



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

KIZILÖTESİ DERİNLİK KAMERASI İLE 3 BOYUTLU
YENİDEN MODELLEME

BERKAN DEMİRCİ

Temmuz 2019

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

KIZILÖTESİ DERİNLİK KAMERASI İLE 3 BOYUTLU
YENİDEN MODELLEME

BERKAN DEMİRCİ

Yüksek Lisans Tezi

Danışman

Dr. Öğr. Üyesi Murat PEKER

Temmuz 2019

Berkan DEMİRCİ tarafından **Dr. Öğr. Üyesi Murat PEKER** danışmanlığında hazırlanan “**KIZILÖTESİ DERİNLİK KAMERASI İLE 3 BOYUTLU YENİDEN MODELLEME**” adlı bu çalışma jürimiz tarafından Niğde Ömer Halisdemir Üniversitesi Fen Bilimleri Enstitüsü **Elektrik-Elektronik Mühendisliği** Ana Bilim Dalı’nda Yüksek Lisans tezi olarak kabul edilmiştir.

Başkan : Dr. Öğr. Üyesi Mehmet Çağrı SOYLU, Kayseri Erciyes Üniversitesi



Üye : Dr. Öğr. Üyesi Murat PEKER, Niğde Ömer Halisdemir Üniversitesi



Üye : Dr. Öğr. Üyesi Engin Cemal MENGÜÇ, Niğde Ömer Halisdemir Üniversitesi



ONAY:

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

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

Doç. Dr. Murat BARUT
MÜDÜR

TEZ BİLDİRİMİ

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

Berkan DEMİRCİ



ÖZET

KIZILÖTESİ DERİNLİK KAMERASI İLE 3 BOYUTLU YENİDEN MODELLEME

DEMİRCİ, Berkan

Niğde Ömer Halisdemir Üniversitesi

Fen Bilimleri Enstitüsü

Elektrik-Elektronik Mühendisliği Ana Bilim Dalı

Danışman

: Dr. Öğr. Üyesi Murat PEKER

Temmuz 2019, 38 Sayfa

Gelişen ve ucuzlayan teknoloji ile birlikte ortaya çıkan Kinect gibi kırmızı yeşil mavi-derinlik (Red Green Blue-Depth, RGB-D) kameralar üç-boyutlu (3B) yeniden oluşturma konusunda yapılan çalışmaların hız kazanmasını sağlamıştır. Kızılötesi algılayıcılar ile derinlik bilgisini görüntü ile birleştiren bu kameralar ile nesnenin farklı açılarından çekilen görüntüleri nesnelerin 3B yeniden oluşturmalarını mümkün kılmaktadır. 3B yeniden oluşturma ile nesnelerin yeniden oluşturulması, ortam haritalama, robotik, arkeoloji, biomedikal, oyun endüstrisi ve sanal gerçeklik gibi bir çok uygulama alanına sahiptir. Bu tez çalışması kapsamında Kinect V1 kullanılarak elde edilmiş olan RGB-D veri kümesi üzerinde; SIFT, SURF ve ORB görüntü eşleştirme algoritmalarının 3B yeniden oluşturma performansı karşılaştırılmıştır. Kinect algılayıcısından alınan derinlik verisi oldukça falza gürültü içerdiğinden ve nokta bulutunda aykırı bireyler oluşturduğundan RANSAC algoritması aykırı noktaların tespiti ve çıkarılması aşamasında eşleştirme algoritmaları ile birlikte kullanılmıştır. ICP algoritması oluşturulmakta olan modele eklenen her yeni görüntünün düzgün ve hassas biçimde hizalanmasının sağlanabilmesi için kullanılmıştır.

Anahtar Sözcükler: Üç boyutlu nesne modelleme, Üç boyutlu yeniden oluşturma, Öznitelik tanımlayıcılar, yinelemeli en yakın nokta, küresel eniyileme, nokta bulutu hizalama.

SUMMARY

3D RECONSTRUCTION WITH INFRARED DEPTH CAMERA

DEMİRCİ, Berkan

Nigde Omer Halisdemir University

Graduate School of Natural and Applied Sciences

Department of Electrical and Electronics Engineering

Supervisor : Assistant Professor Dr. Murat PEKER

July 2019, 38 Pages

Red Green Blue-Depth (RGB-D) cameras such as Kinect, which emerged with developing and inexpensive technology, accelerated the studies on three-dimensional (3D) reconstruction. Combining depth information with images with infrared sensors, these cameras enable images to be reproduced by 3D objects from different angles of the object. It has many applications such as 3D reconstruction, object reconstruction, media mapping, robotics, archeology, biomedical, game industry and virtual reality. Within the scope of this thesis, RGB-D dataset which was obtained by using Kinect V1; The 3D reconstruction performance of SIFT, SURF and ORB image matching algorithms were compared. Since the depth data obtained from the Kinect sensor contains very high noise and creates contradictory individuals in the point cloud, the RANSAC algorithm has been used in conjunction with matching algorithms in the detection and removal of contradictory points. The ICP algorithm is used to ensure that every new image added to the model being created is aligned accurately and precisely.

Keywords: Three-dimensional object modeling, three-dimensional reconstruction, feature descriptors, iterative closest point, global optimization, point cloud alignment.

ÖN SÖZ

Yüksek lisans tez çalışmamın yürütülmesi esnasında, çalışmalarına yön veren, bilgi ve tecrübesini esirgemeyen, sürekli araştırmaya teşvik eden, beraber çalışmaktan büyük mutluluk duyduğum ve bana her türlü desteği sağlayan danışmanım, Sayın Dr. Öğr. Üyesi Murat PEKER'e en içten teşekkürlerimi sunarım.

Bu tezi, sadece bu çalışmam boyunca değil, tüm öğrenim hayatım boyunca maddi ve manevi her türlü desteğini benden esirgemeyen, ilgisini her zaman yanımda hissettiğim kıymetli aileme ithaf ediyorum.

İÇİNDEKİLER

ÖZET	iv
SUMMARY	v
ÖN SÖZ	vi
İÇİNDEKİLER	vii
ŞEKİLLER DİZİNİ	ix
SİMGE VE KISALTMALAR	xi
BÖLÜM I GİRİŞ	1
1.1 Literatür Özeti	4
BÖLÜM II DERİNLİK BİLGİSİ KULLANARAK NESNE YENİDEN OLUŞTURMA.....	9
2.1 Giriş	9
2.2 Derinlik Bilgisi Elde Etme Yöntemleri.....	9
2.2.1 Stereo görüntü.....	10
2.2.2 Lazer üçgenleme (Laser triangulation)	12
2.2.3 Yapılandırılmış ışık kaynağı (Structured light)	12
2.2.4 Uçuş süresi (Time of Flight - ToF) kameralar	13
2.3 RGB-D	14
2.3.1 Kinect V1	14
2.3.2 Kinect V2	16
2.4 3-Boyutlu Yeniden Oluşturma.....	17
2.4.1 Öznitelik çıkarma ve eşleştirme yöntemleri	18
2.4.1.1 SIFT.....	18
2.4.1.2 SURF.....	21
2.4.1.3 ORB.....	22
2.4.2 Yinelemeli en yakın nokta (ICP)	22
2.4.3 Küresel eniyileme	23

BÖLÜM III BENZETİM ÇALIŞMASI ve SONUÇLAR	25
3.1 Veri Kümesi	25
3.2 Benzetim Sonuçları	27
BÖLÜM IV SONUÇLAR	32
KAYNAKLAR	33
ÖZGEÇMİŞ	38

ŞEKİLLER DİZİNİ

Şekil 1.1. 3B yeniden oluşturma akış şeması.	2
Şekil 2.1. Farklı 3B Aktif ve Pasif sensörlerinin gösterimi.....	10
Şekil 2.2. Stereo kamera modeli	11
Şekil 2.3. Lazer üçgenleme modeli	12
Şekil 2.4. Yapılandırılmış ışık kaynağı (Tarin, 1999)	13
Şekil 2.5. Uçuş süresi kamerası çalışma prensibi	14
Şekil 2.6. Kinect algılayıcı yapıtaşları	15
Şekil 2.7. Kinect kızılötesi ışık yapısı	15
Şekil 2.8. Kinect – Üçgenleme	16
Şekil 2.9. 3B yeniden oluşturma akış şeması	17
Şekil 2.10. Gauss farkları işlemi.....	19
Şekil 2.11. Uç noktaların tespiti	20
Şekil 2.12. 3B nesne yeniden oluşturma birleştirme hatası.....	23
Şekil 3.1. Washington RGB-D veri kümesi örnek nesneleri (Lai vd., 2011).....	25
Şekil 3.2. Washington RGB-D veri kümesi: tam çözünürlük sahne görüntüsü (a), nesnenin sahneden kesilmiş görüntüsü (b), tam çözünürlük derinlik görüntüsü (c), nesnenin sahneden kesilmiş derinlik görüntüsü (d), tam çözünürlük görüntü için nesne maskesi (e) ve nesnenin sahneden kesilmiş görüntüsü için nesne maskesi (f)	26
Şekil 3.3. Washington RGB-D veri kümesinde sadece görüntü çiftleri kullanılarak elde edilen eşleştirme sayıları	27
Şekil 3.4. Washington RGB-D veri kümesinde sadece görüntü çiftleri kullanılarak elde edilen eşleştirmelerin RANSAC ile azaltılması sonucu elde edilen yeni eşleştirme sayıları.....	28
Şekil 3.5. Küresel eniyileme olmadan elde edilen 3B nesne modeli SIFT+RANSAC (a), SURF+RANSAC (b) ve ORB+RANSAC (c).....	29
Şekil 3.6. Nesnenin ön yüzüne ait yakın gösterim görüntüler SIFT+RANSAC (a), SURF+RANSAC (b) ve ORB+RANSAC (c)	30
Şekil 3.7. Küresel eniyileme ile elde edilen 3B nesne modeli SIFT (a), SURF (b) ve ORB (c).....	30

Şekil 3.8. Küresel eniyileme ile elde edilen 3B nesne modeli SIFT (a), SURF (b) ve ORB (c).....	31
--	----

SİMGE VE KISALTMALAR

Simgeler

G	Gauss filtresi
I	Görüntü karesi
σ	Ölçek katsayısı

Açıklama

Kısaltmalar

2B	İki Boyutlu
2.5B	İki buçuk Boyutlu
3B	Üç Boyutlu
ICP	Iterative Closest Point
PCL	Point Cloud Library
RANSAC	Random Sample Consensus
RGB-D	Kırmızı, Yeşil, Mavi ve Derinlik Görüntüsü
DoG	Difference of Gaussian
SIFT	Scale Invariant Feature Transform
SURF	Speeded Up Robust Features
ORB	Oriented Fast Rotated Brief

BÖLÜM I

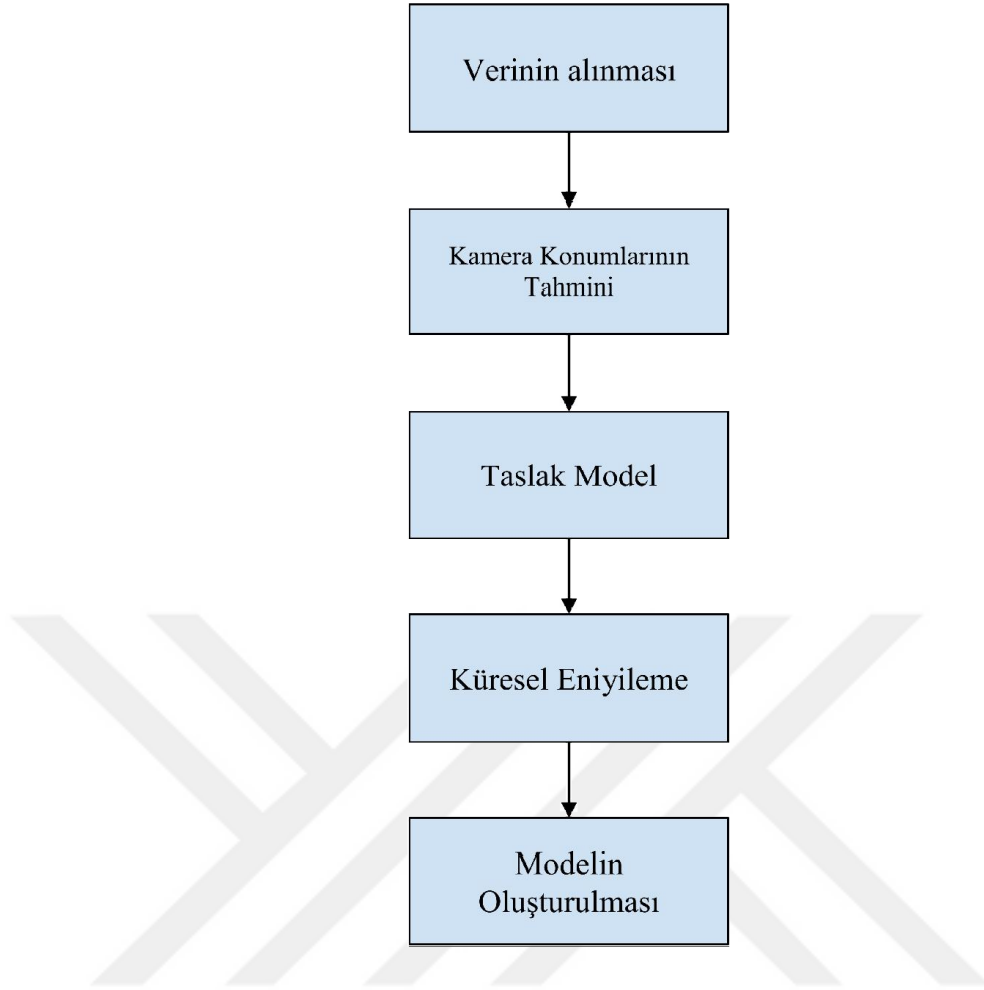
GİRİŞ

Gelişen ve ucuzlayan teknoloji ile birlikte ortaya çıkan Kinect gibi kırmızı yeşil mavi-derinlik (Red Green Blue-Depth, RGB-D) kameralar üç-boyutlu (3B) yeniden oluşturma konusunda yapılan çalışmaların hız kazanmasını sağlamıştır. Kızılötesi algılayıcılar ile derinlik bilgisini görüntü ile birleştiren bu kameralar ile nesnenin farklı açılarından çekilen görüntüleri nesnelerin 3B yeniden oluşturmalarını mümkün kılmaktadır. 3B yeniden oluşturma ile nesnelerin yeniden oluşturulması, ortam haritalama, robotik, arkeoloji, biomedikal, oyun endüstrisi ve sanal gerçeklik gibi birçok uygulama alanına sahiptir. RGB-D kameraların gelişimleri hızla devam etmekte ve her geçen gün daha yüksek çözünürlük ve doğruluğa sahip ulaşılabilir teknolojiler geliştirilmektedir.

Bir nesnenin 3B yeniden oluşturulabilmesi için o nesneye ait farklı konumlardan çekilen RGB-D görüntülere ihtiyaç duyulmaktadır. Çekilen her bir görüntü için görüntünün çekildiği kameraların altı serbestlik derecesine (6 degrees of freedom) sahip konum bilgilerinin yüksek doğrulukta bilinmesi gerekmektedir. Fakat kamera konumları her zaman kesin olarak elde edilemeyebilir. Kamera konumlarının bilinmediği durumlar için araştırmacılar 3B yeniden oluşturma işlemini genellikle aşağıdaki aşamalardan oluşan bir akış ile özetlenebilir. Bunlar;

1. Verinin alınması,
2. Kamera konumlarının yaklaşık olarak tahmin edilmesi,
3. 3B yeniden oluşturma (taslak model),
4. Küresel eniyileme ile kamera konumlarının daha hassas bir şekilde tahmini,
5. Yeni kamera konumları ile 3B yeniden oluşturma (nihai model)

olarak özetlenebilir. Bu akış Şekil 1'de gösterilen akış şeması ile görselleştirilmiştir.



Şekil 0.1. 3B yeniden oluşturma akış şeması.

Bu akış için ilk adım verinin alınmasıdır. Veriler kameralar aracılığıyla elde edilmektedir. Kamera görüntülerine yardımcı olarak derinlik bilgisinin elde edilmesi de mümkündür. Farklı türden veriler kullanılarak 3B yeniden oluşturma işleminin başarımı arttırılmaya çalışılmaktadır.

İkinci adımda, kamera konumlarının yaklaşık olarak tahmin edilmesi gerekmektedir. Kamera konumlarının yaklaşık olarak hesaplanabilmesi için görüntüler arası eşleştirmelerin bulunması gerekmektedir. Eğer görüntüler arası eşleştirme noktaları bulunabilir ise kameraların yer değiştirme değerleri yaklaşık olarak hesaplanabilir. Fakat, görüntü içeriğine bağlı olarak görüntüler arası eşleştirilen noktaların sayısı düşük olabilir. Bu gibi durumlarda hesaplanan kamera konumlarının doğruluğu düşük olacak ve yüksek hatalı işlem yapılmasına yol açılacaktır. Aksi durumlarda ise görüntüler arası doğru eşleştirilen nokta sayısının yüksek olması kamera konumlarının yüksek doğrulukta tahmin edilmesi anlamına gelmektedir. Görüntüler arasında tespit edilen yüksek

eşleştirme sayısı her zaman yüksek doğrulukta kamera konumu tahmini için yeterli olmayabilir. Eşleştirilen nokta sayısının yüksek olduğu durumlarda kamera konumlarının tahmin doğruluğunu düşüren bir diğer hususta yanlış noktaların eşleştirilmesidir. Buna göre, görüntüler arası doğru noktaların eşleştirilmesi ve doğru eşleştirmelerin sayısı, 3B yeniden oluşturma problemi için önemli problemlerdir.

Üçüncü adımda, 3B modelin bir taslağının oluşturulması ele alınmaktadır. 3B modelin oluşturulabilmesi için farklı konumlardan çekilen görüntülerde yer alan eşleştirilmiş noktalar, kamera parametreleri ve kamera konumları kullanılarak derinlik bilgileri birleştirilmektedir. Derinlik bilgisi görüntü çiftleri arasında hesaplanabileceği gibi harici bir algılayıcı aracılığıyla da elde edilebilir. Derinlik bilgisinin harici bir algılayıcıdan alınması daha yoğun ve yüksek doğrulukta modeller oluşturulmasını mümkün kılmaktadır. İkinci adımda elde edilen kamera konum tahminleri her zaman yüksek doğrulukta olmadığı için ilk oluşturulan model taslak model olarak adlandırılmıştır. Bu taslak modelin düzeltilerek daha düzgün bir model elde edilmesi gerekmektedir.

Dördüncü adımda, taslak model ve tahmin edilen kamera konumları ele alınarak eniyileme yöntemleri ile var olan hatalar azaltılmaya çalışılmaktadır. Son adımda ise iyileştirilen kamera konumları ve derinlik bilgileri kullanılarak üçüncü adım tekrar çalıştırılmaktadır. Sonuç olarak elde edilen model daha düzgün bir nihai model olacaktır.

Bu tez kapsamında derinlik bilgisi ve görüntü bilgisinin birlikte üretildiği RGB-D kamera kullanılmıştır. Seçilen kamera ucuz ve çok bulunabilir bir model olan Kinect V1 modelidir. RGB-D kamera kullanımı ile daha yoğun ve daha doğru bir model elde edilmesi amaçlanmıştır. Literatürde yer alan görüntü noktası eşleştirme algoritmaları ve eniyileme yöntemleri kullanılarak bir karşılaştırma ortamı oluşturulmuş ve yöntemler bir arada kullanarak 3B yeniden oluşturma başarımları incelenmiştir. Böylece yapılan çıkarımlar doğrultusunda başarılı bir 3B nesne yeniden oluşturma sistemi elde edilmiştir.

Sonraki alt bölümde 3B yeniden oluşturma ile ilgili literatür taraması verilecektir. Bölüm II’de, derinlik bilgisi kullanılarak 3B yeniden oluşturmada kullanılan cihaz ve yöntemler incelenmiştir. Bölüm III’te, tez kapsamında kullanılan veri kümesi ve yöntemlerin karşılaştırılması incelenmiştir. Bölüm IV’te ise gerçekleştirilen tez çalışmasından elde edilen sonuçlar ve değerlendirmeler verilmiştir.

1.1 Literatür Özeti

3B derinlik bilgisinin algılanma yöntemine göre iki temel grup altında incelenebilir. Bunlar aktif ve pasif yöntemlerdir. Aktif yöntemler, 3B yeniden oluşturma doğrudan derinlik bilgisinin alınmasının mümkün olduğu algılayıcıların kullanılarak gerçekleştirildiği yöntemleri belirtir. Aktif yöntemlerde kullanılan algılayıcılar 3B lazer tarayıcılar, uçuş süresi tabanlı derinlik algılayıcılar ve belirli bir desenin yansıtılarak derinlik bilgisinin doğrudan elde edildiği algılayıcılardır (Horaud vd., 2016). Bu algılayıcı tipinde alınan ortam bilgisi doğrudan derinlik verilerini içerdiğinden derinlik bilgisinin çıkarımı için ek bir işleme ihtiyaç duyulmamaktadır. Ayrıca, aktif yöntemlerde kullanılan algılayıcılar çok pahalı cihazlardır.

Pasif yöntemler, derinlik bilgisinin algılayıcı tarafından doğrudan elde edilemediği ve ek yöntemler kullanılarak derinlik bilgisinin tahmin edildiği yöntemleri tanımlamaktadır. Bu yöntemler, binoküler stereo (binocular stereo) (Cochran ve Medioni, 1992), fotometrik stereo (Clark, 1992), hareketten yapı çıkarımı (Pollefeys vd., 2004) gibi yöntemleri içermektedir. Pasif yöntemlerde kullanılan cihazlar daha uygun maliyetli olmasına rağmen başarımları olarak kabul edilebilir bir düzeyi sağlayabilmeleri için özgün yaklaşımlar gerekmektedir.

Stereo kamera veya çoklu kamera kullanılarak derinlik haritası oluşturulması konusu araştırmacılar tarafından çok çalışılmış bir konudur. Derinlik haritası oluşturulması konusunda yapılan ilk çalışmalardan biri olan (Cochran ve Medioni, 1992), derinlik haritasının başarımına etki eden problemler dört madde halinde belirtilmiştir. Bunlar;

- Fotometrik değişimler: Sahneler arası, ışığın açısından kaynaklanan bazı farklılıklar söz konusu olabilir. Bu nedenle görüntüler arası eşleştirme işleminin başarımını etkileyecek ve oluşturulacak olan derinlik haritası düşük doğruluğa sahip olacaktır.
- Örtüşmeler: Görüntüler arasında, açı farklılıklarından kaynaklanan iki ve görüntüler arasında karşılığı olmayan noktalar oluşabilir. Böyle bölgeler genellikle görüntüler arasında yanlış eşleştirmelere sebebiyet verecektir.

- Tekrar eden desenler: Bir duvarın tuğlaları, kitaplıktaki kitaplar gibi görüntü içerisinde tekrar eden görüntü bölgeleri yanlış eşleşmelere sebep olabilir.
- Desensiz bölgeler: Görüntü içerisinde desen içermeyen geniş bölgelerin eşleştirilmesi mümkün olmayacak veya yanlış eşleştirmeler ortaya çıkacaktır. Bu nedenle elde edilecek derinlik haritası da hata içerecektir.

Cochran ve Medioni çalışmalarında bu sorunların giderilebilmesi için alan ve öznitelik tabanlı yaklaşımların birlikte kullanıldığı bir yöntem önermiştir.

Lotti ve Giraudon (1993) çalışmasında, stereo görüntüler arasında eşleştirme işleminde kullanılan sabit boyutlu pencereye alternatif olarak adaptif bir pencereleme yöntemini önermiştir. Adaptif pencere boyutunun kullanıldığı eşleştirme mekanizması daha yüksek doğrulukta eşleştirme olanağı sağlamış ve derinlik haritası oluşturma başarımını arttırmıştır.

Adjouadi ve Candocia (1994) tarafından sunulan yöntemde stereo eşleştirme işlemi için Walsh dönüşümü tabanlı bir yaklaşım kullanılmıştır. Bu çalışmada, stereo görüntüden derinlik hesabının en temel problemleri olan öznitelik bulma, bulunan öznitelikleri eşleştirme, eşleştirilen özniteliklerin doğruluğunu kontrol etme ve doğru eşleştirilen noktaları kullanarak derinlik hesaplama konularında iyileştirmeler sağlanmıştır.

Pollefeys vd. (2004), stereo el kamerasından alınan görüntüler üzerinden yoğun (dense) 3B yeniden oluşturma işlemi için bir öznitelik takip yaklaşımı kullanarak derinlik haritası çıkartmıştır. Burada öznitelikler takip edilmekte ve paket halinde elde edilen ardışık karelerden oluşturulan kamera konumları ve 3B derinlik haritası düzeltilmektedir. Bu düzeltme işlemi sonucunda elde edilen derinlik haritası ve kamera konumları hareketten yapı çıkarımını (structure from motion - SFM) mümkün kılmaktadır.

Esteban ve Schmitt (2004), ilk olarak her bir kamera açısından bir derinlik haritası hesaplamıştır. Bu derinlik haritaları birleştirilerek bir maliyet yığını (cost volume) oluşturulmuştur. Ardından bu maliyet yığını üzerinde iterasyonel bir algoritma ile en düşük hacme sahip yüzey belirlenmektedir. Bu yüzey belirlenirken görsel silüet özellikleri de kullanılmıştır. Esteban ve Schmitt tarafından yapılan çalışmaya benzer

şekilde, Goesele vd. (2006) maliyet yığınlarını oluşturmuş ve bu maliyetleri VRIP (Curless ve Levoy, 1996) yöntemini kullanarak birleştirmişlerdir.

Furukawa vd. (2008) çalışmasında, stereo eşleştirme için ilk olarak görüntüler arası küresel temel öznitelik noktalarının 3B koordinatları hesaplanmıştır. Görsel bir kabuk model oluşturularak bulunan bu öznitelikler kabuk model üzerine yerleştirilmiştir. Bu elde edilen kabuk model enerji minimizasyon algoritması kullanılarak iyileştirilmiştir.

Stereo veya çoklu kamera kullanılarak gerçekleştirilen sistemlerin başarımları aktif yöntemlere nazaran daha karmaşık ve başarımları düşük olabilmektedir. Bunları göz önünde bulundurarak bu tez kapsamında Kinect V1 cihazı kullanılmıştır. Kinect V1’de kullanılan algılayıcı (US20170004623A1, 2005) patenti ile ortaya atılmıştır. Kinect V1, yapılandırılmış ışık (Albitar vd., 2007) yaklaşımını kullanmaktadır. Lazer tarayıcılara göre nispeten daha uygun maliyetli olması ve pasif yaklaşımlara göre derinlik bilgisini hazır sunması cihazın tercih edilirliliğini arttırmaktadır. Kinect’te kullanılan derinlik haritası 640x480 piksel boyutundadır. Diğer 3B derinlik algılayıcılarının kullanıldığı cihazlar ile karşılaştırıldığında bu çözünürlük değeri yüksek hassasiyete sahip model oluşturmak için oldukça düşüktür. Ayrıca, kinect tarafından sağlanan derinlik haritası yüksek gürültü oranına sahiptir. Gürültülü derinlik haritası, 3B yeniden oluşturma işleminin başarımlarını etkileyen bir diğer faktör olarak ele alınmalıdır.

Literatürde yer alan kinect tabanlı 3B yeniden modelleme uygulamalarında modelin daha yüksek çözünürlüğe ve doğruluğa sahip olmasını sağlamak için çeşitli teknikler kullanılmaktadır. Bu uygulamalardan bir tanesi olan KinectFusion (Newcombe vd., 2011; Izadi vd., 2011), kinect sahnelerini kullanarak yoğun ve hatasız 3B yeniden modelleme yapmayı amaçlamaktadır. Newcombe vd. (2011) karmaşık sahnelerin gerçek zamanlı 3B modellenmesi için yoğun yüzey eşleştirme ve kamera konum takibi işlemlerini gerçekleştirmiştir. Kamera konumları, ardışık sahnelerden alınan tüm derinlik bilgileri ile yeni gelen sahnenin derinlik bilgilerini iteratif en yakın nokta (Iterative Closest Point - ICP) algoritması ile takip edilmektedir. Her sahne arasında takibe göre tüm derinlik bilgisinin kullanıldığı bu yöntem, oluşturulan 3B sahnede daha az kayma ve daha yüksek doğruluk sağlamaktadır. Song vd. (2013) yüzey yeniden oluşturma probleminin çözümü için KinectFusion’a çok benzer bir çalışma gerçekleştirmiştir.

Izadi vd. (2011) ise KinectFusion uygulamasında kullanılan yöntemlerin daha özgün bir GPU çekirdeği ve genişletilmiş GPU işlem yöntemi iyileştirilmesini temel almaktadır. Bu iyileştirmeler ile KinectFusion sisteminin kamera konum takibi ve 3B yeniden oluşturma kapasitesini etkilemeden bölütleme ve kullanıcı etkileşimini yapabilmesini mümkün kılmıştır.

Bu yöntemler nesnenin 3B yeniden oluşturulmasını mümkün kılmaktadır. Fakat oluşturulan nesne halen sahnenin küçük bir parçasıdır ve sahnenin elde edilmek istenen obje dışında kalan görüntüsünü de içermektedir. Chen vd. (2013) çalışmasında basit bir uzaklık eşik değeri kullanarak istenilen nesnenin sahnedan çıkarılmasını kolaylaştırmıştır. Böyle bir işlem yanlış eşleşebilecek bölgeleri azalttığı için nesnenin yeniden oluşturma başarımını da arttırmaktadır. Fakat bu yöntem fazla girinti çıkıntıya sahip nesneler üzerinde iyi çalışmamaktadır.

Bir diğer çalışmada ise Chen vd. (2014) sahneyi yeniden oluşturduktan sonra sahnedan çıkarılmak istenilen nesneyi veri kümesinde yer alan nesneler ile eşleştirerek daha doğru modelleme yapılabilmesini sağlamıştır. Bu yöntem ise veri kümesinde olmayan nesneler için düzgün çalışmamaktadır.

Morana (2014) farklı bir yaklaşım ile 3B sahne yeniden oluşturma çalışması gerçekleştirmiştir. Chen vd. (2013; 2014) çalışmalarından farklı olarak bu çalışmada nesnelerin modellenmesi için süper-quadrikler kullanmıştır. Böylece modellerin düzensiz bir şekle sahip olması durumunda elde edilen sonuçlar iyileştirilmiştir.

Kinect algılayıcısının yüksek gürültü oranına sahip olduğundan bahsedilmişti. ICP algoritması yüksek gürültü içeren veriler üzerinde kullanıldığında başarımı oldukça düşük bir algoritmadır. Bu nedenle, Chatterjee vd. (2012), ICP algoritmasının yüksek gürültüde yanlış sonuçlar üretmesinin önüne geçebilmek adına ICP algoritmasından önce uyarlanabilir çift-yönlü filtre (adaptive bilateral filter) kullanmıştır.

Sturm vd. (2013) tarafından, KinectFusion tabanlı insan figürlerinin taranması ve yazdırılmasını mümkün kılan CopyMe3D adındaki yazılım gerçekleştirilmiştir. Benzer şekilde Rooker vd. (2013) ReconstructMe adında bir uygulama geliştirmiştir.

Bunların yanında Skanect yazılımı Kinect algılayıcı kullanarak 3B nesne oluşturma konusunda başarılı sonuçlar üretmektedir. Bu yazılım kullanıcıdan arka plana ait olan bölgelerin işaretlemesini istemektedir. Karmaşık sahnelerde bu durum işlemi zorlaştırmaktadır.

Valentin vd. (2015) çalışmasında SemanticPaint adında etkileşimli bir 3B sahne yeniden oluşturma yöntemi önermiştir. Bu sistem KinectFusion, çevrimiçi öğrenme, konuşma tanıma ve ortalama alan çıkarımı gibi farklı teknikleri bir araya getirerek farklı nesneleri yeniden oluşturulmuş sahneden gerçek zamanlı olarak ayırabilmektedir.

Pan ve Taubin (2016) çalışmasında, oluşturulan sahneden nesneye ait nokta bulutlarının ayrıştırılabilmesi için bir yöntem önerilmiştir. Bu yöntemde nesnenin sahnenin ortasında olduğu varsayılmaktadır. Ayrıca bu sistem gerçek zamanlı çalışma için uygun değildir.

Jaiswal vd. (2014), geleneksel bir 3B nesne yeniden oluşturma için kullanılan bütün bir çözüm sunmuşlardır. Bu yöntemde, ilk önce, nesnenin farklı görünümlerinden elde edilen 3B kısmi nokta bulutları RANSAC (Random sample consensus) (Fischler ve Bolles, 1981) algoritması ile kabaca oluşturulmaktadır. Ardından, bu ilk taslak, ICP (Besl ve McKay, 1992) algoritması gibi yinelemeli bir algoritma ile iyileştirilmektedir.

Xie vd. (2015) çalışmalarında, ayrık öznitelik tanımlayıcıları kullanarak sahneler arası eşleştirmeler bulmaktadır. Bulunan bu öznitelik noktaları kullanılarak sahnenin taslak bir modeli oluşturulmaktadır. Ardından yeni sahneler bu modele eklenerek nesne yeniden oluşturma işlemi gerçekleştirilmektedir. Bu çalışmada ağırlıklandırılmış ICP algoritması kullanılarak birleştirme adımının başarımı artırılmıştır. Ayrıca bu çalışmada her ICP adımında ayrı noktaların tespiti yapılarak hata oluşumu en aza indirilmektedir.

3B nesne yeniden oluşturma işlemi için Kinect tarafından sağlanan stereo görüntüler de kullanılarak yeniden oluşturma başarımının arttığı literatürde yapılan çalışmalardan da anlaşılmaktadır (Knopp vd., 2010; Lemuz-López ve Arias-Estrada, 2006). Bu tez çalışmasında 3B nesne yeniden oluşturma işlemi için farklı görüntü eşleştirme (SIFT, SURF, ORB) (Lowe, 1999; Bay vd., 2006; Rublee vd., 2011) algoritmalarından elde edilen sonuçlar incelenecektir. Ayrıca küresel eniyileme kullanılan bir yöntem ile sistem başarımının nasıl değişiklik göstereceği incelenecektir.

BÖLÜM II

DERİNLİK BİLGİSİ KULLANARAK NESNE YENİDEN OLUŞTURMA

2.1 Giriş

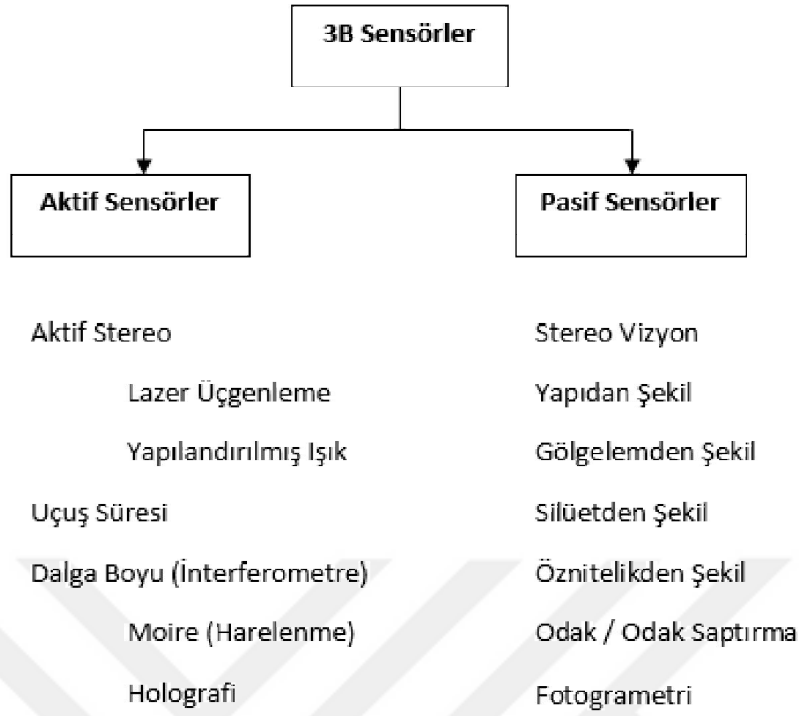
Derinlik algılayıcılar, ortaya çıkışlarından bu yana mühendislik, arkeoloji, biomedikal, oyun endüstrisi ve sanal gerçeklik gibi birçok alanda kullanılmaktadır. Bir sahnenin 3B haritasını veya bir cismin 3B modelini oluşturmak uzun yıllardır bilgisayarla görme alanında çalışılan konulardan biridir.

3B görselleştirme için gerçek dünya nesnelerinin elde edilmesi ve yeniden oluşturması problemi iki ana başlık altında incelenebilir: gerçek dünya nesnelerinin derinlik verilerinin nasıl elde edileceği ve nesnenin birden fazla modelinin tek bir modelde nasıl birleştirileceği. Bu sorunların her ikisi de ayrı ayrı alanlarda araştırmacılar tarafından yoğun biçimde çalışılmaktadır.

Nesnelerinin derinlik verilerinin nasıl elde edileceği alanda yapılan çalışmalarda, derinlik bilgisi farklı yöntemlerle elde edilebileceği ortaya konulmuştur. Bu yöntemlerden bazıları Bölüm 2.2 de incelenecektir.

2.2 Derinlik Bilgisi Elde Etme Yöntemleri

3B nesnenin şeklini sayısal olarak elde etmek için çeşitli literatürde birçok farklı yöntem bulunmaktadır. (Remondino ve El-Hakim, 2006) tarafından yapılan sınıflandırma, 3B yeniden oluşturma işlemlerinde uzaklık ölçme yöntemlerini iki türe ayırır: temaslı ve temassız ölçme. Temaslı çözümler bir algılayıcı (probe) yardımı ile fiziksel olarak nesneye dokunarak mesafe ölçümünü temel alan çözümlerdir. Temassız çözümler ise görüntüler veya ışık kaynakları kullanarak derinlik bilgisini elde etmektedir. Temassız çözümler ayrıca aktif ve pasif olmak üzere iki ana kategoriye ayrılabilir. Bu yöntemlerden bazıları Şekil 2.1’de belirtilmiştir.

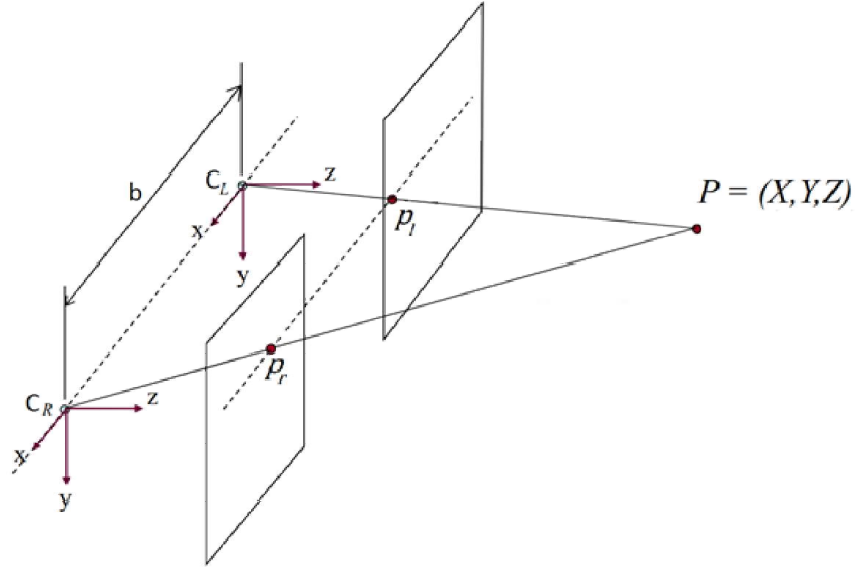


Şekil 2.1. Farklı 3B Aktif ve Pasif sensörlerinin gösterimi

Bu bölümde, mevcut tekniklere genel bir bakış sunulmaktadır. Ayrıca ek olarak, bu tekniklerin çalışma prensiplerinin karşılaştırmalı olarak değerlendirilmiştir.

2.2.1 Stereo görüntü

En yaygın olarak kullanılan 3B görüntü oluşturma tekniklerinden biridir. Çalışma prensibi olarak insan gözünün yapısı taklit edilmeye çalışılmıştır. Şekil 2.2’de gösterildiği gibi aynı düzlemde yerleştirilmiş iki kamera kullanılmaktadır. Böylece sahnenin görüntüsü farklı iki noktadan elde edilmiş olur. Elde edilen iki görüntü üzerinde yatay eksen boyunca eşleşen noktalar tespit edilerek eşitsizlik haritası adı verilen derinlik bilgilerini içeren bir görüntü elde edilir. Daha sonra kalibrasyon işlemi sırasında kamera parametreleri ve kameralar arasındaki mesafe bilindiğinden, eşitsizlik haritası yardımı ile sahnenin veya nesnenin derinlik bilgileri elde edilebilmektedir.



Şekil 2.2. Stereo kamera modeli

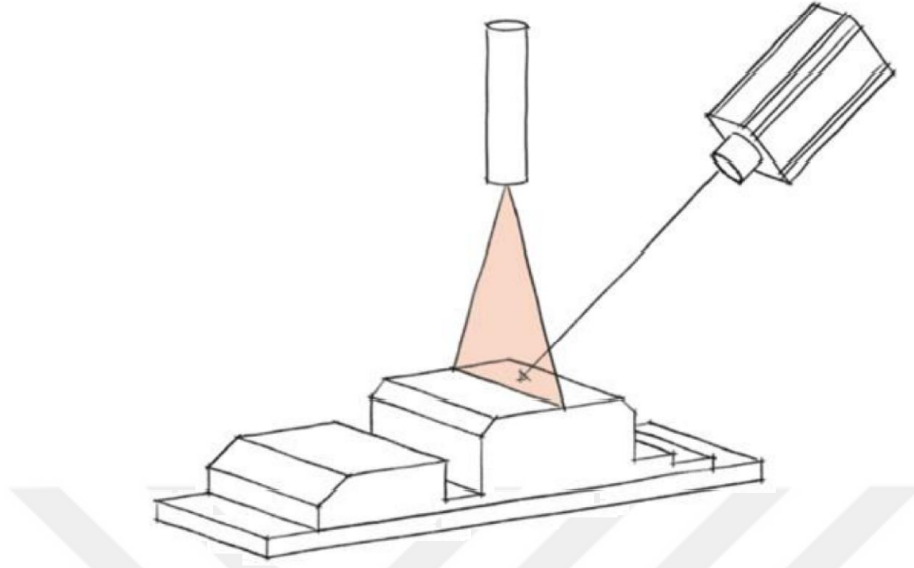
Şekil 2.2’da görüldüğü üzere ve C_L sırasıyla sağ ve sol kamera merkezleridir ve bu iki kamera merkezi arasındaki mesafe ile gösterilmiştir. 3 boyutlu düzlemdeki herhangi bir P noktasının, sol kamera düzlemindeki izdüşümü p_l , ve sağ kamera düzlemindeki izdüşümü olarak adlandırılmaktadır. Stereo kamera modelinde görüntü düzlemleri üzerindeki ve noktaları bulduktan sonra üçgenleme yöntemi ile noktasının derinlik bilgisi hesaplanır.

Piksellerin, iki görüntünün yatay eksenleri arasındaki farkı eşitsizlik (disparity) değeri olarak adlandırılır ve d ile gösterilir. Eşitsizlik değeri elde edildikten sonra, bu piksele karşılık gelen z derinlik değeri Eşitlik 1 kullanılarak hesaplanmaktadır.

$$Z = f (b/d) \quad (1)$$

Bu ifadedeki terimlerden f (focal length) kameranin odak uzaklığını ve b (baseline length) terimi ise de iki kamera merkezi arasındaki mesafeyi temsil etmektedir.

2.2.2 Lazer üçgenleme (Laser triangulation)

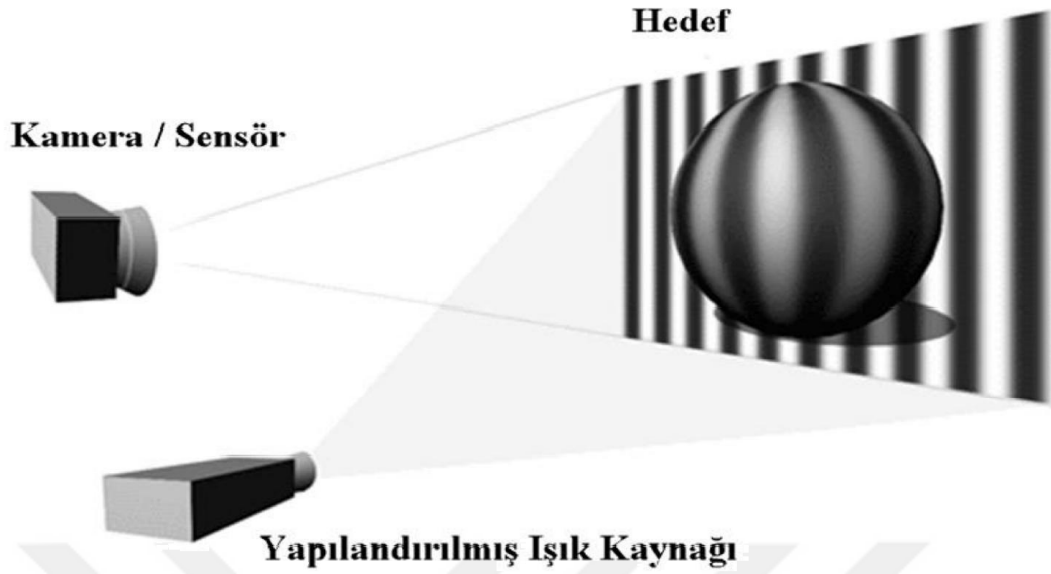


Şekil 2.3. Lazer üçgenleme modeli

Bir lazer kaynağı tarafından ışın noktasal veya şeritler halinde nesne üzerine yansıtılır. Daha sonra 3B modeli elde edilmek istenen nesnenin kontrollü bir şekilde hareket etmesi sağlanır. Bu işlem sırasında nesnenin ne kadar hareket ettiği hesaplanmaktadır veya önceden bilinmektedir. Bu işlem sırada kalibre edilmiş bir kamera ile alınan sürekli görüntülerden faydalanarak lazer noktaları veya çizgilerinin nesnenin üzerindeki hareketinden derinlik bilgisi elde edilir. Genellikle lazer kaynağı ile kamera kontrollü bir ortam oluşturmak için kapalı bir kutu içinde yer almaktadır.

2.2.3 Yapılandırılmış ışık kaynağı (Structured light)

Bu yöntemde, bir projektör yardımı ile 3B modeli elde edilmek istenilen nesne üzerine özel bir desen yansıtılmaktadır. Bu desen genellikle çizgiler şeklinde olmaktadır fakat noktalar veya ızgara şeklinde de kullanılan örnekleri mevcuttur. Nesne üzerinde yansıtılan ışığın oluşturduğu dağılım en az bir kamera (genelde iki kamera) kullanılarak algılanır. Derinlik bilgisi elde edilirken lazer üçgenleme tekniğine benzer bir yöntem kullanılmaktadır. Burada nesneyi hareket ettirmek yerine, ışık kaynağından yansıtılan desen kaydırılmaktadır.

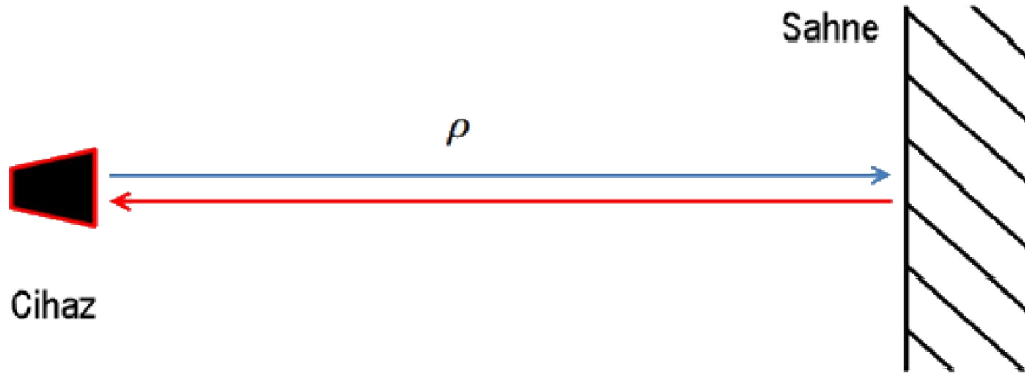


Şekil 2.4. Yapılandırılmış ışık kaynağı (Tarin, 1999)

2.2.4 Uçuş süresi (Time of Flight - ToF) kameralar

Uçuş süresi, bir nesnenin, partikülün veya dalganın bir ortam boyunca ilerlemesi için geçen süre kullanılarak oluşturulan algılayıcılarda kullanılan bir yöntemdir. ToF algılayıcılar, ışının seyahatinin süresini, mesafesini ölçmek veya ortam özelliklerini bulmak için kullanılabilmektedirler.

ToF kameralar, kaynaktan gönderilen ışının ne kadar sürede geriye yansıdığının bilgisini kullanarak nesneye olan mesafeyi ölçen algılayıcılardır. ToF kameralar için iki ana işlem adımı bulunmaktadır. Bunlar aydınlatma ve algılamadır. Aydınlatma aşamasında gözle görülür etkileri önlemek için kızılötesi ışık veya lazer diyot yardımı ile ışınlar gönderilmektedir. Algılama adımı ise yayılan ışınlar bir mercekle ve optik filtre aracılığıyla toplanmakta ve ışının seyahati için geçen zaman, her bir piksel için, şarj bağlı cihaz (charge coupled device - CCD) ilkesine dayanan iki boyutlu bir ışık algılama katmanı dizisi üzerinde ölçülmektedir (Boyle ve Smith, 1970).



Şekil 2.5. Uçuş süresi kamerası çalışma prensibi

2.3 RGB-D

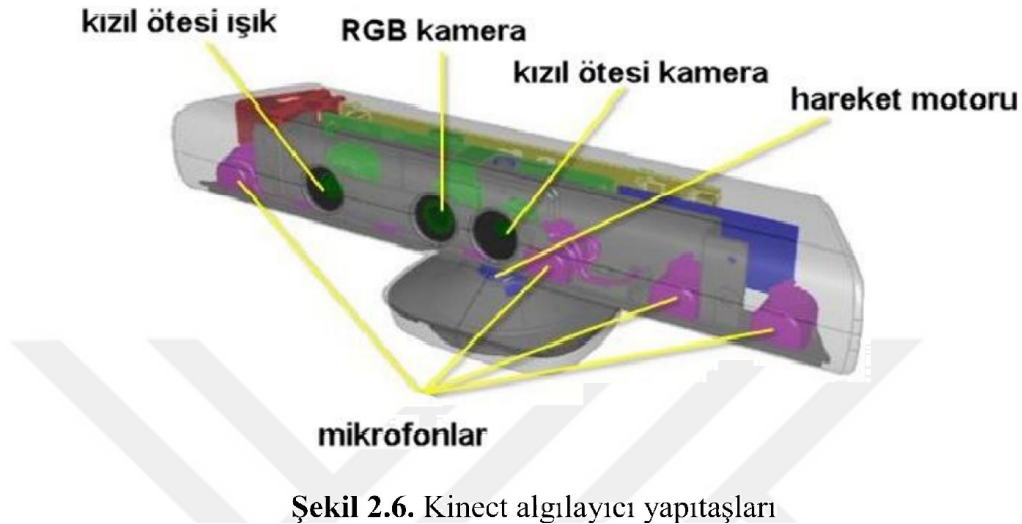
3B yeniden oluşturma yöntemlerinde derinlik bilgisi nesneyi yeniden oluşturmak için tek başına yeterli olmamakta aynı zamanda renk bilgisine de ihtiyaç duyulmaktadır. Bazı derinlik algılayıcılardan elde edilen veriler sadece derinlik bilgisi içerdiğinden dolayı diğer algılayıcılardan elde edilen renk bilgisi ile birleştirilerek sensör füzyonu işlemi gerçekleştirilmiştir. Bu şekilde elde edilen veriler RGB-D ve ya 2.5 B olarak adlandırılmaktadır.

RGB-D veri sağlamak için literatürde en çok kullanılan algılayıcıların başında Microsoft firması tarafından üretilen Kinectler gelmektedir. Kinect'in V1 ve V2 olmak üzere iki farklı versiyonu vardır. İlk sürümü 2010'da çıkmıştır ve akademik çalışmalarda büyük kullanım alanı bulmuştur. Bunun başlıca sebeplerinden biri olarak ucuz ve doğruluğu kabul edilebilir sonuçlar vermesi gösterilmektedir. Bu tez çalışması kapsamında Kinect V1 ile elde edilen veri kümesi kullanılmasına rağmen, bu bölümde V2'den de bahsedilecektir.

2.3.1 Kinect V1

İkinci versiyonunun piyasaya sürülmesinin ardından V1 olarak anılmaya başlanan ilk Kinect; 2010 yılında Microsoft firması tarafından Xbox oyun konsollarında kumanda işlevi görmesi için tasarlanmış bir algılayıcıdır. Bu algılayıcı, oyuncunun fiziksel hareketlerini takip edebilmek adına derinlik bilgisinin elde edilmesinde kullanılan kızılötesi kamera; renkli görüntüyü elde etmek için kullanılan RGB kamera ve kızılötesi kamera için bir adet ışık kaynağı ile donatılmıştır. Ucuz, derinlik bilgisi verebilen bir algılayıcı olduğundan akademik

alışmalar da dahil olmak üzere geniş bir kitle tarafından kullanılmış, hem Microsoft hem de topluluklar tarafından oldukça fazla desteklenmiştir. İnsan takibi, poz ve hareket algılanması, robotik vizyon gibi alanlarda sıklıkla kullanılmaktadır. Şekil 2.6'da Kinect'e ait temel algılayıcılar gösterilmiştir.



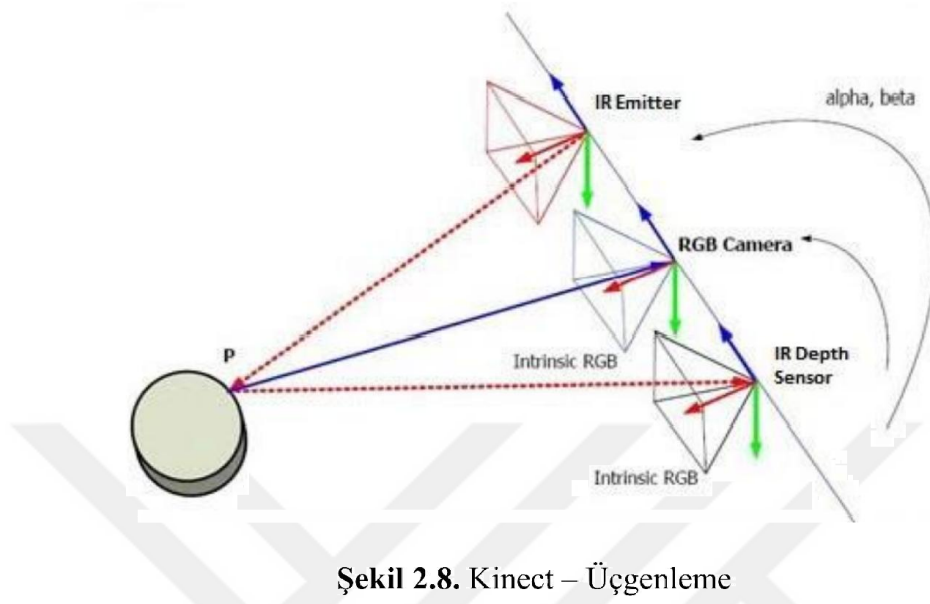
Şekil 2.6. Kinect algılayıcı yapıtaşları

Kinect V1, derinlik haritasının çıkarımında yapılandırılmış ışık tekniğini kullanmaktadır. Kamera çözünürlükleri sırasıyla 1280x960 ve 640x480 piksel olmakla birlikte saniyede 30 kare çekebilmektedirler. 0.5 metreden 8 metreye kadar olan mesafelerde kullanımda güvenilir sonuçlar alınabilmektedir. Kızılötesi ışık kaynağı çizgisel bir desen yerine Şekil 2.7 de görülen yapıya benzer biçimde benek deseni kullanarak yapılandırılmış ışık yayılımı yapmaktadır.



Şekil 2.7. Kinect kızılötesi ışık yapısı

Sonrasında ışık kaynağından yayılan ışın ile kızılötesi kamera ya yansıyan ışın (kaynak, almaç arası mesafe) kullanılarak bütün benekler için üçgenleme uygulanmaktadır (Şekil 2.8).



Şekil 2.8. Kinect – Üçgenleme

V1 in en büyük dezavantajı gürültü bağıışıklığının oldukça düşük oluşudur. Ortamda bulunan diğer ışık kaynakları, Kinect V1 tarafından yayılan desende ve dolayısı ile derinlik bilgisinde kolaylıkla bozulmalara yol açabilmektedirler.

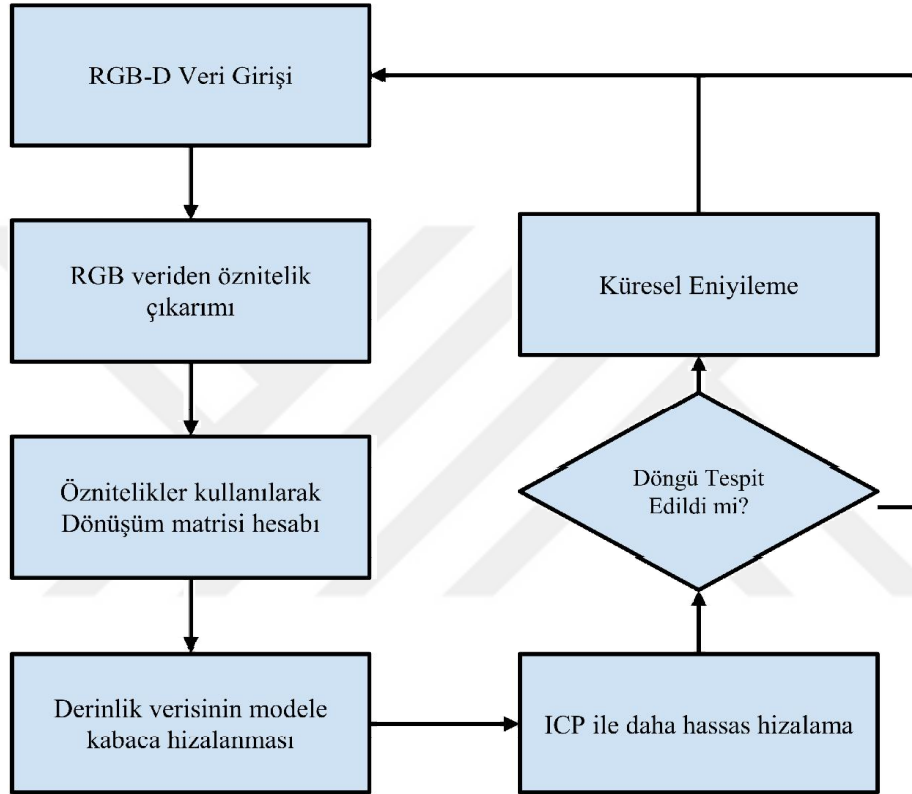
2.3.2 Kinect V2

Kinect V2, derinlik algılama metodu olarak uçuş süresi prensibini kullanmaktadır. Cihaz içerisinde Kinect V1 e benzer şekilde bir kızılötesi ışık kaynağı ile biri RGB diğeri kızılötesi olmak üzere iki kamera yer almaktadır. Kamera çözünürlükleri sırasıyla 1920x1080 ve 512x424 pikseldir ve saniyede 30 kare alınabilmektedir. Kinect V2 için tutarlı kabul edilen çalışma aralığı 0.5 metreden 10 metreye kadar olan aralıktır.

Kinect V1 den farklı olarak V2 de kızılötesi yanıp sönen ışık kaynağı sahneyi aydınlatmakta, sahnede bulunan cisimlerden yansıyan kızılötesi ışınlar, kızılötesi kamera tarafından yakalanmaktadır. Derinlik bilgisi, ışığın gönderilmesi ve kızılötesi kamera tarafından alınması arasında geçen süre kullanılarak hesaplanmaktadır.

2.4 3-Boyutlu Yeniden Oluşturma

3B yeniden oluşturma için Kinect benzeri algılayıcılardan elde edilen çoklu verilerden faydalanılacaktır. Aynı kaynaktan alınmış sabit nesne görüntüleri ve derinlik bilgileri kullanarak ilgili nesnenin 3B yeniden oluşturulmasında kullanılacak olan yönteme ilişkin detaylar bu bölümde verilecektir.



Şekil 2.9. 3B yeniden oluşturma akış şeması

İlk olarak, görüntüler arası eşleştirmeler tespit edilmekte, bu bilgi ile hesaplanan dönüşüm matrisi kullanılarak derinlik bilgileri kabaca hizalanmaktadır. Böylece, hem ICP kullanılarak yapılacak olan eşleştirmenin başarımı artırılacak hem de ICP algoritmasının çalışması için gerekli olan zaman azaltılmış olmaktadır. ICP kullanılarak kabaca hizalanan derinlik bilgileri daha hassas bir şekilde modelle bütünleşmiş olacaktır. Her yeni görüntü için bu işlemler tekrar edilmektedir. Görüntü tabanlı eşleştirmeler bir önceki görüntü karesi referans alınarak yapılmaktadır. Fakat derinlik bilgileri her adımda oluşturulacak olan 3B model referans alınarak ICP algoritması ile hizalanacaktır. En son adımda ise bütün görüntüler işlenmiş ve bir model elde edilmiştir. Fakat bu model bazı hizalama hataları ve gürültüler içermektedir. Gürültülerin giderilmesi için aykırı nokta tespiti kullanılmaktadır. Ayrıca, hizalama hatalarının

iyileştirilmesi için üretilen taslak model ve önceki kamera konumları kullanılarak küresel eniyileme işlemi uygulanmaktadır. Burada anlatılan 3B yeniden oluşturma için izlenecek yöntemin akış diyagramı Şekil 2.9’da verilmiştir.

Sonraki alt bölümlerde kullanılan yöntemlere ait bilgiler verilmektedir. İlk olarak bu tez çalışmasında kullanılan öznitelik tanımlama ve eşleştirme yöntemleri, ardından ICP algoritması ve son olarak küresel eniyileme yöntemi ele alınmıştır.

2.4.1 Öznitelik çıkarma ve eşleştirme yöntemleri

Öznitelik çıkarma ve eşleştirme görüntü çiftleri arasındaki eşleşen noktaların bulunmasını sağlamak için çalışılan konulardandır. Öznitelik çıkarma, görüntü içerisinde önemli ve ayırt edici nokta/bölgelerin belirlenmesi olarak ifade edilebilir. Öznitelik eşleştirme ise bu öznitelik nokta/bölgelerinden çıkarılacak olan tanımlayıcı bilgilerdir. Bu bilgiler, görüntüde oluşabilecek olan bakış açısı, açısal değişimler, gürültü ve boyut değişimleri gibi bozucu etkiler altında da özgün ve doğru eşleştirmeyi sağlayacak nitelikte olmalıdır. Bu bölümde, üç farklı görüntü eşleştirme (SIFT, SURF ve ORB) tekniğinden bahsedilecektir.

2.4.1.1 SIFT

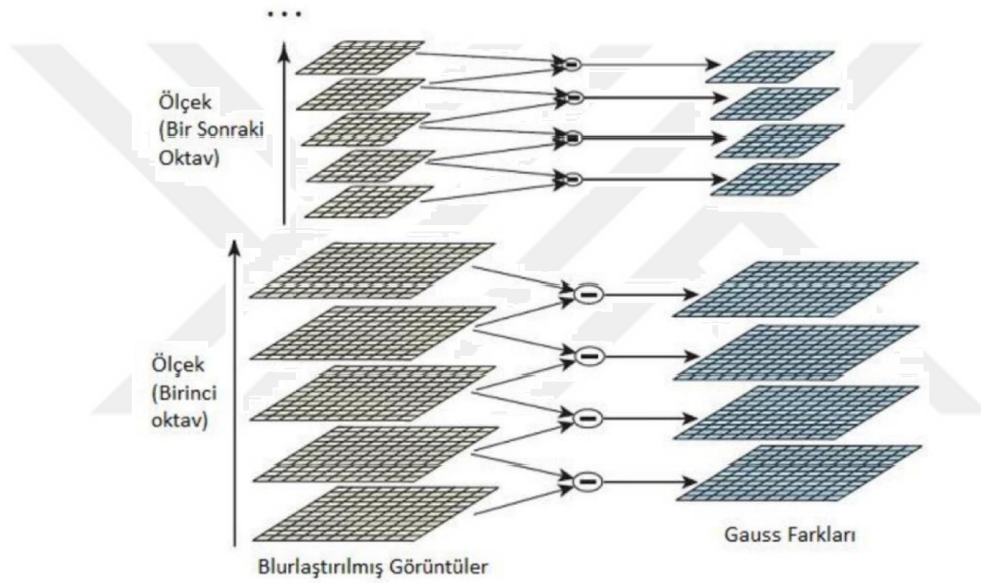
Ölçekten bağımsız öznitelik dönüşümü (SIFT - Scale Invariant Feature Transform) algoritması 2004 yılında Lowe tarafından geliştirilmiştir (Lowe, 2004). SIFT algoritması, görüntüde açıdan ve ölçekten bağımsız öznitelikler üretmeyi amaçlamaktadır. Önerilen bu algoritma, açıdan bağımsızlığı sağlamak için seçilen piksel çevresindeki piksellerin merkez piksele göre eğim yönelimleri ve eğim büyüklükleri kullanmaktadır. Ölçek bazlı değişimlere karşı güvenilirlik sağlamak için ise ölçek uzayı olarak adlandırılan, görüntünün farklı boyutlara ölçeklenmesi ile elde edilen görüntü havuzundan faydalanmıştır. SIFT algoritması dört adımdan oluşmaktadır: Ölçek uzaydaki uç noktaların tespiti, öznitelik konumlanması, yönelim belirlenmesi ve öznitelik tanımlayıcısı.

İlk adım olan ölçek uzaydaki uç noktaların tespiti aşamasında öncelikli olarak görüntü farklı ölçeklerde yeniden boyutlandırılmaktadır. Bir sonraki adımda ise farklı ölçeklerdeki görüntüler Gauss yumuşatma filtresi ile bulanıklaştırılır. Bu durumun matematiksel gösterimi Eşitlik 2’de gösterilmiştir.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (2)$$

Bu ifadede I giriş görüntüsünü, G Gauss konvolüsyon çekirdeğini ve L ise bulanıklaştırılmış görüntüyü temsil etmektedir. σ ise Gauss konvolüsyon çekirdeğinin ölçek katsayısıdır ve Gauss konvolüsyonu çekirdeğini oluşturmak için kullanılan denklem Eşitlik 3’de verilmiştir.

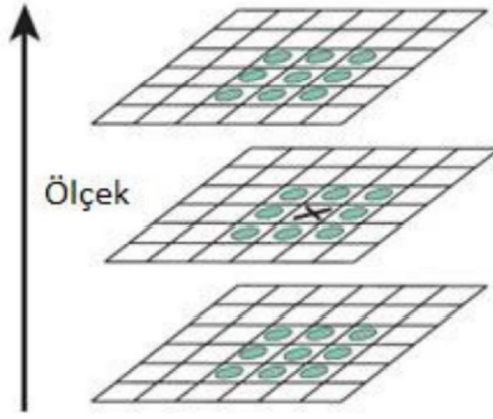
$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (3)$$



Şekil 2.10. Gauss farkları işlemi

Farklı ölçeklerdeki bulanıklaştırılmış görüntülerin sıralı olarak farkları alınarak Gauss farkları (Difference of Gaussian, DoG) işlemi gerçekleştirilir. Bu işlem Şekil 2.10’de gösterilmiştir.

Özellik noktaları (uç noktalar), görüntülerin çeşitli ölçeklerinde DoG havuzundan ölçekler arasındaki en yüksek değer kullanılarak tespit edilir ve bu işlem Şekil 2.11’de gösterilmiştir.



Şekil 2.11. Uç noktaların tespiti

Belirlenen uç noktalar arasından çok düşük ve çok yüksek genliğe sahip olanlar elenerek işleme alınmamaktadır. Bunun sebebi düşük bilgi ve aşırı doyuma ulaşmış noktalardan alınan bilgilerin güvenilirmez olarak görülmesidir. Uç noktalar bulunduktan sonra her bir nokta için konum, ölçek, yönelim gibi bilgiler hesaplanır. Bu bilgiler bize öznitelik noktalarını, bu noktadaki özniteliğin boyutunu ve son olarakta özniteliğin hangi açıya sahip olduğunu göstermektedir.

Bu bilgiler arasında en önemlisi yönelimlerin belirlenmesidir. SIFT algoritması her bir öznitelik bölgesi için yönelimleri hesaplamaktadır. Yönelim hesaplanması için ilgili bölgede yönelim büyüklüklerinin histogramı oluşturulur. Bu histogramın boyutu 36 olarak seçilmiştir ve her bir histogram değeri 36 derecelik bir açıyı kaplayan değerleri belirtmektedir. Ardından bu histogram, içinde en yüksek büyüklük değerini içeren histogram noktası ilgili bölgenin açısını vermektedir. Bu açısal değer yanında açıdan bağımsızlığı arttırmak adına tepe değerinin %80'inden büyük olan iki açısal değer de yönelim açısı olarak ele alınmaktadır. Bir öznitelik bölgesi için toplamda en fazla üç farklı açı atanabilmekte ve her açı değeri için ayrı bir tanımlayıcı hesaplanmaktadır.

Bulunan her bir öznitelik bölgesi ve her bir bölgenin sahip olduğu yönelim açısı kadar tanımlayıcı hesaplanmaktadır. Tanımlayıcılar, yönelim açısı sıfırlanarak açısal değişimlerden bağımsız hale getirilmektedir. Ölçekten bağımsızlık ise uç noktanın yer aldığı ölçek uzayından belirlenmektedir. Böylece elde edilecek değerler hem ölçek hemde açıdan bağımsız olacaktır. Tanımlayıcı çıkarımında ise ölçeğe bağlı olarak seçilen bir boyutta dairesel bir pencere için

hesaplama yapılmaktadır. Ayrıca, tanımlayıcı değerleri, ışık değişimlerinin daha az etkilenmesi için normalize edilmiş ve ağırlıklandırılmış bir görüntü penceresinden hesaplanmaktadır. Tanımlayıcılar, öznitelik bölgesinin parçalara ayrılarak her bir parça içerisinde yönelim büyüklük histogramlarının hesaplanması ve bu değerlerin art arda eklenmesi ile elde edilmektedir.

Böylece elde edilecek olan tanımlayıcı vektörler, başka görüntü bölgelerinden hesaplanmış tanımlayıcı vektörler ile karşılaştırılarak bir görüntü parçasının başka bir görüntü parçası ile eşleştirilmesini belirlemek için kullanılabilir.

2.4.1.2 SURF

SIFT algoritmasının öznitelik tespit ve eşleştirme konularında göstermiş olduğu yüksek performans bu algoritmayı referans hale getirmiştir. Bir çok farklı uygulama ve yaklaşımda bu algoritmada gerçekleştirilen işlemler referans alınmıştır. Bu çalışmalar içerisinde en önemlilerinden bir tanesi de SURF algoritmasıdır. SURF, SIFT algoritmasının hızını arttırmak ve çeşitli görüntü bozulmalarına karşı daha gürbüz bir algoritma elde etmek için önerilmiş bir algoritmadır. SURF algoritmasında, SIFT’te kullanılan adımlara farklı yaklaşımlar ile alternatif üretilmiştir.

Bu yaklaşımlardan ilki, ölçek uzayını oluşturmak için SIFT algoritmasında kullanılan DoG yönteminin aksine kutu filtreler kullanmaktır. Kutu filtreler, integral görüntü (integral image) kullanılarak çok hızlı bir şekilde hesaplanabilir. Ayrıca SURF, boyut uzayını oluştururken görüntü ölçeğini küçültmek yerine filtre boyutunu büyütme ve böylece boyutlandırma ve tekrar integral görüntü oluşturma adımlarında kaybedilecek zamandan tasarruf etmektedir. Elde edilen bu boyut uzayından öznitelik noktalarının tespiti için Hessian görüntü bölgesi bulma tabanlı bir yaklaşım kullanılmıştır. Yönelim hesabında ise yatay ve dikey yönde dalgacık cevabı (wavelet response) gauss ağırlıkları ile birlikte kullanılmaktadır. Bu dalgacık cevabından elde edilen bilgi aynı zamanda öznitelik tanımlayıcısı oluşturulurken de kullanılmaktadır. SIFT’e benzer şekilde öznitelik bölgesi alt parçalara bölünmekte ve dalgacık cevapları öznitelik olarak sıralanmaktadır. Bunlara ilave olarak öznitelik noktalarının tespiti sırasında hesaplanan kenar cevabının işareti göz önünde bulundurulmakta ve sadece eşleşen işarete sahip görüntü parçalarının eşleştirilmesine çalışılmaktadır. Böylece işlem hızı da arttırılmış olmaktadır.

2.4.1.3 ORB

ORB, algoritması FAST (Rosten ve Drummond, 2006) öznitelik tespit ve BRIEF (Hutchison vd., 2010) tanımlayıcılarının biraz değiştirilmiş versiyonlarından oluşturulmuştur (Rublee vd., 2011). İlk olarak öznitelik noktalarını FAST ile belirledikten sonra bu noktalar için Harris köşe değeri hesaplanmaktadır. Bu değerler arasından istenilen N tanesini çekmek için sıralama yapılır. FAST algoritmasında yönelim hesaplanmadığı için bulunan öznitelik noktaları açı değişimlerine bağımlıdır. ORB algoritması bu problemin çözümü için öznitelik bölgesinde ağırlıklandırılmış ağırlık merkezini hesaplamaktadır. Bu noktanın merkez nokta ile yaptığı açı ise öznitelik bölgesinin yönelim değeri olarak alınmaktadır. Bu yönelim açısı kullanılarak bir rotasyon matrisi oluşturulur ve BRIEF öznitelikleri bu matrise göre yönlendirilir. İşlem yükü bakımından ORB, incelenen diğer öznitelik tanımlama algoritmalarına göre daha hafif bir algoritmadır.

2.4.2 Yinelemeli en yakın nokta (ICP)

Yinelemeli en yakın nokta algoritması (Besl ve McKay, 1992), iki nokta bulutu arasındaki uzaklığı en aza indirmeyi amaçlamaktadır. Bu nokta bulutlarından bir tanesi sabit (referans nokta bulutu) tutulmaktadır. Diğer nokta bulutunda yer alan bütün noktalar veya seçilmiş bir kısmı için referans nokta bulutunda bir eşleştirme yapılır. Bu eşleşen noktalar kullanılarak bir dönüşüm matrisi hesaplanır. Bu hesaplamada bütün eşleşmelerin kullanılması doğru olmayabilir. Yanlış eşleşen noktaların hesaba alınmasını engellemek için noktaların ağırlıklandırılması veya aykırı noktaların elenmesi gerekir. Bulunan bu dönüşüm matrisi kullanılarak sabit olmayan nokta bulutunda yer alan noktaların konumları değiştirilir. Ardından bu işlemler belirli bir iterasyon sayısı yada belirli bir hata değerine ulaşana kadar tekrar edilir.

ICP algoritmasında eşleştirmek için noktalar seçilirken hıza veya başarımla göz önünde bulundurularak farklı yöntemler kullanılabilir. Bu yöntemlerden bazıları; Tüm noktaların kullanılması (Besl ve McKay, 1992), var olan noktaların üniform yeniden örneklenmesi (Turk ve Levoy, 1994), rastgele örnekleme (Her iterasyonda farklı sayıda nokta seçimi ile) (Masuda vd., 1996), renk yada yoğunluk bulunduran imgeler için yüksek yoğunluklu noktaların seçilmesi (Weik, 1997), belirtilen yöntemlerin bir nokta bulutu yada her iki nokta bulutu için uygulanması (Godin vd., 1994) sayılabilir. ICP de kullanılan en yakın

nokta hesabı, hata minimizasyonu ve kapalı eğriler oluşturma teknikleri üzerinde de bir çok çalışma mevcuttur.

Bu tez kapsamında, ICP algoritması, görüntü eşleştirme ile yapılan ilk dönüşümünden sonra daha hassas bir hizalama elde edebilmek için kullanılmıştır. ICP içerisinde yanlış eşleştirmelerin dönüşümde hataya sebep olmasını engellemek için RANSAC (Fischler ve Bolles, 1981) algoritması kullanılarak 3B aykırı nokta tespiti yapılmaktadır.

2.4.3 Küresel eniyileme

Ardışıl RGB-D veriler ile hesaplanan kamera konumları ve 3B nesne modeli, her yeni veri ile oluşacak artımlı hatalardan kaynaklanan hatalı hesaplamalar içermektedir. Şekil 2.12’te oluşturulmuş bir 3B nesnede bu etkiden kaynaklanan hatalı kısım gösterilmiştir.



Şekil 2.12. 3B nesne yeniden oluşturma birleştirme hatası

Her yeni gelen veri ile ufakta olsa gerçekleşen bu hatalar toplanarak modelin bozuk bir biçimde elde edilmesine yol açmaktadır. Bu hatalar, küresel bir eniyileme yöntemi ile kamera konumlarının yeniden hesaplanmasından sonra modelin tekrar hizalanması ile giderilebilir. Bunun yapılabilmesi için ardışıl sahnelerden alınan görüntülerin bir noktada döngüyü kapatmaya olanak tanınması gerekmektedir. Döngünün kapatılması, daha önce görülen bir

sahnenin en son gelen sahne ile aynı bölgeyi içeren veriye sahip olması ile mümkün olmaktadır. Böylece bu iki sahne arasında oluşan hata hesaplanabilir ve büyük oranda iyileşme sağlanabilir. Döngünün kapatıldığı sahneler arasında hesaplanan hata, döngü içerisinde yer alan sahneler arasında eşit dağıtılarak kamera konumları ve 3B model yeniden oluşturulmaktadır. Böylece sonuç olarak elde edilecek olan model daha az hata içerecektir. Küresel eniyileme yönteminin sonuçları iyileştirmesi beklense de işlem yükü bakımından ağır bir algoritmadır.

BÖLÜM III

BENZETİM ÇALIŞMASI ve SONUÇLAR

Bu tez kapsamında (Xie, 2015/2019) tarafından sunulan c++ tabanlı program çatısı kullanılmıştır. Bu program çatısı dönüşümler ve hesaplamalar için Point Cloud Library (PCL), Open Source Computer Vision Library (OpenCv) gibi kütüphaneler kullanmaktadır. SIFT, SURF, ORB öznelik tespit ve tanımlayıcıları kullanılarak görüntüler arası eşleştirme sayıları ve görsel sonuçlar ele alınarak elde edilmiştir. Ayrıca, döngünün kapatılması ve küresel optimizasyon sonucunda elde edilen sonuçlar gösterilmiştir.

3.1 Veri Kümesi

Bu tez çalışmasında Kinect V1 ile oluşturulmuş Washington RGB-D veri kümesi kullanılmıştır. Bu veri kümesi meyve, sebze, kutu, şişe vb. 300 objenin yer aldığı ve bu objelerin kategoriler ve örnekler olmak üzere sınıflandırılmış olarak sunulduğu bir çok veri içermektedir. Örneğin veri kümesinde yer alan “elma” kategorisi; kırmızı, yeşil, sarı vb. örneklerle sahiptir. Her örnek için algılayıcının 30,40 ve 60 derece açı ile objeye baktığı ve objenin etrafında tam tur attığı görüntülerden elde edilmiş RGB-D görüntüler bulunmaktadır. Bu veri kümesinde yer alan örnek nesneler Şekil 3.1’de gösterilmiştir.



Şekil 3.1. Washington RGB-D veri kümesi örnek nesneleri (Lai vd., 2011)

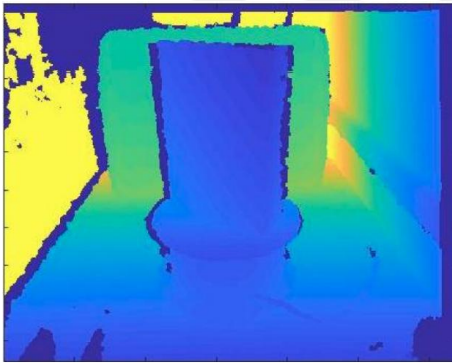
Washington RGB-D veri kümesi döner tabla üzerine yerleştirilmiş nesnelerin belirli açılarda döndürülmesi ile elde edilen veriler içermektedir. Bu verilerden farklı veri kümeleri oluşturulmuştur. Bunlar hem tam çözünürlükte sahneler hem de sadece nesneler görünecek şekilde sahnelerin kesildiği kümelerdir. Şekil 3.2’de tam çözünürlükte, sahnede kesilmiş, bunlara ait derinlik görüntüleri ve bu görüntülere ait maske görüntüleri verilmiştir.



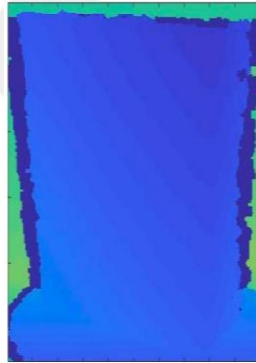
(a)



(b)



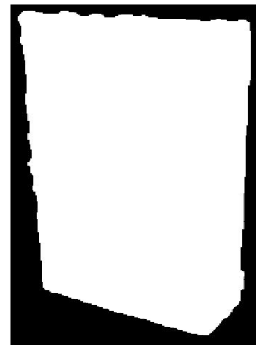
(c)



(d)



(e)

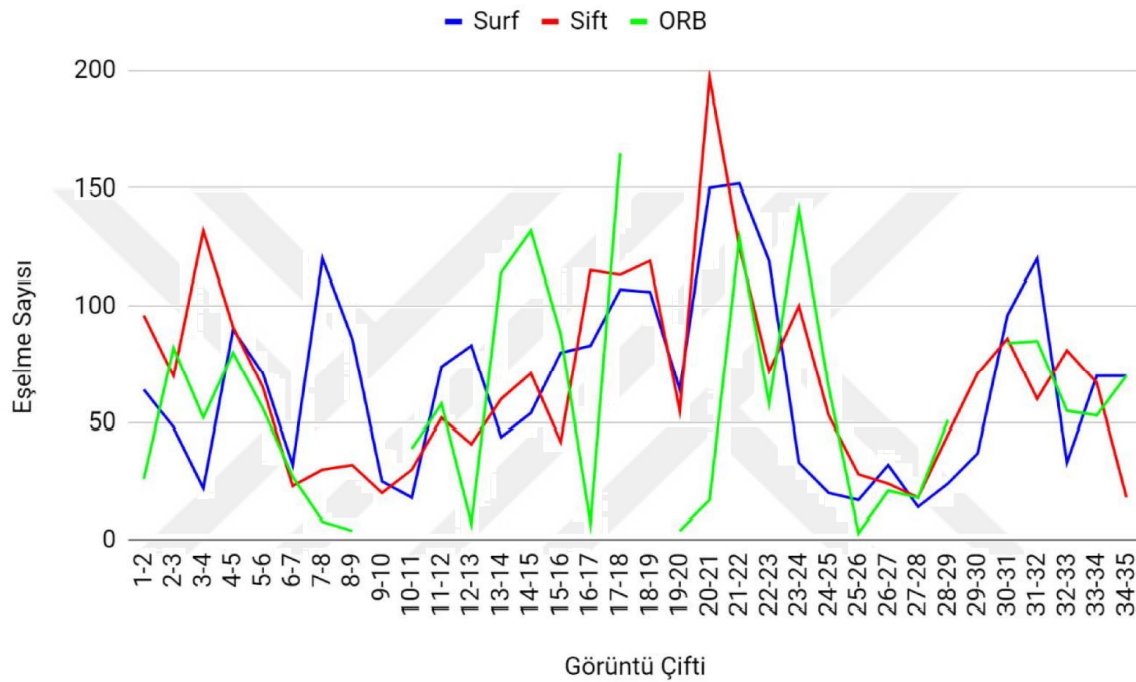


(f)

Şekil 3.2. Washington RGB-D veri kümesi: tam çözünürlük sahne görüntüsü (a), nesnenin sahneden kesilmiş görüntüsü (b), tam çözünürlük derinlik görüntüsü (c), nesnenin sahneden kesilmiş derinlik görüntüsü (d), tam çözünürlük görüntü için nesne maskesi (e) ve nesnenin sahneden kesilmiş görüntüsü için nesne maskesi (f)

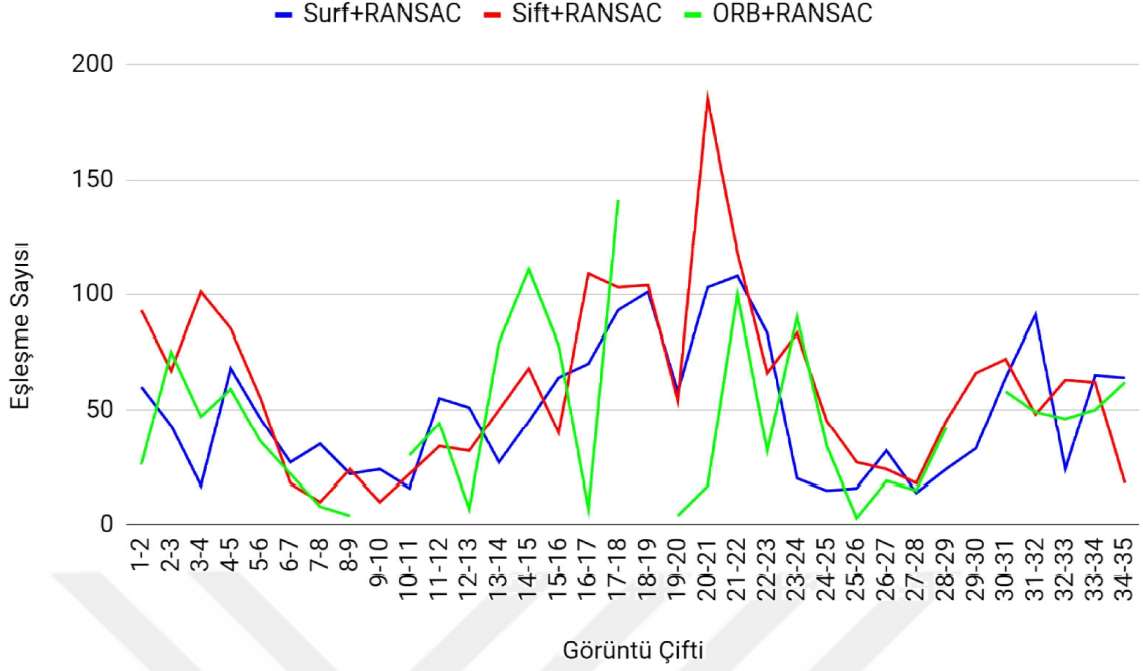
3.2 Benzetim Sonuçları

İlk olarak görüntüler arasında farklı öznitelik tespit ve tanımlayıcı eşleştirme sayıları karşılaştırılmaktadır. Bunun için SIFT, SURF ve ORB algoritmalarının OpenCv kodları kullanılmıştır. Tüm algoritmaların parametreleri değiştirilmeden kaynak değerleri kullanılmıştır. İlk olarak bu algoritmalarından elde edilen ham eşleştirme sayıları Şekil 3.3’de gösterilmiştir.



Şekil 3.3. Washington RGB-D veri kümesinde sadece görüntü çiftleri kullanılarak elde edilen eşleştirme sayıları

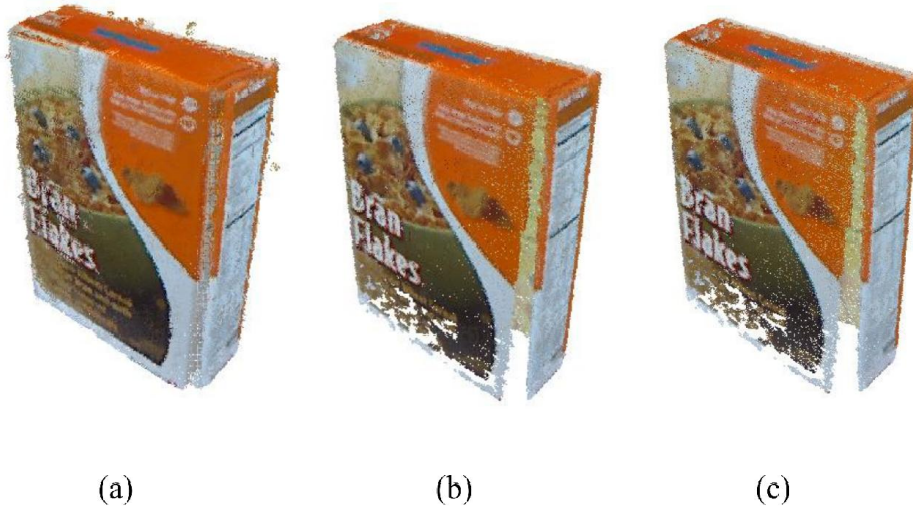
Şekilden görüldüğü gibi farklı konumlardan alınan görüntülerde eşleştirme performansları değişkenlik göstermektedir. Fakat, SIFT ve SURF algoritmaları her görüntü çifti için sırasıyla 18 ve 14 eşleştirme elde edilirken, ORB özellikle tam 90 derecelik dönmelerin olduğu durumlarda hiç eşleştirme bulamayıp veri kaybına yol açmaktadır. Bunun yanında SIFT, SURF ve ORB sırasıyla ortalama 67.7, 66.4 ve 58 adet eşleştirme elde etmiştir. Fakat bu eşleştirmeler içerisinde hatalı eşleştirmeleri de barındırdığından aykırı olan eşleştirmelerin elimine edilmesi gerekmektedir. Bunun için RANSAC algoritması kullanılarak bu özniteliklerin ne kadar güvenilir olduğuna bakılmıştır. RANSAC uygulanan öznitelik eşleştirmeleri Şekil 3.4’te verilmiştir.



Şekil 3.4. Washington RGB-D veri kümesinde sadece görüntü çiftleri kullanılarak elde edilen eşleştirmelerin RANSAC ile azaltılması sonucu elde edilen yeni eşleştirme sayıları

Şekilden görüleceği üzere neredeyse bütün eşleştirme sayılarında bir azalma olmaktadır. Burada yer alan veriler incelendiğinde SIFT+RANSAC, SURF+RANSAC ve ORB+RANSAC ile elde edilen en az eşleştirme sayıları sırasıyla 10, 14 ve 3'tür. SURF ve ORB için en az eşleştirme sayılarında bir değişiklik ortaya çıkmamış fakat SIFT için 18'den 10'a düşmüştür. Bunun yanında ortalama eşleştirme sayıları aynı sırayla 59.3, 49.3 ve 45 olarak ortaya çıkmaktadır. Eşleştirmelerin ortalama elenen sayıları 8.3, 17 ve 13 olarak elde edilmiştir. Buradan SIFT özneliklerinin diğerlerine göre daha güvenilir olduğu anlaşılmaktadır.

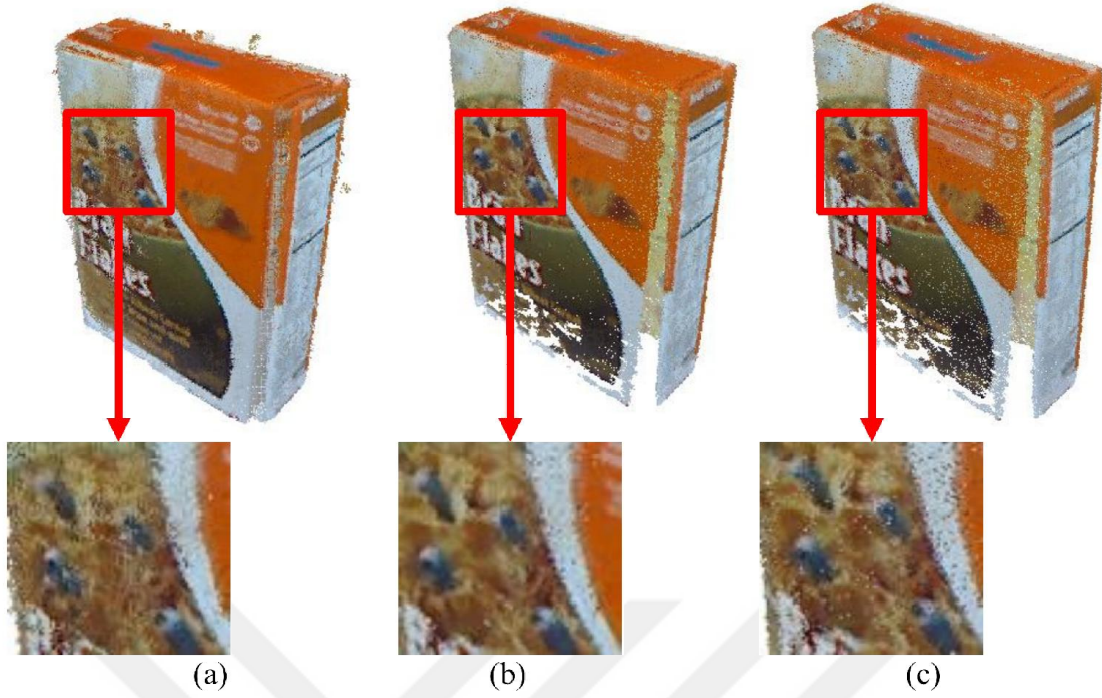
Bu öznelikler kullanılarak 3B yeniden oluşturma yöntemi her bir öznelik tanımlayıcı için küresel eniyileme olmadan çalıştırılmıştır. Elde edilen örnek görüntüler Şekil 3.5'te verilmiştir.



Şekil 3.5. Küresel eniyileme olmadan elde edilen 3B nesne modeli SIFT+RANSAC (a), SURF+RANSAC (b) ve ORB+RANSAC (c)

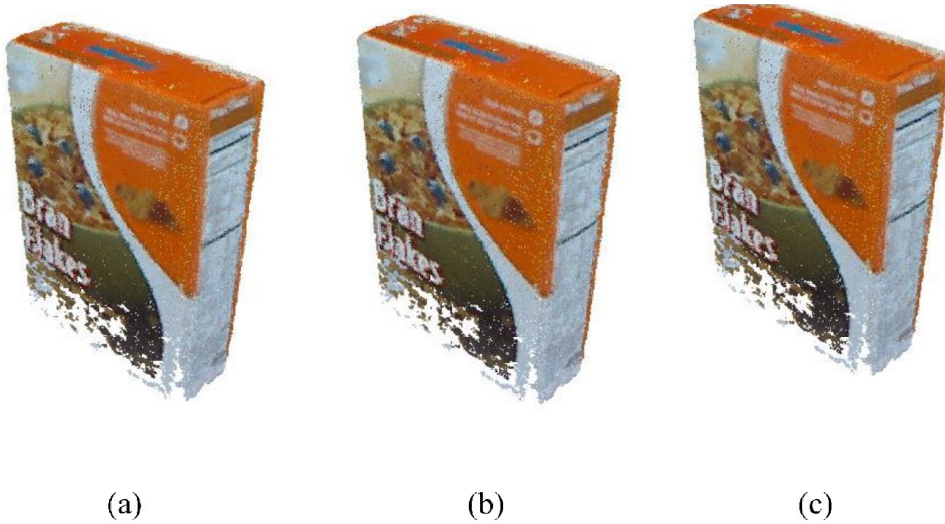
Şekil 3.5 incelendiğinde SIFT+RANSAC ile elde edilen 3B modelin en az hataya sahip olduğu görülmektedir. ORB+RANSAC ve SURF+RANSAC ise öznel eşleştirmelerinin daha çok olduğu önden bakış açısında daha düzgün bir model oluşturulmasını sağlamıştır. Ayrıca SURF+RANSAC ile bulunan eşleştirmelerin önden bakış açısının olduğu görüntülerde en iyi performansı gösterdiği açıktır. Bu çıkarım nesnenin ön yüzü incelendiği zaman daha net görülmektedir (Şekil 3.6).

Şekil 3.6 incelendiğinde SURF+RANSAC'tan elde edilen nokta bulutu hem yoğun hem de diğer yöntemlerden daha doğru bir birleştirme gerçekleştirmiştir. ORB+RANSAC'tan elde edilen sonuç ise daha az nokta içermekte fakat doğru birleştirme açısından SIFT+RANSAC'a göre daha iyi durumdadır. Fakat SIFT+RANSAC yoğun bir nokta bulutuna sahiptir.



Şekil 3.6. Nesnenin ön yüzüne ait yakın gösterim görüntüler SIFT+RANSAC (a), SURF+RANSAC (b) ve ORB+RANSAC (c)

Elde edilen bu sonuçları iyileştirmek için daha önceki bölümde anlatılan küresel eniyileme yöntemi eklenerek döngünün kapatılması ve hataların en aza indirilmesi amaçlanmıştır. Sonuçlar Şekil 3.7’den görülebilmektedir.



Şekil 3.7. Küresel eniyileme ile elde edilen 3B nesne modeli SIFT (a), SURF (b) ve ORB (c)

Bu görüntüden açıkça daha başarılı sonuçlar elde edildiği görülmektedir. Son sahne ile ilk sahne döngüyü kapatan görüntüler içermektedir. Dolayısıyla tüm yöntemlerden elde edilen

sonular benzer ıktılar elde edilecek ekilde iyileştirilmiřtir. Genel grnt itibariyle kresel eniyileme bu veri kmesinde olduka bařarılı sonular vermektedir. Bunun bir sebebi de grntler arası aısal deėiřimlerin uniform olmasıdır. Eėer aısal deėiřimler uniform olmaz ise elde edilen sonular znitelik yntemlerine daha baėımlı olacaktır. Daha detaylı bir inceleme yapabilmek iin řekil 3.8 ile verilen grntler incelenebilir.



řekil 3.8. Kresel eniyileme ile elde edilen 3B nesne modeli SIFT (a), SURF (b) ve ORB (c)

řekil 3.8’den grldė gibi kresel eniyileme sonucunda elde edilen sonular benzer olsa da SIFT yntemi daha yoėun bir nokta bulutu oluřturulmasını saėlamıřtır. SURF algoritması ile ise ORB’ye gre daha yoėun nokta bulutu elde edilmiřtir.

BÖLÜM IV

SONUÇLAR

Bu tez çalışmasında, Kinect algılayıcısı kullanılarak RGB-D veriler üzerinden 3B yeniden oluşturma yöntemleri incelenmiştir. Kinect algılayıcısı maliyet açısından erişilebilir bir cihazdır. Aynı zamanda renkli görüntünün yanında derinlik bilgisi de sağladığından 3B yeniden oluşturma için çok elverişli veriler sunmaktadır. Kinect ile toplanan veriler gürültü ve çeşitli bozulmalar içermektedir. Özellikle ışık değişimlerinden çok etkilenmektedir. Bu tez çalışmasında, Kinect'ten elde edilen veriler 3B yeniden oluşturmada olduğu gibi kullanılamayacağından, literatürde sıklıkla kullanılan yöntemler ve etkinlikleri incelenmiş ve bir karşılaştırma ortamı oluşturulmuştur.

İlk olarak, SIFT, SURF ve ORB öznitelik tespit ve tanımlayıcıları kullanılarak ardışıl görüntüler arasındaki eşleştirmeler tespit edilmiştir. Böylelikle Kinect tarafından sağlanan nokta bulutu daha etkin bir şekilde hizalanabilmektedir. Öznitelikler kullanılarak kabaca hizalanan nokta bulutu ICP ile daha hassas bir şekilde tekrar hizalanmaktadır. Böylece elde edilen nokta bulutu artımlı bir şekilde modeli oluşturmaktadır. Fakat, oluşturulan 3B modelin halen hatalar içerdiği gözlemlenmiştir. Bu problemin iyileştirilebilmesi için sisteme döngünün kapatılması için küresel eniyileme adımı eklenmiş ve sonuçların iyileştiği gözlemlenmiştir.

Küresel eniyileme uygulanan model her öznitelik tanımlayıcısı için başarılı sonuçlar ürettiğinden en an işlem yükü içeren ORB algoritmasının kullanıldığı çözüm daha hızlı çalışabilecektir. Bunun yanında SIFT ve SURF algoritmaları ise daha yoğun model oluşturulmasına olanak sağlamıştır.

Gelecek çalışmalarımızda küresel eniyilemeye ihtiyaç duymadan da benzer sonuçlar elde edilebilmesi için yaklaşımlar denenecektir. Bunlardan bir tanesi kullanılan öznitelik tanımlayıcıların hibrit bir şekilde kullanılması ile algoritmaların bazılarının yetersiz kaldığı durumlarda, diğer algoritmalar kullanılarak başarımın artırılabilirliğinin incelenmesidir.

KAYNAKLAR

Adjouadi, M. and Candocia, F., "A stereo matching paradigm based on the Walsh transformation", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16(12), 1212-1218, 1994.

Albitar, C., Graebbling, P. and Doignon, C., "Robust Structured Light Coding for 3D Reconstruction", *2007 IEEE 11th International Conference on Computer Vision*. Rio de Janeiro, Brazil: s. 1-6, 4-21 October, 2007.

Bay, H., Tuytelaars, T. and Van Gool, L., "SURF: Speeded Up Robust Features", A. Leonardis, H. Bischof, & A. Pinz (Ed.), *Computer Vision – ECCV 2006*. Springer Berlin Heidelberg (ss. 404-417), 7-13 May, 2006.

Besl, P. J. and McKay, N. D., "A method for registration of 3-D shapes", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(2), 239-256, 1992.

Boyle, W. S. and Smith, G. E., "Charge coupled semiconductor devices", *The Bell System Technical Journal* 49(4), 587-593, 1970.

Chatterjee, A., Jain, S. and Govindu, V. M., "A Pipeline for Building 3D Models Using Depth Cameras", *Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing*. New York, NY, USA, s.38:1–38:8, 16-19 December, 2012.

Chen, C., Zou, W. And Wang, J., "3D surface reconstruction based on Kinect", *2013 IEEE Third International Conference on Information Science and Technology (ICIST)* s.986-990, 23-25 March, 2013.

Chen, K., Lai, Y.-K., Wu, Y.-X., Martin, R. and Hu, S.-M., "Automatic semantic modeling of indoor scenes from low-quality RGB-D data using contextual information", *ACM Transactions on Graphics* 33(6), 1-12, 2014.

Clark, J. J., "Active photometric stereo", *Proceedings 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Illinois, USA, s. 29-34, June, 1992.

Cochran, S. D. and Medioni, G., "3-D surface description from binocular stereo", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(10), 981-994, 1992.

Curless, B. and Levoy, M., "A volumetric method for building complex models from range images", *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques - SIGGRAPH '96*. New Orleans, LA, USA, s.303-312, 04 – 09 August, 1996.

Esteban, C. H. and Schmitt, F., "Silhouette and Stereo Fusion for 3D Object Modeling", *Computer Vision and Image Understanding* 367-392, 2004.

Fischler, M. A. and Bolles, R. C., "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", *Commun. ACM* 24(6), 381–395, 1981.

Furukawa, Y., "*High-Fidelity Image-Based Modeling*", *Technical Report* 2008.

Godin, G., Rioux, M. and Baribeau, R., "Three-dimensional registration using range and intensity information", *Videometrics III* 279-290, 1994.

Goesele, M., Curless, B. and Seitz, S. M., "Multi-View Stereo Revisited", *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2 (CVPR'06)*. New York, NY, USA: s.2402-2409, 17-22 June, 2006.

Horaud, R., Hansard, M., Evangelidis, G. and Ménier, C., "An overview of depth cameras and range scanners based on time-of-flight technologies", *Machine Vision and Applications* 27(7), 1005-1020, 2016.

Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Mattern, F., Mitchell, J. C. and Fua, P., "BRIEF: Binary Robust Independent Elementary Features", *Computer Vision –*

ECCV 2010 Heraklion, Crete, Greece, 5-11 September, 2010

Izadi, S., Davison, A., Fitzgibbon, A., Kim, D., Hilliges, O., Molyneaux, D. ve Freeman, D., "KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera", ***Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11***. Santa Barbara, California, USA: s. 559, 16 - 19 October, 2011.

Jaiswal, M., Xie, J. and Sun, M.-T., "3D object modeling with a Kinect camera", ***Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific***. Chiang Mai, s.1-5, 9-12 December, 2014.

Knopp, J., Prasad, M., Willems, G., Timofte, R. and Van Gool, L., "Hough Transform and 3D SURF for Robust Three Dimensional Classification", İçinde K. Daniilidis, P. Maragos, & N. Paragios (Ed.), ***Computer Vision – ECCV 2010*** Heraklion, Crete, Greece, 5-11 September, 2010.

Lai, K., Bo, L., Ren, X. and Fox, D., "RGB-D (Kinect) Object Dataset", <https://rgbd-dataset.cs.washington.edu/index.html>

López, R.L. and Estrada, M.A., "Iterative Closest SIFT Formulation for Robust Feature Matching", : ***Advances in Visual Computing, 2006***, Lake Tahoe, NV, USA, s.502-513 6-8 November, 2006.

Lotti, J.-L. and Giraudon, G., "Adaptive window algorithm for aerial image stereo", ***12th International Conference on Pattern Recognition***, Jerusalem, Israel, Israel, 9-13 Oct. 1994.

Lowe, D. G., "Object recognition from local scale-invariant features", ***Proceedings of the Seventh IEEE International Conference on Computer Vision***, Kerkyra, Greece, s.1150-1157, 20-27 September, 1999.

Masuda, T., Sakaue, K. and Yokoya, N., "Registration and integration of multiple range images for 3-D model construction", ***Proceedings of 13th International Conference on***

Pattern Recognition, Vienna, Austria, s. 879-883, 25-29 Aug. 1996.

Morana, M., "3D Scene Reconstruction Using Kinect", *Advances onto the Internet of Things: How Ontologies Make the Internet of Things Meaningful* 179-190, 2014.

Newcombe, R. A., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A. J., and Fitzgibbon, A., "KinectFusion: Real-time dense surface mapping and tracking", *2011 10th IEEE International Symposium on Mixed and Augmented Reality* Basel, Switzerland, s.127-136, 26-29 October, 2011.

Pan, R. and Taubin, G., "Automatic segmentation of point clouds from multi-view reconstruction using graph-cut", *The Visual Computer*, 32(5), 601-609, 2016.

Pollefeys, M., Van Gool, L., Vergauwen, M., Verbiest, F., Cornelis, K., Tops, J. and Koch, R., "Visual Modeling with a Hand-Held Camera", *International Journal of Computer Vision* 59(3), 207-232, 2004.

Remondino, F. and El-Hakim, S., "Image-based 3D Modelling: A Review: Image-based 3D modelling: a review", *The Photogrammetric Record* 21(115), 269-291, 2006.

Rooker, M., Angerer, A., Capco, J., Heindl, C., Olarra, A., Fuentes, E. and Pichler, A., "Flexible Grasping of Electronic Consumer Goods", *Robotics in Smart Manufacturing*. Porto, Portugal, s.158-169, 26-28 June, 2013.

Rosten, E. and Drummond, T., "Machine Learning for High-Speed Corner Detection", *Computer Vision – ECCV 2006*, Graz, Austria, s. 430-443, 7-13 May, 2006.

Rublee, E., Rabaud, V., Konolige, K. and Bradski, G., "ORB: An efficient alternative to SIFT or SURF", *2011 International Conference on Computer Vision*. Barcelona, Spain: s. 2564-2571, 6-13 Nov. 2011.

Song, T., Zhou, L., Ding, X. and Yi, W., "3D Surface Reconstruction Based on Kinect Sensor", *International Journal of Computer Theory and Engineering*, 567-573, 2013.

Sturm, J., Bylow, E., Kahl, F. And Cremers, D., "CopyMe3D: Scanning and Printing Persons in 3D", *Pattern Recognition* 405-414, 2013.

Tarin, M., "What is Structured Light Imaging?", <http://www.movimed.com/knowledgebase/what-is-structured-light-imaging/>

Turk, G. and Levoy, M., "Zippered polygon meshes from range images", *Proceedings of the 21st annual conference on Computer graphics and interactive techniques - SIGGRAPH '94* New York, NY, USA, s.311-318, 24 – 29 July, 1994.

Valentin, J., Vineet, V., Cheng, M.-M., Kim, D., Shotton, J., Kohli, P. and Torr, P., "SemanticPaint: Interactive 3D Labeling and Learning at Your Fingertips", *ACM Trans. Graph.* 34(5), 154:1–154:17, 2015.

Weik, S., "Registration of 3-D partial surface models using luminance and depth information", *Proceedings. International Conference on Recent Advances in 3-D Digital Imaging and Modeling* Ottawa, Ontario, Canada, s.93-100, 12-15 May, 1997.

Xie, J., "A simple demo for 3D Object Modeling based on an RGB-D camera: ClaireXie/modeling_3d", https://github.com/ClaireXie/modeling_3d

Xie, J., Hsu, Y.-F., Feris, R. S. and Sun, M.-T., "Fine registration of 3D point clouds fusing structural and photometric information using an RGB-D camera", *Journal of Visual Communication and Image Representation*, 32, 194-204, 2015.

Zalevsky, Z., Shpunt, A., Maizels, A. and Garcia, J., Method and system for object reconstruction , United States Patent, No: US20170004623A1dated: 30.03.2005, 2005

ÖZGEÇMİŞ

Berkan DEMİRCİ 01.02.1985 tarihinde KOCAELİ ilinin Gölcük ilçesinde doğdu. İlkokul beşinci sınıfa kadar Denizvler İlkokulunda öğrenim gördükten sonra Ortaokul eğitimine Değirmendere Anadolu Lisesi'nde devam etti. 1999 yılında ailesi ile Ankara ya taşınmasının ardından öğrenimine orta 3. sınıftan itibaren Süleyman Demirel Anadolu Lisesi'nde devam edip orta ve lise öğrenimini tamamladı. 2003 yılında Niğde Ömer Halisdemir Üniversitesi Elektrik Elektronik Mühendisliği bölümünde üniversite öğrenimine başladı. Üniversite öğrenimini tamamlamasının ardından 2011 yılında Niğde Ömer Halisdemir Üniversitesi Elektrik Elektronik Mühendisliği Ana Bilim Dalı'nda yüksek lisans öğrenimine başladı ve yüksek lisans öğrenimine devam etmektedir. İlgi alanı görüntü işleme, mikroişlemciler, kontrol sistemleri ve kablosuz iletişim sistemleridir

