



T.C.
ONDOKUZ MAYIS ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



HARİTA MÜHENDİSLİĞİ ANABİLİM DALI

**RENK VE DERİNLİK VERİLERİ KULLANILARAK İÇ MEKÂNLARIN 3
BOYUTLU MODELLERİNİN OLUŞTURULMASI**

YÜKSEK LİSANS TEZİ

Nihat KARAAHMETOĞLU
(12210234)

Tezin Savuma Tarihi : 25 Ağustos 2015

Tez Danışmanı : Yrd. Doç. Dr. Sedat DOĞAN

Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü

Harita Mühendisliği Anabilim Dalında

Nihat KARAAHMETOĞLU Tarafından Hazırlanan

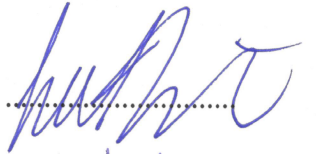
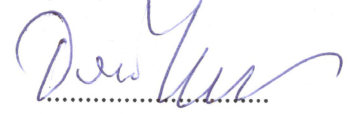
**RENK VE DERİNLİK VERİLERİ KULLANILARAK İÇ
MEKANLARIN 3-BOYUTLU MODELLERİNİN OLUŞTURULMASI**

**başlıklı bu çalışma jürimiz tarafından 25/08/2015 tarihinde yapılan sınav ile
YÜKSEK LİSANS tezi olarak kabul edilmiştir.**

Başkan : Yrd. Doç. Dr. Sedat DOĞAN
Ondokuz Mayıs Üniversitesi

Jüri Üyeleri : Yrd. Doç. Dr. Derya ÖZTÜRK
Ondokuz Mayıs Üniversitesi

Yrd. Doç. Dr. Mahir Serhan TEMİZ
Uşak Üniversitesi


.../.../2015

Prof. Dr. Hüseyin DEMİR

Enstitü Müdürü

ÖNSÖZ

Lisans ve yüksek lisans eğitimim süresince benden ilgi, alaka ve desteğini esirgemeyen danışman hocam sayın Yrd. Doç. Dr. Sedat DOĞAN'a, hayatımın her safhasında yanımda olan, her konuda yardımlarını ve desteklerini hissettiren değerli aileme teşekkürü bir borç bilirim.

Ağustos 2015

Nihat KARAAHMETOĞLU
Araştırma Görevlisi

İÇİNDEKİLER

Sayfa

ÖNSÖZ	v
İÇİNDEKİLER	vii
ÇİZELGELER LİSTESİ.....	ix
ŞEKİLLER LİSTESİ.....	xi
KISALTMALAR	xiii
RENK VE DERİNLİK VERİLERİ KULLANILARAK İÇ MEKÂNLARIN 3 BOYUTLU MODELLERİNİN OLUŞTURULMASI	xv
3D RECONSTRUCTION OF INDOOR ENVIRONMENT USING RGB-D DATA..	xvii
1. GİRİŞ	1
1.1 Literatür Özeti.....	2
2. GENEL BİLGİLER.....	5
2.1 3 Boyutlu Görüntüleme Tekniklerine Genel Bir Bakış.....	5
2.1.1 Lazer üçgenleyiciler.....	7
2.1.2 Yapılandırılmış ışık sensörleri	8
2.1.3 Stereo görüş	9
2.1.4 Fotogrametri.....	9
2.1.5 Uçuş süresi (ToF) tekniği.....	10
2.1.6 İnterferometri	10
2.1.7 Moiré saçakları yöntemi.....	10
2.1.8 Odaklanma yöntemi	11
2.1.9 Gölge bilgisinden yararlanma	11
2.1.10 Doku bilgisinden yararlanma	11
2.1.11 Gölgeleme yöntemi.....	11
3. DERİNLİK SENSÖRLERİ.....	13
3.1 Uçuş Süresi (ToF) Sensörler	13
3.1.1 CW matrisel uçuş süresi (ToF) menzil kameraları	14
3.2 Kinect Kamera	17
3.3 Kamera Kalibrasyonu	19
4. ÖZNETLİK ÇIKARIMI	25

4.1 Harris Köşe Bulucu	26
4.2 Ölçekten Bağımsız Öznitelik Dönüşümü (Scale Invariant Feature Transform-SIFT) .	27
4.3 Hızlandırılmış Robust Öznitelikler (Speeded Up Robust-SURF).....	29
4.4 Normal Hizalanmış Radyal Öznitelikler (Normal Aligned Radial Features-NARF)...	32
4.5 İkili Robust Bağımsız Temel Öznitelik (Binary Robust Independent Elementary Features-BRIEF)	34
4.6 Yönlendirilmiş FAST ve döndürülmüş BRIEF (Oriented FAST and Rotated BRIEF -ORB).....	37
4.7 Nokta Bulutu Kütüphanesinin Yerel Tanımlayıcıları.....	38
4.7.1 Nokta öznitelik histogramı (Point feature histogram-PFH)	39
4.7.2 Yarıçap tabanlı yüzey tanımlayıcı (Radius-based surface descriptor-RSD)	40
4.7.3 3B şekil bağlamı (3D shape context-3DSC)	42
4.7.4 Yönelimler histogramının imzası (Signature of histograms of orientations-SHOT)	43
4.7.5 Dönük imge (Spin image-SI).....	43
4.7.6 Dönüklükten bağımsız öznitelik dönüşümü (Rotation-invariant feature transform - RIFT).....	44
4.7.7 Dönel izdüşüm istatistikleri (Rotational projection statistics-RoPS)	45
5. DÖNÜŞÜM PARAMETRELERİNİN BELİRLENMESİ.....	47
5.1 Rasgele Örneklerin Uzlaşımı (Random Sample Consensus-RANSAC).....	47
5.2 İteratif En Yakın Nokta (Iterative Closest Point-ICP)	48
5.3 Döngü Bulma (Loop Closure) Problemi ve Genel Optimizasyon.....	49
6. UYGULAMA.....	53
6.1 Kinect Kurulumu.....	53
6.2 Kamera Kalibrasyonu.....	54
6.3 Öznitelik Çıkarımı.....	54
6.4 Dönüşüm Parametrelerinin Belirlenmesi	55
6.5 Nokta Bulutunun İşlenmesi	58
7. SONUÇLAR VE ÖNERİLER.....	61
8. KAYNAKLAR.....	63
ÖZGEÇMİŞ	67

ÇİZELGELER LİSTESİ

	<u>Sayfa</u>
Çizelge 2.1. 3B görüntüleme tekniklerinin sınıflandırılması.....	6
Çizelge 4.1. Genel olarak kullanılan öznitelikler.....	26

ŞEKİLLER LİSTESİ

Sayfa

Şekil 2.1. Üçgenleme prensibinin yapısı.....	7
Şekil 2.2. Lazer tarama sisteminin çalışma prensibi.....	8
Şekil 2.3. Yapılandırılmış ışık sistemi	8
Şekil 2.4. Michelson interferometri yapısı.....	10
Şekil 3.1. Uçuş süresi sensörünün çalışma prensibi	13
Şekil 3.2. MESA Imaging SR4000TM ticari ToF kamera	14
Şekil 3.3. Yayılan sinyal SE(t) (mavi) ve alınan sinyal SR(t) (kırmızı).....	15
Şekil 3.4. Saçılma etkisi.....	17
Şekil 3.5. Çoklu yol etkisi: Yayılan ışın (kırmızı) A yüzey noktasında çoklu olarak yansır. Turuncu ışın A noktasından sonra B noktasından sensöre geri yansır.....	17
Şekil 3.6. Kinect for Windows v2.....	18
Şekil 3.7. Kızılötesi sensörün yerleşimi (iFixit.com)	19
Şekil 3.8. Kalibrasyon nesnesi: Satranç tahtası üzerinde köşelerin belirlenmesi	22
Şekil 3.9. Satranç tahtasına göre sensör düzleminin konumu.....	22
Şekil 4.1. Ölçek uzayında Gauss farkları.....	27
Şekil 4.2. SIFT tanımlayıcısı: Anahtar noktaların etrafındaki gradyen yönleri hesaplanır	28
Şekil 4.3. Y ve XY yönündeki ikinci derece Gauss türevleri (sol) ve kutu filtre yakınsakları (sağ)	30
Şekil 4.4. İntegral görüntü kullanılarak dörtgensel bir bölgenin toplamının hesaplanması	31
Şekil 4.5. Yönlendirmeyi belirlemek için kullanılan düşey ve yatay kutu filtreleri ..	32
Şekil 4.6. SURF de kullanılan kayan yönlendirme pencereleri	32
Şekil 4.7. Farklı türde sınır noktaları	33
Şekil 4.8. NARF ile sınır noktalarının belirlenmesi	34
Şekil 4.9. BRIEF özneliği.....	34
Şekil 4.10. Beş örnekleme düzeni.....	36
Şekil 4.11. Beş farklı örnekleme stratejisinin performansları.....	36

Şekil 4.12. İlgi noktası olmaya aday "p" pikseli ve dairesel olarak etrafından seçilen 16 piksel	37
Şekil 4.13. Bir nokta için PFH hesaplanırken saptanan nokta çiftleri	39
Şekil 4.14. Sabit Koordinat Çatısı ve nokta çiftlerden biri için hesaplanan açısal özellikler.....	39
Şekil 4.15. Bir nokta için FPFH hesaplanırken saptanan nokta çiftleri	40
Şekil 4.16. İki nokta ve normalleri ile uyuşan Küre	41
Şekil 4.17. Bulut için RSD değerleri; düz yüzey üzerindeki noktalar maksimum değere sahip olur.....	41
Şekil 4.18. Bir nokta için 3DSC yi hesaplamak için gerekli destek yapısı.....	42
Şekil 4.19. SHOT hesaplamak için destek yapısı. Basit olması için sadece azimuta göre ayrılmış 4 bölüm gösterilmektedir.	43
Şekil 4.20. Modelin 3 noktası için hesaplanan dönük imgeler	44
Şekil 4.21. Tanımlayıcıdaki 3 değişik konumdaki RIFT öznitelikleri.....	45
Şekil 5.1. Graf optimizasyonunun genel iş akışı.....	50
Şekil 5.2. Döngü bulma: (a) Kamera pozisyonları, her biri dönüşümleri temsil eden standart kenarlar ile bağlanmış (b) Daha önce gözlemlenen bir durumla benzerlik kurulması durumunda yeni bir bağlantı oluşur.	51
Şekil 6.1. Satranç tahtasının farklı açılardan resimlerinin çekilmesi	54
Şekil 6.2. Çerçevenin çerçeve ile ve çerçevenin modelle işlenmesi: Çerçevenin bir anahtar çerçeve ile birlikte işlenmesi kayma etkisini azaltmaktadır.	55
Şekil 6.3. Graf optimizasyonu sonucunda kamera pozisyonları daha iyi belirlenir ve kameranın pozisyonlarına bağlı olarak dönüşümler tekrar ayarlanır.	58

KISALTMALAR

2B	: 2 Boyutlu
3B	: 3 Boyutlu
BRIEF	: Binary Robust Independent Elementary Features
FAST	: Features from Accelerated Segment Test
ICP	: Iterative Closest Point
NARF	: Normal Aligned Radial Features
ORB	: Oriented Fast and Rotated BRIEF
PCL	: Point Cloud Library
RGB-D	: RGB – Derinlik verisi
SDK	: Software Development Kit
SIFT	: Scale Invariant Feature Transform
SLAM	: Simultaneous Localization and Mapping
SURF	: Speeded Up Robust Features
ToF	: Time of Flight
WCS	: World Coordinate System

RENK VE DERİNLİK VERİLERİ KULLANILARAK İÇ MEKÂNLARIN 3 BOYUTLU MODELLERİNİN OLUŞTURULMASI

ÖZET

Bu tez çalışmasında renk ve derinlik verileri kullanılarak binaların iç ortamlarının 3-boyutlu modellerinin oluşturulmasını sağlayan bir sistem geliştirmek amaçlanmıştır. Bu amaçla bu tez kapsamında bir RGB-D kamera ile bir binanın iç ortamının ya da bir odanın 3 boyutlu modellenmesi için tasarlanan bir sistem tasarlanmıştır. Oluşturulan bu sistemde açık kaynak kodlu yazılımlar kullanılmıştır.

Bu çalışma Giriş, Genel Bilgiler, Derinlik Sensörleri, Öznitelik Çıkarımı, Dönüşüm Parametrelerinin Belirlenmesi, Uygulama, Sonuç ve Öneriler olmak üzere yedi bölümden oluşmaktadır. İlk üç bölümde temel kavramlar ve teknolojiler hakkında bilgi verilmiştir. Öznitelik Çıkarımı bölümünde 2 boyutlu ve 3 boyutlu öznitelik çıkarım yöntemlerinden bahsedilmiştir. Dönüşüm Parametrelerinin Belirlenmesi bölümünde renk ve derinlik verilerinden elde edilen görüntülerin ve nokta bulutlarının nasıl işleneceği üzerinde durulmuştur. Sistemin uygulama adımları Uygulama bölümünde anlatılmıştır.

Sonuç ve öneriler bölümünde renk ve derinlik verilerinden elde edilen modelin yapısı, artıları ve eksileri anlatılmıştır. Bu sistemle yapılabilecek olan çalışmalar ve sistemin gelişmeye açık yönleri tartışılmıştır.

Anahtar Kelimeler : 3 boyutlu yeniden yapılandırma, Derinlik verisi, Öznitelik çıkarımı, Nokta Bulutu, Kinect

3D RECONSTRUCTION OF INDOOR ENVIRONMENT USING RGB-D DATA

ABSTRACT

In this thesis, it is intended to develop a system that allows the creation of 3D models of interior environments of the buildings by using color and depth data. For this purpose, a system is designed to build a 3D model of a part of building's indoor or a room with an RGB-D camera. Open source software and codes are used in the generated system.

This MSc thesis consists of seven basic chapters, including Introduction, General Information, Depth Sensors, Feature Extraction, Determination of Transformation Parameters, Application, Results, and recommendations. In the first three chapters contains the information about fundamental concepts and technologies that are used in the thesis. The 2D and 3D feature extraction methods are explained in the Feature Extraction section. Following section “Determination of Transformation Parameters” focuses on how to handle and process the data and point clouds that obtained from the color and depth data of the images. Process steps of the system. Are described in the application part.

In the last section, the structure of the model obtained from the color and depth data, its advantages and its disadvantages were described. Finally, future works and how to improve this system are discussed.

Key Words: 3D Reconstruction, RGB-D data, Feature extraction, Point cloud, Kinect

1. GİRİŞ

Gerçek dünyadaki nesnelerin 3 boyutlu (3B) modellerinin oluşturulması ve yeniden yapılandırılması bilgisayarla görme ve bilgisayar grafik uygulamalarının temel bir problemidir. Film özel efektleri, bilgisayar oyunları, sanal gerçeklik gibi birçok alanda önemli uygulamaları vardır. İki veya üç boyutlu bir nesnenin bilgisayar ekranında güzel bir görüntüsünü elde etmek için çözümler bulunması oldukça zahmetli işlemler gerektirir. Bir ressam bu sorunu sezgileri ve deneyimi ile, usta bir fotoğrafçı en iyi açıları, uzaklığı ve perspektifi seçerek çözer. Ancak bilgisayar ekranında bu insani yeteneklerin yerini matematiksel metotlar ve yazılımlar almaktadır.

İnsan gözü özellikleri gereği bakış açısı içerisinde bulunan cisimleri iki boyuttan üç boyuta dönüştürmek zorunda değildir. Bu özellik görsel dünyayı anlamak için insanlık adına olağanüstü bir yetenektir. Bilgisayar görmesi, sanal gerçeklik, nesne tanıma, artırılmış gerçeklik, 3B yeniden yapılandırma gibi bilgisayarlı grafik uygulamaları mantıksal bir yaklaşım ile görsel dünyayı anlamak için insanların bu tür yeteneklerini taklit etmek istemektedir. 3B modelleri elde etmek için çoklu görüş stereo teknikleri (Seitz ve diğ., 2006; Zhang ve diğ., 2009; Bradley ve diğ., 2010) kullanılıyor olsa da 3B modellerin kesinliği ve tamlığı hala sorun teşkil etmektedir. Bu nedenle bilgisayar ortamında gerçek dünya nesnelerinin 3B yeniden yapılandırılması hala çok hantaldır.

Gerçek dünya nesnelerinin 3B modellerinin oluşturulması eski teknolojilerle yüksek maliyetli ve fazla zaman harcanmasını gerektiren güç bir süreç olmakla birlikte, son yıllarda derinlik sensörlerinin gelişimi ile birlikte 3B modelleme ve yapılandırma alanındaki düşük maliyetli derinlik kameraları (Kinect vb.) kullanılarak yapılan araştırmalar hızla artış göstermiştir. Cui ve diğ. (2010), “time of flight” sensör ile elde edilen bir nesnenin sıralı derinlik imgelerinden 3B nesne taraması için bir metot geliştirmişlerdir. Bu çalışmalara ek olarak RGB yöntemi (Vouzounaras ve diğ., 2010; Zheng ve diğ., 2011) ve stereografik yöntemler üç boyutlu nesne tarama tekniği olarak kullanılmıştır (Huang ve Chen, 2011). Farklı nesne tarama teknikleri 3B model oluşturmak üzere kullanılabilir olsa da gerekli cihazların temini masraflı ve

kullanımları için üst seviyede bilgi gerekmektedir. Bunlara ek olarak taranacak nesne hareketli ise tarama esnasında dokuların eksik taranması ve nesnelerin çeşitli şekillerde kendi veya başka nesneler tarafından kapatılmaları (occlusion) gibi problemlerle karşılaşılabilir (Kim ve diğ., 2009). Son yıllarda tarama ve derinlik hesaplama özelliğine sahip yeni bir derinlik kamerası olan Microsoft Kinect Kamera kullanılmaya başlanmıştır. Görüntü ve derinlik yakalamadaki zorlu ışık ve doku durumlarındaki yeteneği, normal bir video kamera olarak kullanılabilmesi, maliyeti ve kolaylıkla elde edilebilir olması Kinect Kamerayı diğer 3B tarayıcılardan daha avantajlı hale getirmiştir. Ayrıca Kinect Kamera kullanılarak yapılan çalışmalar büyük bir artış göstermiştir (Newcombe ve diğ., 2011).

Yakın bir zamana kadar dinamik bir sahnenin geometrik tanımının elde edilmesi, sadece araştırma laboratuvarlarının ve şirketlerin gücünün yettiği en son teknoloji ve aletlere gereksinim duyulan zorlayıcı bir konu olmuştur. Yakın zamanda uçuş süresi (time of flight-ToF) kameralarının ve çok miktarda satılan Microsoft Kinect Kamerasının tanıtılması sonucu derinlik verisi kolaylıkla erişebilir hale getirmiştir. Bilgisayar grafikleri, bilgisayarla görme ve görüntü işleme toplulukları arasında bu tür cihazlara ilgi artmış ve “uçuş süresi (ToF) kamerası nedir?”, “Bu kameralar nasıl çalışır?”, “Uçuş süresi kameranin düşük çözünürlük ve yüksek gürültü karakteristiğini geliştirmenin yolları var mıdır?”, “Nispeten daha ucuza sağlanan bir Kinect kamerasından elde edilen derinlik verisi ile, binlerce liraya mal olan uçuş süresi kameralara oranla ne kadar ileri gidebilirim?” gibi sorular yaygınlaşmıştır.

Bilinmeyen bir ortamda bir robotun otonom hareketi, öncelikle robotun kendi konumunu belirlemesi, daha sonra çevresini modellemesi ve hangi yolu izlediğini modellemesi ile mümkündür. Eşzamanlı konum belirleme ve haritalama (SLAM) (Quang ve Quoc, 2015) metotları, bu probleme çözüm getirebilmek için geliştirilmiştir. Bu problem robotik ve bilgisayarla görme alanlarında çok iyi çalışılmış bir konudur. Derinlik kameralarının maliyetlerinin düşmesi ile birlikte derinlik verileri kullanılarak yapılan çalışmalar da artış göstermiştir.

1.1 Literatür Özeti

Renkli görüntü ve derinlik verisi üreten sensörler (RGB-D), son zamanlarda oldukça popüler hale gelmiştir. Renkli kamera sensörünün derinlik sensörü ile birleştirilmesi ile beraber birbirlerinin açıklarının üstesinden gelebilen bir veri türü ortaya çıkmıştır.

RGB-D Sensörlerin temsilcilerinden olan Kinect piyasaya çıkar çıkmaz bu alanda çalışan araştırmacılar tarafından kullanılmaya başlanmıştır.

Kinect ile yapılan çalışmalar iki aşamadan oluşur;

- i. Veri toplama ve öznitelik noktalarının benzerliklerinin bulunması,
- ii. Lup kapanmasının belirlenmesi ve global eniyileme.

Çeşitli sistemler, sahne hizalanması metotlarını, seyrek öznitelik (sparse feature) olarak da adlandırılan peş peşe görüntü karelerinden ayırık noktaların çıkarılması ve bunlar arasındaki geometrik ilişkilerin bulunması yöntemine dayalı şekilde oluşturmuşlardır. Ana fikir, doğru nokta benzerliklerinden bir dönüşümün tanımının hesaplanmasıdır.

Henry ve diğ. (2010) yapmış oldukları çalışmada seyrek öznitelik eşleştirme yaklaşımını kullanmışlardır. Çalışmalarının yenilikçi tarafı görsel özniteliklerle beraber derinlik haritasından güvenilir bir şekilde algılanan şekil bilgilerini dikkate alan bir birleşim (joint) optimizasyonu algoritması kullanmalarıdır.

Daha sonraları birçok uygulama onların yöntemine yakın metotlar kullanmışlardır. Bu çalışmalarda, ölçekten bağımsız öznitelik dönüşümü (SIFT) temelli öznitelik çıkarımı ve tanımlayıcının yanı sıra (Engelhard ve diğ., 2011), hesap yükü açısından daha verimli olan FAST (Rosten ve Drummond, 2006) öznitelik tanımlayıcısı ve SURF (Bay ve diğ., 2006) tanımlayıcısı da kullanılmıştır. Stücker ve Behnke (2012) , şekil ve doku bilgisinin daha özgün bir şekilde kullanımını göstermişlerdir. Zhang ve diğ. (2012), Henry ve diğ. (2010) da kullanmış olduğu öznitelik eşleştirme yöntemine benzer bir yöntem kullanmışlar ve halı ve perde gibi aynı dokunun tekrar ettiği durumlarda kullanılan yöntemin başarılı olamadığını belirlemişlerdir. Bu sorunun üstesinden gelebilmek için eşleştirme yöntemlerini maksimum benzerlik yöntemi ile modifiye etmişlerdir. Öncül bir bilgi olarak da kameranın hareketinin biraz Gauss gürültüsü altında sabit olduğunu varsaymışlardır.

İki görüntü karesi arasında yoğun nokta takibine dayalı sistemler, hizalanma için görüntüdeki tüm piksellerin bilgisi kullanıldığı için hesap yükü açısından daha maliyetlidir. KinectFusion (Newcombe ve diğ., 2011; Izadi ve diğ., 2011) bu yöntemi gerçekleştirebilmek için iyi bir grafik kartının yardımıyla tüm işlemleri paralel programlamaktadır. Böylece gerçek zamanlı olarak veriyi işleyebilmektedir. ICP algoritması ile üretilen kamera pozisyonu ile birlikte algoritmaları yeni gelen derinlik bilgisini kesik işaretli mesafe fonksiyonu (truncated signed distance function-TSDF) ile 3B yapıya kaynaştırmışlardır.

Meister ve diğ. (2012) , son teknoloji yer kontrol noktaları kullanarak, KinectFusion 3B yeniden yapılandırma yönteminin hangi durumlarda doğru olduğunu değerlendirmeye çalışmışlardır. Sonuçlar sistemin en az 10 mm boyutlarında bir nesneyi çözümleyebileceğini göstermiştir. Bununla birlikte keskin kenarlar ve son derece konkav sahneler hala sistem için problem çıkarmaktadır. KinectFusion sistemi ile, farklı ışık koşullarında ve çevre şartlarında yapılan uygulamalar da bulunmaktadır (Meilland ve diğ., 2011; Meilland ve diğ., 2012; Roth ve Vona, 2012; Whelan ve diğ. 2012).

Quang ve Quoc (2015) yapmış oldukları çalışmalarında renk ve derinlik verilerini kullanarak RGB-D SLAM sistemlerini geliştirmişlerdir. Onlar, sistemlerine Kinect sensöründen gelen verilerin gürültülerini bastırabilmek için bir ön işleme adımı eklemişlerdir. Bunun yanı sıra hareket belirleme aşamasında doğruluğu artırmak için çerçeve-çerçeve eşleştirme yerine anahtar çerçeve yaklaşımını benimsemişlerdir. Buna göre her bir çerçeve, bir referans çerçeve ile eşleştirilmiştir. Eğer robot belli oranda hareket ederse veya çerçeveler arası eşleşme belli bir eşik değerin arasına düşerse referans çerçeve güncellenmektedir. Öznitelik olarak gerçek zamanlı uygulamalara uygun hızlı hesaplama zamanına sahip SURF (Bay ve diğ., 2006) ve ORB özniteliklerini tercih etmişlerdir. Eşleşmelerde güncel çerçeve üzerindeki her bir öznitelik için referans çerçeve üzerinde en yakın iki komşu araştırılır. Etkinliği artırabilmek için Best-Bin-Search denilen tekniğe göre göre en yakın komşunun, en yakın ikinci komşuya oranı belli bir eşik değerin altına düştüğü durumlarda eşleşme reddedilir. Onlar uygulamalarının her adımlarında OpenCV tarafından sağlanan metotları kullanmışlardır. Doğrusal olmayan en küçük kareler probleminin çözümü ile dönüşüm belirlenmektedir. Ancak, öznitelik eşlemenin belirsiz doğasından dolayı tek bir yanlış eşleşme bile önemli bir hataya neden olabileceğinden dolayı uyumsuz ölçüleri (outlier) elemek için RANSAC metodunu kullanmışlardır. En son olarak robotun izlediği yolu belirleyebilmek için graf optimizasyonu kullanmışlardır.

2. GENEL BİLGİLER

3B görüntüleme sensörleri, 3B şekil bilgisini yakalamak için üretilmiş olup, son yıllarda birçok uygulamada ilgi odağı olmuşlardır. Üretimlerinde kullanılan teknolojinin gelişimine bağlı olarak minyatürleştirilmeleri ve optik ve elektronik çeşitli elemanların eklenmesi ile daha etkin ve esnek sensörler ortaya çıkmıştır. Bugün birçok çeşit 3B sensörün bulunmasının yanı sıra diğer sensörlerle kombine edilerek üretilen sensörlerle beraber “sensör füzyonu” yaklaşımları ortaya çıkmıştır. Derinlik sensörlerinin ortaya çıkmasından beri jeoloji, mühendislik, arkeoloji, tersine mühendislik, tıp ve sanal gerçeklik gibi birçok alanda bu sensörlerle uygulamalar yapılmıştır.

2.1 3 Boyutlu Görüntüleme Tekniklerine Genel Bir Bakış

3B görüntüleme sensörleri genellikle iki şekilde çalışır. Birincisi elektromanyetik enerjiyi nesnenin üzerine izdüşürür (aktif) ve nesneden yansıyan enerjiyi kaydeder. İkincisi, nesneden yayılan elektromanyetik enerjiyi toplayarak (pasif) kaydeder. En önemli iletim (transmission) ölçme sistemi, yüksek enerjili x-ray ışınları kullanan ve nesneye iletilen radyasyonu ölçen endüstriyel bilgisayarlı tomografi cihazıdır (Sansoni ve diğ., 2009).

Yansıyan enerjiyi toplayan sensörler optik ve optik olmayan sensörler olarak ikiye ayrılabilir. Optik olmayan sensörlerin içinde akustik (ultrasonik, sismik) sensörler, elektromanyetik (kızılötesi, morötesi, mikrodalga radar vb.) ve diğer sensörler yer alır. Bu tip sensörler, genellikle bir ses veya elektromanyetik enerji atımı ile bu ses veya elektromanyetik enerjinin nesneden geri yansıması arasında geçen zamanı ölçerler (Sansoni ve diğ., 2009). Optik sensörlerde, mesafe bilgisini ışık taşır.

Çok çeşitli 3B teknikler bulunmaktadır. Bu teknikler, çizelge 2.1’de gösterilmiştir.

Çizelge 2.1. 3B görüntüleme tekniklerinin sınıflandırılması

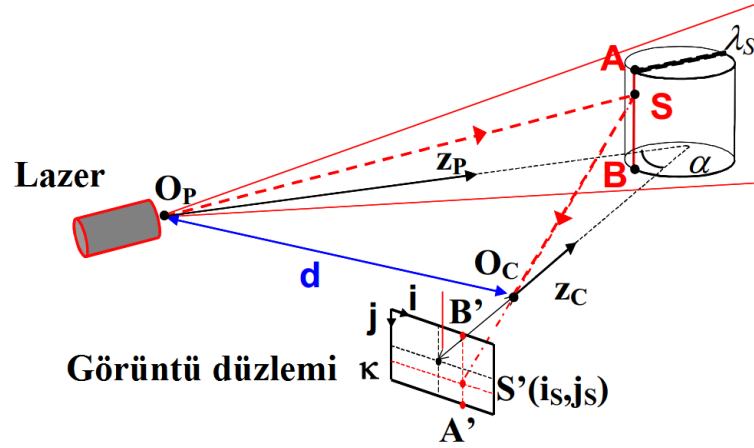
	Üçgenleme	Zaman gecikmesi	Monoküler görüntüler	Pasif	Aktif	Doğrudan	Dolaylı	Uzaklık ölçümü (Range)	Yüzey yönlendirmesi
Lazer üçgenleyiciler	X				X	X		X	
Yapılandırılmış ışık	X				X	X		X	
Stereo görüntüleme	X			X		X		X	
Fotogrametri	X			X		X		X	
Uçuş süresi (ToF)		X			X	X		X	
İnterferometri		X			X	X		X	
Moiré saçak desenli menzil ölçümü			X		X		X	X	
Odaklanma yöntemi			X	X	X		X	X	
Gölge bilgisinden yararlanma			X		X		X	X	
Doku gradyanları			X	X			X		X
Gölgelendirme yöntemi			X		X		X		X
Fotometri ile şekil			X		X		X		X

3B teknikler, optik üçgenleme, zaman gecikmesi ve monoküler görüntülerin kullanıldığı tekniklerden birini temel almaktadır. Bu teknikleri pasif ve aktif metotlar olmak üzere ikiye ayırabiliriz. Pasif metotlarda nesnenin ışığı yansıtma karakteristiği ve sahnenin aydınlanma durumundan yararlanılarak şekil bilgisi elde edilmektedir, aktif bir cihaza ihtiyaç duyulmamaktadır. Aktif cihazların kullanıldığı tekniklerde uygun bir ışık kaynağı, dahili bir bilgi vektörü olarak kullanılır. 3B teknikleri aralarında doğrudan veya dolaylı ölçüm yöntemleri olarak da ayırmak mümkündür. Doğrudan teknikler, bir uzaklık sensörü ile uzaklık verisi (örneğin bilinmeyen bir yüzeyle aralarındaki uzaklık ölçümlerinin kümesi) üretirler. Dolaylı teknikler, hedef

nesnenin özellikleri hakkında bilinen öncül bilgiler ve monoküler görüntüler ile uzaklık verisi ya da yüzey yönelimini ölçen tekniklerdir (Sansoni ve diğ., 2009).

2.1.1 Lazer üçgenleyiciler

Lazer üçgenleyiciler, aktif üçgenleme prensibine dayanırlar. Şekil 2.1 ‘de böyle bir sistemin yapısı görülmektedir.



Şekil 2.1. Üçgenleme prensibinin yapısı

Şekil 2.1’de O_P ve O_C noktaları, lazer kaynağının çıkış ve kameranın giriş noktalarıdır. Bu noktaların arasında bir d baz mesafesi bulunmaktadır. Lazerin Z_P optik eksen ve kameranın Z_C optik eksen α açısı ile kesişmektedir.

Lazer kaynağı tarafından üretilen ışın, nesneye S noktasında çarpar. Geriye saçılan ışın, κ görüntü düzleminde S' noktası olarak görüntülenir. S' noktasının konumunun ölçümü, $\overline{S'O_C}$ gözlem hattını tanımlar ve geometrik olarak S noktasının konumu belirlenebilir (Tek nokta üçgenlemesi). Yüzeyin ölçümü, yüzeyi oluşturan noktaların ölçülmesi yani taranması ile gerçekleştirilir. Klasik üçgenleme mekanizmasında, görüş açısı ile ölçüm çözünürlüğü ve belirsizlik ve gölge efektleri (α açısının aldığı değerin büyümesine bağlı olarak) arasında bir denge bulunmaktadır. Bir taraftan kazanmak için diğer taraftan ödün verilmesi gerekmektedir (Sansoni ve diğ., 2009).

Bu sınırlamanın üstesinden gelebilmek için “senkronize tarama” (Şekil 2.2) teknolojisi denilen sistem geliştirilerek çok daha geniş görüş açısı, küçük α açısı değerlerinde ve uzaklık ölçme doğruluğundan feragat etmeden ölçme işlemi gerçekleştirilebilmiştir (Sansoni ve diğ., 2009).

Şekil 2.2’de görüldüğü üzere iki yüzü ayna konsepti ile birlikte lazer ışının gönderilmesi ve algılanması senkronize edilmiş olmaktadır. Tarama aynası, lazer

The diagram illustrates the optical setup of a laser scanning system. At the bottom, a wavy line represents the **Taranan yüzey** (Scanned surface). A dashed line, labeled **Tarama yapısı** (Scanning structure), shows the path of the laser beam. It starts from a **Lazer kaynağı** (Laser source) at the bottom center, goes up to a **Tarama aynası** (Scanning mirror) which is rotating (indicated by a curved arrow), then reflects to the left to a **Yayılan Lazer Işını** (Diverging laser beam) and to the right to a **Yansıyan Lazer Işını** (Reflecting laser beam). Both beams hit the scanned surface. The reflected beams then travel back up to the scanning mirror and are directed towards a **Lens** and finally to a **CCD dizisi** (CCD array) at the top. A point **P** is marked on the optical axis between the lens and the CCD array.

Tarama aynası döndürüldükçe lazer ışını, taranacak olan yüzey ile kesişen tarama düzlemini süpürür (Feng ve diğ., 2001).

Yapılandırılmış ışık tarayıcıları, bir nesnenin üç boyutlu yapısını, bu nesnenin üzerine belirli bir desene sahip ışıkları iz düşürerek belirlemeye çalışır (Şekil 2.3). Projektörden farklı bir açıdan bakıldığı zaman bir nesnenin üzerine iz düşürülmüş olan bir yapılandırılmış ışık şeridinin deseninde bozulmalar olduğu gözlemlenir. Işık



desenindeki bozulmalardan yararlanılarak şeklin geometrik bilgisi elde edilir (URL-1).

2.1.3 Stereo görüş

Stereo görüş tekniğinde aynı sahne iki (ya da daha fazla) kamera ile eş zamanlı olarak izlenmektedir. Temelde stereo görüş tekniğinin belli başlı işlem adımları: (i) görüntü yakalama, (ii) kamera modellemesi, (iii) öznitelik çıkarımı, (iv) benzerlik analizi, ve (v) üçgenleme aşamalarından oluşur. Başka bir cihaza (örneğin belli bir ışık kaynağına) ya da özel bir iz düşün modeline ihtiyaç yoktur. En önemli avantajı basit ve düşük maliyetli olmasıdır. Stereo yaklaşımın en önemli problemi ise görüntü çiftlerindeki ortak noktaların belirlenmesidir. Bu ancak iyi bilinen bir benzerlik probleminin çözümü ile mümkündür. Şekillerin geometrilerinin belirlenmesi, yüzey dokusunun (yüzey yansıtıcılığının değişiminden etkilenmektedir.) keskinliğine bağlıdır. Stereo görüş yaklaşımının, robotik ve bilgisayarla görme alanlarında, özellikle asıl problemin yüksek kaliteli verinin doğru bir şekilde toplanmasından çok bu verilerin yorumlanmasının daha önemli olduğu durumlarda (hareket planlama, çarpışma engelleme) önemli uygulamaları bulunmaktadır (Sansoni ve diğ., 2009).

Stereo görüş yaklaşımının farklı bir türevinde , bir kameradan elde edilen nesnenin farklı açıdan çekilmiş çok sayıda görüntüsü kullanılmaktadır. Bu tekniğin gerçekleştirilebilmesi için sahnede hareketli bir kısmın bulunmaması gerekir (Sansoni ve diğ., 2009).

Stereo görüş tekniğine dayanan nesne silüetlerinden nesnenin belirlenmesi yaklaşımı çok kullanılan metotlardan biridir. 3B geometri, stereo görüş kullanılarak farklı açılardan elde edilmiş nesne dış hatlarından yararlanılarak bulunabilmektedir (Sansoni ve diğ., 2009).

2.1.4 Fotogrametri

Fotogrametri sayesinde fotoğraflar aracılığı ile güvenilir 3B modeller elde edilebilmektedir. Dijital fotogrametride sayısal görüntüler kullanılır. Kamera kalibrasyonu ve yönlendirmesi, görüntü noktalarının ölçümü, 3B yapının oluşturulması, yüzeyin üretilmesi ve doku ile kaplanması işlemleri önemlidir. Doğru modellerin üretilmesi için kamera kalibrasyonu kritik bir öneme sahiptir. Görüntü ölçümleri otomatik veya yarı otomatik olarak gerçekleştirilebilir. Hava fotogrametrisi ya da yakın resim fotogrametrisi sayesinde şehir modelleme, medikal görüntüleme ,

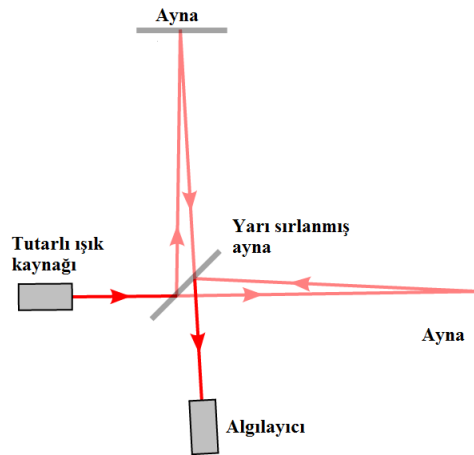
doğal ve kültürel mirasın korunmasına ve envanterinin çıkarılmasına yönelik uygulamalar gerçekleştirilebilmektedir (Sansoni ve diğ., 2009).

2.1.5 Uçuş süresi (ToF) tekniği

Bu tekniğe ait bilgiler bir sonraki bölümde anlatılmıştır.

2.1.6 İnterferometri

İnterferometri yaklaşımları, konumsal veya zamansal olarak periyodik değişen desenlerin yüzey üzerine iz düşürülmesi ve bunu takiben yansıyan ışığın bir referans desen ile karıştırılması ile işlemektedir (Şekil 2.4). Referans desenle sinyal çözümlenerek yüzeydeki değişimler ortaya çıkarılır. Ölçüm çözünürlüğü çok yüksektir (Sansoni ve diğ., 2009). Bu nedenle astronomi, fiber optik, metroloji, optik metroloji, oşinografi, sismoloji, spektroskopi, kuantum mekaniği, nükleer ve parçacık fiziği, plazma fiziği, uzaktan algılama, yüzey profilinin çıkarılması, deformasyon ölçümleri gibi birçok alanda kullanılmaktadır (URL-2).



Şekil 2.4. Michelson interferometri yapısı

2.1.7 Moiré saçakları yöntemi

Bu yöntemde sahne, ızgaradan geçirilerek yansıtılan bir ışık kaynağı ile aydınlatılır, ikinci bir kamera ve ilk ızgara ile özdeş fakat yanal olarak dönmüş bir ızgara ile görüntülenir. Sonuç görüntü eşit uzaklıktaki dış hatları gösterir, fakat karşılık gelen derinlik verileri hakkında bir bilgi yoktur (Sansoni ve diğ., 2009).

2.1.8 Odaklanma yöntemi

Derinlik ve uzaklık, lensin odak parametreleri göz önüne alınarak bulunabilir. Kamera lensi, “görme derinliği” olayı dikkate alınarak bir uzaklık ölçer gibi kullanılabilir (Sansoni ve diğ., 2009).

2.1.9 Gölge bilgisinden yararlanma

Bu teknik, yapılandırılmış ışık yönteminin bir çeşididir. Bilinmeyen bir nesnenin 3B modeli, bilinen bir nesnenin gölgesi nesnenin üzerine yansıtılarak yeniden üretilebilir (Sansoni ve diğ., 2009).

2.1.10 Doku bilgisinden yararlanma

Temel fikir, nesne yüzeyinin yönlendirmesini belirleyebilmek için doku elemanlarının (texture elements-textels) olası dönüşümlerini aramaya dayanır (Sansoni ve diğ., 2009).

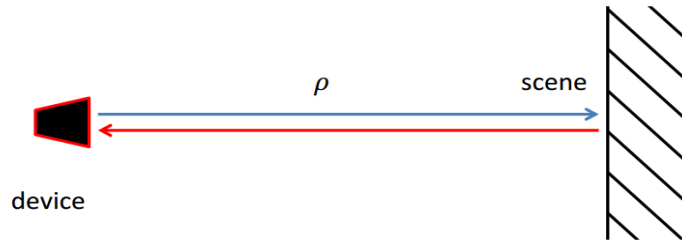
2.1.11 Gölgeleme yöntemi

Bu teknikte, bir nesnenin bir görüş açısından ışık kaynağının konumunun değiştiği ve böylece nesnede farklı gölgelerin oluştuğu görüntülerine ihtiyaç vardır. Bu metodun farklı bir türevi, aydınlanma koşullarının değişimini kullanır. Düşük çözünürlükte veri elde edilen bu yöntemin donanımı çok basit ve düşük maliyetlidir (Sansoni ve diğ., 2009).

3. DERİNLİK SENSÖRLERİ

3.1 Uçuş Süresi (ToF) Sensörler

Bir noktasal uçuş süresi (ToF) sensör, uçuş süresi veya RADAR (Radio Detection and Ranging) prensibine göre sensörün sahne noktasına olan radyal uzaklığını ölçer. Elektromanyetik radyasyon havada ışık hızı (c) ile hareket ettiğine göre optik radyasyonun, τ kadar zamanda kat etmiş olduğu ρ mesafesi $\rho = c\tau$ olarak hesaplanır. Şekil 3.1. 'de uçuş süresi sensörle uzaklık ölçümü görülmektedir: 0 anında soldaki uçuş süresi sensörün vericisinden yayılan radyasyon doğrudan ρ mesafesi boyunca sahneye doğru ilerler, sahne yüzeyinden geri yansır ve ρ mesafesi boyunca geri seyahat eder ve τ zamanında verici ile uygun şekilde konuşlandırılmış uçuş süresi sensörün alıcısına ulaşır. τ zamanda radyasyon 2ρ kadar bir mesafe kat etmiştir (Mutto ve diğ., 2012).



Şekil 3.1. Uçuş süresi sensörünün çalışma prensibi

Bu durumda ρ mesafesi ile τ zamanı arasındaki ilişki

$$\rho = \frac{c\tau}{2} \quad (3.1)$$

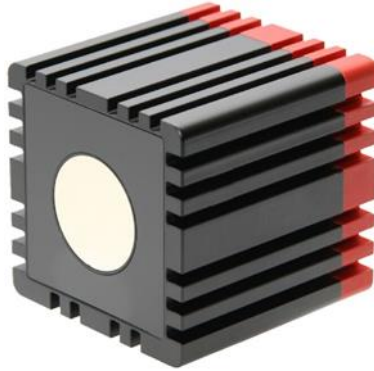
şeklinde tanımlanır.

Sahneyi oluşturan noktalardan çok sahne yüzeylerinin ölçülebilmesi için sahne tarama mekanizması üzerine belli sayıda noktasal ToF sensör monte edilmiştir. Topografik ve arkeolojik çalışmalar için kullanılan havadan arazi izleme LIDAR (Light Detection and Ranging) sistemlerinde ise, doğrusal düzene sahip ToF sensörleri yerine hareketli bir platform üzerine yatay ve düşey yönde düzenli harekete sahip düşey doğrusal bir düzende yerleştirilmiş sensörler tercih edilir. Zamansal sıralı tarama düzeneğinin sahneyi taramak için zamana ihtiyacı olması nedeni ile hareketli

nesnelerin olduđu sahneler gibi dinamik sahnelerin taranması için dođal olarak uygun deđildirler. Zamansal sıralı bir dzenek üzerine monte edilmiř noktasal ToF sensör ile sahne geometrisini elde etmeye çalışan sistemlerden farklı olarak, matrissel ToF kameraları, $N_R \times N_C$ boyutunda her biri bağımsız, ama önlerindeki sahne noktasına olan uzaklığı eş zamanlı olarak ölçen bir sensör matrisi ile sahne geometrisini tek bir çekimde deđerlendirmeye çalışır. ToF kameralar, video görüntüsü hızında derinlik haritaları ya da her bir girişin matris pikseli ile ilgili sahne noktasının arasındaki uzaklığa karşılık geldiđi ölçüm matrisleri üretirler. Basit bir konsepte sahip olmasına rağmen teknolojik olarak gerçekleştirilmesi zordur. Mesafe çözünürlüğünün 1 mm civarında olabilmesi için 5 piko saniyelik zaman adımlarını ölçebilecek bir saate ihtiyaç vardır. ToF sensörlerinde farklı saat teknolojileri kullanılmaktadır. Genel olarak sürekli dalga (CW) yoğunluk modülasyonu (Lange, 2000), optik obtüratör (OS) yaklaşımı (Iddan ve Yahav, 2001) ve tekil foton çıđ diyotları (SPAD) yaklaşımı en çok tercih edilen çözümlerdir (Mutto ve diğ., 2012).

3.1.1 CW matrissel uçuş süresi (ToF) menzil kameraları

Matrissel ToF menzil kameraları, görüntülenen sahnenin 3B geometrisini video hızında (saniyede 50 kareye (fps) kadar) elde edebilme yeteneđine sahip aktif sensörlerdir. MESA Görüntüleme (Şekil 3.2), PMD Teknolojileri ve Optrima SoftKinetic ve Microsoft gibi farklı üreticilerin çeřitli ticari ürünleri bulunmaktadır. Diđer şirketler (Panasonic, IEEE) ve araştırma enstitüleri de ToF kameralar üzerinde çalışmaktadır (Mutto ve diğ., 2012).



Şekil 3.2. MESA Imaging SR4000TM ticari ToF kamera

Sürekli dalga ToF kameralar, sahneye f_{mod} frekansının sinüzoidi ile modüle edilmiř A_E genlikli bir $S_E(t)$ kızılötesi (IR) optik sinyal gönderir.

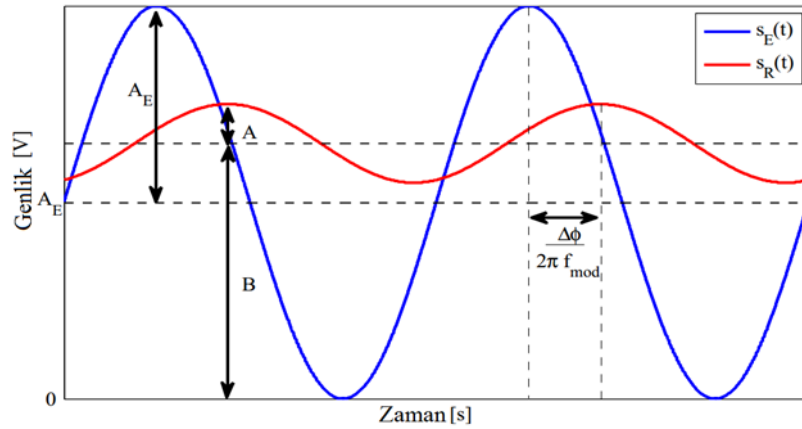
$$S_E(t) = A_E [1 + \sin(2\pi f_{\text{mod}} t)] \quad (3.2)$$

$S_E(t)$ sinyali sahne yüzeyinden geri yansır ve verici ile uygun şekilde yerleştirilmiş alıcıya geri döner. Alıcıya dönen sinyal, genel olarak alıcıya bağlı enerji soğrulması, yayılım zayıflaması (mesafenin karesi ile orantılı) ve sinyaldeki $\Delta\phi$ faz gecikmesi nedeniyle,

$$S_R(t) = A_R [1 + \sin(2\pi f_{\text{mod}} t + \Delta\phi)] + B_R \quad (3.3)$$

şeklinde yazılabilir. Burada A_R dönen sinyalin azalmış genliğini ve B_R alıcıya dönen yayılan sinyalin kızılötesi dalga boyundaki karışım radyasyonunu ifade eder. Yayılan $S_E(t)$ sinyali ve alınan $S_R(t)$ sinyalleri şekil 3.3. 'te gösterilmiştir. A_R (bundan sonra A) niceliği yararlı sinyalin genliği olduğu için genlik olarak isimlendirilir. $A_R + B_R$ (bundan sonra B) niceliği yoğunluk veya ofset olarak isimlendirilir ve alınan sinyalin ortalamasıdır. Buna göre (3.3) eşitliği;

$$S_R(t) = A \sin(2\pi f_{\text{mod}} t + \Delta\phi) + B \quad (3.4)$$



Şekil 3.3. Yayılan sinyal $S_E(t)$ (mavi) ve alınan sinyal $S_R(t)$ (kırmızı)

(3.4) eşitliğinin bilinmeyenleri A ve B kızılötesi radyasyonun volt olarak ölçülen genlikleri ve $\Delta\phi$ birimsiz faz değeridir. En önemli bilinmeyen $\Delta\phi$ faz bilinmeyenidir. CW ToF kameraları, ρ mesafesini $\Delta\phi$ faz bilinmeyi, (3.1) ve (3.3) eşitliklerinden aşağıdaki şekilde çıkarılır:

$$\Delta\phi = 2\pi f_{\text{mod}} \tau = 2\pi f_{\text{mod}} \frac{2\rho}{c} \quad (3.5)$$

ya da

$$\rho = \frac{c}{4\pi f_{\text{mod}}} \Delta\phi \quad (3.6)$$

A ve B bilinmeyenleri sinyal gürültü oranı için önemli parametrelerdir.

A, B ve $\Delta\phi$ bilinmeyenlerini belirleyebilmek için alıcı, $S_R(t)$ sinyalini modülasyon sinyalinin her periyodu için en az 4 kez örnekler (Lange, 2000), Örneğin eğer modülasyon sinyali 30 MHz ise alınan sinyal en az 120 MHz olarak örneklenmelidir. Örnekleme frekansı $F_S = 4f_{\text{mod}}$ ve 4 örneklem $S_R^0 = S_R(t=0)$, $S_R^1 = S_R(t=1/F_S)$, $S_R^2 = S_R(t=2/F_S)$ ve $S_R^3 = S_R(t=3/F_S)$ olsun. Alıcı, $\hat{A}$, $\hat{B}$ ve $\Delta\phi$ bilinmeyenlerini aşağıdaki şekilde hesaplar:

$$(A, B, \Delta\phi) = \arg \min_{A, B, \Delta\phi} \sum_{n=0}^3 \left\{ S_R^n - \left[A \sin\left(\frac{\pi}{2}n + \Delta\phi\right) + B \right] \right\}^2 \quad (3.7)$$

(3.7) eşitliğinden bazı cebirsel işlemlerden sonra,

$$A = \frac{\sqrt{(S_R^0 - S_R^2)^2 + (S_R^1 - S_R^3)^2}}{2} \quad (3.8)$$

$$B = \frac{S_R^0 + S_R^1 + S_R^2 + S_R^3}{4} \quad (3.9)$$

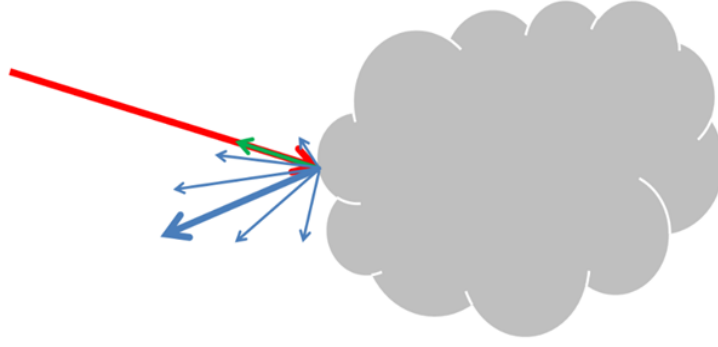
$$\Delta\phi = \arctan 2(S_R^0 - S_R^2, S_R^1 - S_R^3) \quad (3.10)$$

elde edilir. Son olarak ρ mesafesi, (3.6) ve (3.10) eşitliklerinden,

$$\rho = \frac{c}{4\pi f_{\text{mod}}} \Delta\phi \quad (3.11)$$

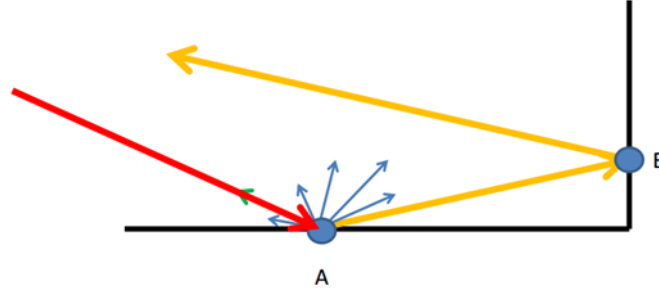
olarak hesaplanır. Pratikte, faz belirsizliği, harmonik distorsiyon, gürültü kaynakları, doygunluk ve hareket bulanıklığı gibi ideal olmayan koşulların etkileri de göz önünde bulundurulmalıdır.

Sinyalin ideal yayılım senaryosu şekil 3.1. 'de gösterildiği gibi olmasına rağmen gerçekte ancak şekil 3.4 'te gösterilen yeşil ışının varlığında söz konusudur. Pratikte şekilde görülen diğer ışınlar (mavi ışınlar) her zaman ihmal edilmez. Özellikle, gelen ışın ile yansıma noktasındaki yüzey normaline göre aynasal yansıma özelliği göstererek yansıyan ışın (kalın mavi ışın) en fazla radyometrik güce sahip olan ışıdır (Mutto ve diğ., 2012).



Şekil 3.4. Saçılma etkisi

Tüm yansıyan ışınlar (mavi ışınlar), diğer sahne noktalarına çarpar ve sensöre geri yansır ve bu nedenle diğer sahne noktalarının ölçüşünü etkiler. Şekil 3.5 'te görüldüğü üzere yayılan ışın (kırmızı ışın), A yüzey noktasına çarptığı anda saçılmaya uğrar. Saçılmaya uğrayan ışınlardan biri (turuncu ışın), diğer bir B sahne noktasına çarpar ve sensöre geri yansır. Sensör pikseli tarafından ölçülen uzaklık iki yansıma yolundan etkilenir. P noktasının kameraya olan radyal mesafesi, kamera ile P noktası arasındaki en kısa zaman uzunluğuna sahip yol ile hesaplanır. Çoklu yol etkisi sahne noktalarının uzaklıklarının yanlış hesaplanmasına etki eder (Mutto ve diğ., 2012).



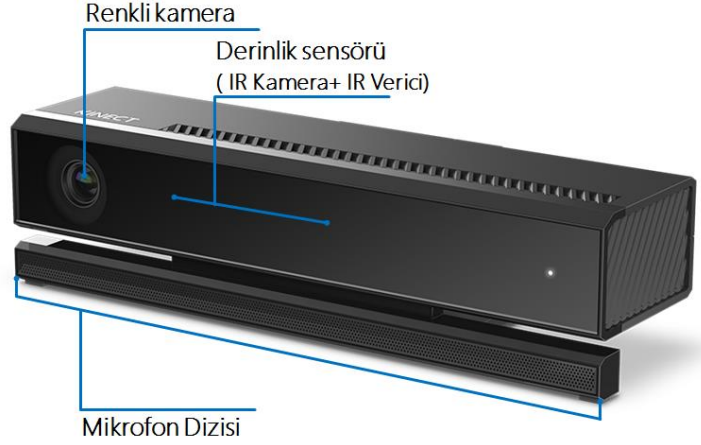
Şekil 3.5. Çoklu yol etkisi: Yayılan ışın (kırmızı) A yüzey noktasında çoklu olarak yansır. Turuncu ışın A noktasından sonra B noktasından sensöre geri yansır.

Çoklu yol etkisi, ToF kameralarının büyük hata kaynaklarından biridir. Bu etki sahneye göre değişim gösterdiği için modellenmesi zordur. Bu hatanın etkisini bastırmak için bir metot bulunmamaktadır (Mutto ve diğ., 2012).

3.2 Kinect Kamera

Kinect, Microsoft firması tarafından piyasaya sürülen öncelikle oyun konsollarında oyuncuların herhangi bir kontrolör olmaksızın oyunla iletişime geçmesi amaçlanmış, üzerinde renkli bir kamera ve kızılötesi sensör bulunan heterojen bir kamera sistemidir. Üzerinde monte halde bulunan standart kamerası ve derinlik verilerini işlemlerini

sağlayan kızılötesi sensörü ile bilgisayarla görme toplulukları tarafından ucuz yollu bir derinlik kamerası olması nedeni ile ilgi görmüş ve bu kamera kullanılarak yapılan çalışmalar son yıllarda artış göstermiştir.



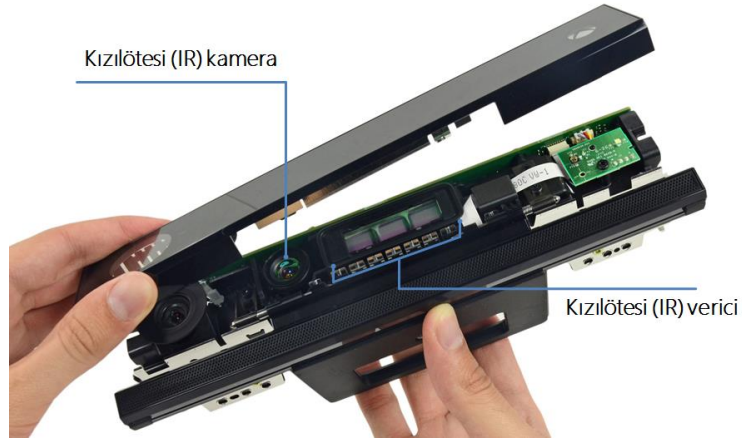
Şekil 3.6. Kinect for Windows v2

İlk olarak 2010 yılında görücüye çıkan Kinect v1, derinlik verilerini elde edebilmek için yapılandırılmış ışık kullanmıştır. Yakın zamanda çıkmış olan Kinect v2 kamerası (Şekil 3.6) ToF sensör kullanarak mevcut teknolojisini son çıkardığı aygıtla geliştirmiş oldu.

Kinect kamerası, derinlik verisi, renk verileri ve özel çipi sayesinde görüş alanında bulunan insanların (6 kişiye kadar) iskelet takibi bilgilerini sağlayabilmektedir. Microsoft'un sunmuş olduğu ücretsiz Kinect SDK veya internette ulaşılabilecek açık kaynak kütüphaneler sayesinde programcılara ve araştırmacılara geniş imkânlar sunmaktadır. Elde edilen akış verileri uygun algoritmalar ve kütüphaneler ile birlikte kullanıldığında görüntü işleme uygulamalarından, hareket takibi ve gözetleme uygulamalarına, yüz tanıma ve biyometrik uygulamalardan, kullanıcı arayüzü uygulamalarına, sanal gerçeklik uygulamalarından, robotik araştırmalarında araştırılan önemli konularından biri olan eşzamanlı konum belirleme ve haritalama (SLAM) uygulamalarına kadar geniş bir yelpazede uygulama olanağı bulunmaktadır.

1920x1080 piksel çözünürlüğe sahip renkli kamera ortamın ışık koşullarına bağlı olarak saniyede 15 ile 30 kare arası görüntü sağlamaktadır. Kinect SDK ile Kinect kamerasından amaca göre, RGBA, YUV, BGRA, BAYER ve YUY2 renk uzaylarında görüntü elde etmek mümkündür. Kinect v2 ile beraber sensörlerin görüş açıları genişletilmiş, renkli kameranın görüş açısı yatayda 84.1 derece, düşeyde 53.8 derece olmuştur (URL-3).

Yatayda 70 derece, düşeyde 60 derece bir görüş açısına sahip olan kızılötesi sensörü sayesinde 512x424 piksel çözünürlüğünde derinlik verisi elde etmek mümkündür. Derinlik verisi saniyede 30 kare hıza kadar veri toplayabilir. Microsoft tarafından açıklanan teknik özelliklere göre derinlik kamerasının faal alanı 0.5 m ile 4.5 m arası olarak belirtilmiştir. Yapılandırılmış ışık teknolojisinden yararlanarak derinlik verisi üreten Kinect v1 'in aksine yeni Kinect v2 (Şekil 3.7) 'de Canesta patentli altında ToF teknolojisine sahip bir sensör kullanılmıştır. Cihazın yeni teknolojisi ile birlikte kullanılan uygulamaların da artacağı düşünülmektedir.



Şekil 3.7. Kızılötesi sensörün yerleşimi (iFixit.com)

RGB-D (Renkli görüntü ve derinlik verisi üreten sensörler) sensörler üzerine yapılan daha önceki çalışmalara göre bu tür sensörlerin, bir ön ısınma zamanının sonuna doğru güvenilir veri ürettiği gözlemlenmiştir. Lachat ve diğ. (2015) yaptıkları çalışmada Kinect sensörünün bu özelliğini denetlemişler ve 30 dakika sonunda sensörün daha istikrarlı veri ürettiğini deneyimlemişlerdir. Dış mekanlarda yapmış oldukları deneylerde sensörün, direk güneş ışığının sensöre yansımadağı koşullarda çalışabildiğini gözlemlemişlerdir. Fakat güçlü arka ışığın olduğu durumlarda sensörün bilgisayarla olan bağlantısında kopukluklar olduğunu rapor etmişlerdir.

3.3 Kamera Kalibrasyonu

Fotoğraf makineleri ve video kameralar gibi renkli görüntüleme aygıtları ve ToF kameralar ve Kinect gibi derinlik görüntüleme aygıtları, ölçüm çalışmaları öncesinde ilk olarak kalibrasyon işlemine ihtiyaç duyarlar. Kalibrasyon işlemi, kameranın iç yöneltme karakteristiğini ve söz konusu cihazın konumsal pozisyonunu dikkate

almalıdır. Anlamli ölçümler elde edilebilmesi için bu işlem doğru ve kesin sonuç vermelidir (Mutto ve diğ., 2012)..

Genel bir cihazın, bir nesnenin bir (ya da daha fazla) a özelliğini ölçtüğünü düşünelim. $\hat{a}$, ölçülen değer ve a^* , $[a_{\min}, a_{\max}]$ aralığında değer alan, ölçülmesi gereken gerçek değer olsun. Ölçüler, iki tür hatanın etkisi altındadırlar:

- Sistematiik hatalar, ölçülerin doğruluğunu azaltırlar.
- Stokastik hatalar, ölçünün hassasiyetini azaltırlar.

Stokastik hataların karakterini ortalama ve varyans değerleri ortaya koyar. Varyans, çok sayıda ölçünün ortalaması alınarak azaltılabilir. Ortalamaya sistematiik hatalar etki etmiş olabilir. Ölçü aletinin kalibrasyonu genellikle sistematiik hataların etkisini azaltmaya yönelik (dolaylı olarak stokastik hataların ortalamasını ve çok sayıda ölçünün ortalaması alınarak varyansı da etkiler.) bir işlemdir (Mutto ve diğ., 2012).

Bir cihazın kalibrasyonu, bir a niceliğinin ölçümü sırasında, $\hat{a}$ ölçülen nicelik ile $a^* \in [a_{\min}, a_{\max}]$ gerçek niceliği arasındaki ilişkinin kestirimidir.

$\hat{a}$ ve a^* arasındaki ilişki iki yaklaşım kullanılarak türetilir.

Parametrik bir fonksiyonla: $a^* = f(\hat{a}, \theta)$, fonksiyonunda θ parametreler kümesini ve $f(\cdot)$ ise $\hat{a}$ ve θ ile a^* arasındaki ilişkiye uygun bir fonksiyonu temsil eder. Bu durumda kalibrasyon, bir parametre kestirim problemi haline gelir ve optimal $\hat{\theta}$ değeri, yer gerçeği (ground truth) ve ölçüm değerlerinin $(a_1^*, \hat{a}_1), (a_2^*, \hat{a}_2), \dots, (a_N^*, \hat{a}_N)$ kümesinden elde edilebilir. Örneğin, kalibrasyon parametrelerini $\hat{\theta}$, karesel ortalama hata yaklaşımı ile elde etmek genel bir yaklaşımdır (Mutto ve diğ., 2012).

$$\theta = \underset{\theta}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^N \left[a_i^* - f(\hat{a}_i, \theta) \right]^2$$

Kaba kuvvet yaklaşımı: Olası özellik değerlerinin kümesi iyi bilindiğinden $(a^* \in [a_{\min}, a_{\max}])$, N tane kontrol noktası $a_1 = a_1^*, a_2 = a_2^*, \dots, a_N = a_N^*$ ile birlikte özellik aralığı örneklenebilir. Her bir a_i^* , $i = 1, 2, \dots, N$ kontrol noktasına karşılık ölçümler gerçekleştirilir ve $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_N$ bağıl ölçümler kümesi elde edilir ve N çift $(a_1^*, \hat{a}_1), (a_2^*, \hat{a}_2), \dots, (a_N^*, \hat{a}_N)$ girdilerinden oluşan tablo sayesinde $a_1^*, a_2^*, \dots, a_N^*$ kontrol noktaları ile $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_N$ ölçümleri arasında ilişki kurulabilir. Elbette ki, daha çok

sayıda aralık örneği (N), ilişkinin daha doğru kurulmasını sağlamasına rağmen daha zahmetli bir kalibrasyon işlemini ihtiyaç duyar. Bu kalibrasyon yaklaşımı, parametrik olmayan veya modelsiz kalibrasyon tanımlanır (Mutto ve diğ., 2012).

Yukarıda anlatılan her iki kalibrasyon yaklaşımının da avantajları ve dezavantajları bulunmaktadır. Model tabanlı kalibrasyon yaklaşımı, her zaman bulunamayabilen ve elde etmesi zor olabilen bir modele ihtiyaç duyar. Ölçüm işleminin modellenmesi, hesaplanması gereken model parametre sayısını azaltmasının yanın da modelsiz kalibrasyona oranla kalibrasyon ölçümleri daha az sayıda N kontrol noktasına ihtiyaç duyar. Her bir kontrol noktası özelliği a^* ve onun bağlı ölçümü $\hat{a}$ arasında analitik benzerliğin bulunması diğer bir avantajdır (Mutto ve diğ., 2012).

Modelsiz yaklaşım açıkça bir modele ihtiyaç duymaz. Fakat örnekleme etkisini azaltmak için çok sayıda kalibrasyon ölçümüne ihtiyaç duyar. Bu yaklaşım, bir önceki yaklaşıma oranla aşırı uydurmaya (overfitting) meyillidir. Çünkü tahmin edilmesi gereken çok sayıda parametreye sahiptir. Kontrol noktası değeri a^* ve onun ölçümü $\hat{a}$ arasında analitik benzerlik bulunmamaktadır. Ayrıca, sadece örneklenmiş kontrol noktası seti için analitik olmasa da direk bir benzerlik vardır. Diğer veriler için benzerliğe, kaçınılmaz örnekleme hatalarının enterpolasyonu ile yaklaşılabilir (Mutto ve diğ., 2012).

Sayısal video kamera gibi standart bir kamerada, kalibrasyon iki bölüme ayrılır.

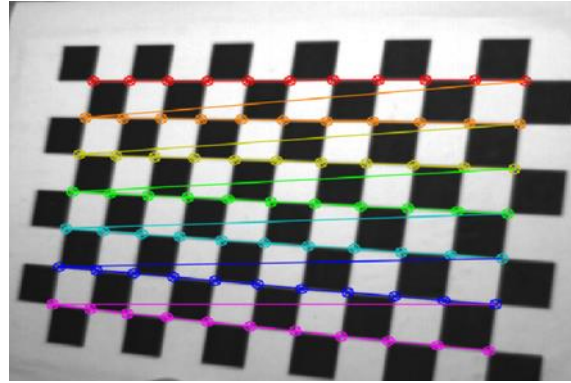
- Geometrik kalibrasyon, örneğin sensör piksellerinin konumu ile bağlı sahne noktaları arasındaki ilişkinin kestirilmesidir.
- Fotometrik kalibrasyon, örneğin güncel sahne noktasının rengi ile karşılık gelen pikselde kamera tarafından ölçülen renk arasındaki ilişkinin kestirimidir.

Fotometrik kalibrasyon daha kapsamlı olarak (Steger ve diğ., 2007; Szeliski, 2010) de bulunabilir. Fotometrik kalibrasyon bu tezin kapsamı dışındadır.

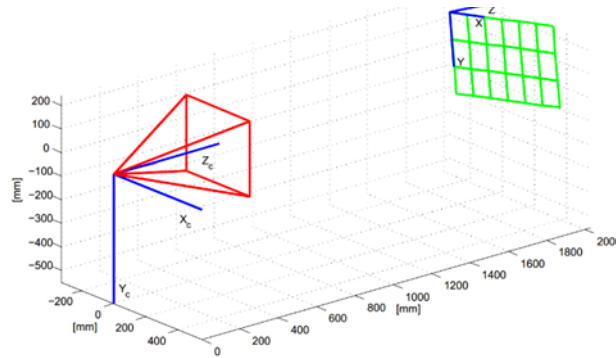
Herhangi bir kalibrasyon tekniği, bilinen büyüklükte, kolaylıkla tanınabilir göze çarpan noktalara sahip kalibrasyon nesnelerini kontrol noktası değeri a^* olarak benimser. Popüler kalibrasyon nesnesi (Heikkila, 1997; Zhang, 1998) desenlerinin boyutları bilinen siyah beyaz satranç tahtasıdır (şekil 3.8). Bu durumda siyah beyaz desenlerin (kolaylıkla belirlenebilen) köşe noktaları şekilde 3.8 'da görüldüğü üzere a^* değerleri olarak kabul edilir (Mutto ve diğ., 2012).

Gerçek dünya koordinat sistemi (World Coordinate System-WCS), satranç tahtasının sol üst köşesi koordinat sisteminin merkezi, x ve y eksenleri satranç

tahtasının düzleminde ve z eksenini ise bu düzleme dik olarak belirlenerek tanımlanabilir (şekil 3.9).



Şekil 3.8. Kalibrasyon nesnesi: Satranç tahtası üzerinde köşelerin belirlenmesi



Şekil 3.9. Satranç tahtasına göre sensör düzleminin konumu

ToF kameralar da kalibrasyon süreci ikiye ayrılır.

- Geometrik kalibrasyon, örneğin sensör pikselinin konumu ile bağlı sahne noktası arasındaki ilişkinin belirlenmesidir.
- Fotometrik kalibrasyon, örneğin sahne noktası ile ona karşılık gelen sensör pikseli arasındaki radyal uzaklık ile ToF kamera tarafından ölçülen uzaklık arasındaki ilişkinin kestirilmesidir.

ToF kameralarda geometrik kalibrasyon işlemi standart kameralara benzer şekilde genlik A_T değerlerinin oluşturduğu görüntüler standart renkler yerine kullanılarak gerçekleştirilebilir (Mutto ve diğ., 2012).

Fotometrik kalibrasyon, radyal mesafe ölçümlerindeki, harmonik distorsiyon ve foton gürültüsü gibi çeşitli faktörlerin etkisini göz önünde bulundurur. Harmonik distorsiyon ölçülecek nesneden uzaklığı bağlı iken, foton gürültüsü alınan sinyalin yoğunluğuna ve genliğine bağlıdır (Mutto ve diğ., 2012).

Kinect gibi farklı doğaya sahip çoklu ölçüm sistemlerinden oluşan sistemlere heterojen görüntüleme sistemleri denir. Heterojen görüntüleme sistemlerinin kalibrasyonu, her bir görüntüleme sisteminin kalibrasyonu ile birlikte sensörlerin göreceli konumlarının da belirlenmesini kapsamaktadır.

4. ÖZİNİTELİK ÇIKARIMI

Bu bölümde, kameradan elde edilen sahnenin bazı parçalarının izlenmesi konusuna odaklanılmıştır. Birçok renk ve derinlik verilerinden yararlanarak 3B modelleme uygulamasının ilk aşamasını ardışık çerçevelerde gelen renk ve derinlik verilerinden çıkarılacak olan nokta bulutlarının birbirleri doğru bir şekilde hizalanabilmesi için gerekli olan anahtar noktaların bulunması oluşturur. Bu noktalar, ardışık imgeler veya nokta bulutları arasındaki doğru dönüşümün belirlenebilmesi için çok önemlidir.

Bilgisayarla görme de özellikle de nesne tanımadada, birçok teknik, yüzey ya da fiziksel nesne üzerinde ilgi noktalarının (anahtar noktalar) belirlenmesine dayanır. Bu belirleme işlemi ise özinitelik çıkarımı ile yapılır. “... Uygun seçilmiş özinitelikler sınıfların karakteristik özelliklerini ifade ederler ve tanıma başarısını olumlu yönde etkilerler. Özinitelik çıkarımının amacı, sınıfa ait bilginin mümkün olan en fazlasını daha küçük boyutlara taşıyabilmektir. Başka bir ifadeyle özinitelik çıkarımının amacı k -boyutlu bir X gözlem uzayından, daha küçük l -boyutlu Y özinitelik uzayına, X uzayını mümkün olan en iyi şekilde ifade edebilen bir T dönüşümü bulabilmektir” (Alpaslan, 2013).

$$X [k \times k] \xrightarrow{T} Y [l \times l] \quad k > l \quad (4.1)$$

Kameranın hareketi esnasında ilgi noktalarının izlenebilmesi için, imgenin konumu, ölçeği ve dönüklüğünden bağımsız öziniteliklere ihtiyaç vardır. Bunlardan bazıları aşağıda listelenmiştir:

- Harris Köşe bulucu (Harris ve Stephens, 1988),
- Ölçekten bağımsız özinitelik dönüşümü (SIFT) (Lowe, 2004),
- Hızlandırılmış robust özinitelikler (SURF) (Bay ve diğ., 2008),
- Normal hizalanmış radyal özinitelikler (NARF) (Steder ve diğ., 2010),
- İkili robust bağımsız temel özinitelik (BRIEF) (Calonder ve diğ., 2010),

Bir özellikle ilgili olarak iki durum söz konusudur: İlgilenilen bir alanı belirleyen bir anahtar noktanın belirlenmesi ve bölgenin karakteristiğini belirleyen tanımlayıcı. Tipik olarak bulucu bir köşe ya da kenar gibi güçlü yoğunluk değişimi olan bir bölgeyi belirler ve bu bölgenin merkezi bir anahtar nokta olarak tasarlanır. Tanımlayıcı genel

olarak çevreleyen noktaların temel yönelimlerini ölçerek verilen anahtar noktayı belirleyen birçok boyutlu öznitelik vektörü oluşturur. Özniteliklerin bir kümesi verildiğinde, bir çift çerçeve arasında anahtar noktaların çiftlerinin bağdaştırılması işlemine eşleme denir.

Yukarıda bahsedilen öznitelikler bir çizelge halinde şöyle özetlenebilir:

Çizelge 4.1. Genel olarak kullanılan öznitelikler

Öznitelik	Bulucu	Tanımlayıcı
Harris köşe operatörü	Var	Yok
SIFT	Var	Var
SURF	Var	Var
NARF	Var	Var
BRIEF	Yok	Var

4.1 Harris Köşe Bulucu

Harris köşe operatörü olarak bilinir. İlk buluculardan bir tanesidir. 1988 de Harris ve Stephens tarafından geliştirilmiştir. Köşe kavramı geniş çaplı olarak düşünülmektedir. Metot sadece köşeleri değil, kenarları ve daha genel olarak anahtar noktaların belirlenebilmesini sağlamaktadır. Bu belirleme işi imge yoğunluklarının ikinci moment matrisinin hesaplanması, yerel varyasyonlarının tanımlanmasıyla yapılır (Högman, 2012).

Dönüklük, ölçek, aydınlanma koşullarının değişimi ve görüntü gürültülerine karşı başarılı olması çok kullanılan bir bulucu olmasına yol açmıştır. Yüzey normallerinin yoğunluk gradyenleri ile birleştirilmesi sonucu Harris Matrisi şu şekilde tanımlanır:

$$C6D_I = \sum_{k \in W} G_k G_k^T \quad (4.2)$$

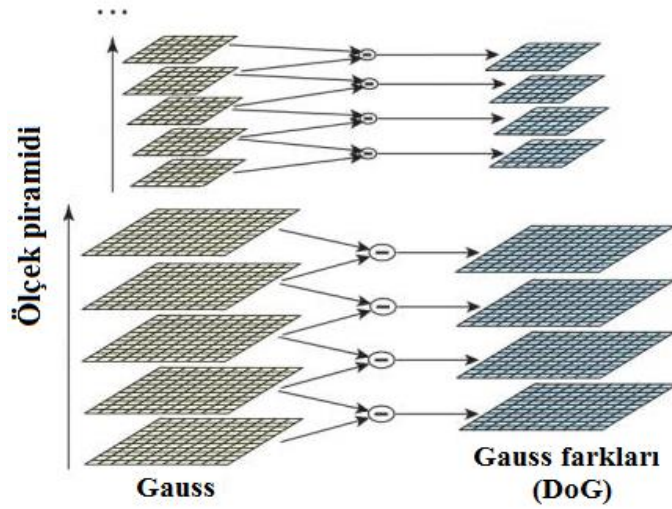
Burada W , i noktasının komşu kümesi ve k noktasının yüzey normali $N_k^T = (N_x, N_y, N_z)$ ve k noktasındaki yoğunluk gradyeni $D_k^T = (D_x, D_y, D_z)$ olmak üzere $G_k = (N_k^T D_k^T)^T = (N_x, N_y, N_z, D_x, D_y, D_z)^T$ dir. D_k^T yüzey normaline ve yerel yoğunluğun en dik yönü doğrultusundaki noktalara ortogondur. Vektörün büyüklüğü yoğunluk değişiminin oranını gösterir (Wang ve diğ., 2014).

Harris operatörünün orijinal biçiminde ölçekten bağımsızlığı sağlamak için matrisin farklı ölçeklerde yeniden hesaplanmasının gerekir.

4.2 Ölçekten Bağımsız Öznitelik Dönüşümü (Scale Invariant Feature Transform-SIFT)

Bilgisayarla görme, görüntü işleme ve robotik alanında sıklıkla kullanılan bir metottur. Nesne belirleme ve tanıma ya da imgeler arasındaki geometrik dönüşümlerin hesaplanması gibi görevlerin yerine getirilmesi için imgeler arasında kolayca eşlenebilen karakteristik değişmez öznitelik noktalarının belirlenmesinde tercih edilmektedir.

SIFT metodunun temel fikri artan karmaşıklıkta aşamalı operasyonlar tanımlamaktır, böylece en maliyetli işlemler en olası örnekler üzerinde gerçekleştirilir (Högman, 2012).

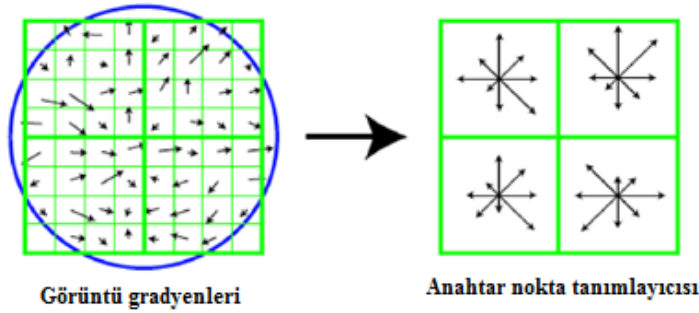


Şekil 4.1. Ölçek uzayında Gauss farkları

SIFT 'in işlem adımları şöyledir (URL-4):

- 1) Ölçek uzayının oluşturulması: İlk başta orijinal görüntünün farklı ölçeklerde gösterimleri oluşturulur (Şekil 4.1). Böylece ölçekten bağımsız olması sağlanır.
- 2) Gauss Laplası (Laplacian of Gaussian-LoG) yakınsaması: Gauss Laplası, görüntülerde anahtar noktalarının bulunması konusunda çok yetenekli olmasına rağmen oldukça hesaplama yükü getirir. Bu nedenle yakın bir sonuç elde edebilmek için hesap yükü açısından daha verimli olan Gauss farkları tercih edilir.

- 3) Anahtar noktaların bulunması: Gauss farkları kullanılarak anahtar noktalar bulunmaya çalışılır. Bunlar, Gauss farklarında elde edilen görüntülerdeki maksimum ve minimum değerlerdir.
- 4) Kötü anahtar noktaların temizlenmesi: Kenarlar ve düşük kontrasta sahip bölgeler kötü anahtar noktalar. Bunların temizlenmesi algoritmanın daha sağlıklı çalışmasını sağlar.
- 5) Anahtar noktalara yönlerin atanması: Her bir anahtar nokta için gradyen yönleri hesaplanır. Bundan sonraki hesaplamalar bu yön bilgisine göre gerçekleştirilir. Böylece dönüklük ve aydınlanma koşullarındaki değişimden öznitelik bilgisinin etkilenmesi engellenmiş olur. Bu sayede yöntemin dönüklükten bağımsız olması sağlanmış olur.



Şekil 4.2. SIFT tanımlayıcısı: Anahtar noktaların etrafındaki gradyen yönleri hesaplanır

Ölçekten bağımsız öznitelik dönüşümünün matematiksel işlem adımlarına göre, öncelikle bir I imgesi ve σ varyanslı G Gaussianının konvolüsyonu ile ölçek uzayı oluşturulur (4.3).

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (4.3)$$

Burada

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (4.4)$$

dir. Gauss farkı (Difference of Gaussian-DoG), σ eksenini boyunca ölçek uzayındaki iki katman arasındaki fark şu şekilde hesaplanır:

$$\begin{aligned} D(x, y, \sigma) &= [G(x, y, k\sigma) - G(x, y, \sigma)] * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (4.5)$$

Gauss farkı, Lindeberg (1994) tarafından gösterildiği gibi Gauss Laplasyanına ($\sigma^2 \nabla^2 G$) yakın sonuçlar üretir (4.7).

$$\sigma \nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma} \quad (4.6)$$

dir ve buradan;

$$G(x, y, k\sigma) - G(x, y, \sigma) \approx (k-1)\sigma^2 \nabla^2 G \quad (4.7)$$

elde edilir. σ^2 çarpanı ölçekten bağımsızlığı sağlar. Anahtar noktaların konumlarının tespit edilmesi, gauss farklarındaki ekstremumların bulunmasıyla yapılır. Lowe kalan $(k-1)$ çarpanının konumu etkilemediğini göstermiştir.

Eşleştirme işlemi için iyi bir taban oluşturmak amacıyla tanımlayıcının oldukça ayırt edici ve ışıktandırmadaki değişim ya da 3B bakış açısı gibi varyasyonlardan etkilenmemesi gerekir. SIFT tanımlayıcısı, anahtar nokta etrafındaki yerel yönlendirilmiş gradyenlerin histogramını (Histogram of Oriented Gradients-HoG) temel alır (Högman, 2012).

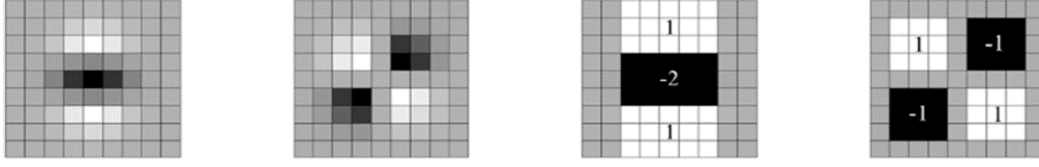
SIFT belirleyicisinin en önemli karakteristiği ölçeğe göre değişmez olmasıdır. Kameranın platformu hareket ettikçe ilgi noktalarını içeren alanlar ölçekteki değişime paralel olarak genişler ya da küçülür. Ölçekten bağımsız olması, anahtar bir noktanın değişik ölçeklerde istikrarlı bir yolla belirlenmesini sağlar (Högman, 2012).

4.3 Hızlandırılmış Robust Öznitelikler (Speeded Up Robust Features-SURF)

Bay ve diğ. (2006) “SURF: Speeded Up Robust Features” adlı çalışmalarında SIFT kadar etkin ve tekrarlanabilir, bazı durumlarda iki kat daha hızlı bir anahtar nokta bulucu ve tanımlayıcı bir yöntem tanıtmışlardır. SURF algoritması karışık sürekli reel değerli fonksiyonlar için dikdörtgensel ayırık tamsayı yakınsaklarını kullandığı için hızlıdır. SURF Hessian matrisinin determinantının maksimumunu kullanır (Brahmbhatt, 2013). Hessian matrisi bir fonksiyonunun ikinci dereceden kısmi türevlerini içeren bir matristir (URL-5). Hessian matrisi:

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (4.8)$$

ile tanımlanır. Burada $L_{xx}(x, \sigma)$, I görüntüsünün, x noktasında Gauss ikinci derece türevinin $\frac{\partial^2}{\partial x^2} g(\sigma)$ konvolüsyonudur. Şekil 4.3 ‘de ikinci derece Gauss türevlerinin nasıl görüldüğü ve bunların SURF dörtgensel tamsayı yakınsakları gösterilmektedir.



Şekil 4.3. Y ve XY yönündeki ikinci derece Gauss türevleri (sol) ve kutu filtre yakınsakları (sağ)

Hessian matrisinin determinantının ekstremumlarının köşe benzeri anahtar noktaları temsil ettiğini daha iyi anlayabilmek için konvolüsyonun birleşmeli olduğu hatırlanmalıdır. Hessian matrisinin elemanları aynı zamanda Gauss filtresi ile yumuşatılmış görüntünün ikinci dereceden uzaysal türevleri olarak da düşünülebilir. Görüntünün ikinci dereceden uzaysal türevleri eğer yoğunlukta ani bir değişim gerçekleşirse bir zirve değerine ulaşır. Bununla birlikte kenarlar ani yoğunluk değişimleri olarak sayılabilirler (Brahmbhatt, 2013). Bay ve diğ. (2006), ölçek ve konum için farklı ölçütler kullanmaktansa Hessian matrisinin determinantına güvenmeyi tercih etmişlerdir. Matrisin determinanı kenarları köşelerden ayırt etmemize yardım olur (Brahmbhatt, 2013). Hessian matrisinin determinanı;

$$\det(H) = L_{xx}(x, \sigma) * L_{yy}(x, \sigma) - L_{xy}(x, \sigma)^2 \quad (4.9)$$

ile belirlenir. Filtrelerin inşasından L_{xx} ve L_{yy} nin sırasıyla yatay ve dikey kenarlara karşılık geldiği açıktır. Ayrıca L_{xy} de köşegen kenarlar tarafından oluşturulan köşelere karşılık gelir. Determinant değeri, dikey ve yatay kenarların bir köşe oluşturacak şekilde kesiştiğinde ya da köşegen kenarlardan bir köşe oluştuğunda maksimum değerine ulaşır (Brahmbhatt, 2013).

SURF algoritmasının kullandığı dikdörtgensel görünümlü yakınsamalara aynı zamanda kutu filtreleri de denir. Kutu filtrelerinin kullanımı konvolüsyon işleminin hızlandırılmasını sağlayan integral görüntülerin kullanımına olanak sağlar. Herhangi bir görüntü için onun integral görüntüsünün herhangi bir pikselinin değeri, orijinden başlamak üzere o piksele kadar ki kümülatif toplamdır (Brahmbhatt, 2013). Matematiksel olarak bir I görüntüsü ve H da onun integral görüntüsü olmak üzere,

H'nin (x, y) pikseli;

$$H(x, y) = \sum_{i < x, j < y} I(i, j) \quad (4.10)$$

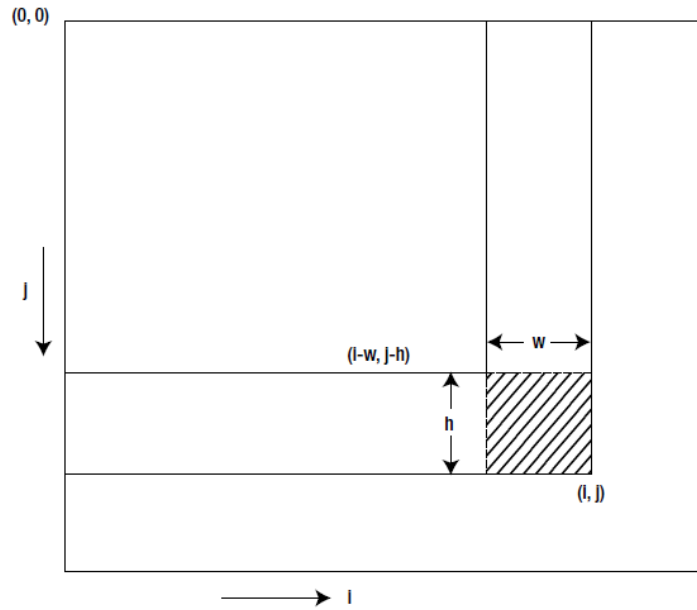
eşitliği ile verilir. İntegral görüntü, orijinal görüntüden aşağıdaki özyineli eşitlik kullanılarak doğrusal bir zamanda hesaplanabilir (Brahmbhatt, 2013).

$$H(x+1, y+1) = I(x+1, y+1) + H(x, y+1) + (x+1, y) - H(x, y) \quad (4.11)$$

İntegral görüntülerin en ilginç özelliklerinden bir tanesi de integral görüntüyü bir kez elde ettiğinizde, orijinal görüntüdeki istenilen piksel boyutuna sahip bir kutunun toplamı sadece 4 işlemde elde edilebilmektedir. Şekil 4.4 'te görülen taralı bölgenin toplamı;

$$\text{Toplam} = H(i, j) - H(i-w, j) - H(i, j-h) + H(i-w, j-h) \quad (4.12)$$

denklemini kullanılarak hesaplanabilir.



Şekil 4.4. İntegral görüntü kullanılarak dörtgensel bir bölgenin toplamının hesaplanması

Çekirdek kullanılarak gerçekleştirilen konvolüsyon işlemi, piksel değerlerinin çekirdek elemanları ile skaler çarpımı ve toplamıdır. Eğer çekirdek elemanları sabit bir değerse iş basitleşir. Çekirdek alanı içindeki piksel değerleri toplanır ve toplam değer, sabit çekirdek değeri ile çarpılır. Kutu filtrelerinin (sabit değerli dikdörtgensel alanlardan oluşan çekirdeklerin) ve integral görüntülerin kullanılması algoritmanın hızlı çalışmasını sağlar (Brahmbhatt, 2013).

SURF algoritması, ölçek uzayı piramidini oluşturmak için imgenin boyutunu düşürmektense Gauss Filtresinin boyutunu yükseltir. Piramit oluşturulduktan sonra Hessian matrisinin değişik ölçeklerdeki determinant değerlerinin maksimumu, piramitteki komşu değerlerin ile kıyaslanması ile bulunur. Böylece SURF anahtar noktaları ve bunların ölçekleri elde edilmiş olur.

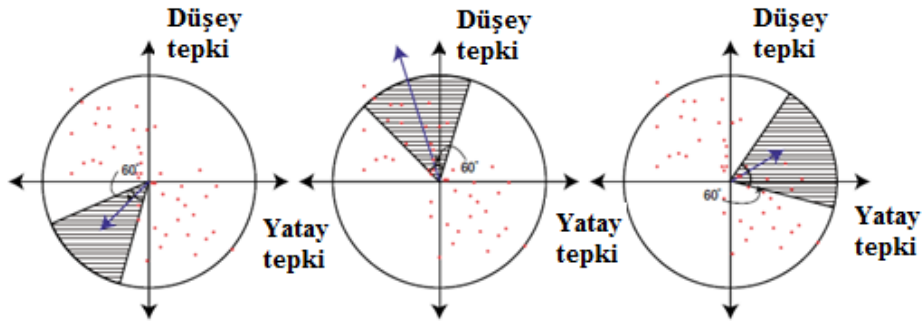
Anahtar nokta yönü, bir anahtar noktanın etrafında anahtar noktanın ölçeğinin 6 katı yarıçaplı bir dairesel komşuluk seçimi ile belirlenir. Bu komşuluktaki her bir



Şekil 4.5. Yönlendirmeyi belirlemek için kullanılan dikey ve yatay kutu filtreleri

noktada dikey ve yatay kutu filtrelerinin (Haar dalgacıkları olarak adlandırılır, şekil 4.5’de gösterilmektedir.) tepkileri kaydedilir (Brahmbhatt, 2013).

Tepki değerleri, bir Gauss ($\sigma = 2.5 \times \text{ölçek}$) fonksiyonu ile ağırlıklandırılır ve yatay tepkilerin x eksenini boyunca, dikey tepkilerin y eksenini boyunca kuvvetli olduğu bir uzay içinde vektörel olarak temsil edilirler. 60 derecelik kayan bir yay parçası tüm bu uzayı tarar (Şekil 4.6).



Şekil 4.6. SURF de kullanılan kayan yönlendirme pencereleri

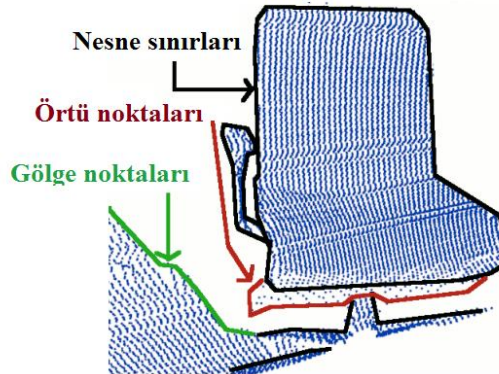
Kayan pencerede içindeki tüm tepkiler yeni bir vektör verecek şekilde toplanır. Kayan pencerenin iterasyonu kadar vektör oluşturulur. Bu vektörlerin en büyüğü anahtar noktanın yönlendirmesi üzerinde etkilidir (Brahmbhatt, 2013).

4.4 Normal Hizalanmış Radyal Öznitelikler (Normal Aligned Radial Features-NARF)

İlgi noktalarının bulunması, önemli yapılara odaklanması ve öznitelik çıkarımı için araştırma uzayının küçültülmesi açısından önem taşır. İlgi noktalarının bulunmasında yüzey yapısı ve sınırlar dikkate alınmalı, nesne farklı açılardan görüntülense dahi ilgi noktaları güvenilir bir şekilde belirlenmelidir. Bulunan ilgi noktaları, yüzey normallerinin hesaplanabilmesi veya tanımlayıcının hesaplanabilmesi için kararlı yüzeyler üzerinde olmalıdır (Steder, 2010).

Normal hizalanmış radyal öznitelikler yaklaşımı 3B nokta bulutlarından öznitelik çıkarımı için kullanılabilir olmasına karşın aynı zamanda, nokta bulutlarından çıkarılan menzil görüntüleri üzerinde de kullanılmaktadır.

Bu yaklaşıma göre bir görüntüde 3 tür sınır bulunmaktadır: Nesnenin etrafını kuşatan nesne sınırları, arka planın bir parçası olan nesne sınırları ve nesne sınırları ile gölge sınırları arasında kalan tarama teknolojisinden kaynaklanan örtü noktalarıdır (Nandanavanam, 2011) (Şekil 4.7).

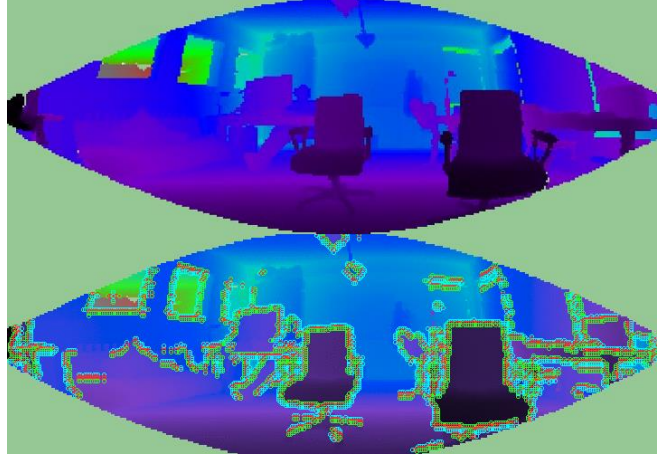


Şekil 4.7. Farklı türde sınır noktaları

Farklı açılardan gözlemlense dahi kararlı ilgi noktalarının etkin bir şekilde aynı yerde belirlenebilmesi için yerel komşuluklarında belirgin yüzey değişimleri olmalıdır. Bu da bu bölgede yüzey değişimlerinden büyük ölçüde farklı baskın yönleri olduğu anlamına gelir. Bunu yakalayabilmek için (Steder, 2010):

- Uzaklık görüntüsünde ön plandan arka plana doğru sürekli olmayan enine çizgiler (non continuous traversal) şeklinde beliren sınırlar (Şekil 4.8), komşu görüntü noktaları arasında yeterli miktardaki 3B mesafe artışlarına bakılarak belirlenir.

- Sınır bilgisi ile beraber her bir görüntü noktasının yerel komşuluklarında ne kadar bir yüzey değişimi olduğu ve bu değişimin baskın yönü belirlenmeye çalışılır.
- Her bir görüntü noktasının etrafındaki baskın yönlere bakılır ve bu yönlerin diğerlerinden ne kadar farklı olduğunu ve o noktadaki yüzeyin ne kadar kararlı olduğunu gösteren bir ilgi değeri hesaplanır.

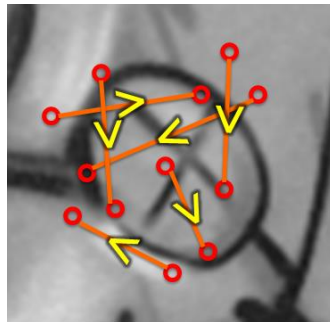


Şekil 4.8. NARF ile sınır noktalarının belirlenmesi

- İlgi değerlerine bir yumuşatma uygulanır.
- Sonuçta ilgi noktalarını bulabilmek için maksimum değerler bastırılır.

4.5 İkili Robust Bağımsız Temel Öznitelik (Binary Robust Independent Elementary Features-BRIEF)

Calonder ve diğ. (2010) tarafından önerilen bu yöntemde özniteliğin üretimi ve eşlemesi çok hızlıdır. Bu öznitelik noktası tanımlayıcısı, görüntü yamaları üzerinden seçilen nokta çiftlerinin yoğunluk değerlerinin kıyaslanması ile oluşturulan bir ikili karakter dizisi şeklindedir (Şekil 4.9). İkili karakter dizileri arasındaki benzerlik Hamming uzaklığından ölçülmektedir.



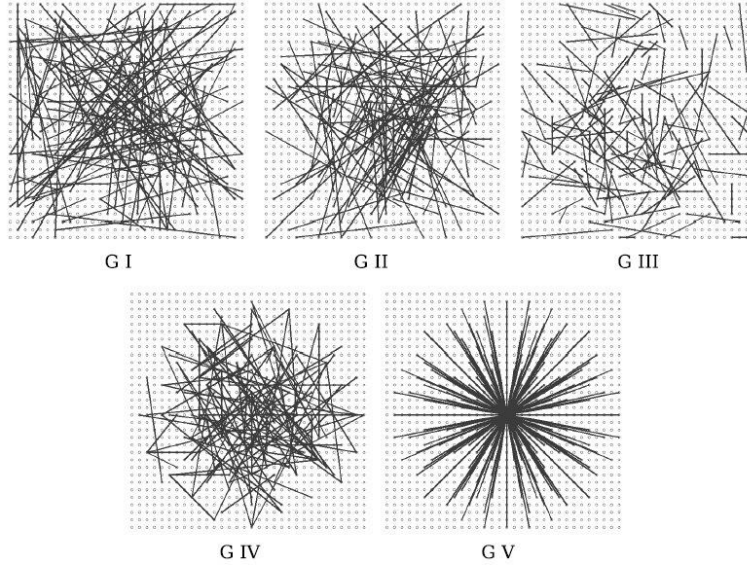
Şekil 4.9. BRIEF özniteliği

Öznitelik çıkarımı yöntemlerinin giderek daha çok veriyi işlemek zorunda kalması ve bu yöntemlerin işlem gücünün daha sınırlı olduğu mobil cihazlarda da kullanılmaya başlamasıyla birlikte hesaplaması hızlı yerel tanımlayıcılara ihtiyaç duyulmuştur. Bu açıdan BRIEF özniteliğinin hesaplanması ve eşlemesi hızlı, hafıza gereksinimi azdır (Calonder ve diğ., 2010).

Tek bir pikselin bilgisi dikkate alındığı durumlarda ilk önce pikselin değerinin gürültülerden etkilenebileceği unutulmamalıdır. Bu nedenle uygulanan yöntemde gürültülerin etkisinin azaltılabilmesi için seçilen görüntü yamalarına yumuşatma uygulanır. Daha sonra görüntü yaması üzerinde piksel çiftleri seçilerek yoğunluk değerleri kıyaslanır. Bu kıyaslama işlemi ile ikili değerlerden oluşan bir vektör elde edilir.

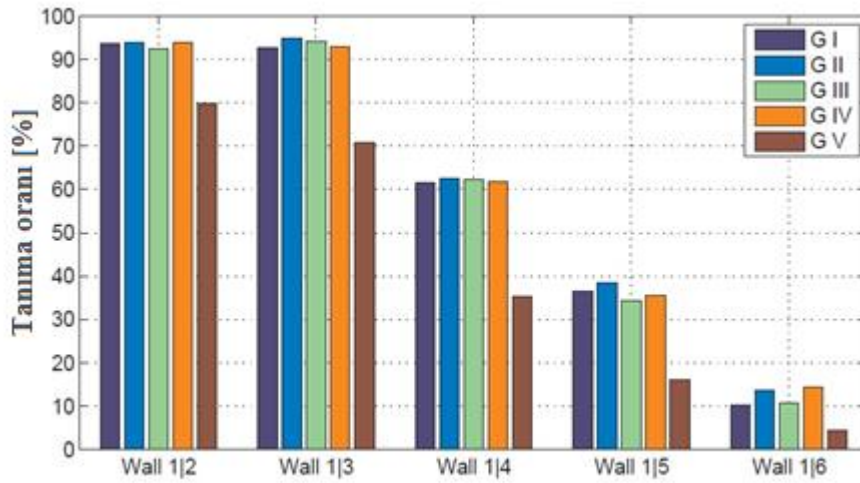
Önerilen yöntemde görüntü yaması üzerinde piksel çiftleri 5 farklı şekilde seçilerek (Şekil 4.10) sonuçları kıyaslanmıştır. Bu seçimler, X ve Y test edilecek olan piksel çiftlerinin konumunu temsil etmek üzere (URL-6);

1. X ve Y, uniform bir şekilde örneklenir.
2. X ve Y nin, Gauss dağılımı kullanılarak rasgele örneklenmesi, böylece seçilen konumlar yamanın merkezine yakın tercih edilmiş olur,
3. X ve Y nin koordinatları belirlenirken ilk olarak X, $0.04 \cdot s^2$ standart sapma ile birlikte Gauss dağılımı kullanılarak rasgele örneklenir. Daha sonra Y_i ler, X_i ortalama ve $0.01 \cdot S^2$ standart sapma ile örneklenir.
4. X ve Y basit kutupsal bir gridin ayrık konumları üzerinden rasgele örneklenir.
5. Her bir i değeri için X_i , (0,0) değerini alırken Y_i , basit kutupsal bir gridin üzerinde tüm olası değerleri alır.



Şekil 4.10. Beş örnekleme düzeni

Bu yöntemin farklı örnekleme stratejilerine göre tanıma performanslarından elde ettiği sonuçlar şekil 4.11’da gösterilmiştir. Görüldüğü üzere en kötü performansı sağlayan beşinci örnekleme yöntemi dışında diğer örnekleme yöntemlerinin yaklaşık aynı sonuçlar verdiği gözlemlenmiştir (URL-6).



Şekil 4.11. Beş farklı örnekleme stratejisinin performansları

Yöntemde benzerlik, eşlenik ikili değer dizileri arasındaki Hamming mesafesi ölçülerek karşılaştırılır. Hamming mesafesinin, XOR operasyonu ile kolaylıkla hesaplanabilmesi metodun hızlı bir şekilde işlemlerini sağlar (Calonder ve diğ., 2010). BRIEF yöntemi, dönüklük ve ölçekten etkilenir ve gürültüye karşı hassastır.

Rosin, yamanın momentini;

$$m_{pq} = \sum_{x,y} x^p y^q I(x,y) \quad (4.13)$$

şeklinde tanımlamıştır. Bu momentlerle yoğunluk merkezi ;

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (4.14)$$

şeklinde bulunur. Köşenin merkezinden (O), yoğunluk merkezine bir $\overrightarrow{OC}$ vektörü oluşturulur. Yamanın yönü

$$\theta = \arctan 2(m_{01}, m_{10}) \quad (4.15)$$

ile hesaplanır. Burada arctan2, arctanjantın quadrant versiyonudur. Rosin (1999), ayrıca köşelerin koyu ya da parlak olduğunu da dikkate almıştır. Ama ORB yönteminde köşe türüne bağlı olarak açı ölçümleri tutarlı olduğu sürece bu bilgi ihmal edilebilir.

BRIEF yöntemi daha önce elde edilen anahtar noktalarının yönleri göz önüne alınarak yönlendirilir. Sağlam ikili öznitelikler, eğitim yaması setlerindeki korelasyonun minimize edilmesi sayesinde öğrenilebilir (URL-7).

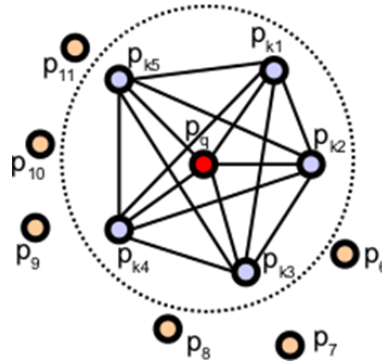
4.7 Nokta Bulutu Kütüphanesinin Yerel Tanımlayıcıları

Nokta bulutu kütüphanesi (Point Cloud Library) (Aldoma, 2012; URL-8), 3B bilgisayarla görme uygulamaları gibi uygulamalarda kullanıcılara, geliştiricilere ve programcılara nokta bulutu verilerini işleyebilmek için algoritmalar sunan açık kaynak kodlu popüler bir kütüphanedir. Derinlik verilerinden üretilen nokta bulutlarının işlenmesinde araştırmacılar tarafından sıklıkla tercih edilmektedir. Literatürde görüntü ve derinlik verileri ile yapılan çalışmalarda derinlik verisinin doğruluğuna bağlı olarak nokta bulutu kütüphanesi, derinlik verisinden elde edilen nokta bulutunun işlenmesi, nokta bulutlarının kaydedilmesi, 3B modelin oluşturulması gibi uygulamaların yanı sıra bu nokta bulutundan öznitelik çıkarımı için de bazı yöntemler sunmaktadır. Düşük maliyetli, pratik derinlik verisi üreten cihazlardan elde edilen nokta bulutları gürültüye karşı oldukça hassas olduğu için nokta bulutu kütüphanesinin sağlamış olduğu özniteliklerin tamamı veya bir kısmı Kinect vb., cihazlarla yapılan çalışmalarda pek tercih edilmemektedir ya da derinlik verileri bir ön işleme tabi tutulmaktadır.

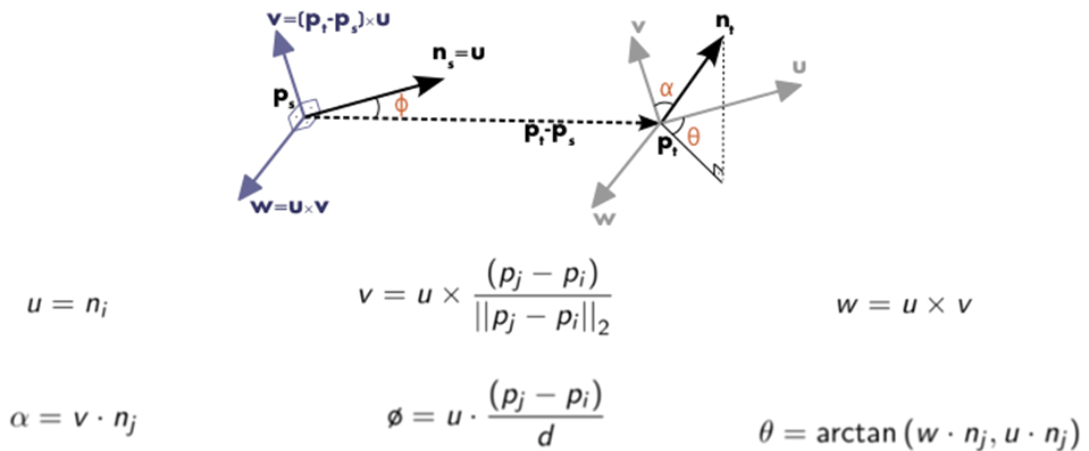
4.7.1 Nokta öznitelik histogramı (Point feature histogram-PFH)

PFH (Nokta Öznitelik Histogramı), Nokta Bulutu Kütüphanesi (PCL) tarafından önerilen en önemli tanımlayıcılardan bir tanesidir ve hızlı nokta öznitelik histogramı (FPFH) için temel oluşturmaktadır. PFH, nokta civarındaki normallerin doğrultuları arasındaki farkı analiz ederek noktayı çevreleyen geometri hakkındaki bilgileri yakalamaya çalışır (bundan dolayı belirlenen normallerin kalitesi tanımlayıcının kalitesini doğrudan etkiler.) (URL-9).

İlk olarak, algoritma civardaki tüm noktaları eşleştirir (sadece seçilen anahtar nokta ile komşuları değil, komşular da birbirleri ile eşleştirilir.) (Şekil 4.13). Her bir nokta çifti için noktaların normallerinden sabit bir koordinat çatısı hesaplanır. Bu koordinat çatısı ile, normaller arasındaki fark 3 açısal değişkenle (α, ϕ, θ) kodlanır (Şekil 4.14). Bu değişkenler ve noktalar arasındaki öklid uzaklık kaydedilir ve tüm çiftler hesaplandığında bir histograma kaydedilir. Sonuçta tanımlayıcı, tüm histogramların birleştirilmesi ile oluşur (URL-9).



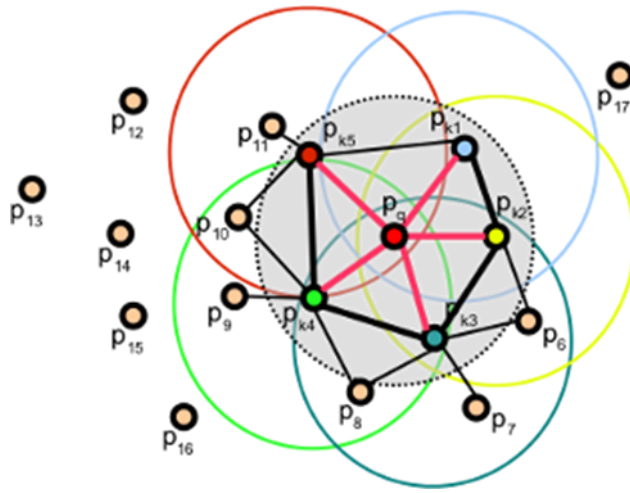
Şekil 4.13. Bir nokta için PFH hesaplanırken saptanan nokta çiftleri



Şekil 4.14. Sabit Koordinat Çatısı ve nokta çiftlerden biri için hesaplanan açısal özellikler

4.7.1.1 Hızlı nokta öznitelik histogramı (Fast point feature histogram-FPFH)

PFH kesin sonuçlar verse de hesap yükü açısından gerçek zamanlı uygulaması pratik değildir. N anahtar noktalı, k komşuluklu bir bulut göz önüne alındığında algoritma, $o(nk^2)$ karmaşıklığına sahip olur. Bundan dolayı, FPFH tanımlayıcısı meydana getirilmiştir. FPFH, komşular arasındaki ek bağlantıları çıkararak güncel anahtar nokta ve komşuları arasındaki direkt bağlantıyı göz önüne almaktadır (Şekil 4.15). Bu durumda karmaşıklık $o(nk)$ olur. Bundan dolayı sonuç histogramı sadeleştirilmiş nokta öznitelik histogramı (SPFH) olarak anılır. Referans çatısı ve açısal değişkenler PFH 'te olduğu gibi hesaplanır (URL-9).



Şekil 4.15. Bir nokta için FPFH hesaplanırken saptanan nokta çiftleri

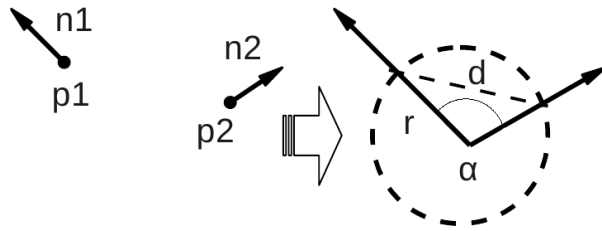
Komşular arasındaki ek bağlantıların dikkate alınmamasından doğan bilgi kaybı nedeniyle tüm histogramlar hesaplandıktan sonra ekstra bir işlem uygulanır. Noktanın komşularının SPFH'ları aralarındaki uzaklığa göre ağırlıklandırılarak hesaplanır. Son olarak 3 değişkene ait tüm histogramlar (PFH 'teki uzaklık bilgisi dikkate alınmaz) birleştirilerek sonuç tanımlayıcı oluşturulur (URL-9).

4.7.2 Yarıçap tabanlı yüzey tanımlayıcı (Radius-based surface descriptor-RSD)

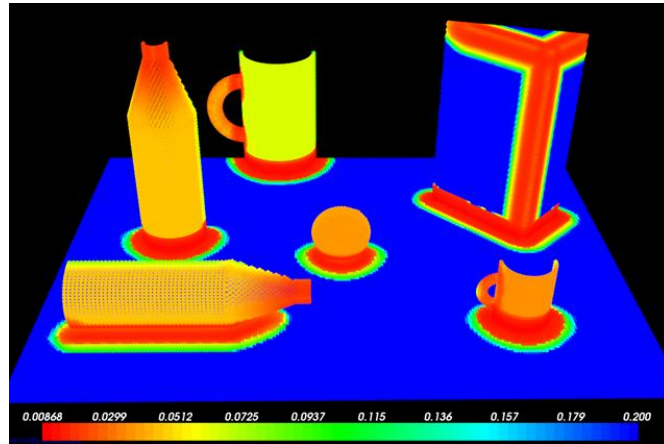
Yarıçap Tabanlı Yüzey Betimleyici (RSD), nokta ve komşusu arasındaki ışınsal ilişkiyi temsil eder. Her bir anahtar nokta ve komşu çifti için algoritma bu noktalar arasındaki uzaklığı ve normaleri arasındaki farkı hesaplar. Bunun üzerine her iki noktanın da bir küre yüzeyi üzerinde yattığı kabul edilerek, noktalar ve normaleri bu kürenin bulunmasına yardımcı olur (aksi takdirde, sonsuz çoklukta olası küre oluşur)

(Şekil 4.16). Sonuç olarak nokta-komşu noktası kürelerinden sadece maksimum ve minimum yarıçaplı olanları alınır ve o noktanın tanımlayıcısı olarak kaydedilir (URL-9).

İki nokta, düz bir yüzey üzerinde bulunduğu zaman kürenin yarıçapı sonsuz olur. Diğer taraftan eğer noktalar bir silindirin eğimli yüzeyinde bulunuyorlarsa yarıçap aşağı yukarı silindirin yarıçapı ile aynı olacaktır. Bu nedenle RSD ile nesnelerin ayrımı yapılabilmektedir. Şekil 4.17 ‘de görüldüğü üzere maksimum yarıçapa sahip noktalar (mavi), bir düzlemin parçası olarak değerlendirilebilmektedir (URL-9).



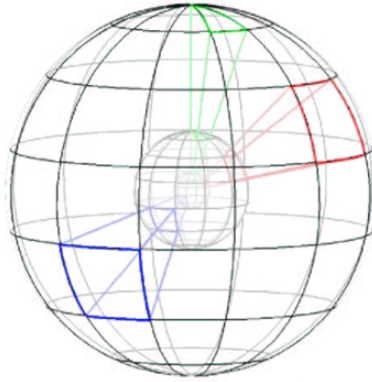
Şekil 4.16. İki nokta ve normalleri ile uyuşan küre



Şekil 4.17. Nokta bulutu için RSD değerleri; düz yüzey üzerindeki noktalar maksimum değere sahip olur.

4.7.3 3B şekil bağlamı (3D shape context-3DSC)

3B Şekil Bağlamı (3DSC), Belongie ve diğ. (2002) tarafından geliştirilen 2B şekil bağlamının üç boyuta uyarlanmış versiyonudur. İlk olarak tanımlayıcının hesaplandığı nokta için, bu noktayı merkez kabul eden bir destek yapısı oluşturulur. Daha doğru sonuçlar elde edebilmek için destek yapısı olarak belli bir yarıçapa sahip küre tercih



Şekil 4.18. Bir nokta için 3DSC yi hesaplamak için gerekli destek yapısı

edilir. Bu kürenin kuzey kutbu ile o noktadaki normal karşılaştırılır. Daha sonra bu küre 3B bölgelere ya da bölümlere ayrılır (Şekil 4.18). Elde edilen bölümler logaritmik olarak bölümlenmiştir ve merkeze doğru küçülmektedirler. Bu nedenle çok küçük bölümleri engellemek için minimum yarıçap açıkça belirtilir, aksi halde yöntem yüzeydeki küçük değişimler için aşırı hassas olur (URL-9).

Her bir bölüm için, bu bölümde yer alan komşu noktalardan hesaplanan birikimli ağırlıklandırılmış bir değer elde edilir. Ağırlık değeri, bölümün hacmine ve yerel nokta yoğunluğuna (mevcut komşu nokta sayısı) bağlıdır. Bu tanımlayıcıya, belli bir dereceye kadar çözünürlükten bağımsız olma özelliği kazandırır (URL-9).

Küre, normal yönünde konumlandırıldığı için iki eksenini sabit kalırken bir eksenini serbest olarak yönlendirilebilir. Bundan dolayı, tanımlayıcı dönüklükle başa çıkamaz. Bu durumun üstesinden gelebilmek için (iki farklı buluttaki aynı nokta aynı değeri alacağından), destek küresi normal etrafından N kez döndürülür ve yukarıda bahsedilen işlemler kürenin her konumu için tekrarlanır. Sonuçta o nokta için N tane tanımlayıcı üretilmiş olur (URL-9).

4.7.3.1 Benzersiz şekil bağlamı (Unique shape context-USC)

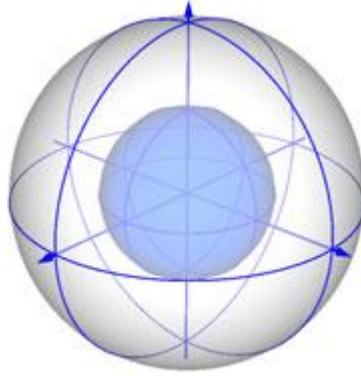
Benzersiz Şekil Bağlamı (USC) tanımlayıcısı, 3DSC'yi, her bir nokta için benzersiz bir yönlendirme sağlamak amacıyla bir yerel referans çatısı tanımlayarak genişletir.

Bu sadece tanımlayıcının kesinliğini geliştirmekle kalmaz, aynı zamanda boyutunu da küçültür ki yönlendirme için çoklu tanımlayıcı (3DSC için N tane) hesaplanmasına gerek kalmaz (URL-9).

4.7.4 Yönelimler histogramlarının imzası (Signature of histograms of orientations-SHOT)

Yönelimler Histogramlarının İmzası (SHOT), 3DSC gibi, küresel destek yapısı içindeki topoloji (yüzey) bilgisini elde etmektedir (Şekil 4.19). Bu küre azimut boyunca 8, yükseklik boyunca 2 ve yarıçap boyunca 2 bölge olmak üzere 32 bölüme veya birime bölünür. Her bir birim için tek boyutlu yerel histogram hesaplanır. Anahtar noktanın normali ile hacim içindeki mevcut noktanın normali arasındaki açı değişken olarak seçilir (URL-9).

Tanımlayıcı, bütün yerel histogramların hesaplanarak birleştirilmeleri sonucunda elde edilir. USC tanımlayıcısında olduğu gibi, SHOT yönteminde tanımlayıcının dönüklükten etkilenmemesi için yerel referans çatısı oluşturur. Bu yöntem aynı zamanda gürültüye karşı da etkilidir (URL-9).

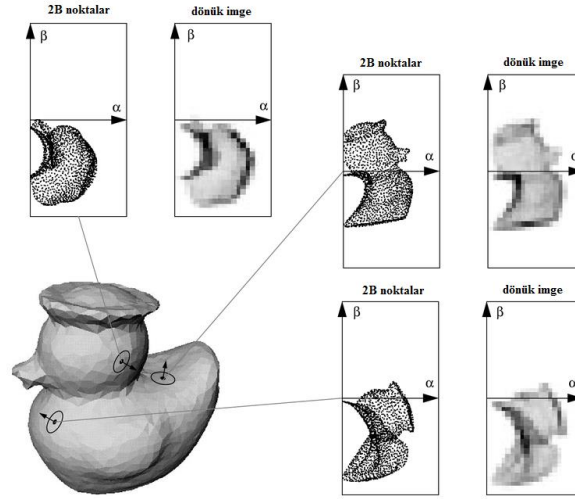


Şekil 4.19. SHOT hesaplamak için destek yapısı. Basit olması için sadece azimuta göre ayrılmış 4 bölüm gösterilmektedir.

4.7.5 Dönük imge (Spin image-SI)

Dönük imge (SI) (Johnson, 1997) en eski tanımlayıcılardan birisidir. 1997 yılından bu zamana kadar sadece belirli uygulamaların bazılarında kullanım alanı bulmuştur. Aslında köşelerden, kenarlardan ve poligonlardan oluşan yüzeyleri tanımlamak için dizayn edilmiş daha sonra nokta bulutlarına uyarlanmıştır (URL-9).

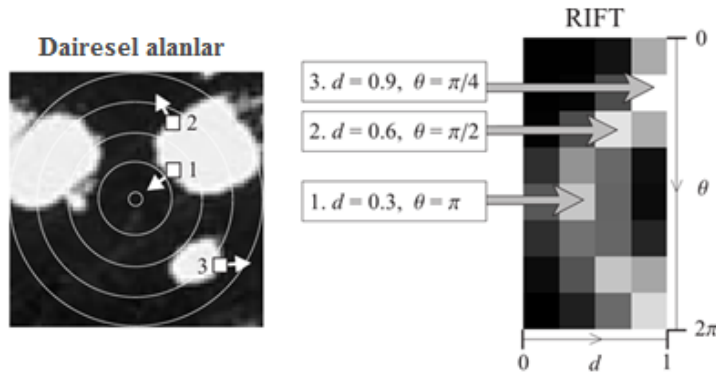
Destek yapısı olarak ilgi noktasını merkez kabul eden, o noktanın normali ile hizalanmış belirli boyutlara sahip bir silindir tercih edilir. Bu silindir radyal ve dikey olarak bölümlere ayrılır. Her bir bölüm için bölüm içinde bulunan noktalar toplanır ve tanımlayıcı üretilir. Sonucun iyileştirilebilmesi için ağırlıklandırma ve enterpolasyon işlemleri kullanılır. Sonuç tanımlayıcı, koyu alanların yüksek nokta yoğunluklu alanlara karşılık geldiği gri ölçek görüntüsü olarak elde edilir (Şekil 4.20) (URL-9).



Şekil 4.20. Modelin 3 noktası için hesaplanan dönük imgeler

4.7.6 Dönüklükten bağımsız öznitelik dönüşümü (Rotation-invariant feature transform-RIFT)

Dönüklükten Bağımsız Öznitelik Dönüşümü (RIFT), SI 'da olduğu gibi 2B öznitelik konseptlerinden yararlanmaktadır (Örneğin, SIFT). RIFT, hesaplanabilmesi için yoğunluk bilgisini gerektiren (bu tezin kapsamında) tek tanımlayıcıdır (yoğunluk değerleri RGB renk değerlerinden elde edilebilir.). Bu nedenle RIFT standart XYZ nokta bulutları ile kullanılmaz, ayrıca dokuya bilgisine ihtiyaç duyulur (URL-9).



Şekil 4.21. Tanımlayıcıdaki 3 değişik konumdaki RIFT öznitelikleri

İlk adımda, noktanın yer aldığı yüzey üzerine (verilen yarıçaplı) dairesel bir alan yerleştirilir. Bu alan, seçilen histogram bölümünü büyüklüğüne bağlı olarak eş merkezli halkalara bölünür. Sonra belirlenen yarıçaplı küre içinde bulunan komşu noktalar bir histogramın içine yerleştirilir (Şekil 4.21). Her bir noktadaki yoğunluk gradyeninin büyüklüğü ve yönlendirmesi değerlendirilir. Dönüklükten etkilenmemesi için, gradyenin yönü ile ilk aşamada belirlenen dairesel alanının merkezinden dışarı yönlendirilmiş vektör arasındaki açı ölçülür (URL-9).

Lazebnik ve diğ. (2005) 'nin orijinal uygulamasında 4 halka ve 8 bölümlü yönelim histogramı kullanılarak 32 boyutlu bir tanımlayıcı üretilmiştir. Hiçbir zaman büyük bir sorun olarak görülmediği halde RIFT, doku yapısının dönüklüğüne karşı etkili değildir (URL-9).

4.7.7 Dönel izdüşüm istatistikleri (Rotational projection statistics-RoPS)

Rotasyonel İzdüşüm İstatistikleri (RoPS) özneliği, üçgen kafes ile çalışması açısından diğer tanımlayıcılardan biraz farklıdır. Dolayısıyla bu yöntemde, öncelikle nokta bulutundan ihtiyaç duyulan kafes modelin üretilebilmesi için bir üçgenleme adımına ihtiyaç duyulur. Bir anahtar noktada RoPS hesaplamak için belirli bir yarıçap değerine göre yerel bir yüzey kırılır. Böylece sadece bu yerel yüzey parçası içerisinde kalan noktalar ve üçgenler hesaba katılır. Daha sonra tanımlayıcıya dönüklükten bağımsızlık özelliği kazandıracak şekilde bir yerel referans çatısı hesaplanır. Noktayı orijin kabul edecek şekilde bir koordinat sistemi oluşturulur ve eksenler yerel referans çatısı ile hizalanır. İlk olarak yerel yüzey mevcut eksen etrafında döndürülür. Sonra yerel yüzeydeki tüm noktalar XY, XZ ve YZ düzlemleri üzerine iz düşürülür. Her bir düzlem için iz düşüm noktalarının dağılımının istatistiksel bilgisi hesaplanır ve son tanımlayıcıyı oluşturmak üzere bu bilgiler birleştirilir (URL-9).

5. DÖNÜŞÜM PARAMETRELERİNİN BELİRLENMESİ

Bir önceki bölümde bahsedilen yöntemlerle görüntüden çıkarılan ilgi noktaları ve bunların birbirleri ile uygun eşleşmelerinden oluşan eşleşme çiftleri ile iki görüntü ya da iki nokta kümesi arasında sağlam bir dönüşümün belirlenmesi mümkündür. Bu dönüşüm ile kaynak görüntü karesindeki her bir noktanın hedef görüntü karesi üzerindeki eşleniği üzerine iz düşürülmesi mümkündür. Bu çalışmada bahsedilen 3 boyutlu dönüşüm matrisi ve öteleme vektörü ile tanımlanan bir perspektif dönüşümdür. Bu dönüşüm 4x4 boyutlarında bir matris şeklinde yazılabilir. Daha sonra $P = (x,y,z,1)^T$ olmak üzere bir P noktasına dönüşüm matrisi uygulanarak eşlenik noktaya iz düşürülebilir (Högman, 2011):

$$P_0 = T_{DÖNÜŞÜM} P \quad (5.1)$$

5.1 Rasgele Örneklerin Uzlaşımı (Random Sample Consensus-RANSAC)

RANSAC, Rasgele örneklerin uzlaşımı için kullanılan bir kısaltımadır. Uyuşumsuz veriler içeren gözlemlenmiş veri kümelerinden elde edilen matematiksel modelin parametrelerini tahmin etmek için kullanılan iteratif bir metottur. Deterministik olmayan bir metottur ve sadece belirli bir olasılık için makul sonuçlar üretmektedir. Daha fazla iterasyona izin verildikçe bu olasılık artar. Algoritma ilk olarak 1981 yılında Fischer ve Bolles tarafından yayınlanmıştır. Dağılımı model parametrelerinin bazı kümeleri ile açıklanabilen verilere uyumlu veriler, modele uymayan verilere de uyuşumsuz veriler denir. Temel kabul şudur ki algoritmanın verileri uyumlu ve uyuşumsuz verilerden oluşur. Bunlara ek olarak veri gürültüye maruz kalmış olabilir. Uyuşumsuz veriler gürültünün aşırı değerlerinden, ya da hatalı ölçümlerden ya da verilerin yorumlanmasıyla ilgili doğru olmayan hipotezlerden gelebilir. RANSAC aynı zamanda verilen uyumlu verilerin genellikle küçük bir kümesinde veriye uyan ya da en uygun şekilde açıklayan bir modelin parametrelerinin tahmin edilmesi için bir prosedürün varlığını kabul eder. RANSAC amacına, orijinal verinin rasgele bir alt kümesini iteratif olarak seçerek ulaşır (Bethencourt ve Jaulin, 2012).

Bu veriler hipotetik uyumlu verilerdir ve hipotez aşağıdaki şekilde test edilir:

- Bir model hipotetik uyumlu verilere uygundur. Başka bir deyişle, modelin bütün serbest parametreleri uyumlu verilerden inşa edilir.
- Diğer bütün veriler uygun modele karşı test edilir ve eğer bir nokta tahmin edilen modele iyi şekilde uygunsa aynı zamanda hipotetik uyumlu veri olarak düşünülebilir.
- Tahmin edilen model eğer yeterli çoklukta hipotetik uyumlu veri olarak sınıflandırılacak noktaya sahipse makul bir şekilde iyi bir modeldir.
- Model hipotetik uyumlu verilen başlangıçtaki kümesinden tahmin edildiğinden, tüm hipotetik uyumlu verilerden yeniden tahmin edilebilir.

Sonuç olarak, model kendisi ile bağıntılı uyumlu verilerin hatalarının tahmin edilmesiyle hesaplanır. Bu prosedür belirli bir sayı kadar tekrarlanır. Bu tekrarların her birinde ya çok az nokta uyumlu nokta olarak sınıflandırıldığından model reddedilir, ya da karşılık gelen hata ölçüsüyle birlikte tam bir model elde edilir. İkinci durumda eğer tam modelin hatası daha önce kaydedilen modelden daha düşük ise tam model saklanır. RANSAC'ın model parametrelerini tam tahmin yapabilmesi bir avantajdır. Başka bir deyişle veri kümesinde belirgin sayıda uyuşumsuz veri bulunsa bile parametreleri yüksek derecede kesinlikle tahmin edebilir. Bu parametrelerin hesaplanmasında geçen zamanının bir üst sınırının olmayışı RANSAC'ın dezavantajıdır. Hesaplanan iterasyonların sayısı sınırlı olduğunda elde edilen çözüm uygun olmayabilir ve veriye iyi bir şekilde uymayabilir. Bu durumda çok fazla sayıda iterasyon hesaplanarak makul bir model üretme olasılığı yükseltilebilir. RANSAC'ın diğer bir dezavantajı problemin belirli bir eşik seviyesinin belirlenmesini gerektirmesidir (Bethencourt ve Jaulin, 2012).

5.2 İteratif En Yakın Nokta (Iterative Closest Point-ICP)

İteratif en yakın nokta (ICP) (Besl ve McKay, 1992; Rusinkiewicz ve Levoy, 2001), iki nokta bulutunun noktaları arasındaki farkları minimize eden bir algoritmadır. İteratif en yakın nokta yöntemi sıklıkla, farklı tarama verilerinden 2B veya 3B yüzeylerin yeniden yapılandırılması, robotların konumlarının belirlenerek optimum güzergah belirlenmesi için kullanılır (Yang ve Medioni, 1992). Oldukça basit ve genellikle gerçek zamanlı uygulamalarda kullanılan bir yöntemdir. Bu yöntemin

modeli, ham tarama verilerindeki noktalar arasındaki mesafenin minimum olabilmesi için iteratif olarak düzenlenir. Algoritma giriş verisi olarak nokta bulutlarını, öncül bir dönüşüm parametresini ve iteratif işlemlerin sonlandırılabilmesi için bir durma kriterini ihtiyaç duyar (Bethencourt ve Jaulin, 2012). Yakınsama teorisine göre en küçük kareler yöntemi ile algoritmanın iteratif adımlarında, eşlenik noktalar arasındaki mesafe azaltılır (Besl ve McKay, 1992). Sonuçta düzeltilmiş dönüşüm bilgisi elde edilir.

Algoritmanın aşamaları şunlardır:

- İlk olarak rasgele noktalar seçilir. Bu örneklem noktalarının seçimi için farklı yöntemler kullanılabilir.
- Seçilen noktaların eşleniği olan en yakın noktalar araştırılır. K-d ağacı kullanılarak bu araştırmanın daha hızlı gerçekleştirilmesi sağlanabilir (Simon, 1996).
- En küçük kareler yöntemi ile eşlenik noktaların dönüşüm parametreleri hesaplanır. Nokta çiftleri ağırlıklandırılabilir ve ağırlık kriterine göre istenmeyen eşlenik nokta çiftleri reddedilebilir.
- Hesaplanan dönüşüm, nokta çiftlerine uygulanır. Bu işlemler iteratif olarak, durma kriterine kadar devam eder.

Basit ve hızlı bir yöntem olması nedeni ile özellikle düşük maliyetli cihazlardan elde edilen derinlik verileri ve bunlardan oluşturulan nokta bulutları için sıklıkla tercih edilen bir yöntemdir.

5.3 Döngü Bulma (Loop Closure) Problemi ve Genel Optimizasyon

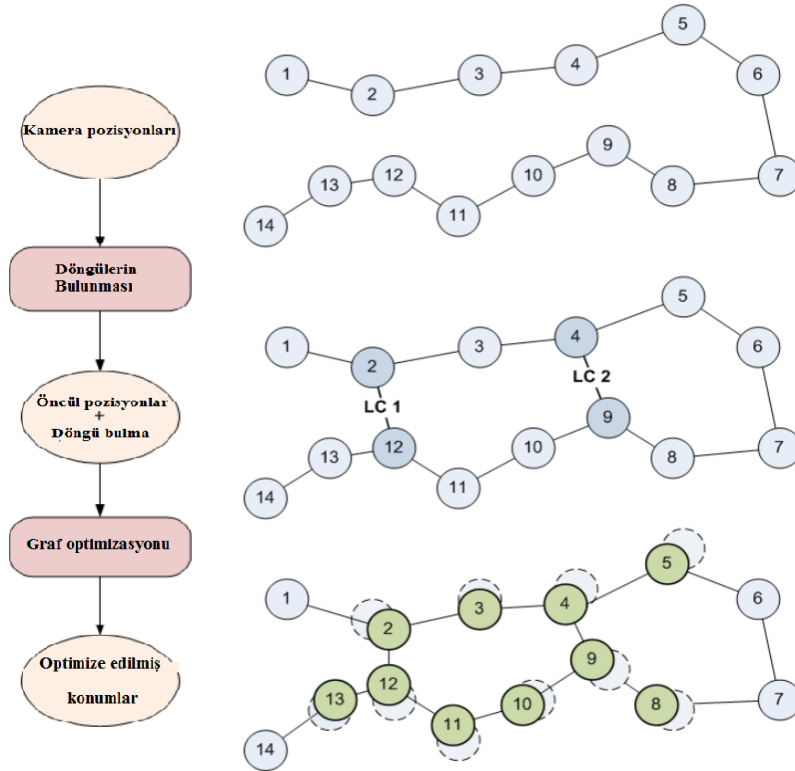
Ardışık görüntü karelerinin birbirlerine göre hizalanması ile birlikte uzun olmayan mesafeler boyunca kamera pozisyonu takip edilebilmektedir. Ancak, belirli görüntü karelerinin hizalanmasındaki hatalar, derinlik verisinin sahip olduğu gürültü gibi nedenlerle, kameranın hareketi boyunca hesaplanan kamera yörüngesi, gerçek kamera yörüngesinden sapmalar gösterir. Bu da haritalamada hatalar oluşmasına yol açar. Görüntü karelerinin hizalanması sırasında ortaya çıkan hataların birikmesi sonucu sonuç modelde aynı bölgenin iki farklı konumda birden ortaya çıkması ile karşılaşılabilir (Henry ve diğ. 2010). Bu problem, kayma (drift) problemi olarak bilinir ve eşzamanlı konum belirleme ve haritalama (SLAM) uygulamalarının ana

problemlerinden biridir. Eğer döngü bulma problemine bir çözüm üretilebilirse kameranın hareketi daha iyi modellenebilir ve kamera pozisyonları iyileştirilebilir.

Döngü bulma probleminin çözülebilmesi için seçilen yöntem, kamera daha önce ziyaret ettiği bir konuma geldiği zaman bunu algılayabilmelidir. Üretilen harita, aynı bölgenin iki farklı konumda ortaya çıkması sonucu oluşan belirsizliği önleyebilmek için düzeltilmelidir.

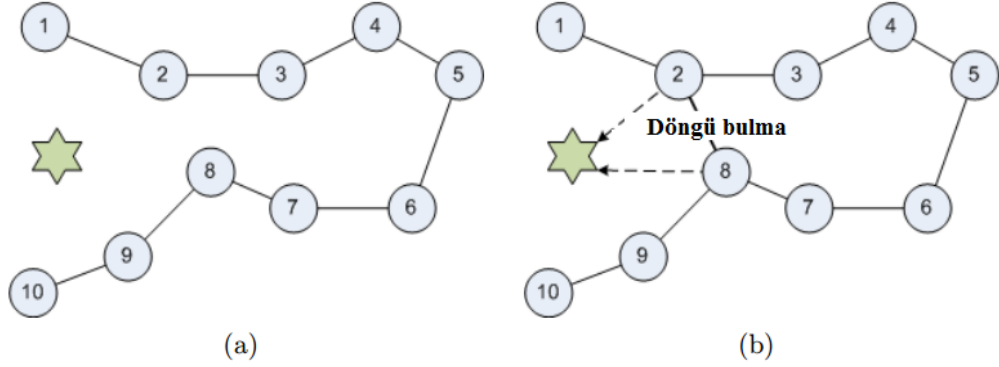
Bir çözüm önerisi, bir graf yapısı ile görüntü kareleri arasındaki koşulların belirlenmesidir. Graf yapısının kenarları görüntü kareleri arasındaki geometrik koşullara karşılık gelir. Art arda gelen görüntü karelerinin hizalanmasından elde edilen bağıl dönüşümler bazı koşullara uyar. Döngü bulma gerçekleşmeden de grafımız, doğrusal bir zincire sahip olur. Döngü bulma, görüntü kareleri arasındaki geçici olmayan koşullar olarak temsil edilebilir (Henry ve diğ. 2010).

Düğümünün, kameranın pozisyonunun durumunu gösteren değişkenler ve kenarların, düğüm çiftini birbirine bağlayan ilişkiler arasındaki ilgili koşulları temsil ettiği böyle bir graf, pozisyon grafi olarak isimlendirilir. Graf gösterimi, problemi kolayca anlaşılan ve özlü bir şekilde temsil eder. Bunun yanı sıra, düğümlerin en olası düzeninin bulunması amacıyla kullanılabilir (Şekil 5.1).



Şekil 5.1. Graf optimizasyonunun genel iş akışı

Döngü bulma (Şekil 5.2), basit bir fikre dayanır. Herhangi bir varsayımda bulunmadan robot güzergâhı izlerken ve geçmiş görüntü karesi verilerini saklar. Yeni gelen görüntü karesi geçmiş görüntülerle kıyaslanır. Eğer güncel gözlemler, hafızada tutulan eski gözlemlerle yeteri kadar benzerlik içeriyor ise bu görüntü kareleri arasında dönüşüm hesaplanabilir ve yeni bir koşul olarak eklenebilir (Högman, 2012).



Şekil 5.2. Döngü bulma: (a) Kamera pozisyonları, her biri dönüşümleri temsil eden standart kenarlar ile bağlanmış (b) Daha önce gözlemlenen bir durumla benzerlik kurulması durumunda yeni bir bağlantı oluşur.

Optimizasyon probleminin çözümü için farklı yaklaşımlar bulunmaktadır:

GraphSLAM: Olasılıksal robot teknolojilerinde (Thrun, 2002), kanı (belief) şartlı olasılık dağılımları boyunca temsil edilebilir. Kanı dağılımı, gerçek durum ile alakalı her bir olası hipoteze bir olasılık (ya da yoğunluk değeri) atar. Kanı dağılımları, durum değişkenleri üzerine soncul olasılıklardır. GraphSLAM (Thrun ve Montemerlo, 2006), bir grafa, hareket modelleri ve ölçümlerinin olasılığı tarafından tanımlanan koşulları temsil eder. Soncul haritanın üretimi, tüm koşulların dikkate alınarak optimizasyon probleminin çözülmesi ile başarılabilir.

TORO (Tree-based network optimizer) Ağaç temelli ağ en iyileyicisi (Grisetti, 2007): En olasılar (maximum likelihood-ML) haritalarını üretebilmek için ilişkiler ağına dayanan en küçük kareler tekniği uygulanır. Bu, gradyenin negatifi ile orantılı bir şekilde hatayı minimize etmek için farklı gradyenler kullanılarak yapılır. Henry ve diğ. (2010) tarafından öne sürülen 3B yoğun (dense) modelleme projesinde kullanılmıştır.

HOG-Man: Aktif 2B ve 3B (Grisetti ve diğ., 2010) haritalama için manifoldlar üzerine hiyerarşik bir optimizasyondur. Bu yaklaşım farklı seviyeleri göz önünde bulundurur. Problemin alt seviyelerde çözümü kısmen üst seviyeleri etkiler, bu nedenle daha büyük ve karmaşık haritaları işleyebilir. Bu yaklaşım RGBD-6D-SLAM (Engelhard ve diğ., 2011) projesinde kullanılmıştır.

g²o (Kümmerle ve diğ., 2011): Graf optimizasyonu için geliştirilmiş açık kaynak bir çatıdır. g²o çatısı, minimize edilmesi gereken hata fonksiyonunu hassas bir şekilde yeniden tanımlama imkânı verir ve çözümün nasıl gerçekleşeceğini belirler. Bu çatı içerisinde 3B SLAM veya demet dengelemesi gibi en genel problemlerin çözümü için bir takım sınıflar kullanıcılara sağlanmaktadır.

6. UYGULAMA

6.1 Kinect Kurulumu

Kinect v2 sensörü ile çalışmaya başlamadan önce kullanılacak olan bilgisayarın aşağıdaki minimum sistem gereksinimlerini karşılaması gereklidir.

- 64-bit (x64) processor
- Dual-core 3.1 GHz ya da daha hızlı bir işlemci
- USB 3.0 (mümkünse süper speed)
- 4B RAM
- DirectX11 desteği olan bir ekran kartı
- Windows 8 ya da 8.1 ya da Windows Embedded 8 işletim sistemi

Kinect v2 sistem gereksinimleri incelendiğinde bir önceki versiyonuna göre, sistem gereksinimleri yükseldiği söylenebilir. Ayrıca Windows 8 ya da 8.1 işletim sistemine sahip bir bilgisayarda Visual Studio 2012 veya 2013 geliştirme ortamına ihtiyaç duyması bakımından bu sensör, kullanıcıyı geliştirme platformu olarak kısıtlamaktadır.

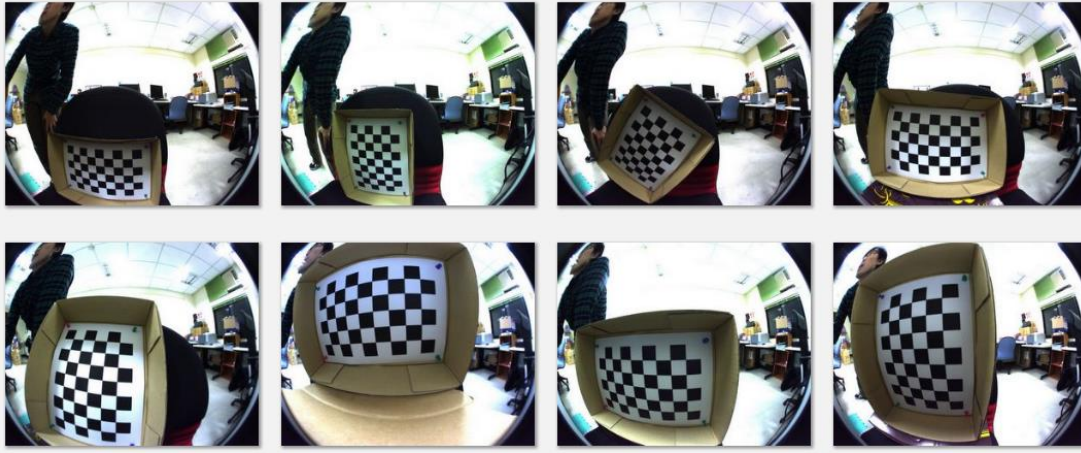
Sensörün kullanılabilmesi için Microsoft'un resmi sitesi üzerinden ücretsiz yazılım geliştirme kiti SDK 2.0 indirilmesi gereklidir. SDK kurulumu sırasında gerekli olan sürücülerin yanı sıra başlangıç olarak kullanılabilecek olan alıştırma dosyaları da gelmektedir. Bu alıştırma dosyaları bilgisayara otomatik olarak kurulan SDKBrowser 2.0 aracı ile ulaşıp tercih edilen örnek uygulamalar istenilen klasörün altına kurulabilmektedir.

Visual Studio programı sayesinde Kinect sensörü kullanılarak C++ ve C# dillerinde uygulama geliştirilebilir. C# programlama dili ile uygulama geliştirilmek istenirse bu tez kapsamında belirtilen bazı kütüphanelerin C# uyumlu versiyonları ya da wrapper kütüphaneler kullanılabilir.

6.2 Kamera Kalibrasyonu

Kamera kalibrasyonu, 3B nokta bulutlarının ve renk bilgisinin hassasiyetinin artırılması ve tam bir modelin üretilmesi için önemli bir konudur. Doğru sonuçlar elde edilebilmesi için RGB ve derinlik kameralarının geometrik olarak kalibre edilmesi gereklidir.

Kalibrasyon işlemi için satranç tahtası şeklinde bir düzlem kullanılır. Satranç tahtasının farklı açılardan görüntüleri çekilir (Şekil 6.1).



Şekil 6.1. Satranç tahtasının farklı açılardan resimlerinin çekilmesi

Renkli görüntülerde satranç tahtasının köşeleri çıkarılır. Derinlik görüntülerinde ise satranç tahtasının siyah beyaz deseni görünmediği için satranç tahtasının dikdörtgen şeklinin köşeleri bu iş için idealdir. Standart kalibrasyon algoritması (Zhang, 1998) kullanılır. Bu çalışmada kamera kalibrasyon uygulaması için açık kaynak kodlu bilgisayarla görme kütüphanesini OpenCV, nin sağlamış olduğu metotlar önerilmiştir.

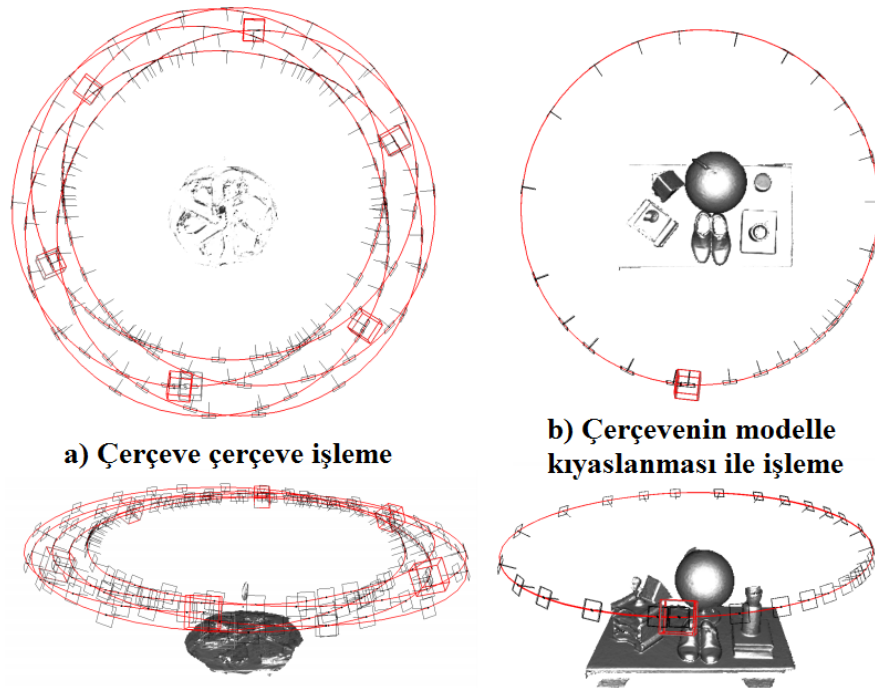
6.3 Öznitelik Çıkarımı

Görüntü ve derinlik verilerinden öznitelik çıkarımı çok önemli bir konudur. Bu tez çalışmasında 4. bölümde incelenen öznitelik çıkarımı yöntemlerinden, çok kullanılan yöntemler olan SIFT ve SURF öznitelikleri tercih edilmiştir. SURF özniteliği gerçek zamanlı uygulamalarda SIFT yöntemine göre daha hızlıdır. NARF özniteliği ve 3B nokta bulutundan hesaplanabilecek olan öznitelikler Kinect derinlik verileri görüntülerle yüklü olduğu için tercih edilmemiştir. Kinect v2 ‘nin çıkması ile birlikte son dönemde nokta bulutu kütüphanesinin 3B yerel betimleyicilerinin de kullanıldığı

alışmalar olduėu bilinmektedir. Bu tez kapsamında nerilen yntem, bu alanda yapılan alışmalarda daha ok kullanılmaktadır.

6.4 Dnřm Parametrelerinin Belirlenmesi

Bu alışmada elde edilmek istenen 3B model tek bir Kinect sensr ile modellenmek istenmektedir. Herhangi bir uzmanlıėı olmasını beklemediėimiz kullanıcı bu sensr i meknda serbeste hareket ettirerek bu iřlemi gerekleřtirebilmelidir. Bu iřlem gerek zamanlı gerekleřtirilmek istenirse sensrn gnderdiėi veriler ereve ereve iřlenmeye alışılır. Her bir ereve iftinin hizalanması sırasında ortaya ıkan kk hataların birikerek, kayma problemi denilen bir etkiye neden olmaktadır. Burada anlatılan modelde bu etkiyi biraz azaltabilmek iin grnt akıřından elde edilen her grnt karesi bir nceki grnt karesi yerine seilen bir anahtar grnt karesi ile deėerlendirilir (řekil 6.2). Bu anahtar grnt karesi, kameranın veya bir SLAM uygulaması sz konusu ise robotun belli miktarda hareket etmesi ya da gncel grnt karesinin anahtar grnt karesi ile benzerliklerinin azalması sonucunda gncellenir. Ya da Newcombe ve diė. (2010) yaptıkları gibi anahtar grnt karesi belli sayıda grnt karesi sonucunda gncellenebilir, rneėin 8 karede bir gibi.



řekil 6.2. erevenin ereve ile ve erevenin modelle iřlenmesi: erevenin bir anahtar ereve ile birlikte iřlenmesi kayma etkisini azaltmaktadır.

Dönüşüm parametrelerinin belirlenmesi için en küçük kareler dengelemesi kullanılabilir (Arun ve diğ., 1987). En az üç eşleşme minimum çözümün gerçekleştirilmesi için yeterlidir. Daha fazla nokta daha iyi sonuçlar üretilmesini sağlar. Fakat şu an için düşük maliyetli derinlik sensörleri ile gerçekleştirilen bu tür çalışmalarda eşleştirme işleminin belirsizliği ve tek bir yanlış eşleşmenin bile sonuç dönüşüm parametrelerinde önemli derecede bir hataya yol açabilmesi nedeniyle uyuşumsuz verileri elemine etmek için RANSAC algoritması kullanılır. Bu algoritmada öncelikle üç eşleşme çifti seçilir ve dönüşüm bu çiftler üzerinden hesaplanır. Daha sonra tüm eşleşme çiftlerine dönüşüm uygulanır ve öklid uzaklıkları belirlenen eşik değerinden fazla olan eşleşme çiftleri elenerek uyumlu veri kümesi elde edilmeye çalışılır. Bu küme ile daha özlü bir dönüşüm elde edilir. Bu işlem hatalar azaltılıncaya kadar iteratif olarak devam ettirilir. RANSAC metodu algoritması;

ALGORİTMA :Değişmez T dönüşümünün bulunması

Gereksinim : İki 3B nokta kümesi S ve S' olmak üzere

```

1: T, eniyiUyumluVeriler  $\leftarrow \emptyset$ 
2: for iterasyon = 0 to N do
3:   örnekle  $\leftarrow$  S ve S' kümelerinden 3 nokta çiftini rasgele seç
4:   örneklerden şuan kiDönüşüm 'ü hesapla
5:   uyumluVeriler  $\leftarrow \emptyset$ 
6:   kareselOrtalamaHata  $\leftarrow 0$ 
7:   for all  $s_i \in S$  do
8:     hata  $\leftarrow \|\text{şuan kiDönüşüm} \cdot s_i - s'_i\|^2$ 
9:     if hata < eşikDeğer then
10:      uyumluVeriler  $\leftarrow$  uyumluVeriler +  $(s_i, s'_i)$ 
11:      kareselOrtalamaHata  $\leftarrow$  kareselOrtalamaHata + hata
12:     end if
13:   end for
14:   kareselOrtalamaHata  $\leftarrow \sqrt{\text{kareselOrtalamaHata} / |\text{uyumluVeriler}|}$ 
15:   if  $|\text{uyumluVeriler}| > \text{uyumluVerilerEşikDeğer}$  then
16:     uyumluVerilerden şuan kiDönüşüm 'ü tekrar hesapla
17:     if şuan kiDönüşüm, T 'den daha iyi ise then
18:       T  $\leftarrow$  şuan kiDönüşüm
19:       eniyiUyumluVeriler  $\leftarrow$  uyumluVeriler
20:     end if
21:   end if
22: end for
23: return T, eniyiUyumluVeriler

```

RANSAC metodundan belirlenen dönüşümler, kamera pozisyonları arasındaki ilişkiyi biçimlendirir. Düşümler, görüntü kareleri yakalandığı andaki kamera pozisyonlarını ve kenarlar, görüntü karesi çiftleri arasındaki dönüşümleri göstermek üzere bir pozisyon grafi oluşturulur. Bu graf, kameranın yörüngesini göstermektedir. Bu grafin büyüklüğünü sınırlı tutmak için aşağıda belirtilen şartlar altında grafa yeni bir düğüm eklenir:

- Referans çerçeve ile güncel çerçeve arasındaki dönüşüm belli bir eşik değerden büyük olduğu zaman. Quang ve Quoc (2015) öteleme için 0.2m ve dönüklük için 15° olarak belirlemişlerdir.
- Uyumlu verilerin sayısı belli bir eşik değerinin altına düşerse, yeterli eşleşmenin sağlanamaması durumunu engellemek için.

Kamera daha önce bulunduğu bir konuma geldiğinde graf üzerinde döngü kapanması oluşur. Döngü kapanması önceki çerçeve ile şu anki çerçeve arasındaki dönüşümün hesaplanması ile bulunur. Grafin içerdiği hata yüzünden kameranın tüm yörüngesinin tutarlılığı korunamaz. Bu nedenle pozisyon grafinin optimize edilmesi gerekir (Quang ve Quoc, 2015). Graf optimizasyonu için aşağıdaki doğrusal maliyet fonksiyonunu minimize eden g^2 o (Kümmerle ve diğ., 2011; Högman, 2012) çatısı kullanılır.

$$F(P) = \sum_{T_{i,j} \in E} \|T_{i,j}P_i - P_j\|^2 \quad (6.1)$$

burada P_i , P_j kamera pozisyonlarını, T_{ij} hesaplanan dönüşümü ve E de graftaki kenarlar kümesini temsil etmektedir.



a) Graf optimizasyonu öncesi 3 boyutlu model

b) Graf optimizasyonu sonrası 3 boyutlu model

Şekil 6.3. Graf optimizasyonu sonucunda kamera pozisyonları daha iyi belirlenir ve kameranin pozisyonlarına bağlı olarak dönüşümler tekrar ayarlanır.

6.5 Nokta Bulutunun İşlenmesi

Çevrenin modelinin üretilebilmesi için sistem tüm nokta bulutlarını 3B koordinat sistemine aktarmalıdır. Bir RGB-D kameradan her bir görüntü karesi için yaklaşık 250,000 nokta civarında bir veri elde edebilir (Henry ve diğ., 2010). Bu açıdan bakıldığında tüm nokta bulutlarının bir araya getirilmesi, çok miktarda belleğe ihtiyaç duyar. Tüm nokta bulutları birbirleriyle bir araya getirilmeden önce nokta bulutlarının boyutu düşürülmelidir (Quang ve Quoc, 2015).

Bunun için literatürde kullanılan bir yaklaşım yüzey elemanları (surfels) (Pfister ve diğ., 2000; Krainin ve diğ., 2011) kullanılabilir. Bir yüzey elemanı, konum, yüzey yönlendirmesi, yama boyutları ve renk bilgisinden oluşmaktadır. Yüzeye elemanları, zamanla farklı açılardan görme ile birlikte artan bir güven ölçütü barındırır. Bu ölçüt, yüzey elemanlarının normallerine bağlı yönlerin oluşturduğu 2B yönlerin histogramı şeklinde temsil edilir. Düşük güvene sahip olan yüzey elemanı silinir.

Her bir RGB-D çerçevesine belirlenen normallere bağlı olarak yüzey elemanlarının normalleri güncellenebilir. Yüzey elemanının rengi, o yüzey elemanının normali ile en çok hizalanan RGB-D çerçevesinden alınır.

Uygulamanın son aşamasında üretilen 3B modeller OpenGL ve Nokta Bulutu kütüphanesinin (PCL) görüntüleme ve kayıt metotlarından yararlanılarak elde edilen

veriler görüntülenmesi sağlanmış ve diğer 3B modelleme yazılımları ve nokta bulutu görüntüleme yazılımlarında açılacak şekilde dosyaya kaydedilmiştir.

7. SONUÇLAR VE ÖNERİLER

Günümüzde lazer tarama sistemleri istenilen doğrulukta 3B yoğun nokta bulutları sağlamaktadır. Buna karşın bina iç ortamlarında bu tür tarama sistemlerinin kullanılması hem çok pahalı hem de veri toplama geometrisinin bina içi ortamlarında çalışmak için uygun olmaması nedeni ile sorunlar ortaya çıkabilmektedir. Pratik ve daha küçük ölçekte yapılan çalışmalar için fazla maliyetli çözümler getirmektedir. Ayrıca bu tarayıcıların bazılarının kullanımı ve yazılımları eğitilmiş bir kullanıcı gerektirebilmektedir.

Diğer yandan bilgisayar donanım ve yazılım teknolojilerindeki gelişmelerinin yanı sıra Sensör teknolojilerinin de gelişmesiyle beraber özellikle derinlik verisi üreten sensörler ve RGB-D kameralarla yapılan uygulamalar oldukça popüler hale gelmiştir. Bu çalışmada düşük maliyetli bir RGB-D kamera olan Microsoft Kinect kamerası kullanılarak bina iç mekânlarının 3B modellenmesi üzerine örnek bir sistemin tasarımı üzerinde durulmuştur. Bu nedenle Kinect sensörü kullanarak yapılan çalışmalarda genel işlem adımlarından bahsedilmiştir. KinectFusion (Izadi ve diğ. 2011) çalışmasından ayrı olarak diğer çalışmalarda tercih edilen seyrek öznitelik eşleştirmesine dayalı bir yaklaşım benimsenmiştir. Yoğun nokta bulutu eşleştirmesine dayalı bir yöntemin getireceği hesap yükünden dolayı bu çalışmada tercih edilmemiştir.

Gerçekleştirilen sistem çalıştırıldığı anda profesyonel eğitimi olmayan sıradan bir kullanıcının kamerayı istediği iç mekânın içinde serbestçe gezdirerek o ortamın 3B modelini oluşturmasını sağlayacak ve daha sonra bu modeli herhangi bir 3B model düzenleme yazılımında düzenlenecek şekilde kayıt edebilmesi amaçlanmıştır. Bu aşamadan bakıldığında sistemin geliştirmesi gereken birçok yönü bulunmaktadır. Kullanılan metotların ve algoritmaların her koşulda aynı sonucu vermemesi, farklı mekânlar için ekstra işlem adımları uygulanması gerektiği görülmüştür.

İleriki çalışmalarda bu çalışmalardaki kullanılan algoritmaların matematik modelleri üzerinde durularak özgün yaklaşımlar üretilmeye çalışılacaktır. Yeni çıkan çalışmalardan uygun görülen ama şu an için bu sistemde kullanılmayan metot ve

algoritmaların çalışmamıza eklenmesi düşünülmektedir. Sistemin gerçek zamanlı performansı geliştirilecektir. Bu amaçla grafik işlemcinin gücünden de yararlanılarak paralel programlama düşünülmektedir.

Geliştirilen sistemle elde edilen 3B modellerin kullanılabileceği alanların ileride daha gelişmesi ve yapılan çalışmaların kapsamlarının artacağı düşünülmektedir.

8. KAYNAKLAR

- Aldoma, A., Marton, Z. C., Tombari, F., Wohlking, W., Potthast, C., Zeisl, B., Rusu, R. B., Gedikli, S., Vincze, M., 2012. Point cloud library. *IEEE Robotics & Automation Magazine*, 1070(9932/12).
- Alpaslan, N., 2013. Gradyan Tabanlı Heterojen Öznitelik Çıkarma Yöntemlerine Yeni Yaklaşımlar, Yüksek Lisans Tezi, İnönü Üniversitesi, Fen Bilimleri Enstitüsü, Malatya, 341555.
- Bay, H., Tuytelaars, T., Van Gool, L., 2006. Surf: Speeded up robust features. In *Computer vision–ECCV 2006*, Springer Berlin Heidelberg, 404-417.
- Belongie, S., Malik, J., Puzicha, J., 2002. Shape matching and object recognition using shape contexts, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 4(2), 509-522.
- Besl, P. J., McKay, N. D., 1992. Method for registration of 3-D shapes. In *Robotics-DL tentative*, *International Society for Optics and Photonics*, Nisan, 586-606
- Bethencourt, A., Jaulin, L., 2012. 3D Reconstruction using interval methods on the kinect device coupled with an IMU, *International Journal of Advanced Robotic Systems*, xxx.
- Brahmbhatt, S., 2013. *Practical OpenCV*, 1 ed., Apress Berkely, CA, USA ISBN13: 978-1-4302-6079-0
- Bradley, D., Heidrich, W., Popa, T., Sheffer, A., 2010. High resolution passive facial performance capture. *ACM Transactions on Graphics*, 29(4), 41.
- Calonder, M., Lepetit, V., Strecha, C., Fua, P., 2010. Brief: Binary robust independent elementary features. *Computer Vision*, 778-792.
- Cui, Y., Schuon, S., Chan, D., Thrun, S., Theobalt, C., 2010. 3D shape scanning with a time-of-flight camera. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on IEEE*, Haziran, 1173-1180.
- Endres, F., Hess, J., Engelhard, N., Sturm, J., Cremers, D., Burgard, W., 2012. An evaluation of the RGB-D SLAM system. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on IEEE*, Mayıs, 1691-1696.
- Engelhard, N., Endres, F., Hess, J., Sturm, J., Burgard, W., 2011. Real-time 3D visual SLAM with a hand-held RGB-D camera. *RGB-D Workshop on 3D Perception in Robotics at the European Robotics Forum*, Vasteras, Sweden (Vol. 180), Nisan.
- Feng, H. Y., Liu, Y., Xi, F., 2001. Analysis of digitizing errors of a laser scanning system. *Precision Engineering*, 25(3), 185-191.
- Grisetti, G., Stachniss, C., Grzonka, S., Burgard, W., 2007. A Tree parameterization for efficiently computing maximum likelihood maps using gradient descent. In *Robotics: Science and Systems*.
- Grisetti, G., Kummerle, R., Stachniss, C., Frese, U., Hertzberg, C., 2010. Hierarchical optimization on manifolds for online 2D and 3D mapping. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on IEEE*, 273-278.
- Harris, C., Stephens, M., 1988. A combined corner and edge detector. *Alvey vision conference*, Vol. 15, 50.

- Heikkilä J., Silven O., 1997. A four-step camera calibration procedure with implicit image correction, *CVPR*.
- Henry, P., Krainin, M., Herbst, E., Ren, X., Fox, D., 2014. RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments. *Experimental robotics*, Springer Berlin Heidelberg, Ocak, 447-491.
- Högman, V., 2012. Building a 3D map from RGB-D sensors, Master's thesis, Royal Institute of Technology (KTH), Computer Vision and Active Perception Laboratory, Stockholm, Sweden
- Huang, Y. Y., Chen, M. Y., 2011. 3D object model recovery from 2D images utilizing corner detection. *System Science and Engineering (ICSSE)*, 2011 International Conference on IEEE, 76-81.
- Iddan, G. J., Yahav, G., 2001. Three-dimensional imaging in the studio and elsewhere. *Photonics West 2001-Electronic Imaging*, International Society for Optics and Photonics, 48-55.
- Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A., Fitzgibbon, A., 2011. KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera. *24th annual ACM symposium on User interface software and technology*, 559-568.
- Johnson, A., 1997. Spin-Images: A Representation for 3-D Surface Matching. PhD Thesis, Carnegie University, Robotics Institute, Pittsburgh, PA.
- Kim, Y. M., Theobalt, C., Diebel, J., Kosecka, J., Matusik, B., Thrun, S., 2009. Multi-view image and tof sensor fusion for dense 3d reconstruction, *Computer Vision Workshops (ICCV Workshops)*, 2009 IEEE 12th International Conference on IEEE, 1542-1549.
- Krainin, M., Henry, P., Ren, X., Fox, D., 2011. Manipulator and object tracking for in-hand 3d object modeling. *The International Journal of Robotics Research*, 30(11), 1311-1327.
- Kümmerle, R., Grisetti, G., Strasdat, H., Konolige, K., Burgard, W., 2011. g²o : A general framework for graph optimization, *Robotics and Automation (ICRA)*, 2011 IEEE International Conference on IEEE, 3607-3613.
- Lachat, E., Macher, H., Mittet, M. A., Landes, T., Grussenmeyer, P., 2015. First experiences with kinect V2 sensor for close range 3D modelling, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS)*.
- Lange, R., 2000. 3D Time-Of-Flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology, PhD Thesis, University of Siegen. Siegen, Germany.
- Lazebnik, S., Schmid, C., Ponce, J., 2005. A Sparse texture representation using local affine regions, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 27(8), 1265-1278.
- Lindeberg, T., 1994. Scale-space theory in computer vision, Kluwer/Springer, Boston.
- Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, 60, 91-110.
- Meilland, M., Comport, A. I., Rives, P., 2011. Real-time dense visual tracking under large lighting variations. *British Machine Vision Conference*. Dundee, Scotland.
- Meilland, M., Rives, P., Comport, A. I., 2012. Dense RGB-D mapping for real-time localisation and navigation, *IV12 Workshop on Navigation Positioning and Mapping*, Alcalá de Henares, Spain.

- Meister, S., Izadi, S., Kohli, P., Hämmerle, M., Rother, C., Kondermann, D., 2012. When can we use KinectFusion for ground truth acquisition?, *Workshop on Color-Depth Camera Fusion in Robotics* (Vol. 2).
- Mutto, C. D., Zanuttigh, P., Cortelazzo, G. M., 2012. *Time-of-flight cameras and microsoft kinect (TM)*, Springer Publishing Company, Incorporated.
- Nandanavanam, M., P., 2011. Hauler detection for an autonomous wheel loader, Masters' Thesis, Department of Technology, Örebro University, Örebro, Sweden
- Newcombe, R. A., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A. J., Kohi, P., Shotton, J., Hodges, S., Fitzgibbon, A., 2011. KinectFusion: Real-time dense surface mapping and tracking. *Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on IEEE*, 127-136.
- Pfister, H., Zwicker, M., Van Baar, J., Gross, M., 2000. Surfels: Surface elements as rendering primitives. *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, ACM Press/Addison-Wesley Publishing Co., 335-342.
- Quang, H. P., Quoc, N. L., 2015. Some improvements in the RGB-D SLAM system. *Computing & Communication Technologies-Research, Innovation, and Vision for the Future (RIVF), 2015 IEEE RIVF International Conference on IEEE*, 112-116.
- Rosin, P. L., 1999. Measuring corner properties. *Computer Vision and Image Understanding*, 73(2), 291-307.
- Rosten, E., Drummond, T., 2006. Machine learning for high-speed corner detection. *Computer Vision-ECCV 2006*, Springer Berlin Heidelberg, 430-443.
- Roth, H., Vona, M., 2012. Moving volume KinectFusion. *BMVC*, September, 1-11.
- Rublee, E., Rabaud, V., Konolige, K., Bradski, G., 2011. ORB: an efficient alternative to SIFT or SURF. *Computer Vision (ICCV), 2011 IEEE International Conference on IEEE*, 2564-2571
- Rusinkiewicz, S., Levoy, M., 2001. Efficient variants of the ICP algorithm. *3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on IEEE*, 145-152.
- Sansoni, G., Trebeschi, M., Docchio, F., 2009. State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation, *Sensors*, 9(1), 568-601.
- Seitz, S. M., Curless, B., Diebel, J., Scharstein, D., Szeliski, R., 2006. A comparison and evaluation of multi-view stereo reconstruction algorithms. *Computer vision and pattern recognition, 2006 IEEE Computer Society Conference on IEEE*, Vol. 1, 519-528.
- Steder, B., Rusu, R. B., Konolige, K., Burgard, W., 2010. NARF: 3D range image features for object recognition. *Workshop on Defining and Solving Realistic Perception Problems in Personal Robotics at the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* (Vol. 44).
- Steger, C., Ulrich, M., Wiedemann, C., 2007. *Machine Vision Algorithms and Applications*. Wiley-VCH.
- Stückler, J., Behnke, S., 2012. Integrating depth and color cues for dense multi-resolution scene mapping using rgb-d cameras. *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2012 IEEE Conference on IEEE*, 162-167.
- Szeliski R., 2010. *Computer Vision: Algorithms and Applications*. New York: Springer.

- Thrun, S., 2002. Probabilistic robotics, *Communications of the ACM*, 45(3), 52-57.
- Thrun, S., Montemerlo, M., 2006. The graph SLAM algorithm with applications to large-scale mapping of urban structures, *The International Journal of Robotics Research*, 25(5-6), 403-429.
- Wang, K., Zhang, G., Bao, H., 2014. Robust 3D Reconstruction with an RGB-D camera. *Image Processing, IEEE Transactions on*, 23(11), 4893-4906.
- Whelan, T., Kaess, M., Fallon, M., Johannsson, H., Leonard, J., McDonald, J., 2012. Kintinuous: Spatially extended kinectfusion.
- Vouzounaras, G., Perez-Moneo Agapito, J. D., Daras, P., Stryntzis, M. G., 2010. 3D reconstruction of indoor and outdoor building scenes from a single image, *Proceedings of the 2010 ACM workshop on Surreal media and virtual cloning ACM*, 63-66.
- Zhang, G., Jia, J., Wong, T. T., Bao, H., 2009. Consistent depth maps recovery from a video sequence, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(6), 974-988.
- Zhang, Y., Luo, C., Liu, J., 2012. Walk&sketch: create floor plans with an rgb-d camera, *2012 ACM Conference on Ubiquitous Computing ACM*, 461-470.
- Zhang, Z., 2000. A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(11), 1330-1334.
- Zheng, H., Yuan, J., Gu, R., 2011. A novel method for 3D reconstruction on uncalibrated images. *Third International Conference on Internet Multimedia Computing and Service ACM*, 138-141.
- URL-1: https://en.wikipedia.org/wiki/Structured-light_3D_scanner, (Ziyaret tarihi: 1 Haziran 2015)
- URL-2: <https://en.wikipedia.org/wiki/Interferometry>, (Ziyaret tarihi: 1 Haziran 2015)
- URL-3: <http://smeenk.com/kinect-field-of-view-comparison/>, (Ziyaret tarihi: 1 Haziran 2015)
- URL-4: <http://www.aishack.in/tutorials/sift-scale-invariant-feature-transform-introduction/>, (Ziyaret tarihi: 1 Haziran 2015)
- URL-5: http://194.27.49.11/makine/ikaymaz/optimizasyon/dosyalar/optimizasyon_bolum3_2009.pdf, (Ziyaret tarihi: 1 Haziran 2015)
- URL-6: <https://gilscvblog.wordpress.com/2013/09/19/a-tutorial-on-binary-descriptors-part-2-the-brief-descriptor/>, (Ziyaret tarihi: 1 Haziran 2015)
- URL-7: http://rpg.ifi.uzh.ch/docs/teaching/05_feature_detection_2.pdf, (Ziyaret tarihi: 1 Haziran 2015)
- URL-8: <http://pointclouds.org/>, (Ziyaret tarihi: 1 Haziran 2015)
- URL-9: [http://robotica.unileon.es/mediawiki/index.php/PCL/OpenNI_tutorial_4:_3D_object_recognition_\(descriptors\)](http://robotica.unileon.es/mediawiki/index.php/PCL/OpenNI_tutorial_4:_3D_object_recognition_(descriptors)), (Ziyaret tarihi: 1 Haziran 2015).

ÖZGEÇMİŞ

Adı Soyadı : Nihat KARAAHMETOĞLU

Doğum Yeri ve Tarihi : GİRESUN 28.12.1983

Adres : Mimar Sinan Mah. 184. Sok. No:23/4
ATAKUM/SAMSUN

E-Posta : nihat.karaahmetoglu@omu.edu.tr

Lisans : Ondokuz Mayıs Üniversitesi
Harita Mühendisliği Bölümü 2011

Yüksek Lisans : -

Mesleki Deneyim ve Ödüller : Araş. Gör. – Ondokuz Mayıs Üniversitesi
Mühendislik Fakültesi Harita Mühendisliği
2012-...

Yayın ve Patent Listesi : -