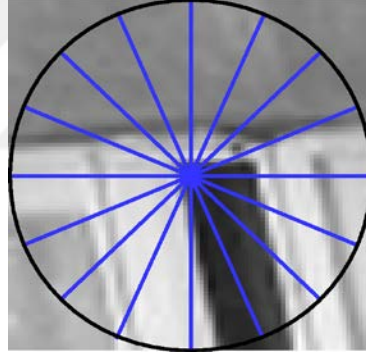
Bu bölümde, bir görüntü bölgesindeki yönelimli genliklerin histogramı kullanılarak gürbüz ve ayrıştırıcı bir tanımlayıcı elde edilmesi amaçlanmaktadır. Dilimlere ayrılmış yönelimlerin histogramı (Segmented Orientation Histogram - SOH) tanımlayıcıları, histogram tabanlı tanımlayıcıların ayırt edici özelliklerinin, yeni bir eşleştirme yöntemi ile kullanılmasından elde edilmektedir. Araştırmacılar tarafından yapılan çalışmalarda histogram tabanlı tanımlayıcıların ayırt edici özellikleri sıkça vurgulanmaktadır. Fakat literatürdeki yöntemler öznitelik bölgesinde rotasyondan kaynaklanabilecek hataları en aza indirmek için baskın açının bulunmasına ihtiyaç duymaktadır. Baskın açı bulunması önceki çalışmalarda da gösterildiği üzere, hataya açık bir işlemdir.

Daha önceki yapılan çalışmalarda eşleştirme yöntemi olarak L1-norm (Manhattan uzaklığı) ve L2-norm (Öklid uzaklığı) uzaklıkları kullanılmaktadır. SOH algoritmasında ise ortalama uzaklık fark fonksiyonu (Average Magnitude Difference Function-AMDF) kullanarak rotasyonel tanımlayıcı bozulmalarının önüne geçilmiştir. AMDF algoritması ses işleme uygulamalarında perde frekansının bulunmasında sıklıkla kullanılmaktadır (Rabiner vd., 1976; Zeng vd., 2003; Hui vd., 2006; Muhammad, 2011). Bu çalışmada ise SOH tanımlayıcı yapısına bağlı olarak eşleştirme fonksiyonu olarak kullanılacaktır.

5.3.1 SOH Tanımlayıcılarının Elde Edilmesi

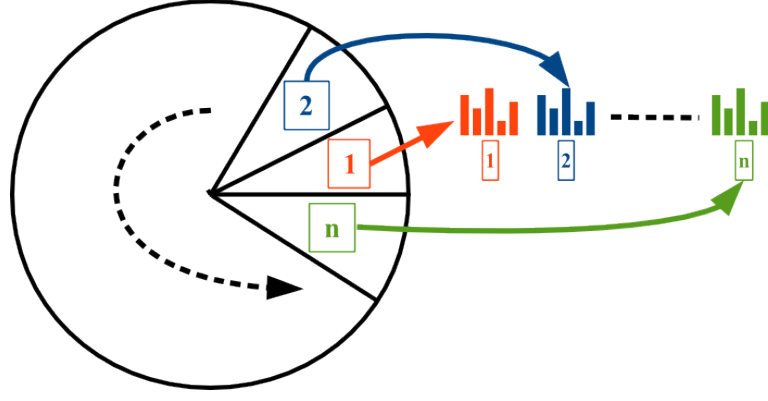
Bu alt bölümde SOH tanımlayıcısının çıkarım detayları incelenecektir. Tanımlayıcılar, ilgi noktalarının belirttiği görüntü bölgelerinden çıkarılmaktadır. Bu bölgeler normalize edilerek kullanılır. Ardından bu görüntü parçaları çevresel gürültüleri azaltmak için Gauss filtresi uygulanarak yumuşatılır. Benzetimlerden elde edilen veriler göz önüne alınarak, uygulanan Gauss filtresinin yumuşatma kat sayısı $\sigma = 2$ olarak kullanılmıştır.

Yumuşatılan bu görüntü parçaları tanımlayıcı çıkarımına hazır durumdadır. Fakat yöntemde, AMDF eşleştirme algoritmasının kullanılabilmesi için tanımlayıcının çıkartılacağı görüntü parçasının iç bölümlerinin iyi seçilmesi gerekir. Seçilen iç bölgeler açısal değişimin periyodik bir yapıda olmasını sağlamalıdır. Dairesel periyodikliği elde edebilmek için önerilen görüntü dilimleri Şekil 5.13’de gösterilmiştir.



Şekil 5.13. Görüntü parçasının dilimlenmesi

Bu görüntü dilimlerinin her birinden Rm (4.10c) ve Rg (4.10d) değerleri kullanılarak histogramlar oluşturulmaktadır. SOH tanımlayıcıları her bir görüntü parçasındaki dilimlerden elde edilen histogramların ardışıl olarak eklenmesi ile elde edilir (Şekil 5.14).



Şekil 5.14. SOH tanımlayıcının oluşturulması

5.3.2 AMDF Kullanarak Tanımlayıcı Eşleştirilmesi

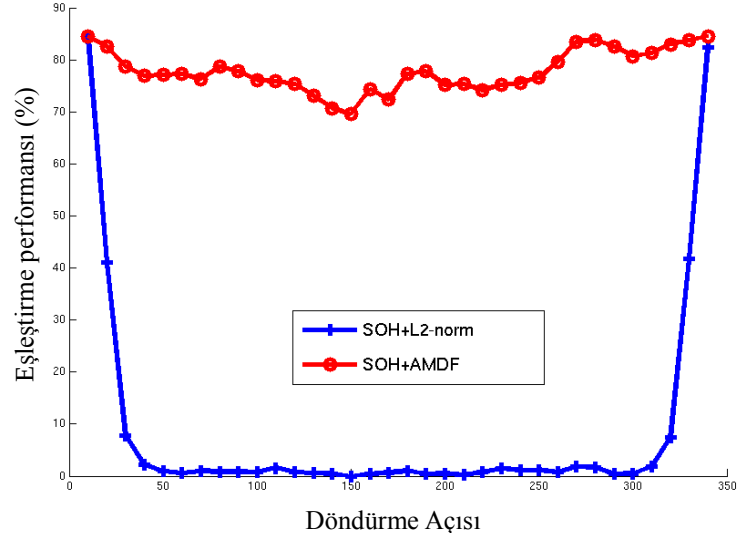
AMDF algoritması, genellikle ses işleme algoritmalarında perde frekansının bulunmasında kullanılmaktadır. Bir ses dalgası ve zamanda ötelenmiş versiyonu arasındaki benzerlik ölçütünü kullanarak bu işlemi gerçekleştirmektedir. Eğer sinyalde bir periyodiklik mevcutsa bu iki sinyal arasındaki farklılık belli bir kayma miktarında en az olmaktadır. Bu minimum değer iki sinyal arasındaki benzerlik miktarı olarak kullanılabilir. AMDF, kayma miktarının (τ) bir fonksiyonu olarak Denklem (5.11) ile hesaplanmaktadır.

$$f_{AMDF}(\tau) = \frac{1}{N} \sum_{i=0}^{N-1} |S_1(i) - S_2(i - \tau)|, \quad i = 0, 1, 2, \dots, N - 1 \quad (5.11)$$

Burada N toplam sinyal boyutu, $S_1(i)$ referans sinyalin i 'ninci değerini, $S_2(i - \tau)$ ise ikinci sinyalin $(i - \tau)$ 'uncu değerini belirtmektedir. AMDF fonksiyonu mümkün olan bütün kayma (τ) değerleri için hesaplanır ve minimum değeri,

$$F = \arg_{\tau} \min (f_{AMDF}(\tau)) \quad (5.12)$$

ile ifade edilir. Literatürde sıklıkla kullanılan L2-norm uzaklığı ve AMDF karşılaştırılacak olur ise Şekil 5.15'te görülen performans grafiği ortaya çıkmaktadır.

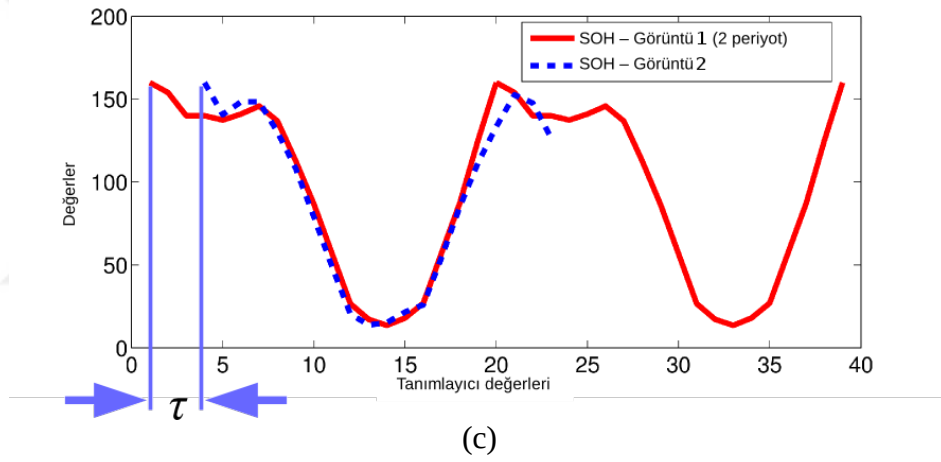
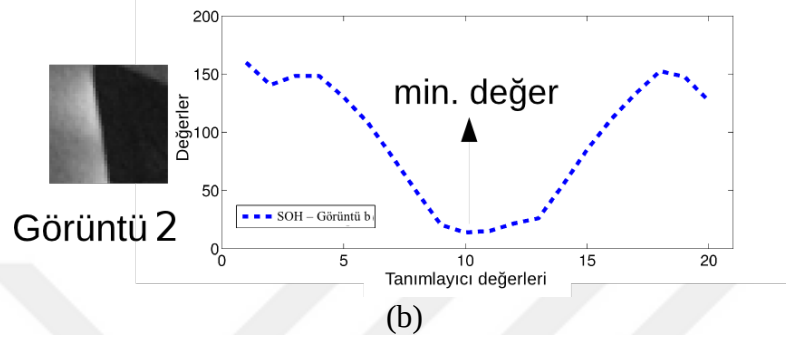
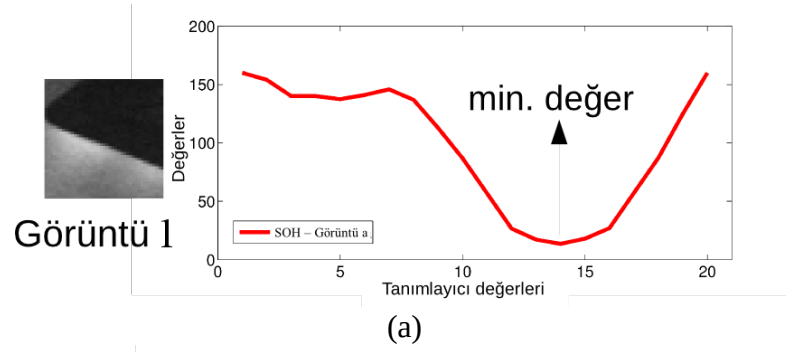


Şekil 5.15. AMDF ve L2-norm uzaklığı kullanılması durumunda elde edilen performans farkı

Şekil 5.15’den de anlaşılacağı üzere, L2-norm uzaklığı kullanılarak elde edilen eşleştirme başarımı, görüntüdeki dönme açısı $\pm 10^\circ$ ’yi geçtikten sonra %40’ın altına düşmektedir. AMDF algoritması ile eşleştirilen görüntüler de ise dönme açısının artan değerlerinde bile yüksek başarımla elde edilmektedir. Sonuç olarak gerçekleştirilen sistemin açısal değişimlerden etkilenmediği görülmektedir.

SOH’da görüntü parçasında meydana gelebilecek açısal değişimlerin neden olacağı dönme kadar tanımlayıcıların değerlerinin dairesel bir yapıda kayması (τ) beklenmektedir. SOH tanımlayıcısında açısal değişim durumunda oluşan kayma Şekil 5.16’da gösterilmiştir.

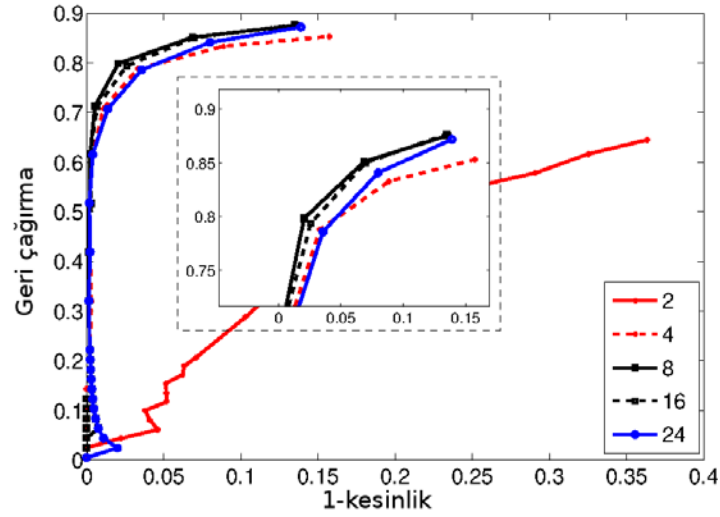
Burada *Görüntü 1* ve *Görüntü 2* birbirinin döndürülmüş halidir. SOH tanımlayıcılarındaki periyodik yapı, AMDF uygulanabilmesi için oldukça elverişlidir. Ayrıca AMDF sadece toplama ve çıkarma işlemleri içerdiğinden işlem yükü bakımından hafif bir algoritmadır.



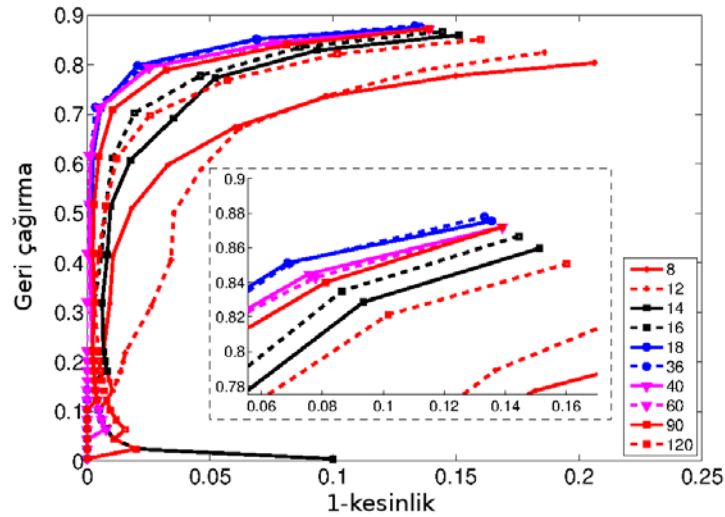
Şekil 5.16. SOH tanımlayıcısında açısal değişim durumunda oluşan kayma (a) Referans görüntü parçası (Görüntü 1) ve SOH tanımlayıcısı (Dilim sayısı 20 ve her dilimdeki histogram boyutu 1 olarak seçilmiştir.), (b) 60 derece döndürülmüş görüntü parçası (Görüntü 2) ve SOH tanımlayıcısı, (c) Görüntü 1 ve Görüntü 2 arasındaki kayma düzeltilmiş durumda.

5.3.3 SOH Tanımlayıcısı için Parametre seçimi

SOH tanımlayıcılarının değiştirilebilir parametreleri, dilim sayısı ve her bir dilimde oluşturulacak olan histogram boyutudur. Dilim sayısı tanımlayıcının ayırt ediciliğini ve açısal değişimlerdeki performansını, histogram boyutu ise doğrudan tanımlayıcı performansını etkilemektedir. En uygun parametre değerlerini elde edebilmek için boyut ve açı değişimlerinin bir arada bulunduğu *Boat* görüntü kümesi kullanılarak farklı parametre değerlerinde testler yapılmıştır. Elde edilen sonuçlar Şekil 5.17’de verilmiştir.



(a)



(b)

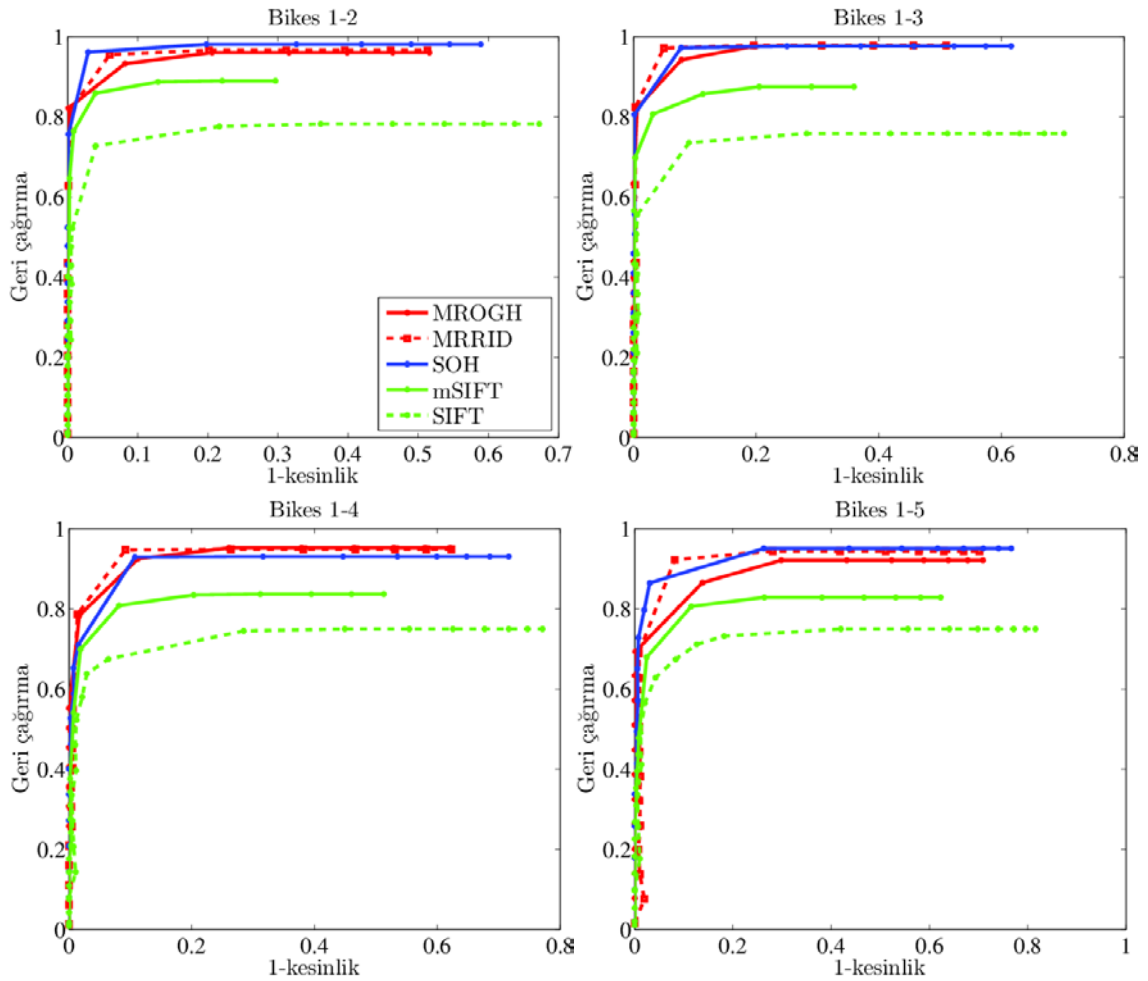
Şekil 5.17. Değişik parametreler için SOH tanımlayıcı performansları (a) Dilim sayısı sabit ve 18 iken histogram boyutunun performansa etkisi, (b) Histogram boyutu sabit ve 8 iken dilim sayısının performansa etkisi.

Şekil 5.17(a), dilim sayısının sabit olduğu ve histogram boyutunun değiştirildiği durumda sistem başarımını göstermektedir. Histogram boyutunun 8 ve 16 olduğu durumlarda eşleştirme başarımı en yüksek değerindedir. Bu iki değerden hangisinin seçileceği tanımlayıcı boyutuna bağlı olup, 8 değerinin seçilmesi daha küçük tanımlayıcı boyutu elde edilmesi açısından daha uygundur. Şekil 5.17(b) grafiği incelendiğinde ise en yüksek performans değerlerinin 18 ve 36 dilim sayısı ile elde edildiği görülür. Bu durumda tanımlayıcı boyutu göz önüne alındığında dilim sayısı için 18 değeri seçilmelidir. SOH tanımlayıcılarında kullanılan dilim sayısı ve histogram

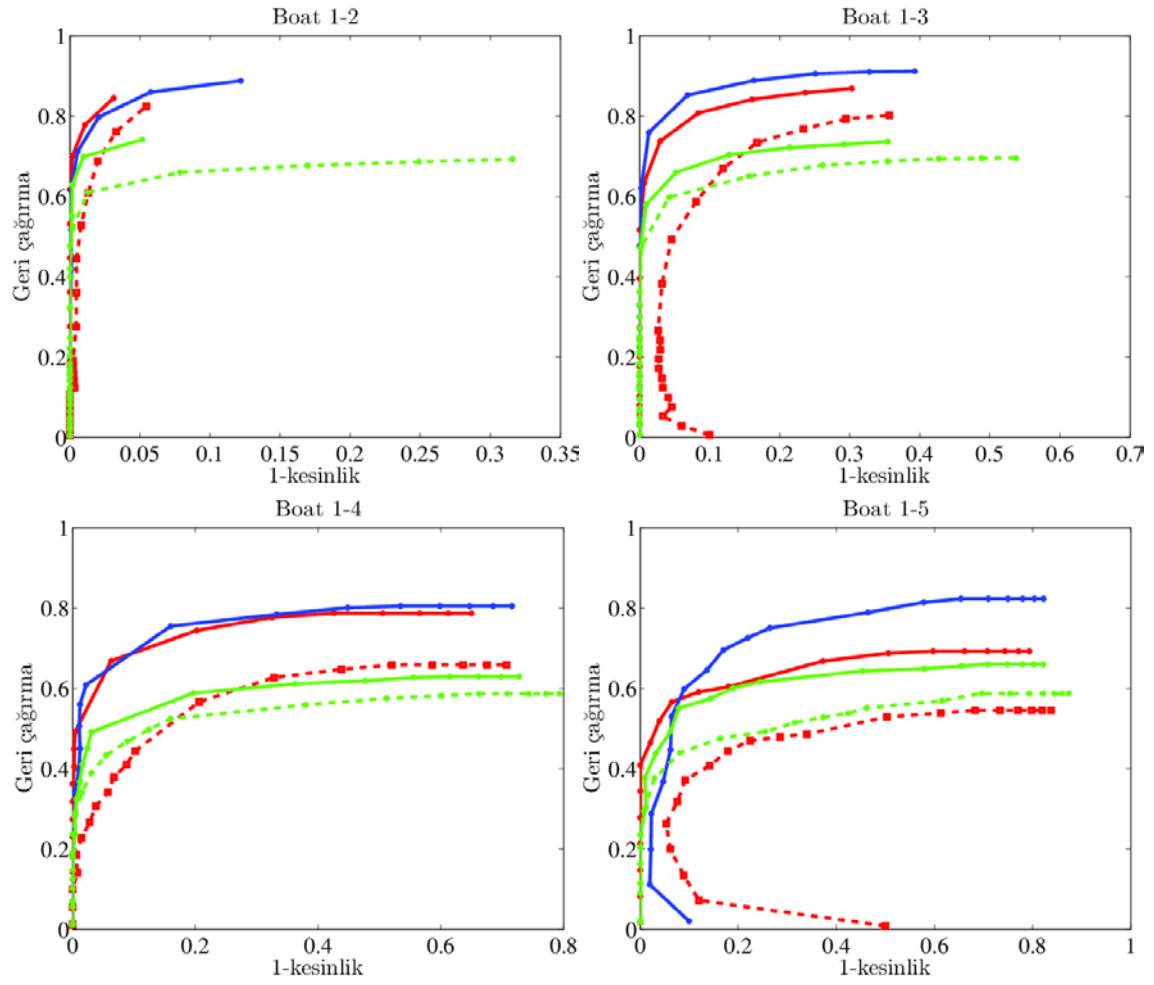
boyutu parametreleri sırasıyla 18 ve 8 olarak alındığında, tanımlayıcı boyutu 144 olmaktadır.

5.3.4 Benzetim Sonuçları

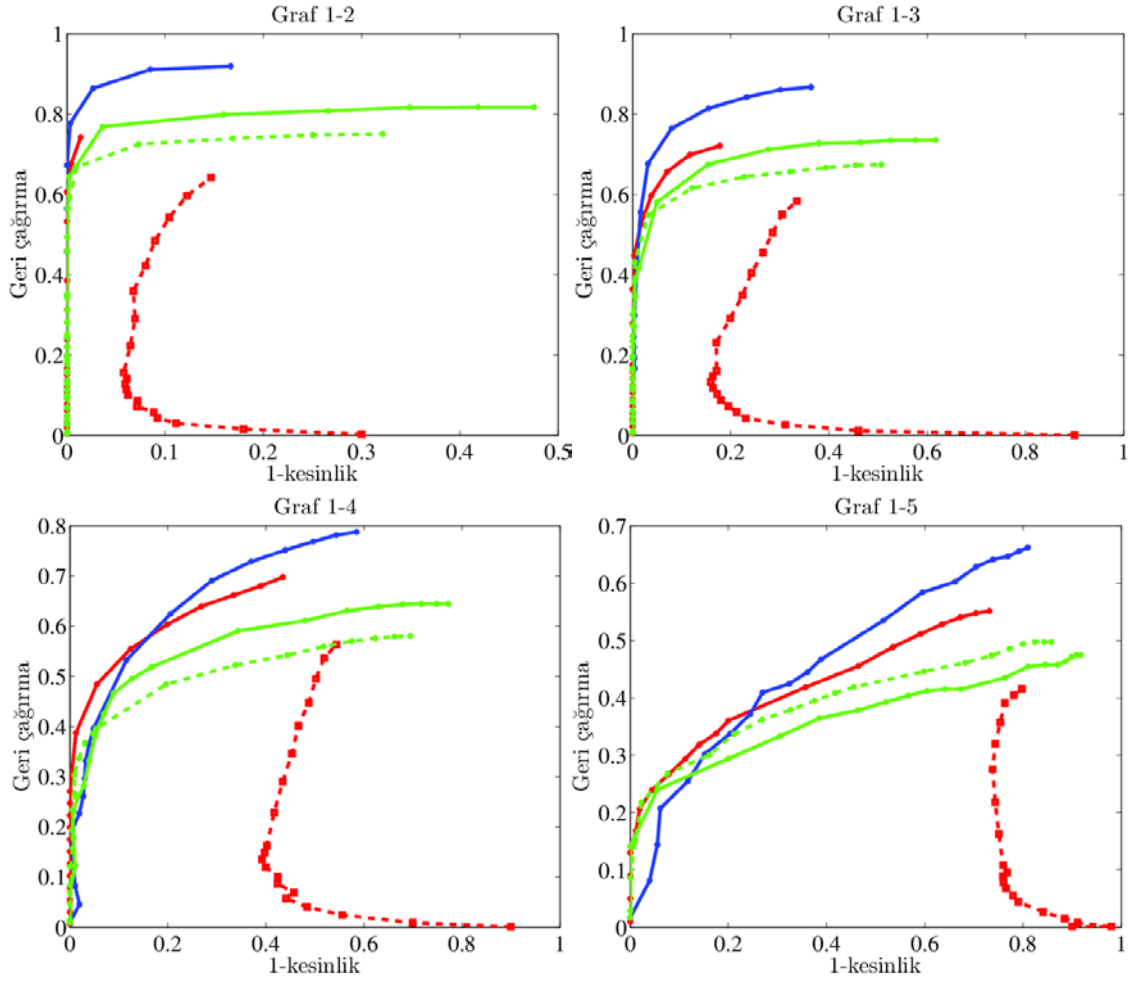
Karşılaştırma amacı ile *Oxford* veri kümesi üzerinde SIFT, MROGH, MRRID ve SOH algoritmaları çalıştırılmış ve elde edilen sonuçlar Şekil 5.18’de verilmiştir. Burada SIFT etiketine sahip veriler, SIFT tanımlayıcılarının, SIFT yönelim bulma algoritması ile tek baskın açı kullanılarak çalıştırılması sonucu elde edilen sonuçları göstermektedir. mSIFT ise SIFT yönelim bulma algoritmasında eşik değeri 0.8’den büyük olan en çok 3 yönelim değerinin kullanıldığı SIFT tanımlayıcılarının performansını göstermektedir. SOH etiketiyle gösterilen sonuçlar ise bu tez kapsamında önerilen tanımlayıcı ve eşleştirme algoritmaları ile elde edilen sonuçlardır.



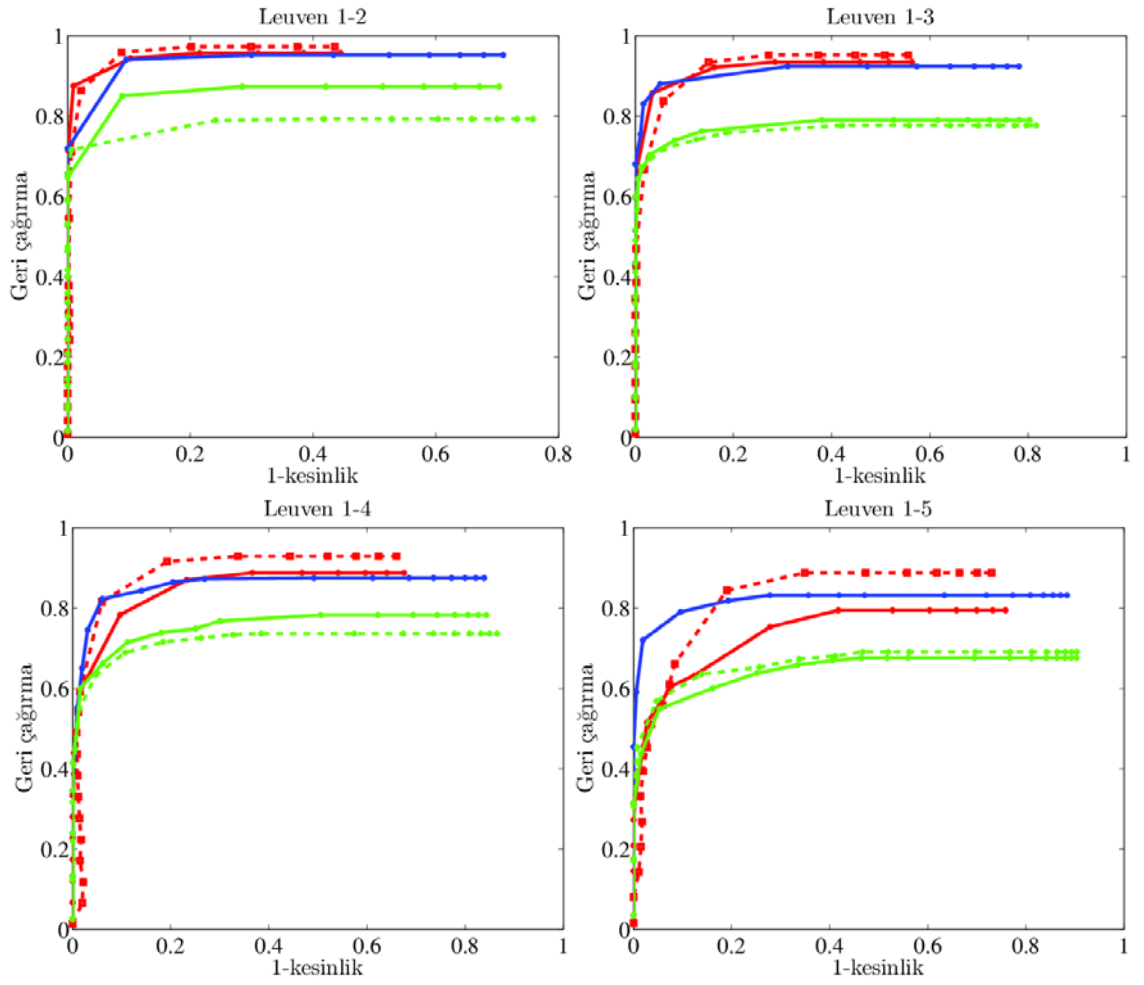
a) Bikes (Yumuşatma)



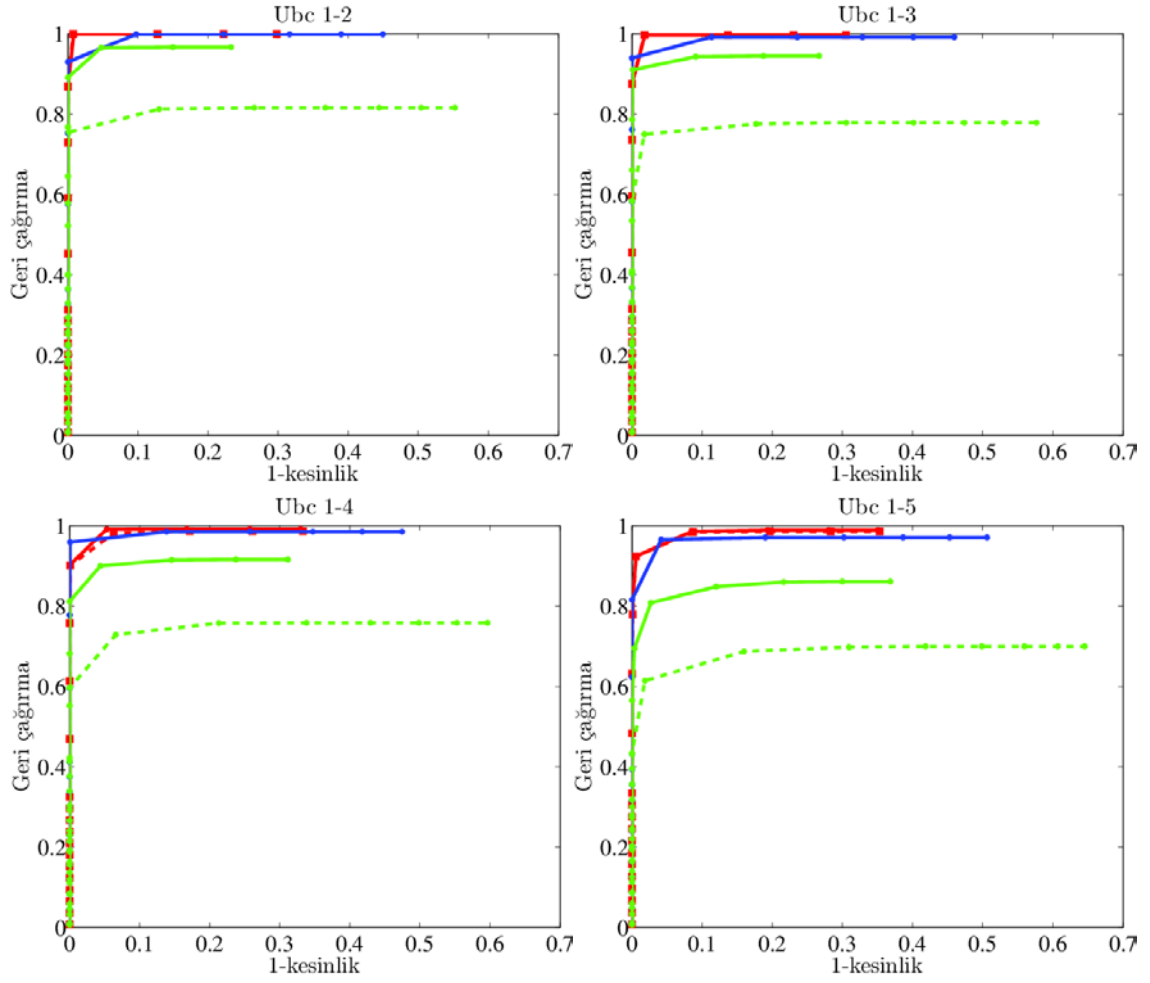
b) *Boat* (Açı ve Boyut)



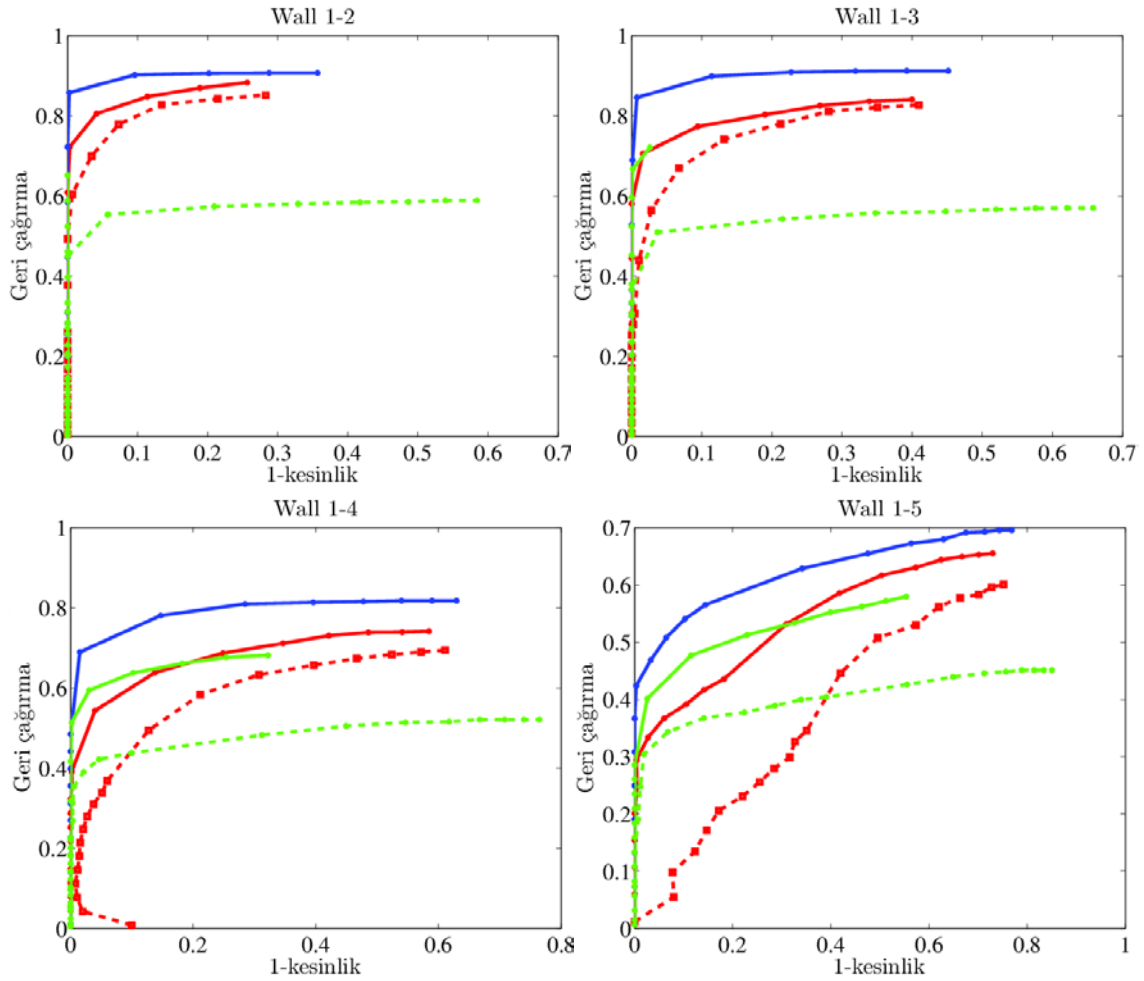
c) *Graffiti* (Açı ve Bakış açısı)



d) *Leuven* (Işık değişimi)



e) *Ubc* (JPEG bozulması)



f) Wall (Bakış açısı)

Şekil 5.18. Yöntemlere ait performans eğrileri

Tek açının kullanıldığı SIFT algoritması *Graffiti* görüntü kümesi haricinde en kötü sonuçları vermiştir. Birden fazla açının kullanıldığı SIFT sonuçları ise tek açılı SIFT algoritmasından daha yüksek performans göstermiş olsa da ışık değişimlerinin olduğu *Leuven* görüntü kümesinde tek açılı SIFT ile benzer performans göstermiştir. MRRID algoritması genel olarak MROGH algoritmasının gerisinde kalmıştır. Özellikle bakış açısı değişimlerinin olduğu görüntü kümelerinde MRRID algoritmasının performansı bozulma derecesi arttıkça daha çok düşmektedir. MRRID algoritması ışık değişimlerinin daha baskın olduğu görüntü kümelerinde ise beklenildiği gibi en iyi performansı vermektedir. MROGH algoritması genel performans olarak MRRID'ye göre daha kararlı sonuç vermektedir. Işık değişimlerinin olduğu durumlar dışında MROGH, MRRID'den daha iyi başarımlar göstermiştir. SOH algoritması ise ışık değişimleri dışında kalan tüm durumlarda en iyi performansı sergilemiştir. Bunun

nedeni ise ışık değişimlerinin olduğu görüntü kümelerinde histogram özellikleri, MRRID’de kullanılan piksel tabanlı özelliklere göre daha çok veri kaybına uğramasıdır. SOH algoritmasının, özellikle hem bakış açısı hem de açı değişiminin olduğu *Wall* görüntü kümesinde diğer algoritmaların oldukça üstünde bir performans sergilediği görülmektedir.

3000 öznitelik üzerinde bu yöntemler çalıştırıldığı zaman elde edilen süreler Çizelge 5.4’te verilmiştir. SOH süre açısından MRRID’den sonraki en kötü ikinci algoritma olmasına rağmen EZKH uygulamalarına yönelik bir avantaja sahiptir.

EZKH uygulamalarında, görüntüler arası dönme açısının değeri sistem dinamiği veya algılayıcılar yardımıyla yaklaşık olarak belirlenebilir. Dönme açısının yaklaşık olarak bilindiği bu gibi durumlarda, AMDF algoritmasındaki τ kayma miktarı, dönme açısına bağlı olarak sınırlandırılabilir. Bu sınırlandırma ile çalışma zamanının yüksek olmasına neden olan çoklu τ kaydırmalarının önüne geçilerek algoritmanın performansından ödün vermeden eşleştirme hızı ciddi oranda artırılabilir. Bu durumda SOH, SIFT algoritmasına yakın bir zamanlama ile çok daha iyi bir performans göstermektedir. SOH algoritmasının diğer algoritmalara (MROGH, MRRID, mSIFT) göre bir diğer üstünlüğü de daha az hafıza gerektirmesidir.

Çizelge 5.4. Algoritmaların iki görüntü arasındaki eşleştirmeleri hesaplamadaki ortalama süre ve hafıza karşılaştırmaları

Algoritmaların Süre ve Hafıza Karşılaştırmaları				
Metot	Tanımlayıcı Boyutu (byte)	Tanımlayıcı Çıkarım Süresi (sn.)	Eşleştirme Süresi (sn.)	Toplam Süre (sn.)
SIFT	128	0.48	1.9	2.38
mSIFT	128*3	0.79	2.4	3.19
MRRID	256	8.41	2.9	11.31
MROGH	196	3.52	2.4	5.92
SOH	144	0.65	9.2	9.85

BÖLÜM VI

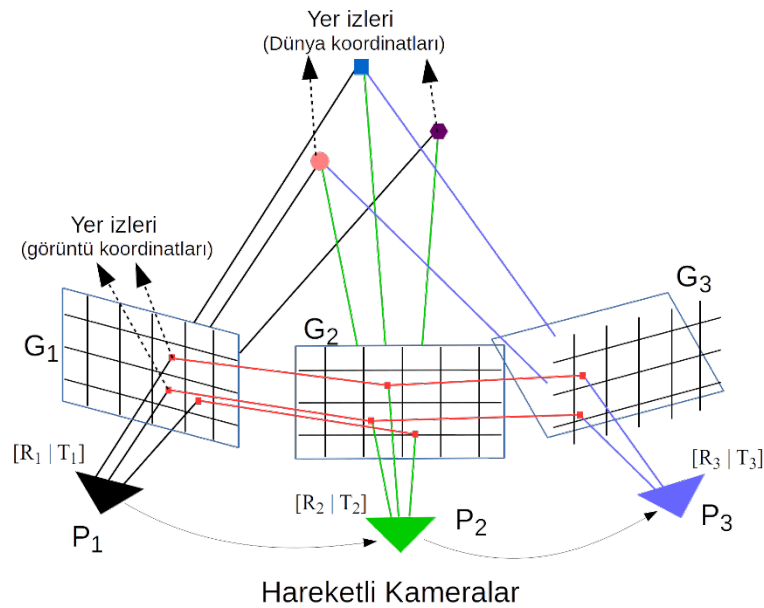
GÖRSEL EŞ ZAMANLI KONUMLAMA VE HARİTALAMA

6.1 Giriş

Tez kapsamında önerilen SOH tanımlayıcıları ve TPDK yer izi tespit yöntemi kullanılarak oluşturulan Kalman filtresi tabanlı bir EZKH sistemi tanıtılacaktır. İlk olarak görsel poz kestirimi için gerekli bilgiler verilecektir. Ardından Kalman filtresi incelenecek ve kullanılan sistem modeli verilecektir. Böylece elde edilen model ile EZKH sistemi oluşturulacaktır.

6.2 Görsel Konum Değişimi Kestirimi

Görsel seyir sistemi, başlangıç noktası bilinen bir aracın yer değiştirme bilgisinin kameralar kullanılarak elde edilmesi ile oluşturulur. Yer değiştirme bilgisinin elde edilebilmesi için Bölüm III’de incelenen yer izi tespiti ve eşleştirilmesi yöntemleri sıklıkla kullanılmaktadır. Görsel seyir ve haritalama için kullanılan sistem genel yapısı Şekil 6.1’de gösterilmiştir.



Şekil 6.1. Kamera hareketinden konum hesaplama

Şekil 6.1’de gösterilen hareket halindeki bir kameradan alınan görüntülerin (G_1, G_2, G_3) ve ilk kamera konumunun (P_1) bilindiği kabul edilmektedir. Buna göre kamera hareket ederken ayrıık zamanlı olarak çekilen görüntülerden çıkartılan yer izleri kullanılarak her bir görüntü anında kameranın konum bilgileri ve buna bağlı olarak yer izlerinin dünya koordinatlarındaki konumu hesaplanabilir.

Bu bölümde görsel seyir sistemi oluşturmak için gerekli olan sistem modelleri incelenecektir. İlk olarak görsel seyir probleminin tanımı yapılacaktır. Sonrasında ise bu problemin kameralardan alınan görüntüler kullanılarak çözümüne ilişkin detaylar incelenecektir. Kullanılan kamera görüntülerinden elde edilen verilerde çevresel etkilerden dolayı oluşacak gürültü etkilerini en aza indirebilmek için Kalman filtresi modeli verilecektir.

6.2.1 Problem Tanımı

$I = [I_1, I_2, I_3, \dots, I_n]$, sabit bir ortam içerisinde alınmış görüntülerden oluşan bir veri kümesidir ve bu kümede yer alan ardışık görüntülerin, örtüşen içeriklere sahip oldukları kabul edilmektedir. Bu görüntüler kullanılarak kamera konum bilgileri elde edilmek istenmektedir. Her bir görüntüden elde edilen yer izlerinin dünya koordinatları $X = [X_1, X_2, X_3, \dots, X_n]$ ile ve her bir görüntünün alındığı kamera konumları $P = [P_1, P_2, P_3, \dots, P_n]$ ile ifade edilir. Konum bilgileri kameraya ait iç parametreler (K) ve kamera konum değişim matrisi (M) kullanılarak,

$$P = KM \quad (6.1)$$

ile ifade edilir. İç parametreler, kameraya ait odak uzaklığı (f), çerçeve oranı (α), eğrilik (skew) açısı (γ) ve optik merkezi (c_x, c_y) gibi fiziksel özellikleri temsil etmektedir. K matrisi Denklem (6.2)’de gösterildiği gibi ifade edilmektedir.

$$K = \begin{bmatrix} f & \gamma & c_x \\ 0 & \alpha f & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (6.2)$$

M matrisinin sahip olduğu bilgi içerisinde hem yer değiştirme hem de her eksene ait dönme açısı yer alır. M matrisi,

$$M = [R \ T] \quad (6.3)$$

ile gösterilir. Burada R matrisi 3×3 boyutundadır ve kameranın dönme açılarını ifade eder. T ise kameranın dünya koordinatlarında konum değişim bilgisini içerir ve 3×1 boyutundadır.

P matrisi kullanılarak görüntüde bulunan yer izlerinin dünya koordinatlarına çevrilmesi mümkündür. Bu tez kapsamında ise sadece görüntüler kullanılarak kamera konumlarının elde edilmesi amaçlanmıştır. İki görüntü arasında konum değişiminin hesaplanması için görüntülerde bulunan yer izlerinin eşleştirilmesi ve eşleştirilen yer izlerinin üçgenleme metodu ile P matrisinin oluşturulması gerekmektedir. P matrisinin hesaplanabilmesi için yer izleri kullanılarak ana matris, esas matris, homografi matrisi vb. yöntemler kullanılabilir.

6.2.1.1 Ana Matris

Kalibrasyon parametreleri bilinmeyen iki kamera arasındaki ilişkiyi ifade etmek için kullanılan genelleştirilmiş bir matristir. Diğer yöntemlerin aksine en az 8 eşleştirilmiş yer izi gerekmektedir. Ana matris,

$$x_i^* F_{i,j} x_j = 0 \quad (6.4)$$

ile hesaplanmaktadır. i ve j , örtüşen içeriklere sahip iki görüntüye ait indisleri ifade etmektedir. x_i ve x_j ise ardışık görüntülerden elde edilen ve eşleştirilen yer izi konumlarını ifade etmektedir.

6.2.1.2 Esas Matris

Esas matris (E) iki görüntü arasındaki göreceli konum değişimini ifade etmektedir. Bu matris iki görüntünün çekildiği kameralar arasındaki konum değişimi ve her bir

eksendeki açı değişim bilgilerini içermektedir. Esas matris matematiksel olarak,

$$E_{i,j} = R[T]_{\times} \quad (6.5)$$

ile gösterilir. P_0 başlangıç konumundan alınan görüntü karesi olmak üzere başlangıç konumu olarak koordinat sisteminin merkezi seçilir ise kameranın konum değişim matrisi $M_0 = [I \ 0]$ olur. Sonraki konum değişimleri ise $M_j = [R \ T]$ ile ifade edilir. T matrisi ile ifade edilen yer değiştirme bilgisi, bilinmeyen bir s ölçek oranına sahiptir. Elde edilen yer değiştirme bilgileri s ile çarpıldığı durumda metrik değerler elde edilebilir. Ölçeğin bulunabilmesi için ölçüleri bilinen bir objenin kullanılması veya görüntülerin alındığı uzaklık bilgisinin bilinmesi gerekir. Yer izleri kullanılarak esas matrisin elde edilebilmesi için,

$$x_i^* E_{i,j} x_j = 0 \quad (6.6)$$

ifadesi kullanılır. Eşleştirilen yer izlerinin en az 8 tane olması durumunda ölçekten bağımsız bir şekilde esas matrisi elde edilebilir. Kamera iç parametrelerinin bilindiği durumda ise 7, 6 veya 5 adet eşleştirme ile esas matris elde edilebilir.

6.2.1.3 Homografi Matrisi

Homografi matrisi (H), iki görüntü arasında eşleştirilmiş yer izlerinden elde edilen ve görüntü düzleminde herhangi bir noktanın, diğer görüntü düzlemine iz düşümünü tanımlayan matristir. Bir görüntü düzleminde diğerine H matrisi ile tanımlanan iz düşüm ilişkisi Denklem (6.7) ile gösterilir.

$$x_j = H_{i,j} x_i \quad (6.7)$$

Homografi matrisi görüntü düzlemlerinde iz düşüm dönüşümünü tanımlamaktadır. Düzlemsel sahnelerde esas matrisin hesaplanmasının mümkün olmadığı durumlarda yardımcı olarak kullanılabilir.

6.2.2 Kamera Konum Değişimi Hesaplama

Sadece yer izi eşlemelerinin bilindiği bir durumda, ilk kamera pozisyonu biliniyor ise, kameranın bir sonraki pozisyonu hesaplanabilir (Hartley ve Zisserman, 2003). Bu bölümde poz elde etmek için Hartley ve Zisserman tarafından ele alınan esas matris kullanımı incelenmiştir. Denklem (6.5)'te esas matrisin iki bileşeni olduğu görülmektedir. Bu bileşenler daha önce de bahsedildiği üzere kamera pozisyonlarının açı değişimlerini ve konum değişimlerini ifade etmektedir. Esas matris hesaplanmış ise, kamera konum değerleri E matrisi ayrıştırılarak elde edilebilir. E matrisinin tekil değer ayrışımı Denklem (6.8)'deki gibi hesaplanır.

$$E = USV^T \quad (6.8)$$

U ve V matrisleri 3x3 boyutunda ortonormal matrislerdir. S matrisi esas matrisin tekil değerlerini ifade eder ve 3x3 boyutunda köşegen bir matristir. Esas matrisin iki tekil değeri birbirine eşit ve son tekil değerinin sıfır olması gerekmektedir. S matrisinin köşegen değerleri, E matrisinin tekil değerleridir ve,

$$S = \begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (6.9)$$

ile gösterilir. Esas matrisi, $S = [T]_{\times}$ olmak üzere $E = SR$ ile tanımlanabilir. Esas matris özelliklerinden faydalananarak W ve Z matrisleri,

$$W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.10a)$$

$$Z = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.10b)$$

olarak tanımlanır. Burada W açı matrisi, Z çarpık-simetrik matristir. Bu matrisler için,

$$ZW = \text{diag}([1,1,0]) \quad (6.11a)$$

$$ZW^T = -\text{diag}([1,1,0]) \quad (6.11b)$$

ifadeleri yazılabilir. Bu denklemleri kullanarak esas matrisin tekil değeri dönüşümü için iki denklem, $E = S_1 R_1$ ve $E = S_2 R_2$ ise,

$$S_1 = -UZU^T, \quad R_1 = UW^T V^T \quad (6.12)$$

$$S_1 = UZU^T, \quad R_2 = UWV^T \quad (6.13)$$

olarak yazılabilir. Bu çözümde R_1 ve R_2 açılı matrislerinin geçerliliğini kontrol etmek için $R_1^T R_1$ çözümü,

$$R_1^T R_1 = (UW^T V^T)^T UW^T V^T = VWU^T UW^T V^T = I \quad (6.14)$$

yapılarak R_1 matrisinin ortogonal olduğu görülür. Ayrıca Denklem (6.15)'e göre R_1 matrisinin açılı matrisi olduğu söylenebilir.

$$\det(R_1) = \det(UW^T V^T) = 1 \quad (6.15)$$

Eğer $\det(UV^T) = -1$ bulunur ise R_1 yansıma barındıran açılı matrisi olur ve bu çözüm geçerli değildir. S_1 çarpık simetrik matrisi,

$$-S_1^T = (UZU^T)^T = UZ^T U^T = -UZU^T = S_1 \quad (6.16)$$

ile hesaplanır. Böylece elde edilen bağıntılar birleştirilir ise,

$$S_1 R_1 = -UZU^T UWV^T = -U(-\text{diag}([1,1,0]))V^T = E \quad (6.17)$$

ifadesi elde edilir. Bu durumda $S_1 R_1$ geçerli bir çözümdür. $S_2 R_2$ içinde aynı adımlar gerçekleştirildiği durumda benzer sonuç elde edilecektir. Esas matrisin tekil değeri ayrışımını elde ettikten sonra kameranın poz değişimi, açılı değişimi ve yer değişimi olarak elde edilebilir. $E = SR$ çözümü ile T yer değiştirme bilgisi,

$$ST = [T]_{\times} T = T \times T = 0 \quad (6.18)$$

olarak bulunur. Buradan anlaşılacağı üzere T vektörü, S matrisinin boş uzayında yer

almaktadır. S_1 ve S_2 matrislerinin boş uzayları U matrisinin üçüncü kolonuna bakılarak elde edilebilir. T vektörü, S matrisinin boş uzayında olduğuna göre katsayı ile çarpılmış λT vektörü ölçek katsayısı olarak kullanılabilir. λ katsayısının işaretine bağlı olarak yer izlerinin kameranın önünde veya arkasında olma olasılığı bulunmaktadır. u_3 , U matrisinin üçüncü kolonu olmak üzere, dört farklı çözüm elde edilebilir. P_1 ilk kamera konumu olsun P_2 olası çözümleri,

$$P_2 = [UWV^T \ u_3], \quad \lambda = 1 \text{ için 1. olası çözüm} \quad (6.19a)$$

$$P_2 = [UW^TV^T \ u_3], \quad \lambda = 1 \text{ için 2. olası çözüm} \quad (6.19b)$$

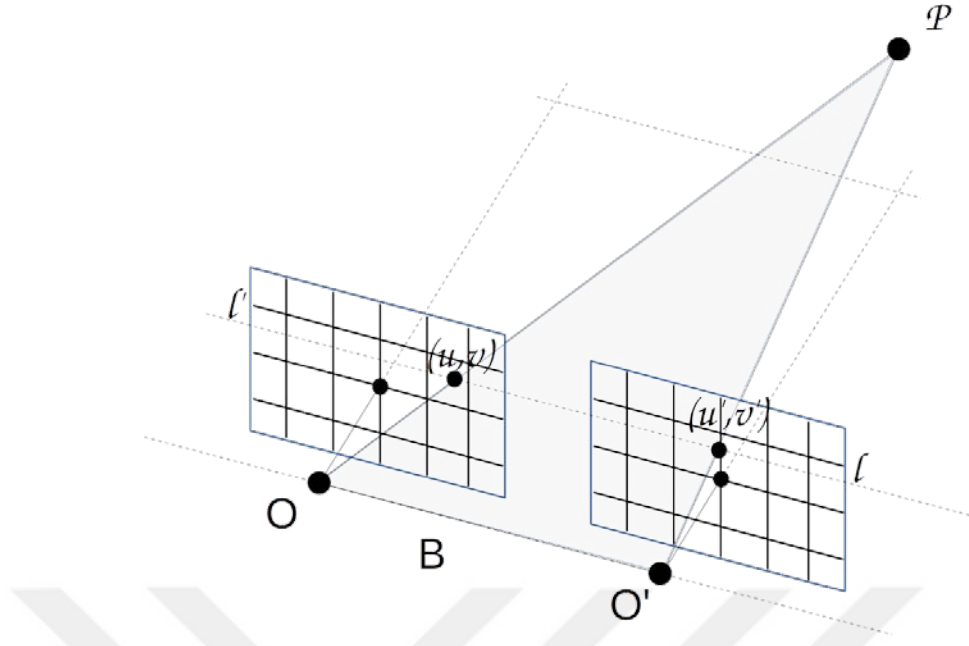
$$P_2 = [UWV^T \ -u_3], \quad \lambda = -1 \text{ için 1. olası çözüm} \quad (6.19c)$$

$$P_2 = [UW^TV^T \ -u_3], \quad \lambda = -1 \text{ için 2. olası çözüm} \quad (6.19d)$$

olarak ifade edilir. Bu sayede kameranın yeni konumu için gerekli değerler elde edilmiş olur. Bu dört olası çözümden, yer izleri ile doğrulama yapılarak, çözümlerin tümünün pozitif derinlikte olduğu kamera konumları kullanılmaktadır.

6.2.3 Stereo Kamera Kullanımı ve Derinlik Bilgisi

Tek kamera kullanıldığı durumlarda ölçek ya da derinlik bilgisinin elde edilmesi için özel şartların gerçekleşmesi gerekmektedir. Bu şartlardan bir tanesi boyutu bilinen bir nesnenin kalibrasyon için kullanılmasıdır. Bir diğeri ise konumu bilinen iki kameranın kullanılmasıdır. Her durumda boyutu bilinen bir kalibrasyon nesnesinin kullanımı mümkün olmayabilir. Bu sebeple son yıllarda stereo kameraların kullanımı yaygınlaşmıştır. Şekil 6.2’de standart bir stereo kamera yapılandırması gösterilmiştir.



Şekil 6.2. Stereo kamera yapılandırması

Şekil 6.2’de görülen O ve O' birbirlerine göre uzaklıkları bilinen iki kameranın odak noktalarını ifade etmektedir. (u, v) ve (u', v') noktaları ise P dünya koordinatlarına sahip yer izinin görüntü düzlemlerine karşılık gelen konumlarını temsil eder. l ve l' ise epipolar çizgiler olarak isimlendirilir. İki odak arası uzaklık ise ana çizgi olarak isimlendirilir ve B ile gösterilmiştir. Burada iki kamera arasındaki konum değişimi sadece tek düzlemde bir öteleme olarak ifade edilir. Bu durumda açıl matrisi de birim matris olarak ele alınabilir. İdeal stereo sahnesinde $v = v'$ olmalıdır. u ve u' arasındaki fark ise $d = u - u'$ iki resim arasındaki farklılıktır. P yer izinin dünya koordinatları (X, Y, Z) üçgenleme metodu,

$$Z = \frac{B}{d} f_u \quad (6.20a)$$

$$X = Z \frac{u'}{f_u} \quad (6.20b)$$

$$Y = Z \frac{v'}{f_u} \quad (6.20c)$$

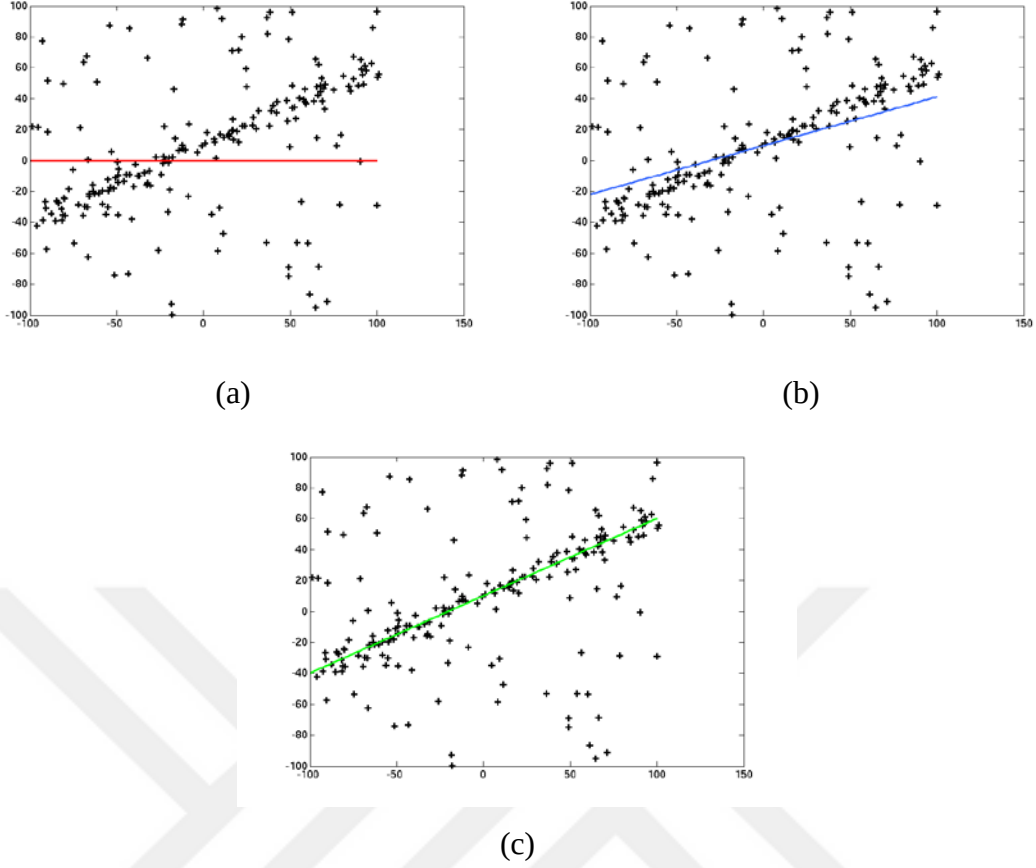
ile hesaplanır. Burada f_u kameraların odak uzaklığını belirtmektedir. Kameraların teknik olarak birbirinin aynı olduğu kabul edilmiştir. Böylece stereo kamera özellikleri kullanılarak derinlik bilgisi elde edilir.

6.2.4 Yanlış Yer İzi Eşleştirmeleri

Daha önce belirtildiği gibi kamera konum değişimi hesabında kullanılan yer izi eşleştirmelerinde yanlış eşleştirme olmadığı varsayılmıştır. Gerçek dünya uygulamalarında ise yanlış eşleştirme olması ihtimali oldukça yüksektir. Dolayısıyla bu yanlış eşleştirmeler konum değişimi hesabında hatalara sebep olacaktır. Yer izi bulunması ve eşleştirilmesi için kullanılan yöntemlerden bazıları Bölüm III ve Bölüm IV'te verilmiştir. Fazla desen içermeyen görüntülerden çıkartılan yer izlerinin eşleştirilmeleri de yer izlerinin tespiti kadar zor olmaktadır.

Genellikle eşleştirmeler, yer izlerinden elde edilen tanımlayıcılar arasındaki uzaklıklara bakılarak tanımlanır. Fakat gerçek dünya görüntülerinde yer alan gürültülerden dolayı literatürde yer alan hiçbir yöntemin başarımı kusursuz olmayacaktır. Bu nedenle eşleştirmelerde hata olduğu kabul edilir. Ayrıca eşleştirme hataları kamera konum değişiminde kullanılan denklemlerde hatalara sebep olur ve zamanla konum takibini imkânsız hale getirebilir. Bu yanlış eşleştirmelerin doğru bir şekilde tespit edilebilmesi için görüntüler arası doğru eşleştirilen yer izlerinin yanlış eşleştirmelerden fazla olması gerekmektedir.

Aykırı değerler içeren veri kümelerinde, aykırı değerleri tespit edebilmek için yöntemler önerilmiştir (Fischler ve Bolles, 1981). Bunlar arasında en çok kullanılan rastgele örnek uyuşumu (Random sample consensus) olarak isimlendirilen RANSAC algoritmasıdır. RANSAC özyinelemeli bir algoritmadır ve veri kümesinde yer alan değerler arasından rastgele seçimler yapar. Bu seçimleri kullanarak $(ax + by + c = 0)$ formunda bir doğru uydurarak veri kümesini bir doğru ile temsil etmeyi amaçlar. Tekrarlamalı bir şekilde seçilen noktalar kullanılarak uydurulan eğrilere örnek olarak Şekil 6.3'te verilen durumlar gösterilebilir.



Şekil 6.3. Veri kümesine doğru uydurma problemi a) uydurulan doğrunun veri kümesinde çok az veriyi işaret etmesi durumu b) uydurulan doğrunun veri kümesinde çoğu veri için doğru olması fakat en iyi çözüm olmaması durumu c) veri kümesini en doğru temsil eden doğrunun bulunduğu durum

RANSAC algoritmasının görüntü tabanlı sisteme uyarlanması için veri kümesinin belirlenmesi gerekmektedir. Görüntüler arası eşleştirmelerden elde edilen esas matris, ana matris veya homografi matrisi kullanılarak RANSAC algoritması uygulanabilir.

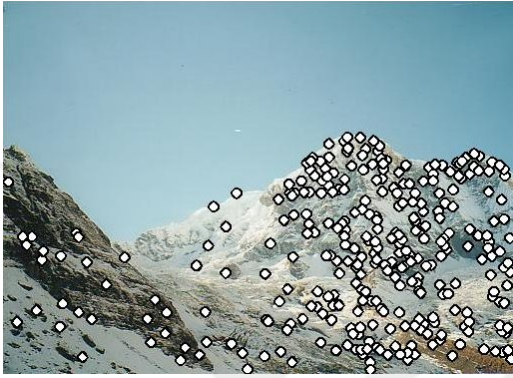
Eşleştirilmiş olan yer izleri arasından rastgele 5, 7 veya 8 adet öznitelik seçilir. Kullanılacak olan öznitelik sayısı seçilen matrise (esas matris, ana matris, homografi matrisi) göre belirlenmektedir. Seçilen bu rastgele yer izleri kullanılarak bir matris oluşturulmakta ve bu matris tekrar kullanılarak görüntüler arası yer izleri eşleştirilmektedir. Eşleştirme sayısı ne kadar çok ise en doğru matrisin bulunduğu karar verilmektedir. Şekil 6.4'te RANSAC kullanılarak yanlış eşleştirmelerden ayrılmış bir örnek gösterilmiştir.



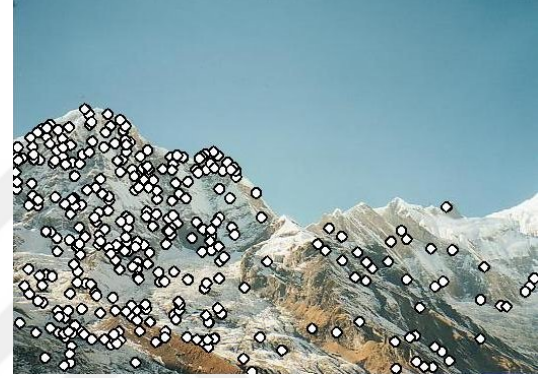
(a)



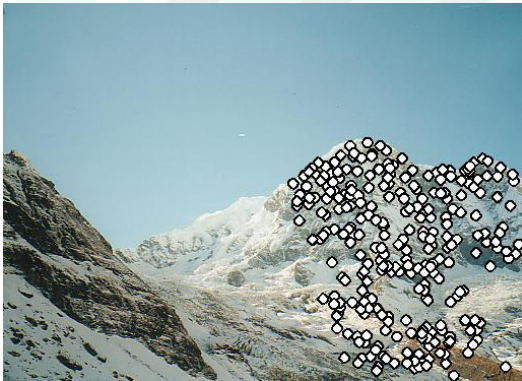
(b)



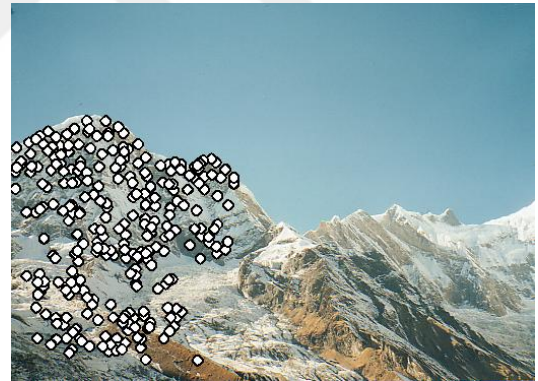
(c)



(d)



(e)



(f)

Şekil 6.4. Yanlış eşleştirmeye neden olan yer izlerinin belirlenmesi a-b) giriş görüntüleri c-d) ilk eşleştirilen yer izleri e-f) RANSAC kullanılarak elde edilen doğru eşleştirme sağlayan yer izleri

RANSAC kullanılarak elde edilen konum değişim değerleri daha yüksek doğruluğa sahiptir. Tez kapsamında aykırı eşleştirmelerin elenmesi için RANSAC algoritması kullanılmıştır.

6.3 Kalman Filtresi

Kalman (1960) tarafından önerilen oldukça popüler bir filtredir. Dinamik bir sistem modelinin önceki durumlarını değerlendirerek bir sonraki durumu tahmin etmede kullanılmaktadır. Kalman filtresi kestirim, konumlama ve gezinim uygulamalarında sıklıkla kullanılan özyinelemeli, olasılıksal ve gürültü bağışıklığı yüksek bir kestiricidir.

6.3.1 Kalman Filtre Modeli

Sistem tarafından t_{k-1} anında gözlemlenen dünya koordinatları $\vec{x}'_{k-1} = (X, Y, Z)^T$ ve hız vektörleri $\vec{v}'_{k-1} = (\dot{X}, \dot{Y}, \dot{Z})^T$ olsun. Kamera hareket ettiği sürece $\vec{x}'$ ve $\vec{v}'$ vektörlerinde de ∇t_k kadar sürede oluşan bir konum değişikliği olmaktadır. Böylece aracın konumu,

$$\vec{x}'_k = \hat{R}_k \vec{x}'_{k-1} + \vec{\hat{d}}_k + \nabla t_k \hat{R}_k \vec{v}'_{k-1} \quad (6.21)$$

ile tanımlanır. Burada $\hat{R}_k$ ve $\vec{\hat{d}}_k$, kameranin rotasyon ve konum değiştirme vektörleridir. Hız vektörü $\vec{v}'_k$ yönünü,

$$\vec{v}'_k = \hat{R}_k \vec{v}'_{k-1} \quad (6.22)$$

ile değiştirmektedir. Konum değişimi ve hız vektörlerinin birleştirilmesi ile oluşturulan durum vektörü $\vec{x}_k = (X, Y, Z, \dot{X}, \dot{Y}, \dot{Z})^T$, doğrusal ayırık bir sistem oluşturur. Bu model,

$$\vec{x}'_k = A_k \vec{x}_{k-1} + B_k + \vec{\omega} \quad (6.23)$$

ile ifade edilir. Böylece oluşan durum geçiş matrisi ve kontrol matrisi sırasıyla,

$$A_k = \begin{bmatrix} \hat{R}_k & \nabla t_k \hat{R}_k \\ 0 & \hat{R}_k \end{bmatrix} \quad (6.24)$$

$$B_k = \begin{bmatrix} \vec{\hat{d}}_k \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (6.25)$$

ile verilmiştir. Gürültü vektörü $\vec{\omega}$ 'nin kovaryans matrisi Q olan Gauss beyaz gürültüsü olduğu kabul edilmiştir.

Ölçüm modeli stereo kameralardan ve öznitelik eşleştirmelerinden elde edilen bilgiler ile oluşturulmaktadır. Başlangıç olarak $(0, yükseklik, 0)^T$ konumunda olduğu kabul edilen bir kameranın, yer ile paralel bir optik açısı bulunmaktadır. Kameranın koordinat sisteminde bir noktanın doğrusal olmayan ölçüm,

$$z = \begin{bmatrix} u \\ v \\ d \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} X & f_u \\ Y & f_v \\ b & f_u \end{bmatrix} + \vec{v} \quad (6.26)$$

ile ifade edilir. (u, v) ilgi noktalarından elde edilen bölgelerin güncellenmiş konumlarını, d ise stereo görüntü ile ölçülen mesafe ve (f_u, f_v) ise sırasıyla genişlik ve yükseklik olmak üzere piksel cinsinden odak uzaklığını belirtmektedir. $\vec{v}$ gürültü terimi ise kovaryans matrisi $\mathcal{R}$ olan Gauss beyaz gürültüsü olduğu kabul edilmiştir.

Böylece stereo özniteliklerden elde edilen derinlik bilgisi ve optik kayma bilgisi birleştirilerek konum ve hız bilgisi için yinelemeli bir kestirim sistemi elde edilir.

BÖLÜM VII

GÖRSEL SEYİR SİSTEMİ BENZETİM SONUÇLARI

7.1 Giriş

Bu bölümde, detayları önceki bölümlerde verilmiş olan görsel seyir sisteminin benzetim sonuçları incelenecektir. Benzetim için literatürdeki veri kümelerinden faydalanılacaktır.

7.2 Benzetimde Kullanılan Veri Kümeleri

Benzetim için *New College* ve *Kitti* veri kümeleri kullanılmıştır (Geiger vd., 2012; Smith vd., 2009). İlk olarak *New College* veri kümesi incelenecektir. Bu veri kümesinin toplam boyutu 23.2 GB'tır. Hem lazer verileri hem de stereo kamera görüntüleri bulunmaktadır. Tez kapsamında sadece stereo kamera görüntüleri kullanılacaktır. Alınan veriler toplam 2.2 km uzunluğunda bir parkurdan elde edilmiştir. Veri kümesi, park ve kolej yerleşkesine ait görüntülerden oluşmaktadır. Toplanan stereo görüntüler saniyede 20 kareden oluşmaktadır. Şekil 7.1'de bu veri kümesine ait güzergah görülmektedir.



Şekil 7.1. *New College* veri kümesinin güzergahı (Smith vd., 2009)

New College veri kümesi iki bölüm olarak incelenmiştir. İlk kısım dairesel bir

yörünge'nin bulunduđu Şekil 7.1'de, 1 ile gösterilmiş olan bölgedir. İkinci kısımda ise veri kümesinin tamamı incelenmiştir. Burada 1 ile gösterilen kısım algoritma ile kestirilen konum değışiminde oluşacak kayma miktarını incelememiz açısından, veri kümesinin tamamı ise bütün senaryo sonunda elde edilecek kayma miktarını görmemiz açısından önem taşımaktadır. *New College* veri kümesine ait örnek kamera görüntüleri Şekil 7.2'de verilmiştir.



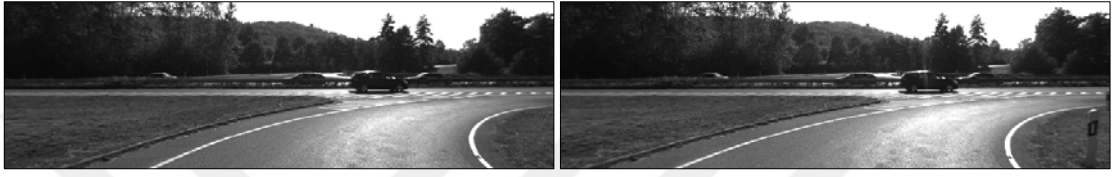
Şekil 7.2. *New College* veri kümesinden örnek görüntü kareleri a) Sol kamera görüntüleri b) Sağ kamera görüntüleri (Smith vd., 2009)

İkinci olarak incelenecek veri kümesi ise *Kitti* veri kümesidir. *Kitti* veri kümesi bir araç üzerine yerleştirilmiş stereo kameralardan alınan görüntüleri içerir. Şekil 7.3'te *Kitti* veri kümesi için kullanılan veri toplama sistemi gösterilmiştir.



Şekil 7.3. *Kitti* Veri kümesinin toplandığı platform (Geiger vd., 2012)

Araç, üzerinde Velodyne HDL-64E lazer tarayıcı ve 2 farklı stereo kamera kümesi ile kayıt yapmaktadır. Ayrıca araç üzerinde GPS ve IMU algılayıcıları da bulunmaktadır. GPS'ten elde edilen konum bilgileri algoritma için konum referansı olarak kullanılmıştır. Araç üzerine yerleştirilmiş kameralar gri ve renkli olmak üzere iki farklı kayıt yapmaktadır. Kamera çözünürlükleri 1.4 mega pikseldir. Kameralar, lazer tarayıcı tarafından tetiklenerek lazer verisi ile eş zamanlı olarak, saniyede 10 kare görüntü alacak şekilde ayarlanmıştır. Şekil 7.4'te *Kitti* veri kümesine ait örnek görüntü karesi verilmiştir.

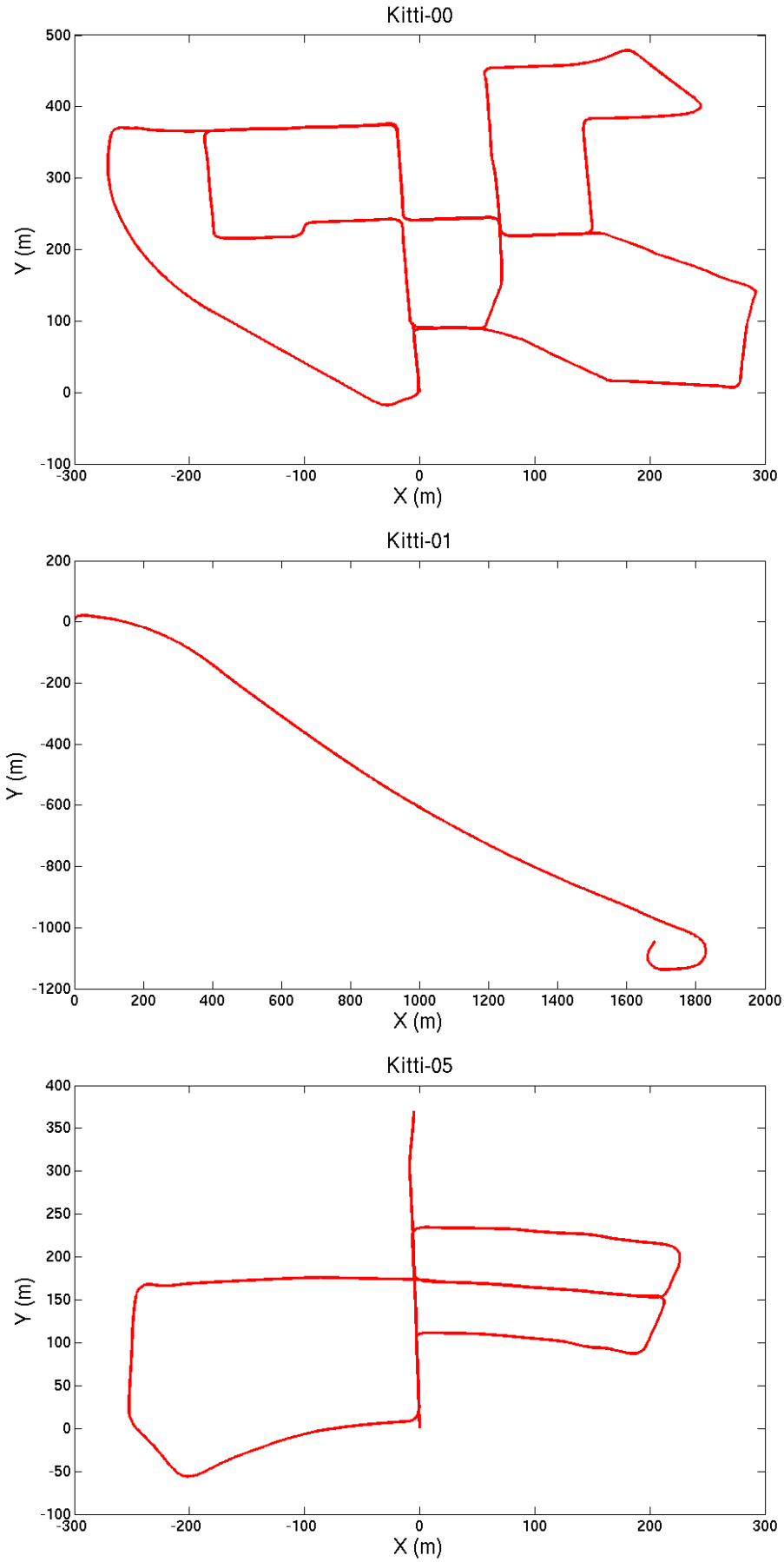


(a)

(b)

Şekil 7.4. *Kitti* veri kümesinden örnek görüntü kareleri a) Sol kamera görüntüleri b) Sağ kamera görüntüleri (Geiger vd., 2012)

Kitti veri kümesi ile araç konum verileri de veri kümesine dahil edildiği için bu veri kümesinde kullanılan yöntemler için karşılaştırmalı olarak hata grafikleri de elde edilebilmektedir. *Kitti* veri kümesi için 3 farklı veri kümesi senaryosu kullanılmıştır. Bu senaryolara ait rotalar Şekil 7.5'te verilmiştir.

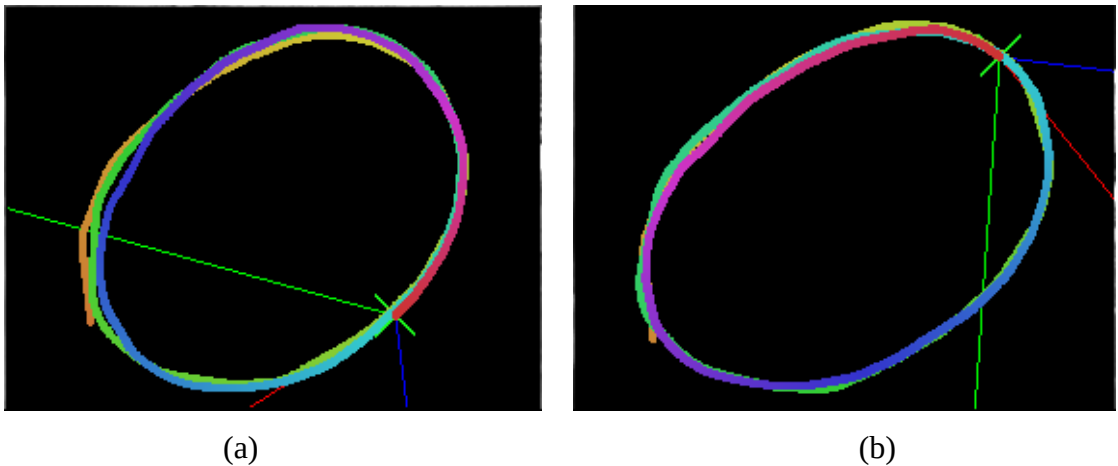


Şekil 7.5. *Kitti* veri kümesinde kullanılan farklı senaryolar için gerçek gezinim rotaları

7.3 Benzetim Sonuçları

Bölüm 6’da detayları verilen EZKB sistemi c/c++ ortamında yazılan bir program kullanılarak Linux ortamında test edilmiştir. Gerçekleştirilen sistem için Badino (2015) tarafından yayınlanan uygulama iskeleti kullanılmıştır. Stereo kameralardan alınan görüntülerden ilk olarak yer izi çıkarımı yapılmıştır. İkinci olarak, her yer izi için tanımlayıcı çıkarımı yapılmıştır. Bu tanımlayıcılar önceki yer izi tanımlayıcıları ile karşılaştırılarak konum değişimi hesaplanmıştır. Hesaplanan konum değişimleri Kalman filtresi kullanılarak filtrelenmiştir. Yer izi çıkarımı için Shi-Tomasi, SIFT ve TPDK yer izi bulma algoritmaları kullanılmıştır. Çıkarılan yer izlerini tanımlamak için sırasıyla FREAK, SIFT ve SOH tanımlayıcıları kullanılmıştır. Veri kümelerinin toplandığı sistemlerin dinamiği bilindiğinden ve veri kümelerindeki görüntüler arası açısal değişimler minimum düzeyde olduğundan SOH algoritmasının τ kayma miktarı sıfır olarak kabul edilmiştir. Benzetim sonuçları Intel Core i7 3.4 GHz işlemciye sahip bir bilgisayar kullanılarak alınmıştır.

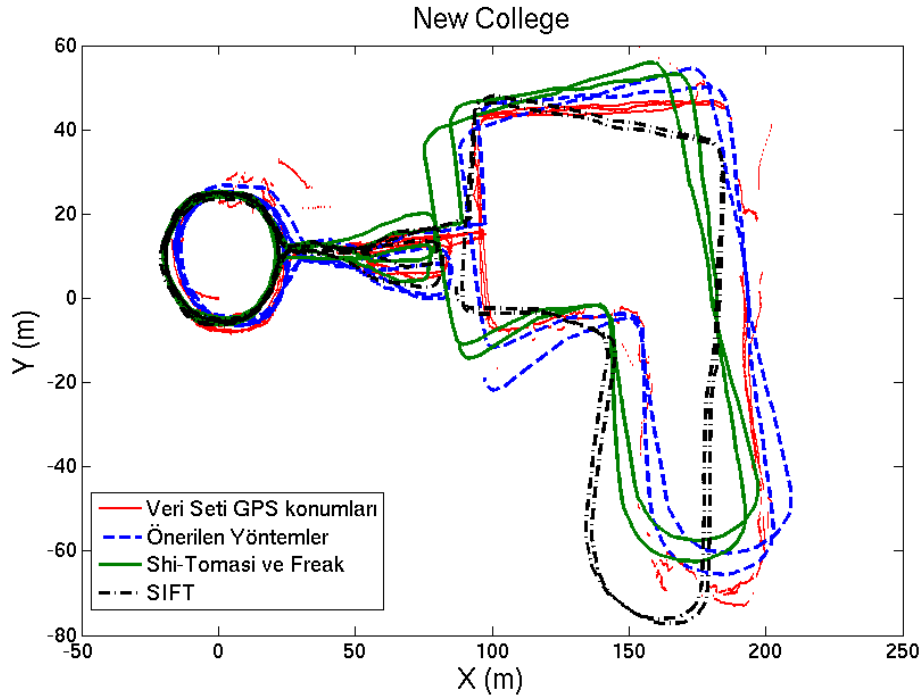
Önerilen yöntemler kullanılarak oluşturulan EZKB sisteminin konum belirleme performansını incelemek için ilk olarak *New College* veri kümesi kullanılmıştır. *New College* veri kümesinde daha önce bahsedilen 1 ile gösterilen kısımda yer alan dairesel rota kullanılmıştır.



Şekil 7.6. *New College* veri kümesi 1. bölgesinden elde edilen sonuçlar a) SIFT kullanılarak elde edilen sonuçlar b) önerilen yöntemler kullanılarak elde edilen sonuçlar

Burada görüldüğü üzere TPDK ve SOH yöntemlerinin kullanıldığı konum değişimi hesaplama sisteminde daha az bir kayma olduğu görülmektedir. Kaymanın az olması

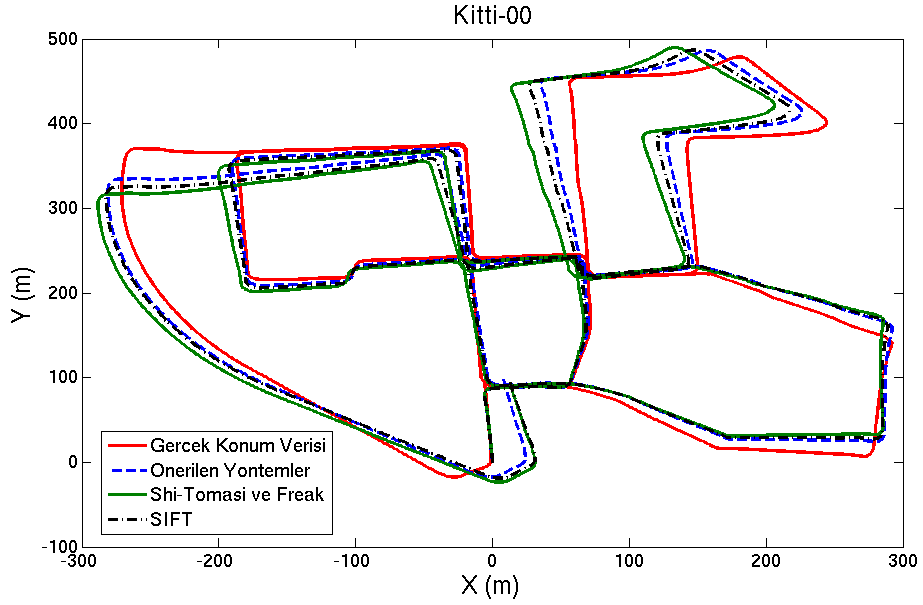
sistemin çıkardığı yer izlerinin tekrar tespit edilebilir ve tanımlayıcılar tarafından eşleştirilebilir olduğunu göstermektedir. *New College* veri kümesinin tamamı test için kullanıldığında ise Şekil 7.7’de gösterilen konum kestirimleri elde edilmektedir.



Şekil 7.7. *New College* veri kümesinden elde edilen konum bilgileri

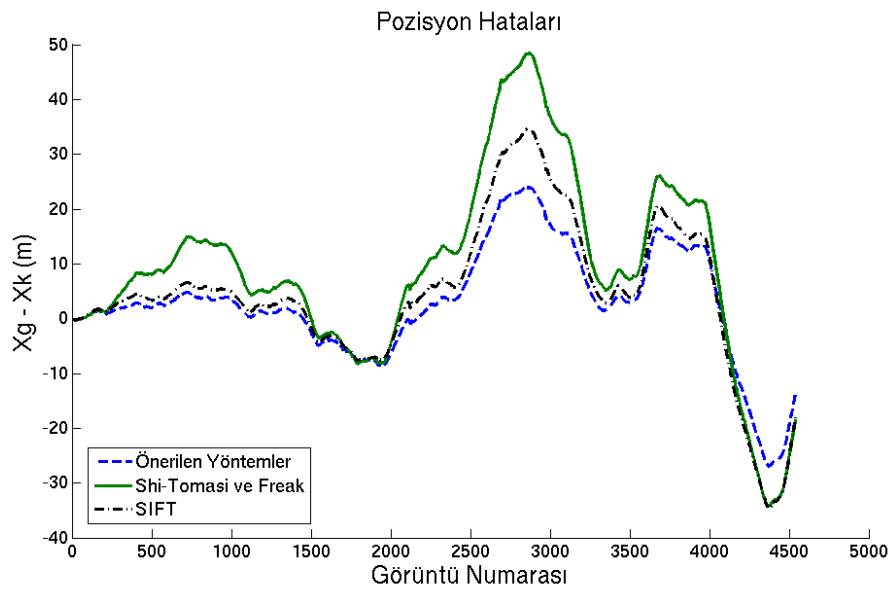
New College veri kümesi ile yayınlanan GPS verileri kırmızı ile gösterilmiştir. Tez kapsamında geliştirilen yöntemlerin kullanıldığı sistem ise mavi ile gösterilmiş olup konum kestiriminde diğer yöntemlere göre daha az kaymaya sahiptir. Burada hata analizi için sağlanmış gerçek konum bilgileri olmadığı için hata verileri incelenememiştir.

Kitti veri kümesinde ise ilk senaryo olan *Kitti-00* senaryosundan elde edilen konum kestirim sonuçları Şekil 7.8’de verilmiştir. X ve Y aracın hareket yönünü ifade etmektedir. Kameralardan alınan görüntülerdeki yatay düzlem X yönünü, görüntü düzleminin içine doğru olan eksen ise Y yönünü ifade etmektedir.

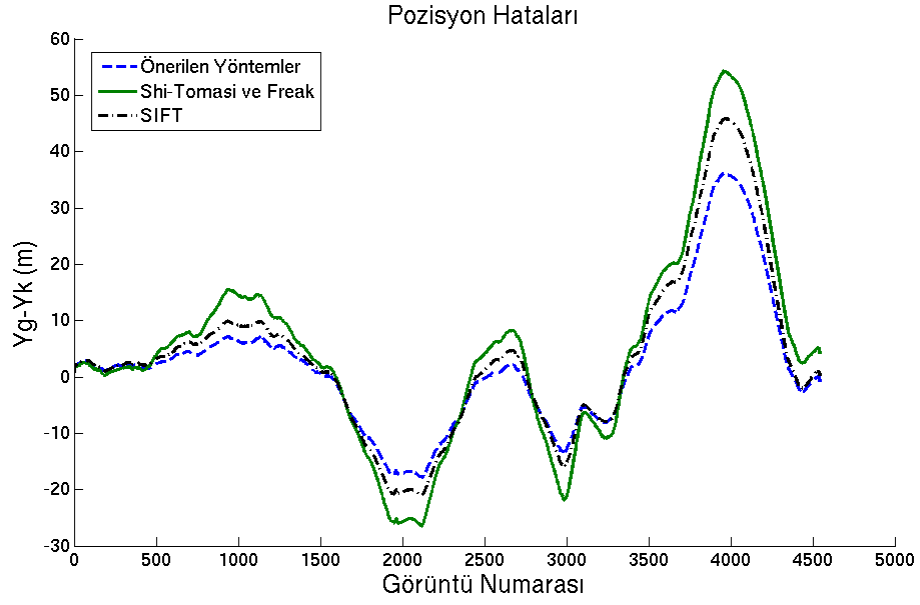


Şekil 7.8. *Kitti-00* senaryosundan elde edilen konum bilgileri

Şekil 7.8'den de görüleceği üzere konum kayması her yöntem için mevcuttur. Kaymanın olması uzun süreli veriler için beklenen bir sonuçtur. Kayma miktarının az olması demek kullanılan yer izlerinin ve tanımlayıcılarının konum değişimi için daha uygun olduğu anlamına gelmektedir. *Kitti-00* senaryosundan elde edilen veriler incelendiği zaman X ve Y yönlerinde elde edilen hata grafikleri de Şekil 7.9 ile verilmiştir. $X_g - X_k$ ve $Y_g - Y_k$, gerçek X_g ve Y_g konum değerleri ile kestirilen X_k ve Y_k konum değerleri arasındaki farkı ifade etmektedir.



(a)



(b)

Şekil 7.9. *Kitti-00* senaryosundan elde edilen hata grafikleri a) X yönündeki hatalar b) Y yönündeki hatalar

Tez kapsamında önerilen yöntemlerle oluşturulan sistemden elde edilen hatalar diğer yöntemlere göre daha düşük çıkmaktadır. Ortalama karesel hatanın karekökü (RMSE) ve ortalama karesel hata (MSE),

$$RMSE = \sqrt{\frac{1}{N} \sum_{1}^N ((Xg - Xk)^2)} \quad (7.1)$$

$$MSE = \frac{1}{N} \sum_{1}^N ((Xg - Xk)^2) \quad (7.2)$$

ile ifade edilmektedir. Hata değerleri incelendiğinde Çizelge 7.1 ve Çizelge 7.2’de verilen sonuçlar elde edilmiştir.

Çizelge 7.1. *Kitti-00* senaryosundan elde edilen X yönündeki ortalama hatalar

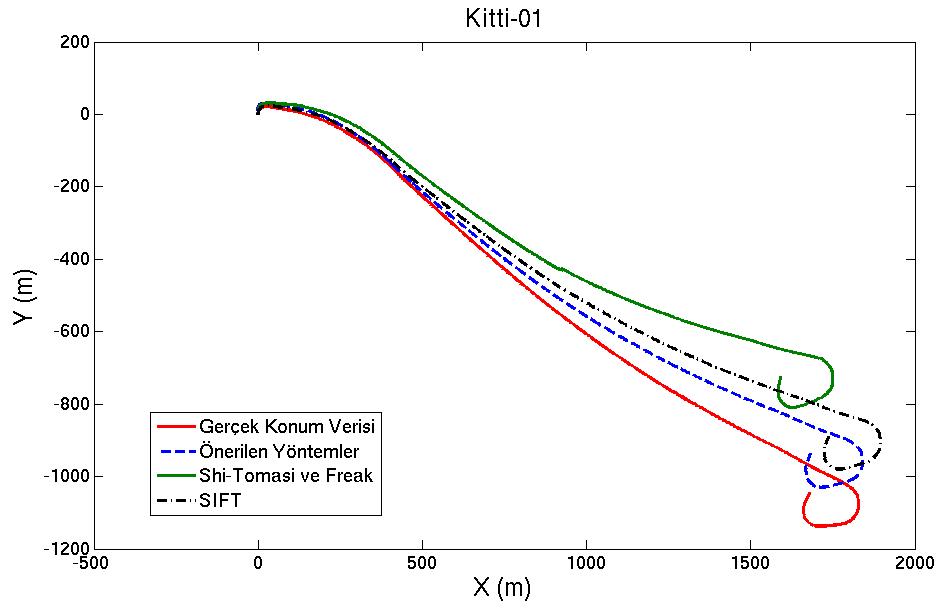
Hata Türü	Yöntem		
	Shi-Tomasi ve Freak	SIFT	Önerilen Yöntemler
MSE	378.40	212.77	116.95
RMSE	19.45	14.59	10.81

Çizelge 7.2. *Kitti-00* senaryosundan elde edilen Y yönündeki ortalama hatalar

Hata Türü	Yöntem		
	Shi-Tomasi ve Freak	SIFT	Önerilen Yöntemler
MSE	354.18	236.60	148.63
RMSE	18.82	15.38	12.19

Çizelge 7.1 ve Çizelge 7.2’de verilen değerlerden önerilen yöntemlerin X ve Y yönünde 10.81 ve 12.19 metre ortalama hataya sahip olduğu görülmektedir. Bu ortalama hatalar SIFT ile elde edilen hatalardan yaklaşık olarak 3-4 metre ve Shi-Tomasi ve FREAK ile elde edilen hatalardan ise yaklaşık olarak 6-9 metre daha iyi olduğu görülmektedir.

İkinci senaryoda ise Şekil 7.10 ile verilen konum kestirim sonuçları elde edilmiştir.



Şekil 7.10. *Kitti-01* senaryosundan elde edilen konum bilgileri

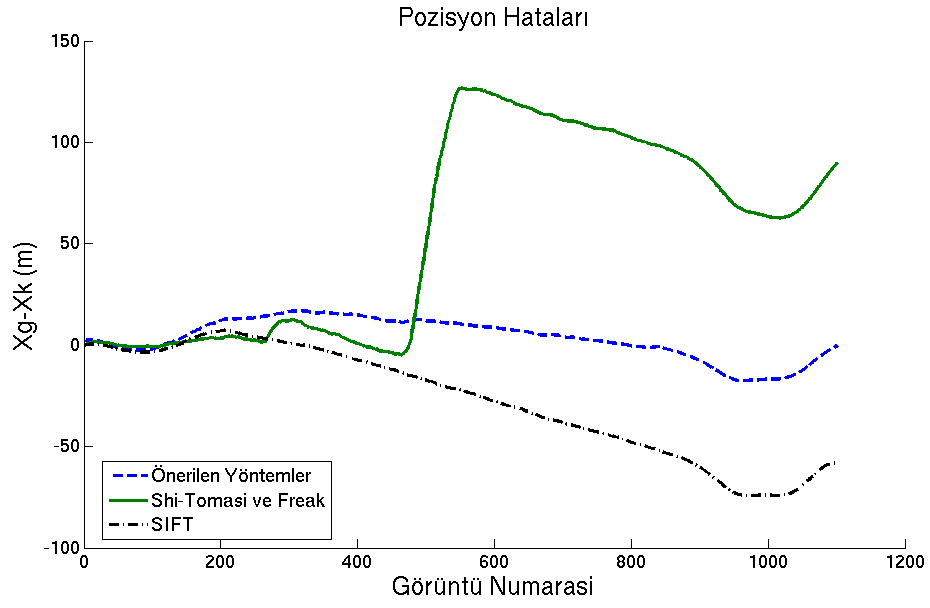
Şekil 7.10’dan da görüleceği üzere konum kayması bu senaryoda daha belirgin görülmektedir. İlk senaryoya göre daha kısa bir rotaya sahip olmasına rağmen burada alınan sonuçlarda Y yönündeki kaymanın daha çok olduğu görülmektedir. Y yönündeki kaymaya sebep olan durum ise bu senaryodan verilebilecek bir görüntü karesi ile incelenebilir. Örnek görüntü kareleri Şekil 7.11’de verilmiştir.



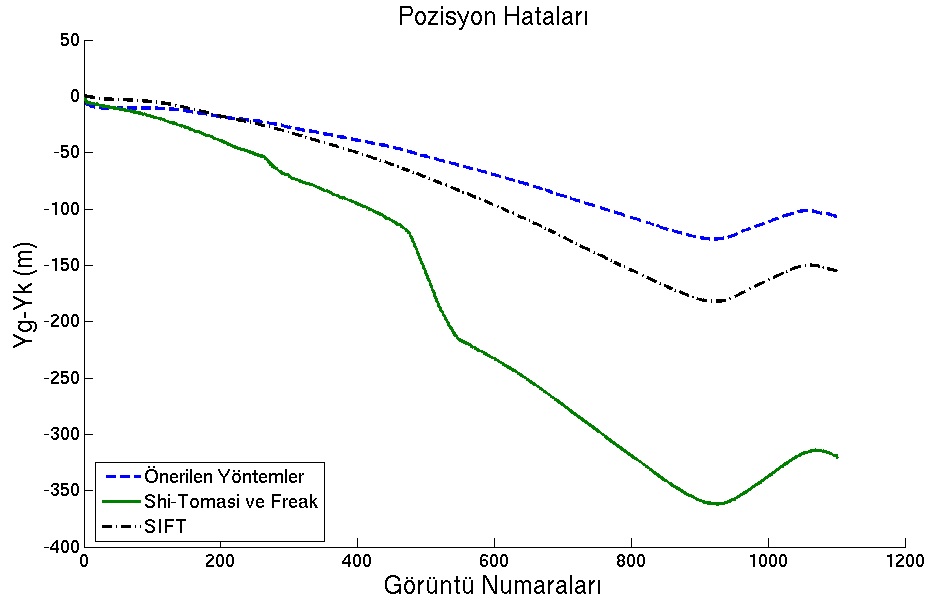
Şekil 7.11. *Kitti-01* senaryosundan örnek görüntü kareleri (Geiger vd., 2012)

Görüntülerden anlaşılacağı üzere çekim yapılan rotada ufuk çizgisinin uzaklığından dolayı çıkartılan öznitelikler kameranın konumuna göre fazla değişim göstermemektedir. Bu durumda elde edilecek olan konum değişimlerinin doğrulukları da daha düşük olmaktadır. Dolayısıyla ilk durumda uzun bir rotada elde edilen hata değerlerine yakın hatalar elde edilmiştir.

Kitti-01 senaryosunda X ve Y yönlerinde elde edilen hata grafikleri Şekil 7.12 ile verilmiştir.



(a)



(b)

Şekil 7.12. *Kitti-01* senaryosundan elde edilen hata grafikleri a) X yönündeki hatalar b) Y yönündeki hatalar

Önerilen yöntemlerden elde edilen sonuçlar grafiklerden de görüleceği gibi diğer yöntemlere göre daha düşük çıkmaktadır. İlk senaryoda olduğu gibi bu verilerden elde edilen RMSE ve MSE değerleri Çizelge 7.3 ve Çizelge 7.4’de verilmiştir

Çizelge 7.3. *Kitti-01* senaryosundan elde edilen X yönündeki ortalama hatalar

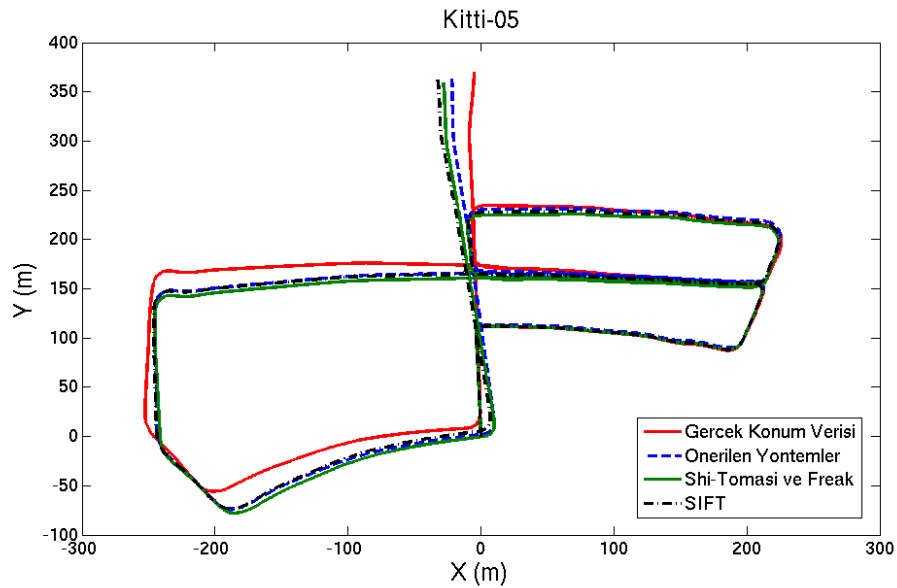
Hata Türü	Yöntem		
	Shi-Tomasi ve Freak	SIFT	Önerilen Yöntemler
MSE	5269.54	1467.40	109.04
RMSE	72.59	38.30	10.44

Çizelge 7.4. *Kitti-01* senaryosundan elde edilen Y yönündeki ortalama hatalar

Hata Türü	Yöntem		
	Shi-Tomasi ve Freak	SIFT	Önerilen Yöntemler
MSE	51502.31	11713.37	5761.31
RMSE	226.94	108.23	75.90

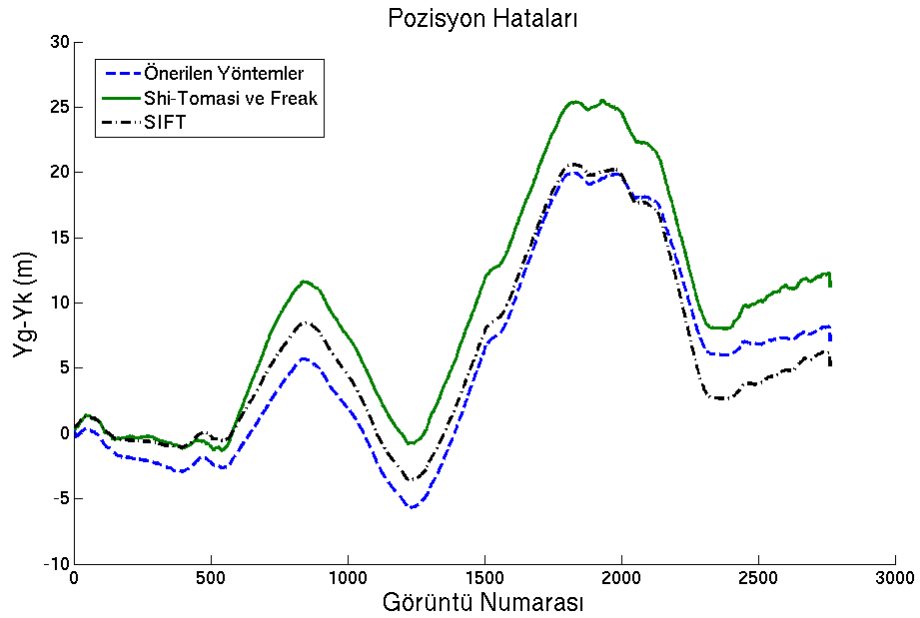
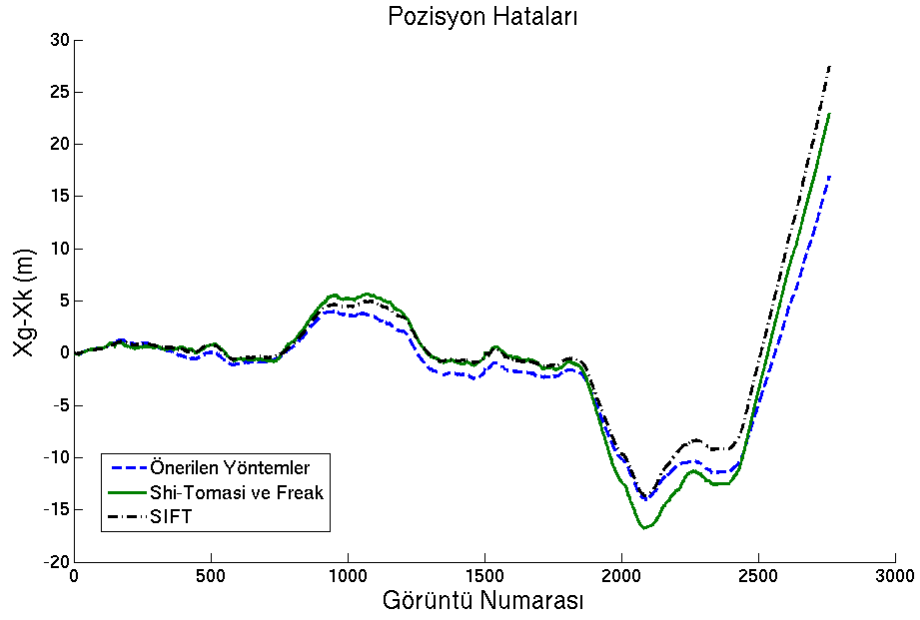
Çizelge 7.3 ve Çizelge 7.4’de verilen değerler incelendiğinde önerilen yöntemlerin x ve y yönünde 10.44 ve 75.90 metre ortalama hataya sahip olduğu görülmektedir. Bu ortalama hatalar SIFT ile elde edilen hatalardan yaklaşık olarak 28-33 metre ve Shi-Tomasi ve Freak ile elde edilen hatalardan ise yaklaşık olarak 62-150 metre daha iyi olduğu görülmektedir.

Son senaryoda ise Şekil 7.13 ile verilen konum kestirim sonuçları elde edilmiştir.



Şekil 7.13. *Kitti-05* senaryosundan elde edilen konum bilgileri

Bu senaryoda ışık değişimlerinin çok olduğu bir rota incelenmiştir. Bu rotada güneşin çok olduğu ve gölge olan alanların bol olması konum kestirimini zorlaştırmaktadır. *Kitti-05* senaryosundan elde edilen veriler incelendiği zaman X ve Y konumlarında elde edilen hata grafikleri Şekil 7.14'te verilmiştir.



Şekil 7.14. *Kitti-05* senaryosundan elde edilen hata grafikleri a) X yönündeki hatalar b) Y yönündeki hatalar

Önerilen yöntemlerden elde edilen sonuçlar grafiklerden de görüleceği gibi diğer yöntemlere X yönünde daha iyi iken Y yönünde senaryonun son kısımlarına doğru bozulmaktadır ve daha kötü sonuç vermektedir. Hata verilerinden elde edilen RMSE ve MSE değerleri ise Çizelge 7.5 ve Çizelge 7.6’da verilmiştir.

Çizelge 7.5. *Kitti-05* senaryosundan elde edilen X yönündeki ortalama hatalar

Hata Türü	Yöntem		
	Shi-Tomasi ve Freak	SIFT	Önerilen Yöntemler
MSE	51.5	44.3	34.3
RMSE	7.18	6.66	5.86

Çizelge 7.6. *Kitti-05* senaryosundan elde edilen Y yönündeki ortalama hatalar

Hata Türü	Yöntem		
	Shi-Tomasi ve Freak	SIFT	Önerilen Yöntemler
MSE	155.12	84.94	86.53
RMSE	12.45	9.22	9.30

Çizelge 7.5 ve Çizelge 7.6’da verilen değerler incelendiğinde önerilen yöntemlerin X ve Y yönünde 5.86 ve 9.30 metre ortalama hataya sahip olduğu görülmektedir. Bu ortalama hatalara göre önerilen yöntemler, X yönünde diğer yöntemlerden daha iyi olmasına rağmen, Y yönünde oluşan 0.08 metrelik bir hata ile SIFT’in gerisinde yer almıştır. Bunun nedeni, veri kümesinin son görüntülerinde ufuk çizgisinin daha uzakta olması ve bu bağlı olarak nispeten daha az öznetelik eşleştirilebilmesidir. Fakat önerilen yöntemler Y yönünde Shi-Tomasi ve FREAK ile alınan sonuçtan daha başarılıdır.

BÖLÜM VIII

SONUÇ VE ÖNERİLER

İnsansız araçlar ile ilgili çalışmalar son zamanlarda çok önem kazanmıştır. Bu araçların her türlü fiziksel zorluk altında hareket edebilmeleri büyük önem arz etmektedir. İnsansız araçların içinde bulundukları ortam ve kendi konumları hakkında bilgi sahibi olmaları gerekmektedir. Görsel algılayıcı sistemler, çevresel farkındalık için bilgi sağlaması ve aracın konumunu belirlemeye imkân sunması nedeniyle insansız araçlarda kullanıma uygundur.

Bu çalışmada, görsel sistemlerden elde edilen bilgiler kullanılarak yapılan hesaplamalar ile seyir halindeki bir insansız aracın konum takibi gerçekleştirilmiştir. Tez çalışmasında önerilen yöntemler literatürdeki en yaygın veri kümeleri kullanılarak test edilmiştir. Testler sonucunda önerilen algoritmaların, mevcut yöntemlerden daha başarılı olduğunu gösterilmiştir. Bu tez kapsamında yer izi tespitinde, tanımlayıcıların yönelim açısı hesabında ve tanımlayıcı eşleştirmede özgün yeni yöntemler önerilmiştir. Ayrıca tanımlayıcı eşleştirme yönteminin başarımını arttıran özgün yeni bir tanımlayıcı deseni önerilmiştir.

Önerilen yer izi tespit yöntemi (TPDK) ile yeniden tespit oranı yüksek bir algoritma elde edilmiştir. Bu yöntem ile havadan çekilmiş olan şehir görüntülerinden elde edilen noktaların eşleştirme sayısı 1468 iken, SIFT ile 781'dir. Şehir görüntüsüne göre daha az desen içeren hava görüntülerinde ise TPDK ile 1496 yer izi eşleştirilmişken, SIFT ile 229 eşleştirme elde edilmiştir. Gerçekleştirilen testlerin sonucunda TPDK'nın, SIFT algoritmasına göre en kötü durumda bile iki kat daha fazla eşleştirme sağladığı gözlemlenmiştir.

Yönelim açısı hesabı ile ilgili olarak, yer izi noktalarından elde edilen görüntü parçalarından, yönelim açılarının etkin ve hızlı bir şekilde bulunmasını sağlayan bir algoritma önerilmiştir. Bu yönelim açısı hesaplama yöntemi (SIH), diğer yönelim açısı hesabı gerektiren algoritmalara kolaylıkla entegre edilebilir. Ayrıca bu algoritmanın, görüntüde oluşabilecek ışık, boyut ve açısal değişimlerden etkilenmediği gibi görüntüde oluşabilecek bulanıklıktan da etkilenmediği gösterilmiştir. SIH algoritması ile SIFT

tanımlayıcılarının başarımı arttırılmış ve çıkarım süresi kısaltılmıştır.

Bu tez kapsamında önerilen diğer bir yenilik ise SOH (Segmented Orientation Histograms) tanımlayıcıları ve bu tanımlayıcıların eşleştirme mekanizmasıdır. Özellikle EZKH (Eş Zamanlı Konumlama ve Haritalama) uygulamalarında, görüntüler arası dönme açısının değeri sistem dinamiği veya algılayıcılar yardımıyla yaklaşık olarak belirlenebilir. Dönme açısının yaklaşık olarak bilindiği bu gibi durumlarda, AMDF algoritmasındaki τ kayma miktarı, dönme açısına bağlı olarak sınırlandırılabilir. Bu sınırlandırma ile çalışma zamanının yüksek olmasına neden olan çoklu τ kaydırmalarının önüne geçilerek algoritmanın performansından ödün vermeden eşleştirme hızı ciddi oranda arttırılmıştır. Ayrıca gerçekleştirilen testlerde de görüldüğü üzere SOH tanımlayıcıları, başarımı yüksek diğer algoritmaların tanımlayıcılarına göre hafızada daha az yer kaplar ve daha başarılı eşleştirme performansı sergilemektedir. Gerçekleştirilen EZKB sisteminden de görüldüğü üzere, SOH algoritmasının performansı, FREAK ve SIFT algoritmalarının performansından daha iyi olduğu görülmüştür.

Tez çalışmasında geliştirilen SOH tanımlayıcılarının performansı yüksek olmasına rağmen, görüntülerdeki dönme açısının bilinmediği veya tahmin edilemediği durumlarda, SOH algoritmasının AMDF eşleştirmesi ile çalışma hızı oldukça yavaş kalmaktadır. Bu problem, AMDF algoritmasının yerine daha hızlı bir algoritma geliştirilerek veya AMDF algoritmasının paralel işlem kabiliyeti olan donanımlarda çalıştırılması ile aşılabılır.

KAYNAKLAR

Abdel-Hakim A.E. and Farag, A.A., "CSIFT: A SIFT Descriptor with Color Invariant Characteristics", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* s. 1978–1983 2006.

Alahi, A., Ortiz, R. and Vandergheynst, P., "Freak: Fast retina keypoint", *Computer Vision and Pattern Recognition (CVPR)* s. 510–517, 2012.

Ambai, M. and Yoshida, Y., "CARD: Compact And Real-time Descriptors", *IEEE International Conference on Computer Vision (ICCV)* s. 97–104, 2011.

Angeli, A., Filliat, D., Doncieux, S. and Meyer, J.-A., "2D Simultaneous Localization And Mapping for Micro Air Vehicles", *European Micro Aerial Vehicles (EMAV)* 2006.

Antsaklis, P.J., Passino, K.M. and Wang, S.J., "An introduction to autonomous control systems", *IEEE Control Systems* 11, 5–13, 1991.

Aulinas, J., Llado, X., Salvi, J. and Petillot, Y.R., "Selective Submap Joining for underwater large scale 6-DOF SLAM", *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* s. 2552–2557, 2010.

Bachrach, A., Prentice, S., He, R., Henry, P., Huang, A.S., Krainin, M., Maturana, D., Fox, D. and Roy, N., "Estimation, planning, and mapping for autonomous flight using an RGB-D camera in GPS-denied environments", *International Journal of Robotics Research* s. 31, 1320–1343, 2012.

QCV v1.03, Badino, H., Computer Vision Framework Library (for Linux) 2015.

Bailey, T., "Mobile Robot Localisation and Mapping in Extensive Outdoor Environments", 2002.

Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L., "Speeded-Up Robust Features (SURF)", *Computer Vision and Image Understanding* 110, 346–359, 2008.

Beall, C., Lawrence, B.J., Ila, V. and Dellaert, F., "3D reconstruction of underwater structures", *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* s. 4418–4423, 2010.

Bellavia, F., Tegolo, D. and Valenti, C., "Keypoint descriptor matching with context-based orientation estimation", *Image and Vision Computing* 32, 559–567, 2014.

Berg, A.C. and Malik, J., "Geometric blur for template matching", *Computer Vision and Pattern Recognition CVPR* s. I–607, 2001.

Besada-Portas, E., de la Torre, L., Moreno, A. and Risco-Martín, J.L., "On the performance comparison of multi-objective evolutionary UAV path planners", *Information Science* 238, 111–125, 2013.

Biederman, I., "Recognition-by-components: a theory of human image understanding.", *Psychological Review* 94, 115, 1987.

Bleiweiss, A., "Parallel compact roadmap construction of 3D virtual environments on the GPU", *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* s. 5007–5013, 2010.

Brown, M. and Lowe, D., "Invariant Features from Interest Point Groups", *British Machine Vision Association*, s. 23.1–23.10, 2002.

Bryson, M. and Sukkarieh, S., "Co-operative Localisation and Mapping for Multiple UAVs in Unknown Environments", *IEEE Aerospace Conference* s. 1–12, 2007.

Burt, P.J. and Adelson, E.H., "The Laplacian Pyramid as a Compact Image Code", *IEEE Transactions on Communications* 31, 532–540, 1983.

Calonder, M., Lepetit, V., Ozuysal, M., Trzcinski, T., Strecha, C. and Fua, P., "BRIEF:

Computing a Local Binary Descriptor Very Fast", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34, 1281–1298, 2012.

Calonder, M., Lepetit, V., Strecha, C. and Fua, P., "BRIEF: Binary Robust Independent Elementary Features", *Lecture Notes in Computer Science Springer Berlin Heidelberg* ss. 778–792, 2010.

Canny, J., "A Computational Approach to Edge Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-8, 679–698, 1986.

Chang, H.J., Lee, C.S.G., Lu, Y.-H. and Hu, Y.C., "P-SLAM: Simultaneous Localization and Mapping With Environmental-Structure Prediction", *IEEE Transactions on Robotics* 23, 281–293, 2007.

Chen, H., Sun, D., Yang, J. and Chen, J., "Localization for Multirobot Formations in Indoor Environment", *IEEE/ASME Transactions on Mechatronics* 15, 561–574, 2010.

Chen, J., Shan, S., He, C., Zhao, G., Pietikainen, M., Chen, X. and Gao, W., "WLD: A Robust Local Image Descriptor", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, 1705–1720, 2010.

Crowley, J.L. and Sanderson, A.C., "Multiple Resolution Representation and Probabilistic Matching of 2-D Gray-Scale Shape", *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-9, 113–121, 1987.

Deng, H., Zhang, W., Mortensen, E., Dietterich, T. and Shapiro, L., "Principal Curvature-Based Region Detector for Object Recognition", *IEEE Conference on Computer Vision and Pattern Recognition CVPR07*. s. 1–8, 2007.

Dubey, S.R., Singh, S.K. and Singh, R.K., "Rotation and Illumination Invariant Interleaved Intensity Order-Based Local Descriptor", *IEEE Transactions on Image Processing* 23, 5323–5333, 2014.

Elinas, P., Sim, R. and Little, J.J., "SLAM: stereo vision SLAM using the Rao-

Blackwellised particle filter and a novel mixture proposal distribution", *IEEE International Conference on Robotics and Automation, ICRA* s. 1564–1570, 2006.

Fan, B., Wu, F. and Hu, Z., "Rotationally Invariant Descriptors Using Intensity Order Pooling", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34, 2031–2045, 2012.

Fan, B., Wu, F. and Hu, Z., "Aggregating gradient distributions into intensity orders: A novel local image descriptor", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* s. 2377–2384, 2011.

Fischler, M.A. and Bolles, R.C., "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", *Commun ACM* 24, 381–395, 1981.

Forssen, P.E. and Lowe, D.G., "Shape Descriptors for Maximally Stable Extremal Regions", *IEEE 11th International Conference on Computer Vision ICCV 2007*. s. 1–8, 2007.

Geiger, A., Lenz, P. and Urtasun, R., "Are we ready for autonomous driving? the kitti vision benchmark suite", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* s. 3354–3361, 2012.

Goerzen, C., Kong, Z. and Mettler, B., "A Survey of Motion Planning Algorithms from the Perspective of Autonomous UAV Guidance", *Journal of Intelligent and Robotic Systems* 57, 65–100, 2010.

Gonzalez, R.C. and Woods, R.E., "Digital image processing", 2nd ed, *Prentice-Hall [u.a.], Upper Saddle River, NJ*, 2002.

Grzonka, S., Grisetti, G. and Burgard, W., "A Fully Autonomous Indoor Quadrotor", *IEEE Transactions on Robotics* 28, 90–100, 2012.

Harris, C. and Stephens, M., "A combined corner and edge detector", *In Proc. of*

Fourth Alvey Vision Conference s. 147–151, 1988.

Hartley, R. and Zisserman, A., "Multiple View Geometry in Computer Vision", ***Cambridge University Press***, 2003.

Havangi, R., Teshnehlal, M., Nekoui, M.A. and Taghirad, H., "An adaptive Neuro-Fuzzy Rao-Blackwellized particle filter for SLAM", ***IEEE International Conference on Mechatronics (ICM)*** s. 487–492, 2011.

Heikkilä, M., Pietikäinen, M. and Schmid, C., "Description of interest regions with local binary patterns", ***Pattern Recognition*** 42, 425–436, 2009.

Heitger, F., Rosenthaler, L., Von Der Heydt, R., Peterhans, E. and Kübler, O., "Simulation of neural contour mechanisms: from simple to end-stopped cells", ***Vision Research*** 32, 963–981, 1992.

Huang, G.P., Mourikis, A.I. and Roumeliotis, S.I., "Analysis and improvement of the consistency of extended Kalman filter based SLAM", ***IEEE International Conference on Robotics and Automation. ICRA 2008*** s. 473–479, 2008.

Huang, S. and Dissanayake, G., "Convergence and Consistency Analysis for Extended Kalman Filter Based SLAM", ***IEEE Transactions on Robotics*** 23, 1036–1049, 2007.

Hui, L., Dai, B.-Q. and Wei, L., "A Pitch Detection Algorithm Based on AMDF and ACF", ***IEEE International Conference on Acoustics Speech and Signal Processing Proceedings***. s. I–I, 2006.

Hwang, S.-Y. and Song, J.-B., "Monocular Vision-Based SLAM in Indoor Environment Using Corner, Lamp, and Door Features From Upward-Looking Camera", ***IEEE Transactions on Industrial Electronics*** 58, 4804–4812, 2011.

Inoue, T., Shibuya, K. and Nagano, A., "Underwater robot with a buoyancy control system based on the spermaceti oil hypothesis development of the depth control system", ***IEEE/RSJ International Conference on Intelligent Robots and Systems***

(**IROS**) s. 1102–1107, 2010.

Jianli, S., Shuang, P., Yugiang, W. and Xibin, W., "Unscented FastSLAM for UAV", ***International Conference on Computer Science and Network Technology (ICCSNT)*** s. 2529–2532, 2011.

Jolliffe, I.T., "Principal component analysis", ***Springer-Verlang*** 1986.

Kallapur, A.G. and Anavatti, S.G., "UAV linear and nonlinear estimation using extended Kalman filter", ***International Conference on Intelligent Agents, Web Technologies and Internet Commerce*** s. 250–250, 2006.

Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems", ***Journal of Basic Engineering*** 82, 35–45, 1960.

Ke, Y. and Sukthankar, R., "PCA-SIFT: a more distinctive representation for local image descriptors", ***Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*** s. II–506–II–513 Vol.2, 2004.

Kim, A. and Eustice, R.M., "Real-Time Visual SLAM for Autonomous Underwater Hull Inspection Using Visual Saliency", ***IEEE Transactions on Robotics***, 2013.

Kok, J., Gonzalez, L.F. and Kelson, N., "FPGA Implementation of an Evolutionary Algorithm for Autonomous Unmanned Aerial Vehicle On-Board Path Planning", ***IEEE Transactions on Evolutionary Computation*** 17, 272–281, 2013.

Kuai, X., Yang, K., Fu, S., Zheng, R. and Yang, G., "Simultaneous localization and mapping (SLAM) for indoor autonomous mobile robot navigation in wireless sensor networks", ***International Conference on Networking, Sensing and Control (ICNSC)*** s. 128–132, 2010.

Lee, H.S. and Lee, K.-M., "Multiswarm Particle Filter for vision based SLAM", ***IEEE/RSJ International Conference on Intelligent Robots and Systems IROS 2009.*** s.

924–929, 2009.

Leonard, J.J. and Durrant-Whyte, H.F., "Simultaneous map building and localization for an autonomous mobile robot", *IEEE/RSJ International Workshop on Intelligent Robots and Systems '91. 'Intelligence for Mechanical Systems, Proceedings IROS '91* s. 1442–1447 vol.3, 1991.

Lindeberg, T., "Feature Detection with Automatic Scale Selection", *International Journal of Computer Vision* 30, 79–116, 1998.

Lindeberg, T., "Scale-space behaviour of local extrema and blobs", *Journal of Mathematical Imaging and Vision* 1, 65–99, 1992.

Liu, Y. and Thrun, S., "Results for outdoor-SLAM using sparse extended information filters", *IEEE International Conference on Robotics and Automation ICRA '03*. s. 1227–1233 vol.1, 2003.

Li, X. and Aouf, N., "Cooperative vSLAM based on UAV application", *IEEE International Conference on Robotics and Biomimetics (ROBIO)* s. 914–919, 2012.

Li, X., Aouf, N. and Nemra, A., "Estimation analysis in VSLAM for UAV application", *IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)* s. 365–370, 2012a.

Li, X., Aouf, N. and Nemra, A., "3D mapping based VSLAM for UAVs", *20th Mediterranean Conference on Control Automation (MED)* s. 348–352, 2012b.

Lowe, D.G., "Distinctive image features from scale-invariant keypoints", *International journal of computer vision* 60, 91–110, 2004.

Lui, W.L.D. and Jarvis, R., "A pure vision-based approach to topological SLAM", *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* s. 3784–3791, 2010.

Mao, G., Drake, S. and Anderson, B.D., "Design of an extended kalman filter for uav localization", ***Information, Decision and Control. IDC'07*** s. 224–229, 2007.

Matas, J., Chum, O., Urban, M. and Pajdla, T., "Robust wide-baseline stereo from maximally stable extremal regions", ***Image and Vision Computing*** 22, 761–767, 2004.

Meng, L., Li, L. and Veres, S.M., "Aerodynamic parameter estimation of an unmanned aerial vehicle based on extended kalman filter and its higher order approach", ***2nd International Conference on Advanced Computer Control (ICACC)*** s. 526–531, 2010.

Mikolajczyk, K. and Schmid, C., "A performance evaluation of local descriptors", ***IEEE Transactions on Pattern Analysis and Machine Intelligence*** 27, 1615–1630, 2005.

Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T. and Gool, L.V., "A Comparison of Affine Region Detectors", ***International Journal of Computer Vision*** 65, 43–72, 2005.

Moravec, H.P., "Visual Mapping by a Robot Rover", ***Proceedings of the 6th International Joint Conference on Artificial Intelligence - Volume 1, IJCAI'79*** s. 598–600, 1979.

Moreels, P. and Perona, P., "Evaluation of Features Detectors and Descriptors based on 3D Objects", ***International Journal of Computer Vision*** 73, 263–284, 2006.

Morel, J. and Yu, G., "ASIFT: A New Framework for Fully Affine Invariant Image Comparison", ***SIAM Journal on Imaging Sciences*** 2, 438–469, 2009.

Morevec, H.P., "Towards Automatic Visual Obstacle Avoidance", ***Proceedings of the 5th International Joint Conference on Artificial Intelligence - Volume 2, IJCAI'77*** s. 584–584, 1977.

Mortensen, E.N., Deng, H. and Shapiro, L., "A SIFT descriptor with global context", ***IEEE Computer Society Conference on Computer Vision and Pattern Recognition***

CVPR 2005 s. 184–190 vol. 1, 2005.

Muhammad, G., "Extended average magnitude difference function based pitch detection", *The International Arab Journal of Information Technology* 8, 197–203, 2011.

Nemra, A. and Aouf, N., "Robust cooperative UAV Visual SLAM", *IEEE 9th International Conference on Cybernetic Intelligent Systems (CIS)* s. 1–6, 2010.

Paz, L.M., Tardos, J.D. and Neira, J., "Divide and Conquer: EKF SLAM in", *IEEE Transactions on Robotics* 24, 1107–1120, 2008.

Prewitt, J.M.S., "Object enhancement and extraction", *Picture Processing and Psychopictorics* 75–149, 1970.

Rabiner, L., Cheng, M., Rosenberg, A. and McGonegal, C., "A comparative performance study of several pitch detection algorithms", *IEEE Transactions on Acoustics, Speech, and Signal Processing* 24, 399–418, 1976.

Rady, S., Kandil, A.A. and Badreddin, E., "A hybrid localization approach for UAV in GPS denied areas", *IEEE/SICE International Symposium on System Integration (SII)* s. 1269–1274, 2011.

Roberts, L.G., "Machine perception of three-dimensional solids". *Massachusetts Institute of Technology*, 1963.

Rosenthaler, L., Heitger, F., Kübler, O. and Heydt, R. von der, "Detection of general edges and keypoints", Sandini, G. (Ed.), *Lecture Notes in Computer Science. Springer Berlin Heidelberg*, s. 78–86 1992.

Rosin, P.L., "Measuring Corner Properties", *Computer Vision and Image Understanding* 73, 291–307, 1999.

Rosten, E. and Drummond, T., "Machine Learning for High-Speed Corner Detection",

Leonardis, A., Bischof, H. and Pinz, A., Lecture Notes in Computer Science. **Springer Berlin Heidelberg**, s. 430–443 2006.

Rublee, E., Rabaud, V., Konolige, K. and Bradski, G., "ORB: An efficient alternative to SIFT or SURF", **IEEE International Conference on Computer Vision (ICCV)** s. 2564–2571, 2011.

Schleicher, D., Bergasa, L.M., Ocana, M., Barea, R. and Lopez, M.E., "Real-Time Hierarchical Outdoor SLAM Based on Stereovision and GPS Fusion", **IEEE Transactions on Intelligent Transportation Systems** 10, 440–452, 2009.

Schmid, C., Dorkó, G., Lazebnik, S., Mikolajczyk, K. and Ponce, J., "Pattern recognition with local invariant features", **Handbook of Pattern recognition and computer vision** 71–92, 2005.

Semsch, E., Jakob, M., Pavlicek, D. and Pechoucek, M., "Autonomous UAV Surveillance in Complex Urban Environments", **IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies WI-IAT '09** s. 82–85, 2009.

Sergi, M., Shankwitz, C. and Donath, M., "LIDAR-based vehicle tracking for a virtual mirror", **Intelligent Vehicles Symposium** s. 333 – 338, 2003.

Shi, J. and Tomasi, C., "Good features to track", **IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR '94** s. 593–600, 1994.

Smith, M., Baldwin, I., Churchill, W., Paul, R. and Newman, P., "The New College Vision and Laser Data Set", **The International Journal of Robotics Research** 28, 595–599, 2009.

Smith, R.C. and Cheeseman, P., "On the Representation and Estimation of Spatial Uncertainty", **The International Journal of Robotics Research** 5, 56–68, 1986.

Smith, R., Self, M. and Cheeseman, P., "Estimating uncertain spatial relationships in

robotics", *Autonomous robot vehicles* 1, 167–193, 1990.

Smith, R., Self, M. and Cheeseman, P., "A stochastic map for uncertain spatial relationships", *Proceedings of the 4th international symposium on Robotics Research* s. 467–474, 1988.

Smith, S.M. and Brady, J.M., "SUSAN - A New Approach to Low Level Image Processing", *International Journal of Computer Vision* 23, 45–78, 1995.

Soleimani, B., Ashtiani, M.-H.Z., Soleimani, B.H. and Moradi, H., "A Disaster Invariant Feature for localization", *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* s. 1096–1101, 2010.

Sunderhauf, N., Lange, S. and Protzel, P., "Using the Unscented Kalman Filter in Mono-SLAM with Inverse Depth Parametrization for Autonomous Airship Control", *IEEE International Workshop on Safety, Security and Rescue Robotics. SSRR 2007* s. 1–6, 2007.

Tardós, J.D., Neira, J., Newman, P.M. and Leonard, J.J., "Robust mapping and localization in indoor environments using sonar data", *International Journal of Robotics Research* 21, 311–330, 2002.

Taylor, S. and Drummond, T., "Binary Histogrammed Intensity Patches for Efficient and Robust Matching", *International Journal of Computer Vision* 94, 241–265, 2011.

Taylor, S., Rosten, E. and Drummond, T., "Robust feature matching in 2.3", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* s. 15–22, 2009.

Templeton, T., Shim, D.H., Geyer, C. and Sastry, S.S., "Autonomous Vision-based Landing and Terrain Mapping Using an MPC-controlled Unmanned Rotorcraft", *IEEE International Conference on Robotics and Automation* s. 1349–1356, 2007.

Thrun, S., Koller, D., Ghahramani, Z., Durrant-Whyte, H. and Ng, A., "Simultaneous

mapping and localization with sparse extended information filters: Theory and initial results", *Algorithmic Foundations of Robotics V* 363–380, 2004.

Tisdale, J., Kim, Z. and Hedrick, J.K., "Autonomous UAV path planning and estimation", *IEEE Robotics Automation Magazine* 16, 35–42, 2009.

Tola, E., Lepetit, V. and Fua, P., "DAISY: An Efficient Dense Descriptor Applied to Wide-Baseline Stereo", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, 815–830, 2010.

Tola, E., Lepetit, V. and Fua, P., "A fast local descriptor for dense matching", *IEEE Conference on Computer Vision and Pattern Recognition CVPR 2008* s. 1–8, 2008.

Tomic, T., Schmid, K., Lutz, P., Domel, A., Kassecker, M., Mair, E., Grix, I.L., Ruess, F., Suppa, M. and Burschka, D., "Toward a Fully Autonomous UAV: Research Platform for Indoor and Outdoor Urban Search and Rescue", *IEEE Robotics Automation Magazine* 19, 46–56, 2012.

Tuna, G., Gulez, K., Gungor, V.C. and Veli Mumcu, T., "Evaluations of different Simultaneous Localization and Mapping (SLAM) algorithms", *IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society*. s. 2693–2698, 2012.

Wang, C., Wang, T., Liang, J., Chen, Y., Zhang, Y. and Wang, C., "Monocular visual SLAM for small UAVs in GPS-denied environments", *IEEE International Conference on Robotics and Biomimetics (ROBIO)* s. 896–901, 2012.

Wang, H., Wei, S. and Chen, Y., "An Improved Rao-Blackwellized Particle Filter for SLAM", *International Symposium on Intelligent Information Technology Application Workshops IITAW '08* s. 515–518, 2008.

Wang, X. and Ma, L., "Sparse extended information filter for feature-based SLAM", *International Conference on Electronics, Communications and Control (ICECC)* s. 2290–2293, 2011.

Wang, Z., Fan, B. and Wu, F., "Local Intensity Order Pattern for feature description", ***IEEE International Conference on Computer Vision (ICCV)*** s. 603–610, 2011.

Welle, J., Schulz, D., Bachran, T. and Cremers, A.B., "Optimization techniques for laser-based 3D particle filter SLAM", ***IEEE International Conference on Robotics and Automation (ICRA)*** s. 3525–3530, 2010.

Williams, P. and Crump, M., "All-source navigation for enhancing UAV operations in GPS-denied environments", ***Proceedings of the 28th International Congress of the Aeronautical Sciences***, 2012.

Williams, S.B., Dissanayake, G. and Durrant-Whyte, H., "An efficient approach to the simultaneous localisation and mapping problem", ***IEEE International Conference on Robotics and Automation, 2002. Proceedings. ICRA '02*** s. 406–411 vol.1, 2002.

Wu, J., Li, Q., He, M. and Zhang, F., "A Terrain Model Generation Method Based on 2D Plan Laser Scanner for Micro UAV Autonomous Flight", Lu, W., Cai, G., Liu, W., Xing, W., *Proceedings of the 2012 International Conference on Information Technology and Software Engineering, Lecture Notes in Electrical Engineering. Springer Berlin Heidelberg* s. 215–225, 2013.

Xu, G., Zhang, Y., Ji, S., Cheng, Y. and Tian, Y., "Research on computer vision-based for UAV autonomous landing on a ship", ***Pattern Recognition Letters*** 30, 600–605, 2009.

Xu, X., Tian, L., Feng, J. and Zhou, J., "OSRI: A Rotationally Invariant Binary Descriptor", ***IEEE Transactions on Image Processing*** 23, 2983–2995, 2014.

Zang, Y., Yuan, K., Zou, W. and Hu, H., "A Two-Step Particle Filter for SLAM of Corridor Environment", ***IEEE International Conference on Information Acquisition*** s. 370–375, 2006.

Zeng, Y.-M., Wu, Z.-Y., Liu, H.-B. and Zhou, L., "Modified AMDF pitch detection algorithm", ***International Conference on Machine Learning and Cybernetics*** s. 470–

473 Vol.1, 2003.

Zhang, S., Tian, Q., Lu, K., Huang, Q. and Gao, W., "Edge-SIFT: Discriminative Binary Descriptor for Scalable Partial-Duplicate Mobile Search", *IEEE Transactions on Image Processing* 22, 2889–2902, 2013.

Zhou, W., Wang, C., Xiao, B. and Zhang, Z., "SLD: A Novel Robust Descriptor for Image Matching", *IEEE Signal Processing Letters* 21, 339–342, 2014.



ÖZ GEÇMİŞ

Murat Peker, 22.11.1985 tarihinde Konya’da doğdu. İlk orta ve lise eğitimini Alanya’da tamamladı. 2003 yılında girdiği Niğde Üniversitesi Elektrik-Elektronik Mühendisliği Bölümü’nden Haziran 2007’de mezun oldu. 2007 Temmuz ayında yüksek lisans öğrenimine başladı ve 2007-09 yılları arasında Y-Vizyon Sinyalizasyon firmasında yazılım mühendisi olarak çalıştı. 2009-2010 Öğretim yılında Niğde Üniversitesi Elektrik-Elektronik Mühendisliği Bölümü’ne araştırma görevlisi olarak atandı. Haziran 2010’da yüksek lisans öğrenimini tamamladı. O tarihten beri yine aynı bölümde araştırma görevlisi olarak çalışmakta ve doktora öğrenimine devam etmektedir. Bilim dalındaki ilgi alanları görüntü, video ve ses işleme, sinyal işleme ve görüntü işleme donanımlarıdır (FPGA, DSP, GPU).

TEZ ÇALIŞMASINDAN ÜRETİLEN ESERLER

Bu tez çalışmasından, 2 (iki) adet uluslararası makale üretilmiştir. Bu üretilen çalışmalar aşağıda sunulmuştur.

Peker, M. and Karakaya, F., “SIH: segmented intensity histogram for orientation estimation in image matching”, *Signal, Image and Video Processing*, 1863-1711, 1-8, DOI: 10.1007/s11760-016-0869-6, 2016

Peker, M., Altun, H and Karakaya, F., “Hardware Implementation of Scale and Rotation Invariant Object Detection Algorithm on FPGA For Robotic Vision Applications”, *Turkish Journal of Electrical Engineering & Computer Sciences*, DOI: 10.3906/elk-1408-187, 2015