GÖRME ENGELLİLERİN NAVİGASYON SİSTEMLERİNİ DESTEKLEMEK İÇİN NESNE HARİTASI OLUŞTURMADA YENİ BİR YAKLAŞIM

Yıldıray YILMAZ
131459001

DOKTORA TEZİ

Bilgisayar Mühendisliği Anabilim Dalı
Bilgisayar Mühendisliği Doktora Programı
Danışman: Dr. Öğr. Üyesi Erdal GÜVENOĞLU

İstanbul
T.C. Maltepe Üniversitesi
Lisansüstü Eğitim Enstitüsü
Şubat, 2022

GÖRME ENGELLİLERİN NAVİGASYON SİSTEMLERİNİ DESTEKLEMEK İÇİN NESNE HARİTASI OLUŞTURMADA YENİ BİR YAKLAŞIM

Yıldıray YILMAZ
131459001
Orcid: 0000-0003-0571-3897

DOKTORA TEZİ
Bilgisayar Mühendisliği Anabilim Dalı
Bilgisayar Mühendisliği Doktora Programı
Danışman: Dr. Öğr. Üyesi Erdal GÜVENOĞLU

İstanbul
T.C. Maltepe Üniversitesi
Lisansüstü Eğitim Enstitüsü
Şubat, 2022



JÜRİ VE ENSTİTÜ ONAYI

Bu belge, Yükseköğretim Kurulu tarafından 19.01.2021 tarihli "*Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge*" ile bildirilen 6689 Sayılı Kişisel Verilerin Korunması Kanunu kapsamında gizlenmiştir.

ETİK İLKE VE KURALLARA UYUM BEYANI

Bu belge, Yükseköğretim Kurulu tarafından 19.01.2021 tarihli "*Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge*" ile bildirilen 6689 Sayılı Kişisel Verilerin Korunması Kanunu kapsamında gizlenmiştir.

TEŞEKKÜR

Doktora eğitimim ve tez çalışmamın her aşamasında bana desteğini esirgemeyen, fikirleriyle ve bilgisiyle bana ilham kaynağı olan, zorlu problemlere karşı beni cesaretlendiren değerli danışmanım Sayın Dr. Öğr. Üyesi Erdal Güvenoğlu'na teşekkür ederim. Tez izleme jürimde yer alarak yaptıkları yorumlar ve verdikleri bilimsel katkı sayesinde araştırmamızın gelişmesine yardımcı olan Sayın Prof. Dr. Emin Murat Esin'e ve Sayın Dr. Öğr. Üyesi Fatma Sarıca'ya çok teşekkür ederim.

Hayatım boyunca her zaman beni destekleyen ve sevgisini esirgemeyen kıymetli aileme ve sevgili dostlarım Melih Ahmet Karaman, Eren İldem, Onur Aktaş ve Bora Gül'e teşekkür ederim.

Yıldıray YILMAZ

Aralık 2022

ÖZ

GÖRME ENGELLİLERİN NAVİGASYON SİSTEMLERİNİ DESTEKLEMEK İÇİN NESNE HARİTASI OLUŞTURMADA YENİ BİR YAKLAŞIM

Yıldıray YILMAZ
Doktora Tezi
Bilgisayar Mühendisliği Anabilim Dalı
Bilgisayar Mühendisliği Doktora Programı
Danışman: Dr. Öğr. Üyesi Erdal GÜVENOĞLU
Maltepe Üniversitesi Lisansüstü Eğitim Enstitüsü, 2022

2017 yılında Dünya Sağlık Örgütü (DSÖ) tarafından gerçekleştirilen araştırmaya göre dünya genelinde 253 milyon insan görme engelli olarak sınıflandırılmaktadır. 217 milyon kişinin ağır görme bozukluğu ve 36 milyon kişinin ise körlük problemi bulunmaktadır. Görme engelli navigasyon sistemleri (duyusal ikame araçları) bireylerin bağımsızlıklarını arttırmak ve sosyal hayata erişimlerini kolaylaştırmak için onların mobilizasyonunu sağlamayı hedeflemektedir. Görme engelli bireyler açısından gezinti sırasında güvenli bir şekilde hareketliliğinin sağlanabilmesi, onların bağımsız yaşamalarının önemli bir bileşenini oluşturmaktadır. Bilgisayarlı görme alanındaki gelişmeler sayesinde görme engelli bireylerin ihtiyaçları her geçen gün daha fazla karşılanabilir hale gelmiştir.

Son on sene içeresinde otonom sürüş alanında büyük ilerlemeler kaydedilmiştir. Buna karşın görme engelli bireylerin otonom navigasyonunu amaçlayan çalışmalar otonom sürüş alanındaki çalışmalarla kıyaslandığında nispeten daha az sayıda yapılmaktadır ve çözümlenememiş birçok problem bulunmaktadır. Bu tez çalışması kapsamında geliştirdiğimiz görme engelli navigasyon sistemi prototipi sayesinde otonom sürüş alanında aktif olarak kullanılan dörtgen süper piksel (stixel) temsilinin çeşitli modifikasyonlar ile görme engelli bireylerin navigasyonu için başarılı bir şekilde uygulanabildiğini göstermekteyiz.

Bu tez çalışmasında geliştirilen prototip, engel tespit ve akustik geri bildirim olmak üzere iki alt modülden meydana gelmektedir. Engel tespit modülünde dörtgen süper piksel ayrıştırmasını ve uv-disparity analizini baz alan ve gerçek zamanlı çalışan yeni bir engel tespit yöntemi önerilmiştir. Önerdiğimiz yöntem, görüntülerdeki engellerin boyutu ve konumu hakkında herhangi bir ön bilgiye ihtiyaç duymamaktadır. Akustik geri bildirim modülünde ise engel tespit modülünde tespit edilen engeller, sesli geri bildirim yoluyla görme engelli kullanıcıya iletilmektedir. Bunu sağlayabilmek için tespit edilen

engelleri çeşitli özelliklerine göre ses dalgalarına çeviren bir seslendirme şeması önerilmiştir.

Geliştirdiğimiz prototip donanım seviyesinde stereo kameraya sahip mobil telefon veya tabletler gibi yaygın kullanılan elektronik cihazlara ve basit bir stereo kulaklığa ihtiyaç duymaktadır. Bu nedenle geliştirildiğimiz prototip kolay giyilebilirdir, vücudu işgal etmez ve düşük maliyetlidir. Ayrıca önerdiğimiz engel tespit algoritması modüler ve entegre edilebilir bir yapıda olduğu için başka görme engelli navigasyon sistemlerine de entegre edilebilir veya bu sistemleri destekleyici ek bir yazılım olarak da kullanılabilir.

Anahtar Sözcükler: 1. Görme Engelli Navigasyonu; 2. Duyusal İkame; 3. Engel Tespiti; 4. Nesne Haritalandırma

ABSTRACT

NEW METHOD FOR OBJECT MAPPING TO SUPPORT NAVIGATION SYSTEMS FOR VISUALLY IMPAIRED PEOPLE

Yıldıray YILMAZ
Ph.D. Thesis
Computer Engineering Department
Computer Engineering Doctoral Programme
Thesis Advisor: Asst. Prof. Dr. Erdal GÜVENOĞLU
Maltepe University, Institute of Science, 2022

According to WHO (World Health Organization) statistics, published in 2017, more than 253 million people worldwide suffer from various visual impairment problems, with 36 million blinds and 217 million people with low vision. Navigation systems for visually impaired people (sensory substitution devices) aim to mobilize individuals to increase their independence and facilitate their access to social life. Providing secure mobilization of visually impaired people is one of the significant components of their independent living. Thanks to advances in the field of computer vision, it is possible to meet the richer needs of visually impaired people.

Over the last decade, significant progress has been made in the field of autonomous driving. On the other hand, studies dedicated to autonomous navigation of visually impaired individuals are relatively few compared to studies in autonomous driving. There are still too many unresolved problems in the domain of visually impaired navigation. In this research, we develop a prototype that shows that the Stixel approach, an actively used technique in autonomous driving, can be successfully applied with various modifications to the navigation systems of visually impaired individuals.

The prototype developed in this research consists of two submodules: obstacle detection and acoustic feedback. A new real-time obstacle detection method is proposed in the obstacle detection module. The proposed method is based on stixel representation and uv-disparity analysis. The method requires no prior information about the size or location of obstacles. The obstacles detected in the obstacle detection module are transmitted to the visually impaired user via audio feedback in the acoustic feedback module. The sonification scheme has been proposed that converts the detected obstacles into sound waves according to their various properties.

At the hardware level, our prototype only needs a mobile device (equipped with a stereo camera) and a stereo headset. Therefore, our prototype is easy to wear, non-invasive and low-cost. In addition, since the proposed obstacle detection algorithm is modular and integrable, it can be integrated into other navigation systems dedicated to

visually impaired people or can be used as additional software to support these kinds of systems.

Keywords: 1. Visually Impaired Navigation; 2. Sensory Substitution; 3. Obstacle Detection; 4. Object Mapping

İÇİNDEKİLER

JÜRİ VE ENSTİTÜ ONAYI i
ETİK İLKE VE KURALLARA UYUM BEYANI ii
TEŞEKKÜR iii
ÖZ iv
ABSTRACT vi
TABLOLAR LİSTESİ x
ŞEKİLLER LİSTESİ xi
KISALTMALAR xiii
ÖZGEÇMİŞ xv
BÖLÜM 1. GİRİŞ 1
BÖLÜM 2. LİTERATÜR ARAŞTIRMASI 6
2.1. Stereo Görme 6
2.1.1. Epipolar Geometri 8
2.1.1.1. Temel Matris (F Matris) 10
2.1.1.2. Temel Matris Tahmini 11
2.1.1.3. Görüntü Rektifikasyonu 13
2.1.2. Stereo Eşleme 15
2.1.2.1. Seyrek Eşleme 19
2.1.2.2. Yoğun Eşleme 20
2.1.2.3. Benzerlik Ölçüleri 21
2.1.2.4. Lokal Yöntemler 23
2.1.2.5. Yarı Global Yöntemler 24
2.1.2.6. Global Yöntemler 27
2.2. Engel Tespiti ile İlgili Literatür Araştırması 28
2.2.1. Aykırılık Uzayı Görüntüleri 30
2.2.2. İşgal Gridleri 31
2.2.3. Dijital Yükselti Haritaları 32
2.2.4. Dörtgen Süper Piksel (Stixel) 33
2.2.5. Engel Tespitinde Diğer Yaklaşımlar 35
2.3. Görme Engellilerin Navigasyonu ile İlgili Literatür Araştırması 37

2.3.1. Sensör Ağları Tabanlı Sistemler 38
2.3.2. Bilgisayarlı Görme Tabanlı Sistemler 42
2.3.2.1. Monoküler Kamera Tabanlı Sistemler 42
2.3.2.2. Stereo Kamera Tabanlı Sistemler 45
2.3.2.3. Derinlik (RGB-D) Kamerası Tabanlı Sistemler 51
BÖLÜM 3. MATERYAL VE YÖNTEM 54
3.1. Kullanılan Materyal 54
3.2. Önerilen Yöntem 55
3.2.1. Stereo Kamera Kalibrasyonu 56
3.2.2. Görüntülerdeki Geometrik Bozulmaların Giderilmesi 60
3.2.3. Görüntülere Rektifikasyon Uygulanması 61
3.2.4. Aykırılık Haritası Tahmini 62
3.2.5. Zemin Profilinin Belirlenmesi 64
3.2.5.1. V-Disparity Haritasının Hesaplanması 64
3.2.5.2. Sahnedeki Zemin Profilinin Tespiti 65
3.2.6. Engellerin Taban Noktalarının Tespiti 66
3.2.6.1. U-Disparity Haritasının Hesaplanması 66
3.2.6.2. Boş Alanların Hesaplanması 68
3.2.7. Engellerin Yüksekliklerinin Tespiti 71
3.2.8. Sahnedeki Engellerin Ayrıştırılması 73
3.2.9. Derinlik Çıkarımı 75
3.2.10. Sahnedeki Engel Haritasının Oluşturulması 76
3.2.11. Akustik Geri Bildirim Sağlanması 78
BÖLÜM 4. DENEYSEL SONUÇLAR 80
BÖLÜM 5. SONUÇ 96
KAYNAKÇA 99

TABLOLAR LİSTESİ

Tablo 1. Akustik geri bildirim oluşturma kuralları 78
Tablo 2. Hata Matrisi 82
Tablo 3. Geliştirilen prototipin engel tespit performansı 83
Tablo 4. Prototiplerin sunması gereken işlevsel özellikler 84
Tablo 5. Stereo kamera tabanlı sistemlerin karşılaştırması 87
Tablo 6. Özellikler için belirlenen ağırlık değerleri 88
Tablo 7. Stereo kamera tabanlı sistemlerin performans değerlendirmesi 90
Tablo 8. Tespit edilen engeller için uzaklık tahminleri ve hata oranları 93
Tablo 9. Geliştirilen sistemdeki ana aşamaların çalışma süreleri 95

ŞEKİLLER LİSTESİ

Şekil 1.1. Görme engelli bireylerin geleneksel olarak kullandıkları teçhizatlar 2
Şekil 2.1. Stereo görme geometrisi 7
Şekil 2.2. Epipolar geometri 8
Şekil 2.3. Normalleştirilmiş sekiz nokta algoritması kullanılarak epipolar geometrinin tahmin edilmesi örneği 13
Şekil 2.4. Geometrik düzeltme işleminin mantığı 14
Şekil 2.5. Geometrik düzeltme işlemi sonucu 15
Şekil 2.6. Fotogrametrik Stereo Çizici 16
Şekil 2.7. Örnek bir görüntü ve aykırılık haritası 17
Şekil 2.8. Epipolar doğru kısıdının gösterimi 19
Şekil 2.9. Referans görüntüde tanımlanan pencere ile hedef görüntü üzerinde eşleşmenin aranması 23
Şekil 2.10. Açıklık probleminin gösterimi 25
Şekil 2.11. Sensör ağları tabanlı görme engelli navigasyon sistemleri 41
Şekil 2.12. Monoküler kamera tabanlı görme engelli navigasyon sistemleri 44
Şekil 2.13. Stereo kamera tabanlı görme engelli navigasyon sistemleri 50
Şekil 2.14. Derinlik kamerası tabanlı görme engelli navigasyon sistemleri 53
Şekil 3.1. Deneyde kullanılan stereo kamera ekipmanı ve askı aparatı 54
Şekil 3.2. Kamera kalibrasyonu işleminde kullanılan satranç tahtası deseni 55
Şekil 3.3. Duyusal ikame cihazının akış şeması 56
Şekil 3.4. Piksel eğriliği 57
Şekil 3.5. Kalibrasyon işleminde kullanılan görüntülerden örnekler 58
Şekil 3.6. Kalibrasyon işlemi sonucunda görüntü çiftleri için hesaplanan ortalama yeniden yansıtma hatası grafiği 59
Şekil 3.7. Kullanılan kalibrasyon desenlerinin kameraya göre pozisyonlarının 3 boyutlu görünümü 59
Şekil 3.8. Radyal bozulmalar 60
Şekil 3.9. Teğetsel bozulma 61
Şekil 3.10. Görüntü rektifikasyonu uygulanmış bir stereo görüntü çiftinin anaglifi 62
Şekil 3.11. Deneylerde kullanılan bir görüntü ve stereo aykırılık haritası 63
Şekil 3.12. V-Disparity haritası ve orijinal sol görüntü 65

Şekil 3.13. V-Disparity haritasında zemin profilini gösteren en küçük kesi (yeşil çizgi) 66
Şekil 3.14. Orijinal sol görüntü ve görüntü için oluşturulan U-Disparity haritası 68
Şekil 3.15. U-Disparity haritasının işgal gridi olarak gösterimi 69
Şekil 3.16. U-Disparity haritasına uygulanan çizge kesme işlemi sorası ortaya çıkan en küçük kesi (yeşil çizgi) 70
Şekil 3.17. Orijinal sol görüntü ve boş alan analizi sonucu edilen boş alan maskesi (beyaz bölge) 70
Şekil 3.18. Örnek bir görüntü ve bu görüntü için elde edilen üyelik değerleri 72
Şekil 3.19. Yükseklik ayrıştırması sonucu (sarı çizgi) 73
Şekil 3.20. Görüntüde tespit edilen dörtgen süper pikseller 74
Şekil 3.21. Görme engelli kullanıcının yönelebileceği sahnedeki boş alanlar 75
Şekil 3.22. Deneyde kullanılan bazı örnek görüntüler ve bunlar için oluşturulan sahne haritaları 77
Şekil 4.1. Gerçekleştirilen deneyden örnek görüntüler 81
Şekil 4.2. Performans değerlendirmesinde elde edilen global skorlar 91
Şekil 4.3. Tespit edilen engellerin kameraya olan uzaklıkları 92
Şekil 4.4. Yürüyüş patikası üzerinde bulunan belirli uzaklıktaki engellerin görüntüden çıkarımı ve tespit edilen zemin profili. 92
Şekil 4.5. Derinlik tahmini hatası grafiği 94

KISALTMALAR

Kısaltma	İngilizcesi	Türkçesi
SAD	Sum of Absolute Differences	Mutlak Farkların Toplamı
SSD	Sum of Square Differences	Kare Farkların Toplamı
NCC	Normalized Cross Correlation	Normalleştirilmiş Çapraz Korelasyon
RANSAC	Random Sample Consensus	Rastgele Örneklem Uzlaşması
SGM	Semi-Global Matching	Yarı-Global Eşleme
DSI	Disparity Space Images	Aykırılık Uzayı Görüntüleri
DEM	Digital Elevation Maps	Dijital Yükselti Haritaları
ETA	Electronic Travel Aid	Elektronik Seyahat Teçhizatı
GPS	Global Positioning System	Global Konumlama Sistemi
RFID	Radio Frequency Identification	Radyo Frekansı ile Tanımlama
GIS	Geographic Information System	Coğrafi Bilgi Sistemi
SLAM	Simultaneous Localization and Mapping	Eşzamanlı Konumlandırma ve Haritalandırma
3D	Three-Dimensional Space	Üç Boyutlu Uzay

IMU	Inertial Measurement Unit	Atalet Ölçü Birimi
RGB	Red Green Blue	Kırmızı Yeşil Mavi
RGB-D	Red Green Blue-Depth	Kırmızı Yeşil Mavi-Derinlik
SVM	Support Vector Machines	Destek Vektör Makinesi
FPS	Frame Per Second	Saniyede Kare

ÖZGEÇMİŞ

Yıldıray Yılmaz

Bilgisayar Mühendisliği Anabilim Dalı

Eğitim

Derece	*Yıl*	*Üniversite, Enstitü, Anabilim/Anasanat Dalı*
Y.Ls.	2012	İstanbul Teknik Üniversitesi, Fen Bilimler Enstitüsü Endüstri Mühendisliği Anabilim Dalı
Ls.	2009	Bahçeşehir Üniversitesi, Mühendislik Fakültesi Bilgisayar Mühendisliği Anabilim Dalı
Lise	2004	Bilfen Koleji

İş/İstihdam

Yıl	*Görev*
2017 -	Kurucu Ortak ve Teknoloji Geliştirme Müdürü, AssistBox
2014 - 17	Yazılım Geliştirme Müdürü, Obss
2012 - 14	Yazılım Geliştirme Ekip Lideri, Agito
2011 - 12	Yazılım Mühendisi, SoftTech
2009 - 11	Yazılım Mühendisi, Yapı Kredi Sigorta

Yayınlar ve Diğer Bilimsel/Sanatsal Faaliyetler

Y. Yılmaz and E. Güvenoğlu, "Free Space Detection for Visually Impaired People", 4. Çukurova Uluslararası Bilimsel Araştırmalar Kongresi, 2020.

BÖLÜM 1. GİRİŞ

Otonom sürüş sitemleri, mobil robotlar ve görme engelli bireylerin yardımcı sistemleri gibi çeşitli alanlardaki bilgisayarlı görme tabanlı uygulamalarda sahne anlama uzun zamandır üstünde çalışılan ve önem teşkil eden bir araştırma konusudur. İnsanlar sahne anlama işlevini yerine getirmek için genellikle görme duyularını kullanırlar. Mobil robotlar, otonom arabalar ve görme engelli bireyler hareketlilik gerektiren işleri gerçekleştirmek amacıyla etraflarındaki çevre hakkında bilgi çıkarımını yapabilmek için yapay görmeye dayalı karmaşık sistemlere ve süreçlere ihtiyaç duyarlar.

2017 yılında Dünya Sağlık Örgütü (DSÖ) tarafından gerçekleştirilen araştırmaya [1] göre dünya genelinde 253 milyon insan görme engelli olarak sınıflandırılmaktadır. 217 milyon kişinin ağır görme bozukluğu ve 36 milyon kişinin ise körlük problemi bulunmaktadır. Aynı araştırmaya göre görme engelli insanların %90'ı düşük gelir düzeyine sahiptir. Ayrıca görme problemi yaşayan insanların %82'si 50 yaş ve üstü kişilerden oluşmaktadır.

Görme engelli sınıfına giren kişilerin günlük aktivitelerinde karşılaştığı birçok zorluk bulunmaktadır. Kapalı ve açık alanlarda güvenli bir şekilde gezinti yapabilmek görme engelli bireylerin karşılaştıkları en kritik problemlerden biridir. Görme engellilerin yürüyüş yollarında engeller, duran veya hareket eden araçlar, başka yayalar veya çukurlar olabilmektedir. Görme engelli bireyler, gezinti yolları üzerindeki bu tür engelleri tespit etmek ve bu engellerden kaçınmak için geleneksel olarak genellikle beyaz baston kullanmaktadırlar. Ayrıca özel eğitimli rehber köpek kullanımı da gezinti için görme engellilerin başvurduğu yollardan biridir. Son yıllarda lazerli baston, ultrasonik engel algılama gözlükleri ve konum bazlı çalışan mobil uygulamalar gibi görme engelli kişilerin mobilizasyonuna yardımcı olacak birçok elektronik yardımcı cihaz ve yazılım da geliştirilmiştir. Şekil 1.1'de görme engelli bireylerin yaygın olarak kullandıkları teçhizatlar gösterilmektedir.



(a) Baston



(b) Özel eğitimli rehber köpek



(c) Ultrasonik engel algılama gözlükleri



(d) Lazerli baston

Şekil 1.1. Görme engelli bireylerin geleneksel olarak kullandıkları teçhizatlar

Geliştirilen yardımcı teçhizat ve yazılımların temel amacı ağır görme problemi veya körlük yaşayan insanların daha fazla bağımsız ve kendi kendilerine yetebilir bir şekilde yaşamalarını sağlamaktır. Ancak geleneksel olarak kullanılan bu teçhizat ve yazılımların görme engelli bireylerin hayatını kolaylaştırıcı kısmi bir etkisi bulunmaktadır. Görme engelli bireyler için gelişmiş ve verimli bir gezinti aracı üretmek, sadece bu insanların çektiği sıkıntıları azaltarak bağımsız yaşamalarını sağlamakla kalmayacak aynı zamanda bu insanların istihdam edilmesinin de önünü açarak toplumun bu insanların yeteneklerinden faydalanabilmesine olanak sağlayacaktır. Bu nedenle görme engelli bireylerin güvenli bir şekilde gezintisine yardımcı olmak için çözümler geliştirmek, son on yılda giderek daha fazla ilgi çeken bir konu haline gelmiştir. Bu alanda yapılan çalışmalar elektronik seyahat yardım sistemleri ve konum bulma sistemleri olmak üzere iki ana gruba ayrılabilir.

Görme engelli bireylerin gezintisi için bazı araştırmacılar global konumlama sistemi veya kablosuz bağlantı sistemlerinin kullanıldığı çeşitli yöntemler geliştirmişlerdir. Her ne kadar geliştirilen bu yöntemler arasında başarılı ve ticarileşmiş çalışmalar olsa da bu tip yöntemler bireylerin gezinti rotasını belirleme veya yön bulmasını sağlama üzerine kurgulanan yöntemler oldukları için bireylerin gezinti yolu üzerindeki engeller tespit edilmemektedir. Bu nedenle global konumlama sistemi ve kablosuz bağlantı sistemlerinin kullanıldığı yöntemler görme engelli bir bireyin baston gibi geleneksel teçhizatlara olan bağımlılığının önüne geçmemektedir.

Günümüzde kamera teknolojisindeki gelişim kamera sensörlerinin küçülmesine olanak sağlamıştır. Cep telefonu ve bilgisayar gibi birçok elektronik cihaz içerisinde birden fazla kameranın bulunması standart hale gelmeye başlamıştır. Bu teknolojik gelişmeye paralel olarak görme engelli bireylerin güvenli ve bağımsız bir şekilde gezintisini sağlayabilmek için araştırmacılar bilgisayarlı görme alanında daha fazla sayıda yöntem geliştirmeye başlamışlardır. Bilgisayarlı görme, dijital görüntülerden bilgi çıkarımını amaçlayan bilimsel bir alandır. Görüntülerin içeriklerini anlamak ve işlemek için teknikler geliştirmek bilgisayarlı görme alanının konusudur. Geliştirilen teknikler genellikle insan görme sistemini taklit etmeye yöneliktir. Bilgisayarlı görme teknikleri ile görüntülerden elde edilen bilgiler; nesne tanımlama, otonom navigasyon, derinlik ölçümü, engel tespiti, arttırılmış gerçeklik ve robotik gibi birçok alanda kullanılmaktadır. Bu açıdan bakıldığında bilgisayarlı görme alanındaki yaklaşımlar sayesinde görme engelli bireylerin bağımsız ve güvenilir bir şekilde hareket etmesi için gezinti yolları üzerinde bulunan engellerin tespitinin yapılabileceği ve elde edilen bilgiler doğrultusunda görme engelli bireyleri yönlendirecek gerekli çıktıların sağlanabileceği görülmektedir.

Herhangi bir ortamın değişken doğası sebebiyle görüntü işleme ve bilgisayarlı görme tekniklerini görme engelli insanların gezintisi sırasında engel tespiti yapabilmek için kullanmak zordur. Elde edilen görüntülerdeki sahnelerin, içerisinde sadece zemin düzlemindeki engellerin bulunduğu bir kroki gibi görsel olarak tanımlanabilir bir biçimde işlenmesi gerekmektedir. Görme engelli bireylerin yürüyüş patikaları üzerinde bulunan engellerin varlığını ve boyutlarını başarılı bir şekilde tespit etmeye ve tespit edilen bu engelleri bireye başka bir duyusu üzerinden iletmeye ihtiyaç vardır.

İnsanlardaki derinlik algısını sağlayan binoküler sistemin bilgisayar sistemleri kullanılarak taklit edilmesi stereo görme probleminin konusudur. Stereo görme; bir sahnenin 2 boyutlu görüntüleri kullanılarak sahne içerisindeki 3 boyutlu bilgilerin çıkarılması işlemidir. Bir sahne içerisindeki 3 boyutlu bilgiler, aynı sahnenin iki veya daha fazla açıdan alınmış görüntülerindeki noktaların birbirlerine olan göreceli uzaklıkları kullanılarak tahmin edilebilir. Stereo görme teknikleri; sürücü destek sistemleri, robot navigasyonu, görme engelli duyusal ikame cihazları, arttırılmış gerçeklik, sanal gerçeklik ve haritalandırma gibi birçok alanda nesnelerin kameraya olan mesafelerinin tahmini için kullanılmaktadır. Stereo görüntüler kullanılarak sahnede bulunan engellerin tespiti, görme engelli bireylerin gezintisi için sıklıkla kullanılan yaklaşımlardan biridir. Stereo görme teknikleri vasıtasıyla görüntülerdeki nesnelerin kameraya olan uzaklıklarının tespit edilebilmesi, aynı zamanda nesnelerin pozisyonlarının ve boyutlarının da tespit edilebilmesine ve böylece sahnenin nesne haritasının oluşturulmasına olanak vermektedir.

Görme engelli bireylerin gezintisi için yapılan engel tespit çalışmalarında araştırmacılar stereo görme tekniklerinden sıklıkla faydalanmaktadırlar. Kamera sensörü dışında ek olarak derinlik ölçüm sensörü kullanan veya aktif stereo görme tekniklerinin kullanıldığı çalışmaların başarı seviyelerinin yüksek olduğu görülmektedir. Ancak derinlik ölçüm sensörü veya aktif stereo görme tekniklerinin kullanıldığı yöntemlerin birçoğunda ek teçhizat ihtiyacı bulunmaktadır. Bu ek teçhizatlar büyüklükleri, ağırlıkları ve maliyetleri nedeniyle günlük yaşamda pratik bir şekilde kullanılabilecek cihazlar değillerdir. Ayrıca görme engelli bireylerin büyük bir çoğunluğunun düşük gelir düzeyine sahip bireyler oldukları göz önünde bulundurulduğunda, genellikle pahalı ekipmanların kullanıldığı bu tip araştırmalardaki yöntemlerin görme engelli bireyler açısından erişilebilirliği düşük seviyede kalacaktır. Günümüzde cep telefonu, tablet ve mini bilgisayarların donanım kapasitelerinin artması ve birçoğunun stereo kameralar ile donatılmış olmaları, görme engelli navigasyon sistemlerinde engel tespiti için ek teçhizat ihtiyacını ortadan kaldırabilecek gelişmelerdir. Bu nedenle bireylerin gezintisi sırasında pasif stereo görüş sistemi kullanılarak engel tespiti problemi hala üzerinde çalışma yapılması gereken aktif bir araştırma alanıdır. Bu çalışma kapsamında stereo kameraya sahip cep telefonu ve tablet bilgisayar gibi görme engelli bireylerin kolayca erişebileceği düşük maliyetli cihazlar üzerinde gerçek zamanlı olarak çalışabilecek, pasif stereo görme

tekniklerine dayanan bir duyusal ikame cihazı prototipi önerilmiştir. Geliştirilen prototip, donanım seviyesinde sadece stereo kameralı bir mobil cihaz ve stereo kulaklıktan oluşmaktadır.

Bu tez çalışmasında geliştirilen prototip, engel tespit ve akustik geri bildirim olmak üzere iki alt modülden meydana gelmektedir. Engel tespit modülünde dörtgen süper piksel (stixel) ayrıştırmasını ve uv-disparity analizini baz alan yeni bir engel tespit yöntemi önerilmiştir. Stereo görme teknikleri kullanılarak sahnedeki engellerin pozisyonu ve boyutları elde edilmiştir. Önerdiğimiz yöntem, görüntülerdeki engellerin boyutu ve konumu hakkında herhangi bir ön bilgiye ihtiyaç duymamaktadır ve gerçek zamanlı çalışmaktadır. Engellerin yüksekliklerinin tespiti için Badino, Franke ve Pfeiffer [2] tarafından önerilen dörtgen süper piksel tekniğindekinden farklı olarak yeni bir üyelik fonksiyonu önerilmiştir. Akustik geri bildirim modülünde ise tespit edilen engeller sesli geri bildirim yoluyla görme engelli kullanıcıya iletilmektedir. Bunu sağlayabilmek için tespit edilen engelleri çeşitli özelliklerine göre ses dalgalarına çeviren bir seslendirme şeması önerilmiştir. Geliştirilen duyusal ikame cihazı prototipi şehir görüntülerinden oluşan veri seti ile test edilmiştir. Geliştirilen prototipin engel tespit performansı duyarlılık, kesinlik, doğruluk ve F1-skoru metrikleriyle değerlendirilmiştir. Ayrıca Tapu ve diğerlerinin gerçekleştirdikleri tarama çalışmasında [3] kullandıkları değerlendirme yöntemi kullanılarak geliştirilen prototipin nitel değerlendirmesi yapılmıştır. Böylelikle geliştirdiğimiz prototipin literatürde öne çıkan benzer çalışmalarla sistemsel olarak karşılaştırması yapılmıştır.

BÖLÜM 2. LİTERATÜR ARAŞTIRMASI

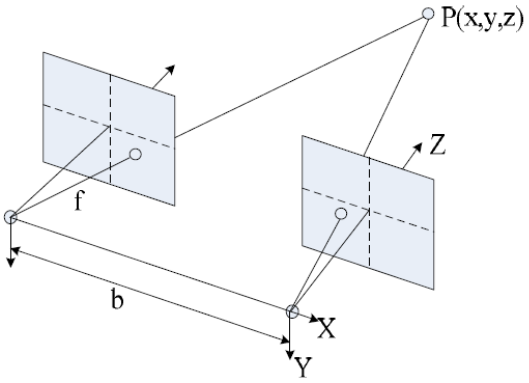
Literatür araştırmasının detayları stereo görme, engel tespiti ile ilgili literatür araştırması ve görme engelli bireylerin navigasyonu ile ilgili literatür araştırması ana başlıkları altında paylaşılmaktadır.

2.1. Stereo Görme

Stereopsis, bir sahnenin binoküler görme sistemi olan herhangi bir canlı tarafından iki ayrı gözle görülmesi ile oluşan derinlik algısıdır. Binoküler görme sistemi, gözlerin kafatasındaki farklı konumlarından dolayı aynı sahnenin iki ayrı görüntüsünü oluşturmaktadır. Sahne içerisinde bulunan her bir nokta binoküler görme sistemi sayesinde oluşan farklı görüntüler içerisinde ayrı konumlarda bulunmaktır. Binoküler aykırılık olarak isimlendirilen bu farklılıklar beynin sahnedeki derinliği hesaplamak için kullanabileceği bilgiler sağlamaktadır.

Bilgisayarlı görme teknikleri, çeşitli matematiksel yöntemleri kullanarak karmaşık bir görsel ortamı modellemeye çalışmaktadır. Bilgisayarlı görme alanının amaçlarından biri, bir veya daha fazla görüntüye dayanarak gördüğümüz dünyayı tanımlamak ve sahnenin ışıklandırma, şekil ve renk dağılımları gibi özelliklerini yeniden yapılandırmaktır. Stereo görme, bilgisayarlı görme alanındaki önemli bir araştırma problemi olan derinlik tahmini için üç boyutlu nokta koordinatlarının yeniden yapılandırılmasını amaçlayan bir alandır.

Bir stereo görme sistemi en az bir tane stereo kameradan, yani yatay veya dikey olarak yerleştirilmiş iki kameradan oluşur. Bu kameralar tarafından aynı anda yakalanan iki görüntü daha sonra görsel derinlik bilgisinin elde edilmesi için işlenir. Aralarında belirli bir mesafe bulunan iki ayrı kamera, tıpkı insan görme sisteminde olduğu gibi aynı sahneye ait noktaları her iki görüntü üzerinde ayırmaktadır. Eşleşen noktaların iki görüntü arasındaki mesafesi aykırılık (disparity) olarak adlandırılmaktadır. Eşleşen noktalar arasındaki aykırılık miktarı, derinlik hesaplaması için kullanılmaktadır. Şekil 2.1'de temel stereo görme geometrisi gösterilmektedir.



Şekil 2.1. Stereo görme geometrisi

Şekil 2.1'de kameralar arasındaki uzaklık b ve kameraların odak uzaklığı f olarak temsil edilmektedir. Bu bilgiler kullanılarak $P(x, y, z)$ noktasının 3 boyutlu uzaydaki pozisyonu aşağıdaki eşitlikler ile elde edilmektedir.

$$X = \frac{b(u_l - u_0)}{d} \tag{1}$$

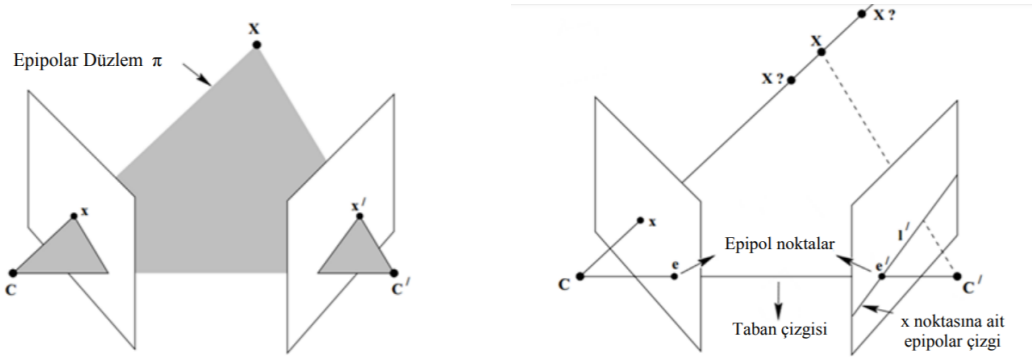
$$Y = \frac{b(v_l - v_0)\alpha_x}{\alpha_y d} \tag{2}$$

$$Z = \frac{b\alpha_x}{d} \tag{3}$$

Eşitliklerde kullanılan (u_0, v_0) terimi kamera optik merkezini, (u_r, v_r) terimi sağ kameradaki görüntü koordinatını ve (u_l, v_l) terimi ise (u_r, v_r) koordinatına sol kamerada karşılık gelen koordinatı ifade etmektedir. Aykırılık değeri $d = u_r - u_l$ eşitliği ile hesaplanmaktadır. s_x ve s_y değerleri x ve y eksenindeki piksel uzunluklarıdır. $\alpha_x = f/s_x$ ve $\alpha_y = f/s_y$ eşitlikleri kullanılmaktadır. Bu eşitlikler kullanılarak görüntülerden 3 boyutlu bilgi çıkarımı yapılabilmektedir.

2.1.1. Epipolar Geometri

Epipolar geometri, bir sahnenin farklı açılardan elde edilmiş görüntüleri arasındaki geometrik ilişkiyi tanımlamaktadır. İki ayrı kamera 3 boyutlu bir sahneyi iki ayrı konumdan görüntülediğinde, sahnedeki 3 boyutlu noktalar ve bu noktaların 2 boyutlu görüntüler üzerindeki izdüşümleri arasında belirli kısıtlamalara yol açan bir dizi geometrik ilişki bulunmaktadır.



Şekil 2.2. Epipolar geometri

Şekil 2.2'de c ve c' noktaları 3 boyutlu bir sahneyi görüntüleyen kameraların optik merkezleridir. Sahnede bulunan X noktasının birinci ve ikinci görüntü düzlemi üzerindeki izdüşümleri Eşitlik (4) ve Eşitlik (5)'deki homojen vektörler ile temsil edilmektedir. Eşleşen bu iki nokta arasındaki ilişkiyi temel matris (F matris) kurmaktadır (Eşitlik (6)).

$$x = (x, y, 1)^T \tag{4}$$

$$x' = (x', y', 1)^T \tag{5}$$

$$x'^T F\, x = 0 \tag{6}$$

Görüntü çiftinin elde edildiği kameraların optik merkezlerinden başlayarak görüntüler üzerindeki x ve x' noktalarından geçecek şekilde 3 boyutlu sahneye doğru ışınlar çizildiğinde bu ışınlar X noktasında kesişmektedir. Bu ışınların çizimi sonucunda oluşan, kameraların optik merkezleri c ve c', 3 boyutlu uzaydaki X noktası ve bu noktanın

görüntü çiftindeki izdüşümleri olan x ve x' noktalarını içeren düzleme epipolar düzlem adı verilmektedir [4].

Görüntü çiftinin elde edildiği kameraların optik merkezleri arasındaki uzaklığa taban çizgisi (baseline) adı verilmektedir. Görüntü çiftinin elde edildiği kamera merkezlerinin arasındaki uzaklık olan taban çizgisi ile görüntü düzlemlerinin kesiştiği noktalar epipol (e, e') olarak adlandırılmaktadır [5]. Epipol noktalar ile temel matrisinin çarpımı sıfıra eşittir (Eşitlik (7) ve Eşitlik (8)) [4].

$$Fe = 0 \tag{7}$$

$$F^T e' = 0 \tag{8}$$

Epipol noktalardan başlayarak görüntü üzerinde çizilen ve epipolar kısıtı sağlayan doğruya ise epipolar çizgi denilmektedir [5]. Epipolar düzlem ile görüntü düzleminin kesişimi sonucu da epipolar çizgiler oluşmaktadır. Eş noktalar eş epipolar çizgiler üzerinde yer almaktadırlar. Şekil 2.2'de birinci kamera tarafından elde edilen görüntüdeki x noktası ile eşleşen ikinci kameranın elde ettiği görüntüdeki x' noktası, l' epipolar çizgisi üzerinde olmalıdır [4]. Epipolar çizgiler matematiksel olarak aşağıdaki gibi ifade edilebilir (Eşitlik (9) ve Eşitlik (10)).

$$l' = Fx \tag{9}$$

$$l = F^T x' \tag{10}$$

Stereo eşleme işlemi sırasında epipolar kısıttan dolayı l' çizgisi, x noktasının eşdeğer noktası olan x' noktasının aranacağı epipolar çizgidir. Bu nedenle x' noktasının ikinci görüntünün tamamı içerisinde aranması yerine sadece l' epipolar çizgisi üzerinde aranması yeterlidir. Bu da arama iki boyutlu aramayı tek boyuta düşürerek arama maliyetini azaltmaktadır.

Üç boyutlu uzayda bulunan bir sahnedeki herhangi bir X noktasının görüntüsü alındığında bu noktaya ait derinlik bilgisi kaybolmakta ve üç boyutlu nokta koordinatı iki boyuta dönüşmektedir. X noktasına ait derinlik bilgisinin elde edilmesi için aynı noktaya

bakan iki farklı açıdan alınmış görüntünün olması gerekmektedir. Görüntü çiftindeki eşleşen noktalar için aşağıdaki adımlar uygulanarak derinlik bilgisi elde edilebilir [4].

1. Görüntülerdeki eşleşen nokta çiftleri kullanılarak temel matrisin hesaplanması
2. Temel matris kullanılarak kameralara ait kamera matrislerinin (P ve P') hesaplanması
3. 3 boyutlu sahnedeki X noktasının iki görüntü üzerindeki izdüşümleri olan x ve x' noktaları kullanılarak üçgenleme yöntemiyle derinlik bilgisinin hesaplanması.

P ve P' kamera matrisleri temel matris kullanılarak hesaplanabilmektedir. Birinci kameranın konumu referans alınarak ve görüntülerdeki geometrik ilişkiler kullanılarak ikinci kameranın konumu tahmin edilebilmektedir. Böylelikle ikinci kameranın birinci kameraya göre ne kadarlık bir kayma (translation) ve kaç derecelik bir dönme (rotation) hareketi yaptığı hesaplanabilir.

2.1.1.1. Temel Matris (F Matris)

Stereo görüntülerde bir görüntü içerisindeki bir noktaya karşılık gelen diğer görüntü içerisindeki epipolar çizgi temel matris kullanılarak hesaplanabilir. 3x3'lük temel matris F, eşleşen herhangi iki nokta arasındaki ilişkiyi kesin olarak ifade etmektedir. I_1 ve I_2 stereo görüntü çiftinde herhangi bir $\{(x, y) \in I_1,\ (x', y') \in I_2\}$ stereo eşleşmesi aşağıdaki eşitliği sağlamalıdır (Eşitlik (11)).

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}^T F \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0 \tag{11}$$

I_1 görüntüsündeki sabit bir (x, y) noktasına I_2 görüntüsünde karşılık gelen noktanın bulunduğu epipolar çizginin denklemi yukarıdaki eşitlikten elde edilebilir.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}^T \left(F \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \right) = 0 \tag{12}$$

Yani $\left(F\begin{bmatrix}x\\y\\1\end{bmatrix}\right)$ ifadesi ile x', y' ve 1'in katsayıları hesaplanabilmektedir. Aynı şekilde Eşitlik (12)'de (x', y') sabitlenerek I_1 görüntüsü içerisindeki epipolar çizgilerin denklemleri de elde edilebilir.

Temel matris iki kameranın farklı pozisyonda olduğu görüntü çiftleri için tanımlanmaktadır. Yani aynı kamera merkezini paylaşan görüntü çiftleri için temel matris tanımı yapılmamaktadır.

2.1.1.2. Temel Matris Tahmini

Temel matris iki görüntü arasındaki özellik noktası eşlemelerinden tahmin edilebilir. Birinci görüntü düzleminde $\{(x_1, y_1), \dots, (x_n, y_n)\}$ konumlarında ve ikinci görüntü düzleminde $\{(x_1', y_1'), \dots, (x_n', y_n')\}$ konumlarında eşleşen özellik noktalarının olduğunu düşünelim. Eşleşen her bir özellik noktası için Eşitlik (13)'deki gibi bir lineer eşitlik tanımlanabilir.

$$[x_i'x_i \;\; x_i'y_i \;\; x_i' \;\; y_i'x_i \;\; y_i'y_i \;\; y_i' \;\; x_i \;\; y_i \; 1][f_{11}\, f_{12}\, f_{13} \;\; f_{21}\, f_{22}\, f_{23} \;\; f_{31}\, f_{32}\, f_{33}]^T = 0 \qquad (13)$$

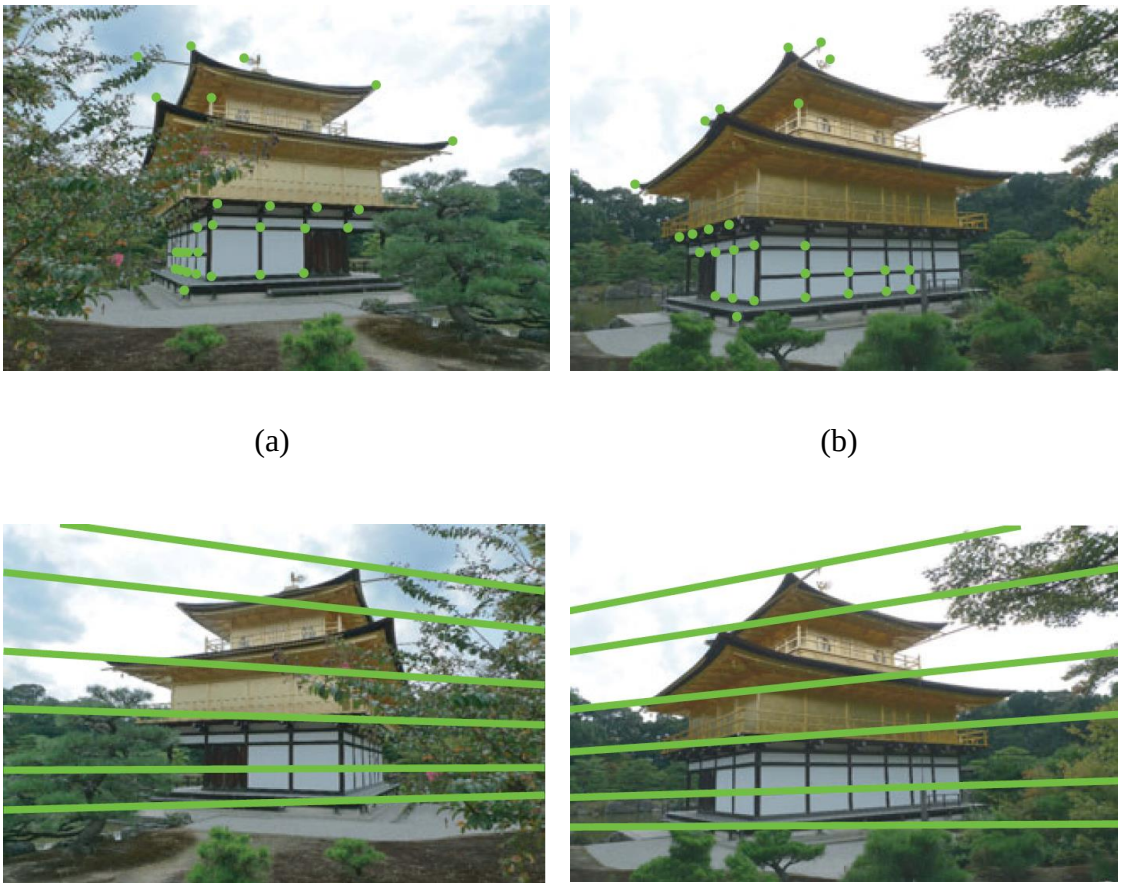
Her bir nokta için doğrusal denklemlerin toplanması $n \times 9$ bir lineer sistem üretmektedir. Bu lineer sistemin çözümü için birçok algoritma vardır. Bu lineer sistem normalleştirilmiş sekiz nokta algoritması [6] ile çözülebilir. Algoritma basamakları aşağıdaki gibidir:

1. Birinci görüntü düzlemindeki $\{(x_1, y_1), \dots, (x_n, y_n)\}$ ve ikinci görüntü düzlemindeki $\{(x_1', y_1'), \dots, (x_n', y_n')\}$ özellik noktaları algoritma girdisidir. Her bir özellik eşleşmesi kümesi ortalaması 0 ve orijinden uzaklığı $\sqrt{2}$ olacak şekilde normalize edilir. Bu normalleştirme işlemi noktaların koordinatlarına 3×3'lük T ve T'matrisleri olarak temsil edilen bir çift benzerlik dönüşümü uygulanarak gerçekleştirilebilir.
2. Her bir özellik eşlemesinin Eşitlik (13) ile bir satırını oluşturduğu $n \times 9$ boyutunda A matrisi oluşturulur.
3. A matrisinin tekil değer çözüşümü hesaplanır. Yani U sol özvektörlerden ve V sağ özvektörlerden oluşan ortogonal matrisler, D ise A matrisinin tekil değerlerini

içeren köşegen bir matris olmak üzere $A = UDV^T$ olarak çarpanlarına ayrıştırılır. D, üst soldan alt sağa doğru azalan pozitif elemanlar içeren 9×9 köşegen matris olacaktır.

4. V matrisinin son kolonu 9×1'lik f vektörü olarak alınır ve bu f vektörünün elemanları her bir satırında soldan sağa sıralanacak şekilde 3×3'lük $\hat{F}$ matrisi oluşturulur.
5. $F = T'^T \hat{F}\, T$ eşitliği kullanılarak temel matris F bulunur.

Şekil 2.3'de normalleştirilmiş sekiz nokta algoritması kullanılarak gerçek bir görüntü çifti için temel matris tahmininin (ve dolayısıyla epipolar geometrinin) sonucunu gösterilmektedir. Şekil 2.3-a ve Şekil 2.3-b aynı sahnenin farklı perspektiflerden görünümünün üzerine görüntülerdeki özellik noktalarının eşleşmelerinin eklenmiş hali gösterilmektedir. Şekil 2.3-c ve Şekil 2.3-d'de bulunan yeşil çizgiler ise temel matris tahmini ile hesaplanmış epipolar çizgilerdir.



(a) (b)

(c) (d)

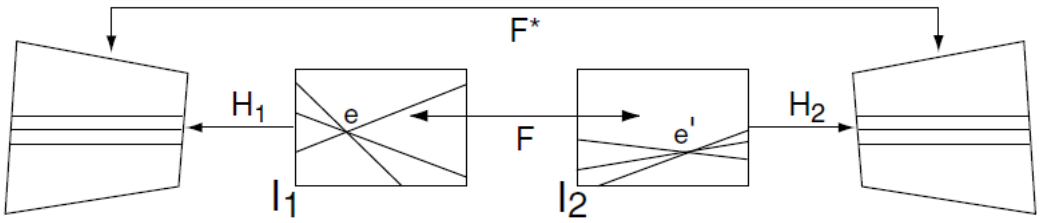
Şekil 2.3. Normalleştirilmiş sekiz nokta algoritması kullanılarak epipolar geometrinin tahmin edilmesi örneği

Hartley ve Zisserman [4] RANSAC algoritması ile özellik eşlemelerindeki aykırı değerleri tespit edip temizleyerek sekiz nokta algoritması ile daha güvenilir temel matris tahminleri yapmışlardır.

2.1.1.3. Görüntü Rektifikasyonu

Şekil 2.3'de görüldüğü gibi epipolar çizgiler genellikle eğik olarak çıkmaktadır. Epipolar çizgilerin eğik olması iki görüntü içerisindeki eşleşen noktaların tahminini zorlaştırmaktadır. Görüntüler içerisinde birbirine karşılık gelen stereo eşleşmeleri tahmin etmeden önce görüntülere geometrik düzeltme işlemi uygulamak yaygın bir yöntemdir.

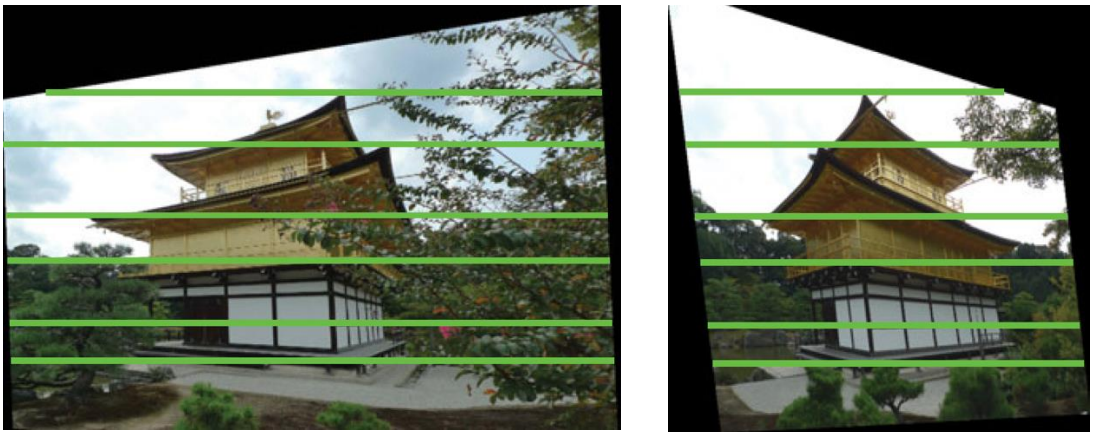
Söz konusu bu geometrik düzeltme görüntü çifti içerisindeki her bir görüntüye izdüşüm dönüşümü uygulanarak gerçekleştirilmektedir. Bunun sonucunda görüntü çiftindeki eşlenik epipolar çizgiler hizalanmış ve paralel hale gelmektedir. Böylelikle yoğun stereo eşleme yapılırken kolaylık sağlanmaktadır.



Şekil 2.4. Geometrik düzeltme işleminin mantığı

Şekil 2.4'de geometrik düzeltme işleminin mantığını göstermektedir. Geometrik düzeltmeden önce I_1 ve I_2 görüntülerindeki epipolar doğrular eğimlidir. Bu nedenle görüntüler üzerinde yoğun eşleme yapılması zordur. Görüntülere uygun izdüşüm dönüşümü uygulanarak geometrik düzeltme yapıldıktan sonra epipolar doğruların eşleşen görüntü satırlarına yerleşmesi sebebiyle yoğun eşleşme kolaylaşmaktadır. Geometrik olarak düzeltilmiş bir görüntü çiftinde epipolar doğrular paraleldir. Bu nedenle epipol noktalarının x eksenine sonsuz uzaklıktaki $(\infty, 0)$ noktasında olduğu söylenebilir.

Literatürde birçok geometrik düzeltme algoritması bulunmaktadır. Şekil 2.5'de Hartley [7] tarafından önerilen geometrik düzeltme algoritmasının Şekil 2.3'deki görüntülere uygulanmasının sonucu gösterilmektedir. Oluşan yeni epipolar doğruların yatay ve hizalanmış olduğu görülmektedir.



(a) (b)

Şekil 2.5. Geometrik düzeltme işlemi sonucu

2.1.2. Stereo Eşleme

Stereo eşleme, iki veya daha fazla görüntüdeki eşleşen piksellerin tespit edilerek bu piksellerin 2 boyutlu konumlarının 3 boyutlu derinlik bilgisine dönüştürülmesi yoluyla sahnenin 3 boyutlu modelini tahmin etme işlemidir. İlk stereo eşleme algoritmaları havadan elde edilmiş üst üste binen görüntüler kullanılarak otomatik olarak topografya haritaları oluşturmak amacıyla fotogrametri alanında geliştirilmiştir. Bunun öncesinde operatörler, her bir göze görüntülerin kaydırılmış versiyonlarını gösteren fotogrametrik stereo çiziciler kullanıyorlardı (Şekil 2.6). Tam otomatik stereo eşleme algoritmalarının geliştirilmesi fotogrametri alanında büyük bir ilerleme sağlamıştır. Bu algoritmalar sayesinde havadan elde edilen görüntülerin çok daha hızlı ve daha az maliyetli bir şekilde işlenmesi mümkün kılınmıştır [8, 9].



Şekil 2.6. Fotogrametrik Stereo Çizici

Bilgisayarlı görme alanında stereo eşleme konusu, üzerinde en çok çalışılan ve en temel problemlerden biri olmuştur [10-15] ve hala aktif araştırma alanlarından biri olmaya devam etmektedir. Stereo eşleme problemi, fotogrametri alanında esas olarak havadan elde edilen görüntülere odaklanırken, bilgisayarlı görme uygulamalarında ise insan görme sisteminin modellenmesine [16], robotik navigasyon [17-19], görme engelli bireylerin navigasyonu, 3 boyutlu modelleme ve otonom araçların geliştirilmesi gibi birçok konuya odaklanmaktadır.

Stereo eşleme işleminde eşleşen noktaların iki görüntü arasındaki mesafesi aykırılık (disparity) miktarını vermektedir. Aykırılık miktarlarının görüntü düzleminde temsil edilmesine aykırılık haritası (disparity map) adı verilmektedir. Şekil 2.7'de örnek bir görüntü ve o görüntüye ait elde edilen aykırılık haritası gösterilmektedir. Eşleşen noktalar arasındaki aykırılık miktarı derinlik hesabı için kullanılmaktadır.



(a) Orijinal görüntü (b) Aykırılık Haritası

Şekil 2.7. Örnek bir görüntü ve aykırılık haritası

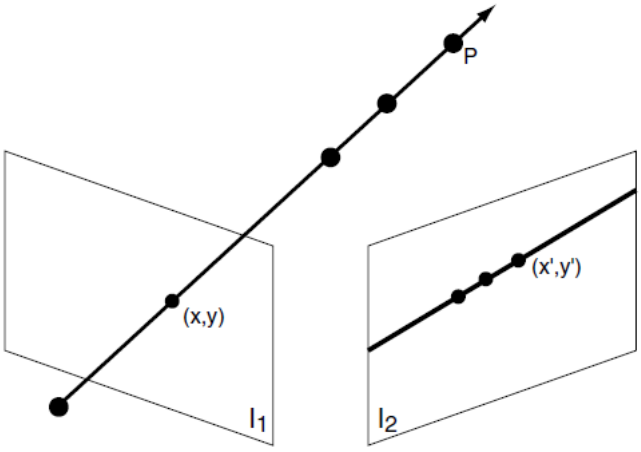
Stereo görme ve eşleme problemi üzerine ilk çalışmaları gerçekleştiren Marr ve Poggio, yaptıkları çalışmada [20] stereo görme problemi için insan görme sisteminin değiştirilmiş halini önermişlerdir. Piksel yoğunluklarının stereo görüntülerde karşılıklı eşleşen piksellerde bile değişebileceğini düşündükleri için sadece kenar ve köşe noktaları gibi keskin yoğunluk değişimlerinin olduğu noktaların eşleşmesine dayanan bir yaklaşım önermişlerdir. Eşleşen bu noktalar kullanılarak seyrek (sparse) aykırılık haritası üretmişlerdir. Ayrıca günümüzde de hala geçerliliğini koruyan ve eşleme probleminin hassas bir şekilde formülasyonunu sağlayan teklik ve süreklilik kısıtlarını tanımlamışlardır. Sonraki yıllarda yapılan çeşitli araştırmalarda [21-23] hatalı eşleşmeleri en aza indirmek ve algoritma karmaşıklığını azaltmak amacıyla yeni kısıtlar da öne sürülmüştür. Günümüzde stereo eşleme işlemi sırasında sıkça kullanılan bazı kısıtların listesi bulunmaktadır.

a) ***Teklik:*** Sahnedeki bir uzaysal noktanın bir resim içerisinde en fazla bir nokta ile temsil edilebileceği varsayımından yola çıkarak, bir görüntü içerisindeki nokta diğer görüntü içerisindeki en fazla bir nokta ile eşleştirilebilir. Teklik kısıtlaması çizgi segmenti tabanlı eşleme yaklaşımlar için geçerli olmayabilir ve bu kısıt sahnede saydam nesneler varsa başarısız olabilir. Ayrıca stereo görüntülerde ortak olmayan alanların da bulunması sebebiyle bir görüntüde bulunan bir piksel ya da özellik için diğer görüntüde ona karşılık gelen piksel ya da özellik bulunamayabilir.
b) ***Süreklilik:*** Bir yüzeyin hemen hemen her yerinde farkın çok yavaş değiştiğini kabul eder. Bu kısıt derin kenarlar için ihlal edilir.

c) ***Uygunluk:*** Yoğunluk tabanlı eşleme yaklaşımlarında eşleşen piksellerin veya piksel gruplarının benzer yoğunluk değerine sahip olması (belirli bir eşik değerinden düşük olması) veya piksel gruplarının yüksek oranda ilişkili olması gerekmektedir. Özellik tabanlı eşleme yaklaşımlarında ise eşleşen özelliklerin benzer özellik değerlerine sahip olması gerekmektedir.

d) ***Pürüzsüzlük:*** Aykırılık değerlerindeki ufak değişiklikler yerel alanlardaki ufak değişiklikleri gösterirken, büyük değişimler gerçek dünyadaki ani derinlik değişimlerinden kaynaklanır. Yani aykırılık değerlerindeki ani değişimler, nesneler arasında geçisin olduğu bölgelere işaret etmektedir.

e) ***Epipolar Doğru Kısıtı:*** Şekil 2.8'de iki kameranın görüntü düzlemleri, sahnedeki bir P noktası ve bu noktanın ilk görüntü (I_1) düzlemi üzerindeki projeksiyonu olan (x, y) noktası gösterilmektedir. I_1 görüntüsünü elde eden kameranın optik merkezinden sahnedeki P noktasına bir hayali çizgi çizildiğinde bu çizgi üzerindeki bütün noktaların konumu görüntü üzerindeki aynı (x, y) konumu olarak çıkacaktır. Bu hayali çizgiye ikinci kameradan bakıldığında epipolar doğru kısıtı oluşmaktadır. Yani stereo eşleşme (x', y') noktası ikinci görüntüdeki bu çizgi üzerinde bir yerde gerçekleşmelidir. Benzer şekilde ikinci görüntüdeki herhangi bir noktanın ilk görüntüdeki eşleşmesi de bir epipolar çizgi üzerinde olmalıdır. Bu kısıt, stereo eşlemeleri arama alanını iki boyuttan tek boyuta indirgemektedir. Diğer tüm kısıtlamaların aksine epipolar doğru kısıtı hiçbir zaman başarısız olmaz ve epipolar geometri bilindiği takdirde güvenilir bir şekilde uygulanabilir. Şekil 2.8'de epipolar doğru kısıtı gösterilmektedir. Tam olarak aynı anda iki görüntü yakalandığında sahnedeki P noktasına karşılık gelen ilk görüntüdeki (x, y) noktasının ikinci görüntüdeki eşleşmesi, ikinci görüntü üzerinde özel bir doğru (epipolar doğru) boyunca gerçekleşmelidir.



Şekil 2.8. Epipolar doğru kısıdının gösterimi

f) ***Yumuşatma:*** Yanlış eşleşmeleri ortadan kaldırmak için global eşleme kısıtıdır. Stereo eşleme işleminde gevşeme uygulamanın bir örneği Barnard ve diğerleri tarafından gerçekleştirilen çalışmada uygulanmıştır [23]. Bu çalışmada her bir aday eşleşme için belirledikleri iyi eşleşme kriterlerine göre bir olasılık değeri atanmıştır. Her aday eşleşme için bu olasılık değerini tekrarlamalı olarak güncellemişlerdir ve olasılık değeri belirli bir eşiğin altında olan eşleşmeleri temizlemişlerdir. Her güncelleme tekrarında belirli bir aday eşleşmenin olasılığını o eşleşmenin tutarlı aykırılık değerine sahip komşu eşleşmelerinin sayısı kadar arttırmışlardır.

2.1.2.1. Seyrek Eşleme

Literatürdeki ilk stereo eşleme algoritmaları özellik tabanlı algoritmalardır. Özellik tabanlı algoritmalarda, kenar ve köşe gibi ayırt edici özellikleri olan ve potansiyel eşleşme şansı bulunan noktalar görüntüden çıkarılarak diğer görüntüde bu noktalara karşılık gelen konumlar tespit edilmeye çalışılmaktadır [8, 9, 22, 24-31]. Bu çalışmalardaki seyrek eşleme yaklaşımı, hem o zamanlar kullanılan hesaplama kaynaklarının kısıtlı olmasından kaynaklanmakta hem de üretilen algoritmaların yüksek kalitede stereo eşleşme yapması amacını taşımaktadır. Bazı çalışmalarda ise [32] çok farklı ışıklandırma özelliklerine sahip sahneler arasında stereo eşleme yapabilmek için

tek kararlı özelliğin kenar noktaları olabileceği yaklaşımı bulunmaktadır. Bu alanda yapılan daha yeni çalışmalar ise görüntülerdeki güvenilir özellik noktalarını çıkarmaya daha sonra çıkarılan bu noktaları ek eşlemeler oluşturmak için tohum olarak kullanmaya odaklanmışlardır [33, 34]. Benzer yaklaşımlar çok görüntülü stereo görme ve 3 boyutlu yapılandırma probleminde de kullanılmaktadır [35-37].

Seyrek eşleşmenin bir başka örneği ise nesnelerin sınırlarında veya kavisli yüzeye sahip cisimlerde meydana gelen profil eğriliklerinin görüntüler arasında eşleştirilmesidir. Profil eğrilerinin eşleştirilmesindeki zorluk genel olarak profil eğrilerinin yerlerinin kamera bakış açısının bir fonksiyonu olarak değişmesinden kaynaklanmaktadır. Bu nedenle eğrileri doğrudan iki görüntüde eşleştirmek ve daha sonra bu eşleşmeleri üçgenleştirmek hatalı ölçümlere sebep olabilmektedir. Kavisli bir cisme bakan üç veya daha fazla görüntü mevcutsa dairesel bir yayı, karşılık gelen kenarların konumlarına yerleştirerek kavisli yüzeylerin yeniden yapılandırılmasını gerçekleştiren teknikler geliştirilmiştir [38-42].

2.1.2.2. Yoğun Eşleme

Günümüzde görüntü tabanlı görsel kaplama veya 3 boyutlu modelleme gibi uygulamalar genellikle yoğun stereo eşlemeye odaklanmaktadır. Yoğun eşleme problemi seyrek eşleme problemi ile kıyaslandığında daha zordur. Çünkü zayıf dokulu yüzerlerde derinlik değerlerinin anlaşılması için belirli tahminler yapma ihtiyacı bulunmaktadır. Yoğun eşleme algoritmaları genellikle dört temel adımdan oluşmaktadır:

a) Eşleme maliyeti hesaplanması
b) Maliyetlerin toplanması
c) Aykırılık değerlerinin hesaplanması ve optimizasyonu
d) Aykırılık değerlerinin iyileştirilmesi

Birinci adım olan eşleme maliyeti hesaplanmasında belirlenen maliyet fonksiyonu ile kullanılan görüntülerdeki karşılıklı piksellerin birbirleri ile ne kadar iyi eşleştiği hesaplanır. İkinci adım olan maliyetlerin toplanması adımında bir eşleme sırası için toplam maliyet hesabı yapılır. Üçüncü adımda en düşük maliyetli eşleme sırası için aykırılık değerleri hesaplanır. Bu adımların tüm piksellere uygulanmasıyla aykırılık

haritası elde edilmektedir. Son adımda tanımlanan çeşitli fonksiyonlarla elde edilen aykırılık haritasının iyileştirilmesi sağlanır.

2.1.2.3. Benzerlik Ölçüleri

Herhangi bir yoğun stereo eşleme algoritmasının ilk adımı piksellerin eşleşme ihtimallerini belirlemek için piksel değerlerini görüntüler arasında karşılaştıran benzerlik ölçüsüdür. Literatürde pek çok benzerlik ölçüsü önerilmiş olsa da stereo eşleme algoritmalarında mutlak farkların toplamı (SAD), kare farkların toplamı (SSD) ve normalleştirilmiş çapraz korelasyon (NCC) benzerlik ölçüleri yapılan çalışmalarda daha yaygın olarak kullanılmaktadır. Bu bölümde sıkça kullanılan benzerlik ölçüleri incelenmiştir.

a) **Mutlak Farkların Toplamı (SAD):** Referans görüntü içerisinde karşılaştırma yapılacak bloktaki piksellerin gri seviye yoğunluk değerleri ile karşılaştırma yapılan görüntüde ilgili bloktaki piksellerin gri seviye yoğunluk değerlerinin farklarının mutlak değeri alındıktan sonra çıkan sonuç değerlerinin toplanması ile hesaplanmaktadır [43]. Arama aralığındaki bloklar ile karşılaştırma yapıldıktan sonra bulunan en ufak değer referans görüntüdeki blok ile en benzer bölge olarak kabul edilmektedir. Mutlak farkların toplamı benzerlik ölçüsü dijital görüntü işlemede nesne tanımlama, stereo eşleme ve video sıkıştırma için hareket tahmini gibi problemlerde sıklıkla kullanılan bir yöntemdir. Eşitlik (14)'de I_1 referans görüntüdeki içerisindeki karşılaştırma yapılacak bloğu, I_2 ise karşılaştırma yapılan görüntüdeki bloğu temsil etmektedir. d ise aykırılık değeridir.

$$\sum_{u,v} |I_1(u,v) - I_2(u+d,v)| \tag{14}$$

b) **Kare Farkların Toplamı (SSD):** Kare farkların toplamı benzerlik ölçüsü hesaplama yönteminin mutlak farkların toplamı benzerlik ölçüsü hesaplama yönteminden farkı blok içerisindeki gri seviye yoğunluk değerlerinin farklarının karesinin alınmasıyla bulunmasıdır. Eşitlik (15)'de I_1 referans görüntüdeki

içerisindeki karşılaştırma yapılacak bloğu, I_2 ise karşılaştırma yapılan görüntüdeki bloğu temsil etmektedir. d ise aykırılık değeridir.

$$\sum_{u,v}[I_1(u,v) - I_2(u+d,v)]^2 \tag{15}$$

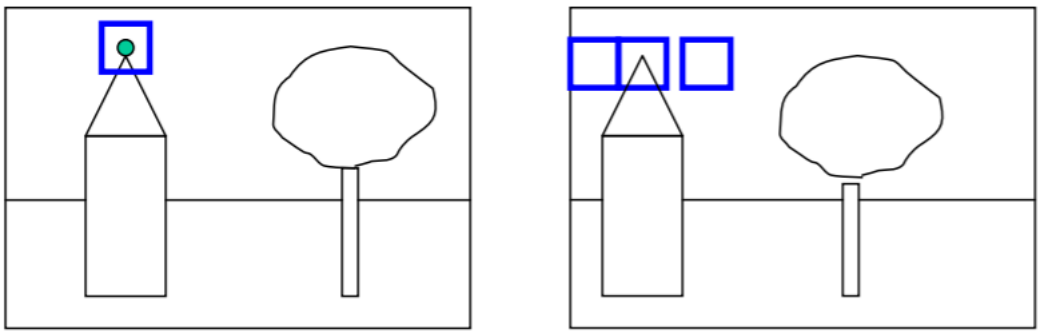
c) **Normalleştirilmiş Çapraz Korelasyon (NCC):** Stereo eşleme yoluyla yoğun ve doğru bir eşitsizlik haritası elde etmedeki zorluklardan bazıları stereo eşleme işleminde kullanılan görüntülerdeki radyometrik faktörlerdeki değişimler, kapanmalar, zayıf dokulu homojen yüzeyler ve fazla dokulu yüzeylerdir. Mutlak farkların toplamı (SAD) ve kare farkların toplamı (SSD) gibi benzerlik ölçüleri karşılaştırma yapılan görüntülerdeki piksel yoğunluğu değerlerinin değişmediği veya az değiştiği varsayımına dayanmaktadır. Ayrıca bu benzerlik ölçüleri tekrar eden veya zayıf dokulu yüzeylerde yanlış eşleşmelere sebebiyet verebilmektedir. Bu tarz patolojik durumların olduğu görüntülerde normalleştirilmiş çapraz korelasyon (NCC) gibi bir benzerlik ölçüsünün kullanılması bu yanlış eşleşmelerden kaçınılmasına olanak vermektedir [44]. Normalleştirilmiş çapraz korelasyon benzerlik ölçüsünün matematiksel ifadesi Eşitlik (16)'da verilmiştir. Yukarıdaki eşitlikte γ korelasyon çıktısını, f korelasyon bulunmaya çalışılan büyük görüntüyü, t karşılaştırma yapılacak pencereyi, $\bar{t}$ pencere içerisindeki değerlerin ortalamasını ve $\bar{f}_{u,v}$ ise pencere ile karşılaştırma yapılan $f(x,y)$ bölgesinin içerisindeki değerlerin ortalamasını vermektedir. Normalleştirilmiş çapraz korelasyon (NCC) içerisindeki normalizasyon işlemi, ortalamayı yoğunluk değerinden çıkararak iki görüntü arasındaki yoğunluk değerleri değişimlerinin etkisini azaltmaktadır. Standart sapmaya bölmek ise korelasyon çıktısı değerinin [-1, 1] aralığında olmasını sağlamaktadır.

$$\gamma(u,v) = \frac{\sum_{x,y}[f(x,y) - \bar{f}_{u,v}][t(x-u,y-v) - \bar{t}]}{\sqrt{\left(\sum_{x,y}[f(x,y) - \bar{f}_{u,v}]^2 \sum_{x,y}[t(x-u,y-v) - \bar{t}]^2\right)}} \tag{16}$$

2.1.2.4. Lokal Yöntemler

Stereo eşlemede lokal yöntemler, “en iyi eşleşme kazanır” (winner-take-all) yaklaşımı kullanarak stereo görüntülerdeki iki boyutlu pencereleri (blokları) eşleştirmeye çalışır. Lokal yöntemler eşleme maliyetini hesaplama (hangi eşleşme ölçütünü kullandıklarına) ve maliyeti toplama yöntemlerine göre değişiklik göstermektedirler. Bu yöntemler eşleşme maliyetini hesaplama biçimlerinden dolayı değil eşleme probleminin çözüm biçiminden dolayı lokaldirler. Referans görüntüdeki her bir piksel için hedef görüntüde diğer piksellerden bağımsız olarak eşleşme aranmaktadır.

Lokal yöntemlerde bir pikselin eşleşmesini bulmak için genellikle piksel üzerinde o pikselin komşu piksellerini de içeren dikdörtgen şeklinde bir pencere tanımlanır (Şekil 2.9-a). Tanımlanan bu pencere ile hedef görüntüde o piksele karşılık gelen pikselin komşu pikselleri ile arasındaki ilişkiye bakılarak eşleşmeye karar verilir (Şekil 2.9-b). Tanımlanan pencerenin hedef görüntüdeki eşleşmesine normalleştirilmiş çapraz korelasyon (NCC), kare farkların toplamı (SSD), normalleştirilmiş SSD ve mutlak farkların toplamı (SAD) gibi eşleme ölçütleri kullanılarak karar verilmektedir.



(a) (b)

Şekil 2.9. Referans görüntüde tanımlanan pencere ile hedef görüntü üzerinde eşleşmenin aranması

Yerel yöntemler, global yöntemler ile kıyaslandığında pikselin aykırılık değerine görüntü üzerinde dolaştırılan lokal pencere içerisindeki az sayıda hesaplama ile karar vermesinden dolayı genellikle hesaplama maliyeti daha azdır. Ancak yerel yöntemler

tekrar eden veya zayıf dokulu yüzeyler, kapanmalar ve homojen bölgeler gibi görüntü içerisinde belirsizlikler yaratan bölgelerden kolaylıkla etkilenmektedir. Bu nedenle böyle görüntülerde daha kalitesiz bir aykırılık haritası oluşmasına sebebiyet vermektedir.

Pencere tabanlı eşleme algoritmalarında kritik problemlerden biri pencere büyüklüğünün seçimidir. Pencere büyüklüğünün güvenilir bir eşleme yapmak için yeterli görüntü detayını içerecek kadar büyük olması gerekirken, izdüşümsel bozulmalardan etkilenmeyecek kadar da küçük olması gerekmektedir.

Stereo görüntülerin paralel geometrisi kullanılarak gözlemlenen belirli bir nokta için kameralar arası uzaklık ne kadar fazlaysa aykırılık o kadar büyük olur. Kameralar arasındaki mesafenin uzak veya yakın olma durumlarının her birinin bazı avantajları ve dezavantajları bulunmaktadır. Örneğin kameralar arası mesafe kısaysa görüntüler arasında bir eşleşme bulmak daha kolaydır. Çünkü aykırılık ve yoğunluk değişimleri daha azdır. Fakat bu durumda nesnelerin 3 boyutlu uzayda tahmin edilen mesafelerindeki hatalar küçük aykırılık değerleri ve olası tahmin hatalarından dolayı artacaktır. Kameralar arasındaki mesafe uzak olduğu durumda ise görüntü üzerinde daha uzun bir mesafede eşleme yapılmaya çalışılacak ve bu nedenle de hatalı eşleme olasılığı artacaktır. Okutumi ve diğerleri yaptıkları çalışmada [45] stereo eşleme elde edebilmek için çeşitli kamera mesafeleri ile yakalanmış çoklu stereo görüntü çiftlerinin kullanıldığı bir yöntem önermişlerdir.

2.1.2.5. Yarı Global Yöntemler

Yarı Global Eşleme (Semi Global Matching) global maliyet hesabını sadece lokal bilgileri kullanarak yapabilmek için Hirschmuller [46] tarafından önerilmiştir. Pencere tabanlı yaklaşımlara benzer çalışma zamanlarında çalışabilen ancak stereo eşlemede açıklık probleminin (aperture problem) etkisini azaltan etkili bir yaklaşımdır. Şekil 2.10'da açıklık problemi gösterilmektedir. Yeşil dörtgenler arama penceresi içerisindeki komşu pikselleri göstermektedir. Şekil 2.10-b'deki iki yeşil dörtgenden hangisinin sol görüntüdeki Şekil 2.10-a'daki yeşil dörtgene karşılık geldiğini belirlemek mümkün değildir.



(a) (b)

Şekil 2.10. Açıklık probleminin gösterimi

Yarı global eşlemede global maliyet hesabı aşağıdaki enerji fonksiyonu ile yapılmaktadır.

$$E_{SGM}(d) = \sum_{p \in P} \left(C(p, d_p) + \sum_{q \in N(p)} T(|d_p - d_q|) \right) \tag{17}$$

Eşitlikte P değeri görüntüdeki pikselleri, $C(p, d_p)$ herhangi bir benzerlik ölçüsü fonksiyonunu, $N(p)$ eşleme araması sırasında kullanılan penceredeki p pikselinin komşu piksel grubunu temsil etmektedir. $T(x)$ fonksiyonu ise aşağıdaki gibi tanımlanmıştır.

$$T(x) = \begin{cases} P_1, & x = 1 \\ P_2, & x > 1 \\ 0, & diğer \end{cases} \tag{18}$$

Yukarıdaki eşitlikte P_1 aykırılık değişimi miktarı 1 olduğundaki ceza maliyetini, P_2 ise daha büyük aykırılık değişimleri olduğu durumlardaki ceza maliyetini temsil etmektedir. Enerji fonksiyonunu minimize etmek için eşleşme maliyetleri her yönden 1 boyutlu olacak şekilde toplanmaktadır. r doğrultusundaki L_r maliyeti aşağıdaki gibi hesaplanmaktadır.

$$
\begin{aligned}
L_r(p,d) = \; & C\left(p,d_p\right) \\
& + \min\Big\{L_r(p-r,d), L_r(p-r,d-1) \\
& + \; P_1, L_r(p-r,d+1) + \; P_1, \\
& \min_i L_r(p-r,i) + \; P_2\Big\} - \; \min_k L_r(p-r,k)
\end{aligned}
\tag{19}
$$

Bu şekilde eşleşme sırasında sıralı olmama durumu cezalandırarak aykırılık haritasını verimli bir şekilde elde etmek mümkündür.

2.1.2.6. Global Yöntemler

Stereo eşlemede global yöntemler, bir enerji fonksiyonunun minimize edilmesine dayanmaktadır. Bu enerji fonksiyonu genellikle veri enerjisi ve pürüzsüzlük enerjisi fonksiyonlarının birleşimi şeklinde tanımlanmaktadır. Amaç global enerjiyi minimize eden aykırılık fonksiyonunu bulmaktır (Eşitlik (20)).

$$E(d) = E_{data}(d) + l\, E_{smooth}(d) \tag{20}$$

Global yöntemler olası aykırılık haritaları ile ilişkilendirilmiş iyi bir enerji fonksiyonu tanımlayama çalışır. Ardından bu enerji fonksiyonu minimize edilerek doğru olan aykırılık haritası seçilir. Enerji fonksiyonun görüntüdeki bütün pikseller için çalışmasından dolayı global yöntemler yerel belirsizliklerde (homojen yüzeyler, kapanmalar vs.) yerel yöntemlerden daha iyi sonuç vermektedir.

Literatürdeki ilk optimizasyon tabanlı global stereo eşleme çalışmalardan biri olan çalışmada [47] enerji fonksiyonu Eşitlik (21) ve Eşitlik (22)'deki gibi uygulanmıştır.

$$E_{data} = \sum_{i,j} \left| I_L(i,j) - I_R\big(i, j + d(i,j)\big) \right| \tag{21}$$

$$E_{smooth} = \sum_{i,j} \left[\sum_{\substack{k_i=-1,0,1 \\ k_j=-1,0,1}} \left| d(i,j) - d\big(i + k_i, j + k_j\big) \right| \right] \tag{22}$$

Yapılan çalışmada veri enerjisi için 1×1 boyutunda, pürüzsüzlük enerjisi için ise 3×3 boyutunda bir pencere görüntüler üzerinde dolaştırılarak ve mutlak farkların toplamı (SAD) eşleme kriteri kullanılarak enerji optimizasyonu yapılmıştır. E_{data} ve E_{smooth} enerji fonksiyonlarının sonuçları Eşitlik (20)'de kullanılarak aykırılık değerleri elde edilmiştir.

Veri enerjisi $\big(E_{data}(d)\big)$ genellikle normalleştirilmiş çapraz korelasyon (NCC), kare farkların toplamı (SSD) veya mutlak farkların toplamı (SAD) gibi eşleme metrikleri

kullanılarak lokal eşleme yöntemleri ile elde edilmektedir. Pürüzsüzlük enerjisi $\left(E_{smooth}(d)\right)$ ise pürüzsüz olmayan aykırılık değeri dağılımlarını cezalandırmak için kullanılmaktadır.

Global stereo eşleme üzerine yapılan çalışmalarda stereo eşleme probleminin doğasına uygunluğundan dolayı dinamik programlama tabanlı eşleme yaklaşımları sıklıkla önerilmiştir. Ayrıca son yıllarda görüntü piksellerini bir çizgenin düğümleri gibi ele alıp bu çizge üzerinde en düşük maliyetli kesiyi bulmayı hedefleyen algoritmalar olan çizge kesi (graph cut) tabanlı yöntemler de oldukça etkileyici sonuçlar üretmiştir.

Dinamik programlama, bir problemin çözümünün alınan alt kararların bir sonucu olarak görüldüğü zaman kullanılabilecek bir algoritma tasarım yöntemidir [48]. Baker ve Binford [22] stereo eşleme problemini, stereo görüntü çiftinde bir tarama çizgisi boyunca korunan kenarların soldan sağa sıralanmasına dayalı olarak yinelemeli olarak bölümlendirmek için bir dinamik programlama algoritması olan Viterbi algoritmasını kullanmışlardır. Literatürde stereo eşleme üzerine gerçekleştirilen başka birçok önemli çalışmada da [49-52] dinamik programlama tekniklerinin kullanıldığı görülmektedir.

2.2. Engel Tespiti ile İlgili Literatür Araştırması

Bir varlığın gezinti sırasında çarpışma olmadan hareket edebileceği alanların belirlenmesi engel tespiti olarak adlandırılmaktadır. Engel tespiti, gezinti sırasında hareket alanının serbest ya da serbest olmayan bölgeler olmak üzere sınıflandırılması olarak da görülebilir. Serbest bölgeler hareketli bir varlığın herhangi bir çarpışma yaşamadan hareket edebileceği, serbest olmayan bölgeler ise engellerin bulunduğu bölgeler anlamına gelir. Bu tez çalışmasında hareket alanı üzerindeki asgari olarak belirli bir yükseklikten fazla olan nesneler engel olarak adlandırılmaktadır. Yukarıdan aşağı doğru asılı duran nesneler veya hareket alanındaki çukurlar dikkate alınmamaktadır.

Görme engellilerin navigasyonu bağlamında çevrelerinin haritalandırılmasının yapılabilmesi için farklı sensörler önerilmiştir. LIDAR teknolojisinin dış mekân senaryolarında başarılı olduğu kanıtlanmıştır, ancak maliyetlidir ve görme engelli bir bireyin hayatında pratik olarak kullanılabilecek bir sensör değildir. Bu tez çalışmasında

pasif stereo kamera sistemi tabanlı duyusal ikame cihazı geliştirilmesine odaklanılmıştır. Görme engelli bireyler için kameraların düşük maliyetli olması, kolay bulunabilirliği ve kameraların insan gözüyle olan benzerliği bu tercihin yapılmasındaki başlıca nedenler arasındadır.

Bilgisayarlı görmeye dayalı engel tespiti yaklaşımlarında kamera tarafından elde edilen görüntülerde bulunan nesnelerin kameralara olan uzaklığı hesaplanmaktadır. Yalnızca tek bir kameranın kullanıldığı monoküler yaklaşımlar elde edilen görüntülerin ölçeğini gerçek dünyayla ilişkilendirmenin bir yolunu gerektirmektedir. Bunu başarmak için Gonçalves ve diğerleri [53] yaptıkları çalışmada sonar sensörden gelen bilgileri kullanmaktadırlar. Pollard ve diğerleri [54] monoküler kamerayı sabit tutarak ölçek bilgisini elde etmişlerdir. Sebesta ve Baillieul [55] ise ölçek bilgisini elde etmek için sahneyi sabit tutmuşlardır. Lin ve diğerleri [56] yaptıkları çalışmada ise görüntüde bulunan ve özellikleri bilinen nesnelerin sahneden sahneye değişimlerini tanımlamaya çalışarak engel tespiti yapmayı amaçlamışlardır.

Görme engelli bireylerin navigasyonu bağlamında dünya ve kamera dinamiktir. Bu nedenle monoküler yaklaşımların bu bağlamda uygulanması güçtür. Stereo kameralar ise görüntülerden üç boyutlu bilgi çıkarımı yapabilmek için kullanılabilmektedir. Stereo görüntülerden aykırılık haritası elde edildikten sonra görüntülerdeki nesnelerin kameraya olan mesafelerinin hesaplanması üçgenleme yöntemi ile başarılabilmektedir. Bu tez çalışması kapsamında bu yaklaşım kullanılmaktadır.

Üretilen yoğun aykırılık haritasının üçgenlenmesi ile yoğun nokta bulutu kümesi elde edilebilir. Elde edilen nokta bulutu kümesi LIDAR veya lazer tabanlı yaklaşımlarla elde edilen nokta bulutu kümeleri ile benzer yapıdadır. Nokta bulutu kümeleri arasındaki bu benzerlik LIDAR kullanılarak engel tespiti yapılan çeşitli algoritmaların kullanılmasına olanak tanımaktadır. Yoğun nokta bulutu kümeleri engel tespiti sırasında sahneleri iyi bir şekilde temsil etmesine rağmen büyük miktarda veri içermesi sebebiyle işlenmesi için yüksek miktarda kaynak kullanımı gerektirmektedir. Bu kaynak ihtiyacını azaltmak ve işlemi hızlandırmak amacıyla sahneyi veya engelleri tespit etmek için aykırılık haritası (disparity map) tabanlı yaklaşımlar, işgal gridleri (occupancy grid), yükselti haritası (elevation map) ve stixel gibi birçok farklı alternatif önerilmiştir.

2.2.1. Aykırılık Uzayı Görüntüleri

Aykırılık Uzayı Görüntüleri üçgenleme işlemi gerçekleştirmeden üç boyutlu dünyanın özelliklerini tanımlamak için yeterli bilgi içermektedir. Aykırılık görüntülerinin üç boyutlu dünya koordinatlarına projeksiyonu, geometrik düzeltme adımındaki belirsizlikler nedeniyle hatalara eğilimlidir ve kameradan uzak nesneler için bir seyreklik etkisi yaratmaktadır. Aykırılık haritasına dayalı işleme, üç boyutlu dünyadaki nesneleri vokseller, nesne modelleri veya gridler olarak temsil etme gereksinimini ortadan kaldırmaktadır.

Aykırılık haritasından elde edilen, kolonlardaki ve satırlardaki aykırılık değerlerinin birikimli olarak toplanması ile oluşturulan histogramlara sırasıyla u-disparity ve v-disparity haritası adı verilmektedir. V-disparity haritası Labayrade ve diğerleri tarafından önerilmiştir [57]. V-disparity haritası, görüntünün yatay ekseninde aynı aykırılık değerine sahip piksellerinin toplanmasıyla elde edilmektedir. Yaptıkları çalışmada engel tespiti işlemi sırasında sahnedeki zemin profilini ayrıştırmak için v-disparity haritasına doğru uyarlama yaklaşımını kullanmışlardır. Araştırmacılar kameranın Z ve Y eksenlerindeki hareketinin asgari düzeyde olduğunu varsaymışlardır ve v-disparity haritasına Hough dönüşümü uygulayarak doğru uyarlamışlardır. Uyarlanan bu doğrunun eğimi zemin düzlemini ve kameranın yüksekliğini belirlemek için kullanılmıştır. Daha sonra v-disparity görüntüsündeki dikey çizgiler tespit edilerek yol ve engeller arasındaki sınırlar belirlenmiştir. Araştırmacıların önerdiği bu yöntem sahnedeki engelleri tespit etmek için oldukça etkili olsa da birbirine yakın engelleri bir araya getirme eğilimindedir. Sahnedeki zemin profilinin ayrıştırılması için oldukça popüler bir yöntemdir ve diğer yaklaşımlarla birlikte sıklıkla kullanılmaktadır.

U-disparity haritası Hu ve diğerleri [58] tarafından önerilmiştir. Görüntünün dikey ekseni boyunca aynı aykırılık değerine sahip piksellerinin toplanmasıyla oluşturulur. Yaptıkları çalışmada zemin düzlemini tahmin etmek için v-disparity haritasını kullanmışlardır ve engelleri tanımlamak için de u-disparity haritasını önermişlerdir. Sahnedeki engellerin sabit bir aykırılık değerine sahip olma eğiliminde olmalarından dolayı u-disparity haritası üzerinde yatay çizgiye yakın bir görünüm oluşturdukları için u-disparity haritasını kullanarak sahnedeki engelleri tanımlamanın

mümkün olduğunu fark etmişlerdir. Bu yaklaşımın etkili olduğu kanıtlanmıştır ve engel tespiti için yaygın olarak kullanılmaktadır. Bu yaklaşımın ana dezavantajı, yatay bir doğrunun sahnedeki bir engeli mi yoksa yolun bir parçası mı olduğunu belirlemek için bir eşik değeri belirleme gerekliliğidir.

Broggi diğerleri [59] engel tespiti için u-disparity ve v-disparity haritalarını görüntüdeki kenar özelliklerini kullanarak oluşturmuşlardır. Sadece yüksek gradyent değere sahip piksellerden elde ettikleri aykırılık haritasını kullanarak v-disparity haritasını oluşturmuşlardır. Yanlış aday pikselleri elemek için ise gradyent açısını kullanmışlardır ve bu şekilde sahnedeki zemin düzlemini tahmin etmişlerdir. Zemin düzlemini tahmin ettikten sonra arama uzayını v-disparity haritasındaki yer düzlemini temsil eden çizgideki aykırılık değeri ile sınırlayarak yoğun aykırılık haritası oluşturmuşlardır. Böylece engel tespiti için gerekli olan hesaplama sayısı azaltılmıştır. Bu çalışma, görüntüdeki kenar özelliklerinin kullanılarak engellerin tespitinin yapılabileceğini göstermesi açısından önemlidir.

Soquet ve diğerleri yaptıkları çalışmada [60] v-disparity görüntüsü üzerindeki zemin profili tahminini ve zemin ile engeller arasındaki sınırların tespitini iyileştirmek için yayılma algoritması kullanmışlardır. Perollaz ve diğerleri yaptıkları iki araştırmada [61, 62] ise u-disparity ve v-disparity görüntülerinden işgal gridleri oluşturmuşlardır. Fakhfakh ve diğerleri [63] sahnedeki zemin profilini ve engelleri ayrıştırmak amacıyla dikey pencereler kullanarak v-disparity ve u-disparity haritalarının ağırlıklı bir varyasyonunu oluşturmuşlardır. Iloie diğerleri yaptıkları çalışmada [64] sahnedeki yayalarla ilişkili pikselleri tespit etmek ve kümelemek için v-disparity ve u-disparity haritalarını ve görüntü ayrıştırması yaklaşımlarını kullanmışlardır.

2.2.2. İşgal Gridleri

İşgal gridleri olasılık tabanlı bir yaklaşım kullanarak sahnedeki nesnelerin işgal ettiği bölgeleri yansıtan ızgaralarla sahneyi temsil etmektedir. İlk olarak Elfes tarafından sonar verilerinin gösterimi ve navigasyon için önerilmiştir [65]. Yguel ve diğerleri [66] LIDAR gibi mesafe bulucu sensörler ile grafik işlemci birimi tabanlı hesaplama yaparak

işgal gridleri oluşturmuşlardır. Matthies ve diğerleri [67] sonar ve stereo kamera gibi farklı sensör türleri ile veriler toplayarak işgal gridleri oluşturmuşlardır.

İşgal gridleri sensörlerden gelen verilerdeki belirsizliğin formülasyonunu gerektirmektedir. Bu belirsizliği Matthies ve diğerleri [67] üçgenleme hatasını modelleyerek, Brandão ve diğerleri [68] maliyet fonksiyonunu entegre ederek ve Perrollaz ve diğerleri ise [69] eşleşme güvenirliliği parametresini kullanarak formülasyonunu yapmışlardır.

Nokta bulutu elde etmek için aykırılık görüntülerindeki aykırılık değerlerinin üçgenlenmesi yöntemi kullanıldığında, kameradan uzak nesneler çok az sayıda nokta üretmektedir. Bu seyreklik engel tespit sistemleri için arzu edilen bir durum değildir. Bu problemi çözmek amacıyla Badino ve diğerleri [70] kutupsal (polar) işgal gridlerini önermişlerdir. Zemin düzlemindeki nesnelerin sınırlarını tespit etmek için dinamik programlama yaklaşımı ile kutupsal işgal gridleri oluşturarak hızlı ve doğru sonuçlar elde etmişlerdir.

2.2.3. Dijital Yükselti Haritaları

Dijital yükselti haritaları ilk olarak Zhang tarafından [71] önerilmiştir. 3 boyutlu dünyanın yukarıdan görünümünün grid şeklinde bir temsilini sağlamaktadır. Gridin hücrelerindeki değerler sahnedeki zeminin veya nesnelerin yüksekliğine karşılık gelmektedir. Yapılan çalışmada [71] stereo eşleme ile elde edilen nokta bulutu verilerinin uyarlanması için düzlemsel ve karesel yüzeyler kullanmıştır. Daha sonra nokta ve yüzey arasındaki mesafeleri iteratif olarak eşikleyerek ve yüzeyleri yeniden hesaplayarak nokta bulutundaki hatalı noktaları elemiştir. Sahnedeki engeller, yükselti haritasına eşikleme uygulanarak tespit edilmiştir. Dijital yükselti haritaları kullanılan sensörden bağımsız olarak nokta bulutları kullanılarak hesaplanabilmektedir.

Oniga ve diğerleri gerçekleştirdikleri çalışmada [72] nokta bulutu verilerini dijital yükselti haritası hücrelerinin yoğunluğuna göre zemin veya engel olarak sınıflandırmıştır. RANSAC algoritmasını ile alan genişletme (region growing) tekniği uygulayarak

sahnedeki zemine bir karesel yüzey uyarlamışlardır. Bu zemin modelini kullanarak nesnelere karşılık gelen noktaları tespit etmeye çalışmışlardır.

Üçgenleme işleminden kaçınmak için Vergauwen ve diğerleri [73] her iki kamerada da sıfır yükseklikte bir nokta seçip o noktada bulunan dikey bir çizgiyi simüle ederek aykırılık haritalarından dijital yükselti haritaları elde etmişlerdir. Çalışmada sol kameraya bir ışık kaynağı yerleştirilmiştir ve bu ışık kaynağının gölgesinin sağ görüntüdeki izdüşümü simüle edilen dikey çizgiyle kesiştirilmiştir. Bu yolla sol görüntüde bulunan hangi alanların sağ görüntüde görünür olduğu veya hangi alanlarda kapanma olduğu tespit edilmiştir.

Dijital yükselti haritaları yoğun stereo verilerden engel tespiti yapılmasına olanak sağlasa da nokta bulutu kümesindeki her bir nokta için işlem yapma gereksinimi bulunmaktadır. Bu nedenle yüksek çözünürlüklü görüntülerden elde edilen yoğun nokta bulutları için gereksinim duyulan işlem kapasitesi gömülü ve gerçek zamanlı uygulamaları kısıtlamaktadır.

2.2.4. Dörtgen Süper Piksel (Stixel)

Dörtgen süper pikseller 3 boyutlu dünyayı stereo kamera tarafından çekilen görüntülere çizilen ve nesnenin uzaklığı ile ilişkili bir aykırılık değerine sahip dikey dikdörtgenlerle temsil etme tekniğidir. Bir dörtgen süper pikseli temsil etmek için sadece iki nokta ve aykırılık değerine ihtiyaç bulunmaktadır. Bir dörtgen süper pikselin çok sayıda piksel içermesi beklendiği için görüntüyü saklamak için gereken kaynak ihtiyacı önemli ölçüde düşüktür.

Rebut ve diğerleri [74] tarafından sahnedeki dolu alanların temsili için dikey doğruların kullanıldığı bir teknik önerilmiştir. Kubota ve diğerlerinin [75] çalışmasında aykırılık haritasındaki engellerin temsili için dikey sütunlar kullanılmıştır. Bu çalışmada dinamik programlama yaklaşımıyla engeller ve zemin arasındaki sınırın tespitine odaklanılarak engellerin yüksekliği göz ardı edilmiştir.

Dörtgen süper piksel temsili ilk olarak Badino ve diğerleri [2] tarafından stereo görüntülerden elde edilen yoğun aykırılık haritalarının verimli bir şekilde analizini gerçekleştirmek için önerilmiştir. Yoğun aykırılık haritası kullanarak dörtgen süper piksel oluşturmayı, gerçek dünya koordinatlarını elde etmek için üçgenlemeyi ve engelleri tespit etmek için polar işgal gridleri kullanmayı önermişlerdir. Gerçekleştirdikleri çalışmada ilk olarak yarı global eşleme ile yoğun aykırılık haritası oluşturmuşlardır. Daha sonra sahnedeki boş alanları tespit etmek için polar işgal gridi oluşturmuşlardır. Boş alanları tespit ettikten sonra sadece ön plandaki nesneleri işleyebilmek amacıyla görüntüye arka plan çıkarımı işlemi uygulamışlardır. Arka plan çıkarımı işlemi sırasında nesnelerin üst sınırları ile arka planı iyi bir şekilde ayrıştırmak için dinamik programlama yaklaşımı kullanmışlardır. Son aşamada elde edilen bilgilere aykırı değer ve gürültü temizliği uygulayarak dörtgen süper pikselleri oluşturmuşlardır. Bu temsil, kent görüntülerindeki sahneyi tanımlamak için araştırmacılara yeterli bilgi sağlamıştır.

Benenson ve diğerleri [76] sahnedeki yayaları tanımlamak için dörtgen süper piksel yaklaşımını kullanmışlardır. Ancak çalışmalarında yoğun aykırılık haritaları yerine maliyet hacmi yaklaşımını kullanmışlardır. Böylelikle hem işlemi hızlandırmışlar hem de üçgenleme aşamasındaki olası hataları azaltmışlardır. Çalışmalarında ilk olarak maliyet hacmini lokal yaklaşım ile hesaplamışlardır. Daha sonra hesaplanan maliyet hacminden v-disparity görüntüsünü elde etmişlerdir. Doğru uyarlama yöntemiyle sahnedeki zemin profilini tahmin etmişlerdir. Dinamik programlama yaklaşımı ile u-disparity haritasından dörtgen süper piksellerin çıkarımını yapmışlardır. Dörtgen süper piksellerin yüksekliği ise her dörtgen süper piksele atanan aykırılığa karşılık gelen maliyet hacmindeki yerel minimumun belirlenmesi ile tahmin edilmiştir.

Benenson ve diğerleri [77] nesne tespitini hızlandırmak için dörtgen süper piksel temsilinin alternatif bir formülasyonunu önermişlerdir. Ufuk çizgisi altındaki satırlar için maliyet hacminden elde edilen v-disparity haritası kullanılarak zemin profili tahmin edilmiştir. Dörtgen süper piksel mesafesinin tahmini görüntüyü birden çok satır bandına dikey olarak bölerek yapılmıştır ve her görüntü sütunu için maksimum yatay gradyent seçilmiştir. Araştırmacılar bu çalışmada çok çekirdekli işlemcinin bulunduğu bir dizüstü bilgisayarda dörtgen süper piksel hesaplaması için 260 fps performansa ulaşabildiklerini belirtmişlerdir.

Carillo ve diğerleri [78] kenar eşlemelerine dayalı seyrek aykırılık haritalarını kullanarak dörtgen süper piksel tekniği ile engel tespiti yapmışlardır. Seyrek aykırılık haritaları kullanılması sebebiyle önerilen yöntem oldukça verimlidir.

Literatürdeki dörtgen süper piksel tabanlı yaklaşımların hızlı ve başarı oranının yüksek olduğu görülmektedir. Otonom sürüş alanında sıklıkla kullanılan bu yaklaşımlardan faydalanılarak görme engelli bireylerinin mobilizasyonunun sağlanması hala açık bir problemdir ve gelecekte daha fazla araştırmanın odağı olabilir.

2.2.5. Engel Tespitinde Diğer Yaklaşımlar

Engel tespitindeki diğer yaklaşımlar nokta bulutlarındaki her bir noktanın bağımsız olarak analiz edilmesi, nokta bulutlarına yüzey uyarlanması ve akış alanlarının kullanılması şeklindedir.

Literatürde nokta bulutunda bulunan bütün noktaları analiz eden ve her bir noktayı engel veya boş alan olarak etiketleyen yaklaşımlarda çeşitli teknikler kullanılmıştır. Williamson ve diğerleri [79] zemine ait olmayan her bir noktayı engel olarak kabul etmişlerdir. Shade ve diğerleri [80] nokta bulutuna yüzey uyarlamak için piksellerin komşuluk ilişkisinden faydalanmışlardır. Bleyer ve diğerleri [81] nokta bulutundaki noktalara iki boyutlu ayrıştırma ve yüzey uyarlama uygulamışlardır. Kramm ve diğerleri [82] özellik çıkarımı ve kümeleme teknikleri ile oluşturdukları seyrek nokta bulutlarını kullanarak engel tespiti yapmışlardır.

2B ve 3B akış alanları sahnenin içeriğinin yer değiştirmesini belirleyerek engel tespiti için alternatif bir yöntem sağlamaktadır. 2B akış alanları optik akış olarak bilinmektedir. Optik akış, 2 boyutlu bir görüntüdeki piksellerin belirgin bir şekilde yer değiştirmesini temsil etmektedir [83, 84]. Sebesta ve diğerleri [55] optik akışta yakın nesnelerin uzak nesnelerden daha büyük bir yer değiştirme ürettiği gerçeğini dikkate alarak engel tespiti için bir yöntem önermişlerdir.

3B akış alanları sahne akışı olarak da adlandırılmaktadır. Sahne akışı nokta bulutu dizisindeki noktaların üç boyutlu yer değiştirmesini temsil etmektedir [85]. Sahne akışını

hesaplamak için çeşitli yaklaşımlar önerilmiştir. Bu yaklaşımlar sahnedeki düzlemsel unsurları kullanarak engel tespiti yapmaktadırlar [86, 87].

2.3. Görme Engellilerin Navigasyonu ile İlgili Literatür Araştırması

Görsel duyu, insanların istenen hedefe ulaşmasında, iç veya dış mekânlarda doğru yolu bulmasında ve gezinmesine yol göstermede önemli bir rol oynamaktadır. Ne yazık ki görme engelli bireyler bu tür faaliyetleri gerçekleştirirken büyük zorluklar yaşamaktadır. Görme eksikliği insanların kişisel, mesleki ve çevresel ilişkilerini etkilemektedir ve günlük yaşam rutinlerini gerçekleştirmede birçok zorluğa yol açmaktadır.

Görme engelli bireylerin hareketliliği hakkındaki çeşitli çalışmalar [88], mekânsal haritalama ve yönlendirme becerileri edinme desteğinin algısal ve kavramsal olmak üzere iki düzeyde sağlanması gerektiğini göstermektedir. Algısal düzeyde, görme engelli bireyin görsel kanalındaki eksiklik, akustik veya dokunma gibi diğer duyularla değiştirilmelidir [89]. Bu tarz duyusal ikame yöntemleri gerçek görüntüleri ve uzaklıkları akustik sinyallere veya elektriksel uyaranlara dönüştürmektedir. Kavramsal düzeyde, görme engelli bir bireye yönelik geliştirilmiş herhangi bir yardımcı cihaz, kullanıcıyı navigasyon sırasında ortaya çıkabilecek öngörülemeyen koşullara uyarlamak için sahne haritalama dahil olmak üzere çeşitli yönlendirme stratejilerinin geliştirilmesine odaklanmalıdır [90].

Literatürde görme engelli bireyler için insan görüşüne duyusal bir ikame sağlamak amacıyla tasarlanmış çok çeşitli elektronik seyahat teçhizatları (Electronic Travel Aid, ETA) bulunmaktadır. Görme engelli kullanıcıların hareketliliğini artırmayı ve yakın çevresi hakkında ek bilgi sağlamayı amaçlayan bu teçhizatlar sensör ağları tabanlı sistemler (aktif sistemler) ve bilgisayarlı görme tabanlı sistemler (pasif sistemler) olmak üzere iki ana kategoriye ayrılmaktadır. Bu bölümde görme engelli bireyler için geliştirilen seyahat sistemleri incelenmektedir.

2.3.1. Sensör Ağları Tabanlı Sistemler

Sensör ağları tabanlı sistemler, çevredeki bilgileri toplayarak görme engelli kullanıcıya akustik veya dokunsal yolla iletmek için tasarlanmışlardır. Bu sistemler genellikle ultrason [91-97], kızılötesi [98], global konumlama sistemi (GPS) [99, 100] ve radyo frekanslı tanımlama (RFID) [99, 101, 102] teknolojilerini kullanmaktadır.

Pressey tarafından önerilen ve literatürdeki öncü sistemlerden biri olan çalışmada [92] görme engelli kullanıcı ile sahnedeki çeşitli engeller arasındaki mesafeyi tahmin eden Mowat sensörlerinden yararlanılmıştır. Geliştirilen sistemi kullanıcının elinde tutarak kullanması gerekmektedir. Sahnedeki tespit edilen engeller, görme engelli kullanıcıya bip sesi desenleri olarak iletilmektedir. Bu açıdan sistem kullanılmadan önce kullanıcının sistem hakkında kapsamlı bir eğitimden geçirilmesi gerekmektedir.

Kay tarafından gerçekleştirilen çalışmada [91] Sonicguide ve Trisensor olarak adlandırılan sonar tabanlı sensörler önerilmiştir. Bu sensörler vasıtasıyla sahnedeki engelleri tespit ederek kullanıcıya tespit edilen engeller hakkında uzamsal bilgi sağlanmıştır. Ancak üretilen iki prototip de düşük çözünürlüklü sensör kullanmaktadır ve bu sebeple 3B uzayda tespit edilen mesafenin hata payı yüksektir.

Ross ve Lightman tarafından geliştirilen Talking Braille [98] isimli sistemde kapalı kamusal alanlarda görme engelli bireylerin yön bulmalarını sağlamak için kablosuz ağ tabanlı teknoloji kullanılmıştır. Sistemin kullanımı, erişim noktalarının önceden tanımlandığı bilinen mekanlarla sınırlıdır. Geliştirilen sistemde kullanılan kızılötesi teknolojisinin güneş ışığına ve aydınlatma koşullarına karşı yüksek duyarlılığı nedeniyle verimli bir şekilde çalışmama olasılığı bulunmaktadır.

Kuc tarafından sunulan stereofonik sonar sistemi [93] görme engelli kullanıcının etrafındaki rastgele konumlarda bulunan nesneleri başarılı bir şekilde tespit etmiştir. Geliştirilen sistemde kullanıcıya geri bildirim dokunsal yolla titreşim olarak sunulmuştur. GuideCane [94], SmartCane [103], UltraCane [104], Batcane [105] ve Necklace Cane [106] sistemleri, görme engelliler tarafından sıklıkla kullanılan sıradan beyaz bastonun yeteneklerini ultrasonik özellikler sayesinde geliştirmeyi ve zenginleştirmeyi amaçlamıştır.

Literatürdeki sonar ve ultrason tabanlı sistemler optimal yürüyüş yolunu tahmin etmeye ve görme engelli kullanıcıyı dokunsal uyarım yoluyla bilgilendirmeyi amaçlamaktadır. Ancak gündelik yaşam senaryolarında test edildiklerinde özellikle kullanıcının dizinin üstünde bulunan nesnelerin, yukarıdan sarkan engellerin ve kaldırım basamaklarının tespitinde oldukça yetersiz kaldıkları görülmektedir.

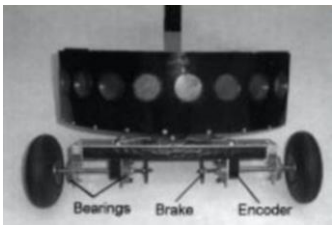
Yelamarthi, Haas ve diğerleri tarafından geliştirilen Smart-Robot [99] cihazı RFID ve GPS lokalizasyonunu entegre eden bir görme engelli navigasyon sistemidir. Önerilen yöntem görme engelli kullanıcının önceden tanımlanmış bir varış noktasına ulaşmasına yardımcı olmak için hem iç hem de dış ortamlarda performans gösterebilmektedir. Kullanıcı geri bildirimi, giyilen bir eldiven yardımıyla iletilen titreşim desenlerine ve akustik mesajlara dayanmaktadır. Benzer bir yaklaşım MOST-NNG [100] adı verilen cep telefonu tabanlı navigasyon sisteminde de önerilmiştir. Bu tür GPS tabanlı sistemlerin sinyal kaybına büyük ölçüde bağlı olduğu ve kullanıcının konumunun tahmin edilmesinde düşük bir doğruluk oranı sağladığı kanıtlanmıştır. Bu durumla özellikle binaların yoğunluğunun fazla olduğu kent ortamlarında sıklıkla karşılaşılmaktadır.

GPS teknolojisine dayanmayan iç ve dış mekân konumlama teknikleri genellikle ek teknolojilerin dahil edilmesini gerektirmektedir. Bu çerçevede temel çözümler genellikle RFID teknolojisine dayanmakta olup iç mekanlarda kullanıcı navigasyonunu kolaylaştırmayı amaçlamaktadır. Kulyukin ve diğerlerinin geliştirdiği RoboCart [107] isimli prototip, görme engelli kullanıcıların belirtilen bir hedefe ulaşmalarına yardımcı olmak üzere tasarlanmış bir RFID sistemidir. Yapılan başka bir çalışmada görme engelli kullanıcının bir market içerisinde alışverişini yapabilmesini sağlamak amacıyla RFID teknolojisi ve QR kodlar beraber kullanılmıştır [108]. Bazı çalışmalarda [109, 110] görme engelli bireylerin müzelerde gezinebilmeleri için RFID teknolojisinin kullanıldığı görülmektedir. RFID teknolojisindeki temel kısıt; kullanılan mekân içerisinde öncelikle birtakım düzenlemelerin yapılması gerekliliğidir. Bu tür sistemler, kullanılan alınan ortamın her yerine etiket yerleştirme gereği nedeniyle çok çeşitli özellikler sunamamaktadırlar. Ayrıca bu tür sistemler kişinin üstünde RFID taşıması gerektiği için çok kullanışlı değillerdir.

Sensör ağları tabanlı sistemler yürüme yolundaki engeller ile görme engelli kullanıcı arasındaki gerçek mesafeyi belirlemede sağladığı yüksek başarım seviyesinden dolayı bu kişilerin navigasyonuna yardımcı olmada oldukça faydalıdır. Ancak yukarıda analiz edilen sistemlerin hiçbiri bilinmeyen bir ortamdaki nesneleri tanımlayamaz veya göreceli tehlike derecelerini belirleyemez. Ayrıca ultrasonik sensörler son derece hassastır ve ortamdaki diğer sensörlerden dolayı veya başka sinyaller sebebiyle girişim problemine sebep olabilir. Lazer tarayıcılar ise hassasiyetleri ve yüksek çözünürlükleri nedeniyle mobil robot navigasyonunda sıklıkla kullanılmasına rağmen görme engelli bir bireyin navigasyonunda kullanmak için hem pahalı hem de ağırlığı fazladır. Şekil 2.11'de literatürde öne çıkan sensör ağları tabanlı görme engelli navigasyon sistemleri gösterilmektedir.



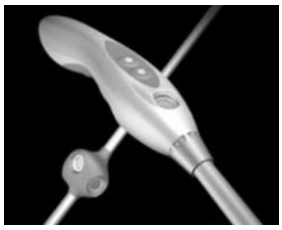
(a) Binaural [93]



(b) Guidecane [94]



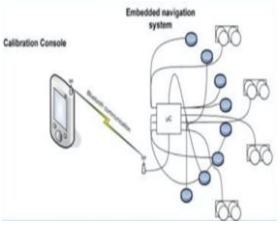
(c) SmartCane [103]



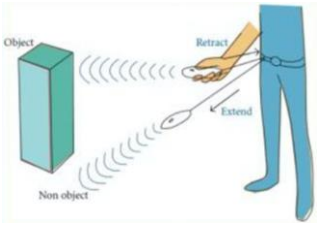
(d) Batcane [105]



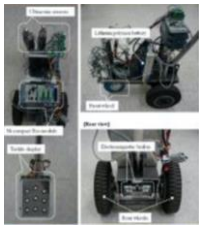
(e) Necklace Cane [106]



(f) EPFL [95]



(g) CyARM [96]



(h) Jeong v.d. [97]



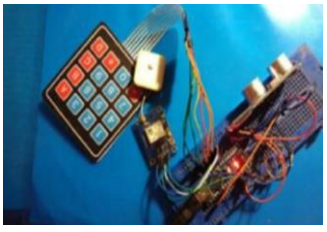
(i) Smart-Robot [99]



(j) RoboCart [107]



(k) Chumkamon v.d. [109]



(l) MuseumTourer [110]

Şekil 2.11. Sensör ağları tabanlı görme engelli navigasyon sistemleri

2.3.2. Bilgisayarlı Görme Tabanlı Sistemler

Bilgisayar donanımlarının işleme kapasitesinin artması ve bilgisayarlı görme algoritmalarındaki gelişimler sayesinde görme engelli bireyler için tasarlanmış çeşitli sistemler bilgisayarlı görme tekniklerine dayanmaktadır. Bilgisayarlı görme tabanlı yöntemler kullanılan kamera sensörünün tipine bağlı olarak monoküler, stereoskopik ve RGB-D olmak üzere üç kategoriye ayrılabilir.

2.3.2.1. Monoküler Kamera Tabanlı Sistemler

Literatürde geliştirilen çeşitli görme engelli navigasyon sistemleri tek kamera kullanarak bu işlevi yerine getirmeyi amaçlamıştır. Meijer'in çalışmasında monoküler kamera ekipmanı görme engellilerin navigasyonunda engel tespiti için bir sensör olarak kullanılmıştır [111]. Bu çalışmada, bir başlık üzerine monte edilen tekli kamera ile görüntüler elde edilmiştir. Elde edilen görüntüler soldan sağa doğru taranarak pikselin parlaklığına göre değişen yükseklikte ses üretimi yapılmıştır. Herhangi bir görüntü işleme tekniği ile görüntüye iyileştirme uygulanmadan pikselin yoğunluğu direkt olarak sese çevrilmiştir. Bu sebeple görme engelli bir insanın navigasyonu için uygulanan bu teknik karmaşık sahnelerde yetersizdir.

Sainarayanan tarafından gerçekleştirilen monoküler kamera tabanlı çalışmada da tekli kamera sistemi kullanmıştır [112]. Bu çalışmada tekli kamera vasıtasıyla elde edilen görüntüler 32x32 ölçeğine düşürülmüştür ve görüntülerin gri ölçeği 4 seviyeye indirgenmiştir. Görüntü işleme tekniklerinin yardımıyla sahnedeki nesneler ve arka plan ayrıştırılmaya çalışılmıştır. İşlenen görüntüdeki piksellerin yoğunluğu ve dikey yönelimleri ile orantılı olarak değişik frekans ve yükseklikte ses sinyalleri üretilmiştir. Bu iki çalışmada da görüntülerdeki nesnelerin kameraya olan uzaklıkları görüntü sensörü vasıtasıyla elde edilmemiştir.

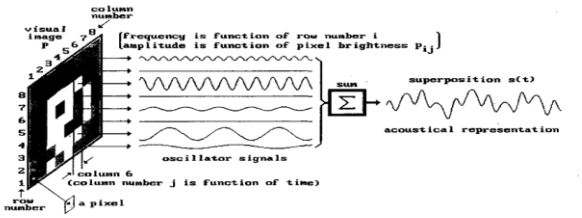
Croce ve diğerleri [113] geliştirdikleri Arianna isimli düşük maliyetli arttırılmış gerçeklik tabanlı sistemde kapalı alanlarda görme engelli kullanıcı için güvenli yürüme yolunun tespit edilmesi amaçlanmıştır. Geliştirilen sistem donanım olarak akıllı telefona entegre edilmiş bir kameradan oluşmaktadır. Kullanıcıya geri bildirim titreşim desenleri

ile yapılmıştır. Geliştirilen sistemde güvenli yürüme yolu, QR kodlar veya zemin üzerine yapılan çizimler takip edilerek belirlenmektedir. Bir diğer işaret tabanlı yaklaşım olan Manduchi'nin çalışmasında [114] Mobile Vision isimli prototip sunulmuştur. Sunulan sistem bir akıllı telefon ile entegre olarak çalışmaktadır ve kapalı alanlarda belirli renkteki işaretlerden faydalanarak çalışmaktadır. Kırmızı, yeşil ve mavi renkteki işaretleri kullanarak görme engelli kullanıcının tuvaletler, asansörler veya çıkışlar gibi bazı ilgi önemli yerleri bulmasını amaçlamaktadır. Görme engelli kullanıcıya metinden konuşmaya çeviri yoluyla sesli geri bildirim mesajları sağlanmıştır. İşaret tabanlı çalışan bu iki yöntemde de ortam hakkında ön bilgi gerekmektedir. Bu nedenle bilinmeyen sahnelerde uygulanamaz.

SmartVision [115] görme engelli navigasyon prototipi dış mekan senaryoları için tasarlanmıştır. GPS, Wi-Fi konumlama, Coğrafi Bilgi Sistemi (GIS) [116], pasif RFID etiketleri ve bilgisayarlı görme teknikleri bütünleşik olarak kullanılarak geliştirilmiştir. Sistem, beyaz bastona tamamlayıcı olarak tasarlanmıştır ve görme engelli kullanıcıyı yaklaşmakta olan tehlikeler konusunda uyarmak tasarlanmıştır. Potansiyel nesnelerin (Asansör, bitkiler, bankamatik, vs.) bulunduğu bir veri tabanı hazırlanmıştır. Test aşamasında önceden kaydedilmiş olan referans görüntüler, kamera tarafından yakalanan video görüntüleri içinde aranmaktadır. Ancak yöntem kamera hareketine karşı oldukça duyarlıdır ve başarısı büyük ölçüde eğitim veri kümesinin boyutuna bağlıdır. Ayrıca ölçeklenebilirlik sorunları bulunmaktadır çünkü daha büyük bir veri tabanı için hesaplama süresi önemli ölçüde artmaktadır.

Tapu ve diğerleri herhangi bir ek teçhizat ihtiyacı olmayan sadece sıradan akıllı telefonun yeteneklerini kullanarak çalışan bir engel tespit ve sınıflandırma sistemi önermişlerdir [117]. Geliştirilen sistem hem iç hem de dış mekanlarda görme engelli bireyin navigasyonunu kolaylaştırmak için tasarlanmıştır. Yapılan çalışmada araştırmacılar, birbirini izleyen karelerden ilgi noktalarını çıkararak Lucas-Kanade algoritmasıyla [118] engelin yerini belirlemeyi önermektedir. Çalışmada nesnelerin hareketi, RANSAC algoritması [119] uygulanarak sınıflandırılan bir dizi homografik dönüşüm yardımıyla kamera hareketinden ayırt edilmektedir. Sistem genel olarak iyi sonuçlar vermesine rağmen büyük ve düz yapıları tespit edememektedir. Ayrıca görme engelli kullanıcı ile bir engel arasındaki mesafeyi yüksek başarım ile tahmin

edememektedir. Şekil 2.12'de literatürde öne çıkan bazı monoküler kamera tabanlı görme engelli navigasyon sistemleri sunulmuştur.



(a) Meijer [111]



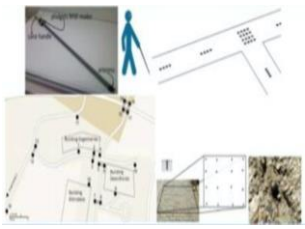
(b) Manduchi v.d. [114]



(c) Arianna [113]



(d) Tapu v.d. [117]



(e) SmartVision [115]

Şekil 2.12. Monoküler kamera tabanlı görme engelli navigasyon sistemleri

Monoküler kamera sistemlerine dayalı geliştirilen sistemler sensör ağları tabanlı sistemlere göre daha kompakt ve bakımı kolaydır. Ancak görme engelli birey ile tespit edilen bir engel arasındaki gerçek mesafeyi tahmin ederken çok düşük bir başarıma sahiptirler. Herhangi bir monoküler sistemin en büyük dezavantajı tek bir görüntüye dayalı olarak engel boyutlarını ve engelin kameraya olan uzaklığını tahmin edememesidir. Ayrıca monoküler sistemler arka plan ve ön plan nesneleri arasındaki ayrıştırmayı da gerçekleştiremezler. Ayrıca kamera hareketinden faydalanarak statik nesneleri tespit etmek için gösterilen çabalara rağmen bu tür sistemler yalnızca bilinen görünüm modellerine sahip engellerin bulunduğu sınırlı senaryolarda çalışmaktadır.

2.3.2.2. Stereo Kamera Tabanlı Sistemler

Monoküler tabanlı sistemlerinin sahip olduğu çeşitli kısıtlar sebebiyle görme engelli navigasyonunda stereo kamera kullanımı birçok araştırmacının gerçekleştirdiği çalışmada önerilmiştir. Bu bölümde literatürde öne çıkan ve stereo kameraların kullanıldığı görme engelli navigasyon sistemleri incelenmektedir.

Capp ve Picton tarafından geliştirilen Optophone isimli sistemde derinlik haritası elde etmek için stereo kamera ile elde edilen görüntülere kenar tespit etme rutini uygulamışlardır [120]. Her iki kameradan elde edilen görüntülerin içerisindeki sahnede bulunan kenar özellikleri kullanılarak görüntülerdeki aykırılık hesaplanmıştır. Çalışmada elde edilen derinlik haritası daha sonra ses sinyallerine dönüştürülerek kullanıcıya iletilmektedir. Bu sistemde görüntülerdeki tüm kenar özellikleri için aykırılık haritası elde edilmektedir. Bu nedenle görme engelli kullanıcının bilmesine ihtiyacı olmayan nesnelerin kenar özellikleri için de sinyal verilmektedir. Ayrıca sadece kenar özellikleri kullanımı ortamdaki nesnelerin algılanması için yeterli değildir.

Zelek ve diğerleri tarafından pasif stereo görüş sistemi kullanılarak gerçekleştirilen bir başka çalışmada ise görme engelli bireye çevre hakkındaki bilgiyi dokunma duyusu yoluyla geri bildirim olarak verecek bir sistem tasarlanmıştır [121]. Sistem bir dizüstü bilgisayar, pasif stereo kamera ve kullanıcının giydiği dokunsal geri bildirim verebilen bir eldivenden oluşur. Bu çalışmada, engellerin konumlarına önem verilmemektedir ve stereo kameraların uygun olmayan konumlanması nedeniyle stereo eşleme adımında problemler yaşanmaktadır.

Stereo kamera tabanlı bir diğer önemli çalışma Kawai ve Tomita tarafından geliştirilen görsel destek sistemidir [122]. Geliştirilen prototip bir dizüstü bilgisayar, üç kameralı stereo görüş sistemi ve bir mikrofonlu kulaklıktan oluşmaktadır. Kameralar tarafından çekilen görüntüler 3 boyutlu yapılandırma elde etmek için analiz edilir ve nesne tanımlama işlemi gerçekleştirilir. Sonuçlar daha sonra ses yoluyla kullanıcıya iletilir. Bu çalışmada görüntü işleme adımı derinlik hesaplamasından sonra gerçekleştirilir ve bu da işlem süresinin oldukça artmasına neden olmaktadır. Ayrıca geliştirilen prototip yalnızca kısıtlamaları olan iç ortamlarda uygulanabilmektedir.

Balakrishnan ve diğerleri tarafından geliştirilen prototip bir başlığa monte edilmiş bir gömülü bilgisayar ve stereo kamera sisteminden oluşmaktadır [123, 124]. Yapılan çalışmalarda stereo kamera sistemi ile elde edilen görüntüler üzerinde kenar tespiti uygulanarak görüntülerde bulunan nesnelerin sınırları tespit edilmiştir. Sınırları tespit edilen nesnelerin görüntü içerisinden çıkarımı sağlanmıştır. Her iki görüntü içerisinde izole edilmiş ve görüntüden ayrıştırılmış nesneleri eşleştirmek için kare farkların toplamı yöntemiyle stereo eşleme gerçekleştirilmiştir. Sonrasında aykırılık hesaplaması yapılmış ve üçgenleme tekniği ile nesnelerin kameraya olan uzaklıkları hesaplanmıştır. Sonuçlar kullanıcıya çeşitli müzik notaları kullanılarak akustik yolla iletilmiştir. Önerilen yöntem çok az nesnenin bulunduğu sahneler üzerinde test edilmiştir. Karmaşık sahnelerde veya kenar özellikleri tespit edilemeyen engellerin bulunduğu görüntülerde başarılı sonuç vermemiştir. Ayrıca görme engelli kullanıcının sesli geri bildirimi anlamlandırabilmesi için öncelikle kullanıcıya bir eğitim verilmesi ihtiyacı bulunmaktadır.

Martinez ve diğerleri yaptıkları çalışmada görme engelliler için havada asılı duran engelleri (ağaç dalı veya güneşlik gibi) tespit etmeyi amaçlayan bir prototip geliştirmişlerdir [125]. Geliştirilen prototipte kullanıcının omzuna monte ettikleri bir stereo görme sistemi kullanmışlardır. Çalışmada görme engelli kullanıcıların bir sonraki hareketini tahmin etmek ve kullanıcıların bulunduğu çevrenin 3 boyutlu haritasını oluşturmak amaçlanmaktadır. Bu 3 boyutlu harita bilgisi, kullanıcıların navigasyonu sırasında olası baş üstü engelleri değerlendirmek ve tespit etmek için kullanılır. Çalışmanın ana katkılarından biri kameranın hareketinden kaynaklı görüntüler arasındaki açı bozulmalarını stabilize etmesidir.

Rodrigez ve diğerleri tarafından gerçekleştirilen çalışmada görme engelliler için stereo kamera kullanılan bir engel tespit sistemi önerilmiştir [126]. Bu çalışmada yoğun aykırılık haritalarından faydalanılmıştır. RANSAC algoritması [119] kullanılarak düzlem uyarlama yöntemi ile yer düzlemi tahmin edilmiş ve bu yer düzlemi üzerinde bulunan engeller tespit edilmiştir. Çeşitli frekanslarda akustik geri bildirim yaparak yaklaşan engeller hakkında kullanıcıya bilgilendirme yapılmıştır. Ancak bu çalışmada önerilen yöntem, görüntülerdeki sahnenin alt bölümünde bulunan en büyük alanın yer düzlemi olacağı varsayımına dayanmaktadır. Bu nedenle birçok sahnede dikey engeller de yer düzlemi olarak algılanmıştır.

Cloix ve diğerleri tarafından sunulan Smart Walker sistemi [127] hem iç hem de dış mekanlarda tehlikeli engelleri ve zemin değişikliklerini tespit etmeyi amaçlamaktadır. Smart Walker, seyrek 3B bilgilerini kullanan gerçek zamanlı stereo engel algılama algoritmasına dayanmaktadır. Geliştirilen sistemin direkt olarak elde edilen 3B noktalar üzerinde çalışması sebebiyle ortamın 3B haritasını oluşturan yöntemlere göre hesaplama maliyeti daha düşüktür. Ancak sistemin dezavantajları ise 4 fps işlem kapasitesinin olması ve engel tespitinde %65'ten daha düşük bir başarıma sahip olmasıdır. Bu başarım seviyesi görme engelli bireylerin navigasyonu bağlamında yeterince güvenilir değildir.

Pradeep ve diğerleri görme engelli kullanıcılar için başa takılarak kullanılan stereo görme tabanlı navigasyon yardım cihazı olan Robot Vision isimli sistemi sunmuşlardır [128]. Sistem yönlendirme bilgisini sağlamak için görsel odometriyi ve özellik tabanlı topolojik eşzamanlı konumlandırma ve haritalandırma (Simultaneous Localization and Mapping, SLAM) yaklaşımlarını kullanmaktadır. Yöntem 3B sahneyi stereo üçgenleme yoluyla yeniden yapılandırmaktadır ve kullanıcının bulunduğu ortamının bir haritasını oluşturmaktadır. Oluşturulan bu harita kullanıcıyı potansiyel engellerin olduğu bir yol boyunca yönlendirmek için kullanılmaktadır. Geliştirilen prototipi görme engelli kullanıcının başına takarak kullanılması gerektiği için yeterince ergonomik olmadığı söylenebilir. Ayrıca sistemdeki 3B harita oluşturulması ihtiyacı nedeniyle navigasyon sırasında kullanıcı tarafından giyilmesi gereken güçlü bir işlem birimini gerektirmektedir.

Everding ve diğerleri tarafından önerilen giyilebilir ve hafif ağırlıklı navigasyon cihazı sabit bir çerçeve frekansında görüntünün tamamını aktarmak yerine yalnızca bir sahnede hareketin neden olduğu yerel değişiklikleri ileten dinamik görüntü sensörlerine dayanmaktadır [129]. Bu şekilde sahnedeki görsel bilgi asenkron olarak oluşturulan hareketli piksel akışı olarak temsil edilmektedir. Rapor edilen deneysel sonuçlara göre sistemin ana sınırlamasının görme engelli kullanıcının baş hizasında bulunan nesnelerin tespiti ile ilgili olduğu görülmektedir. Buna ek olarak normal kulaklıkların kullanımı kullanıcının işitmesini engellediği için optimum değildir.

Yüksek derecede dinamik sahnelerde kameranın hareketini tahmin eden görsel bir navigasyon asistanı Leung ve Medioni tarafından önerilmiştir [130]. Sahnelerden edilen özellik noktalarına uygulanan üçgenleme sonucu hesaplanan 3B koordinatlar kullanılarak sahnedeki zemin düzlemi tahmin edilmiştir. Ardından zemin düzleminin normal vektörü

tahmin edilmiştir. Deneyler sentetik ve gerçek görüntüler üzerinde gerçekleştirilmiştir. Geliştirilen sistem donanım seviyesinde güneş gözlüğüne entegre edilmiş bir stereo kamera ve masaüstü bilgisayardan oluşmaktadır. Görme engelli bir bireyin navigasyonu bağlamında bakıldığında taşınabilir olmaması en önemli dezavantajıdır.

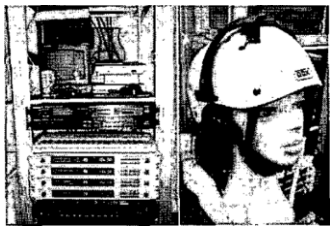
Lin ve diğerleri görme engelli kullanıcılar için giyilebilir bir stereo görüş sistemi önermişlerdir [131]. Cihaz, stereo kameranın monte edildiği bir gözlük ve gömülü bilgisayardan oluşmaktadır. Geliştirilen cihazda video akışı 3G ağı kullanılarak yapılmaktadır. Sistemin en öne çıkan başarısı aykırılık haritasını yerel olarak hesaplaması ve çeşitli engelleri tespit etme kapasitesidir.

Görme engelli kullanıcının baş bölgesine taktığı bir stereo kamera yardımıyla çevreyi algılamasına olanak tanıyan giyilebilir başka bir cihaz da Schwarze ve diğerleri tarafından önerilmiştir [132]. Yöntem bir sahne modeli oluşturur ve kullanıcı ortam içinde hareket ederken onu günceller. Ön plandaki nesneler arka plandan yoğun bir aykırılık tahmini yöntemiyle beraber görsel odometri ve IMU (Inertial Measurement Unit) kullanılarak ayrıştırılmaktadır. Geliştirilen sistem vasıtasıyla görme engelli kullanıcının dış mekân sahnelerinde navigasyonu amaçlanmıştır. Akustik geri bildirim kemikten ses iletimi yapan kulaklıklar aracılığıyla yapılmaktadır.

Literatürde yapılan görme engelli navigasyonu bağlamında gerçekleştirilen ve stereo kamera ekipmanının kullanıldığı bazı çalışmalarda ise kullanıcının gezinti yolu üzerindeki engeller görsel belirginlik haritası yaklaşımı kullanılarak tespit edilmiştir [133, 134]. Görsel belirginlik haritası yaklaşımı görüntülerde bulunan sahnelerdeki nesneleri renk yoğunluklarına ve köşe noktalarına göre tespit ettiği için patika üzerindeki engellerin tespitinde özellikle renk farklılıklarının fazla olmadığı sahnelerde yüksek başarım oranı sağlama şansı bulunmamaktadır. Çalışmaların deneysel sonuçları incelendiğinde duvar, cam ve kapı gibi engellerin tespit edilmediği ve sadece belirgin renk farklılığı bulunan nesnelerin engel olarak tespit edildiği gözlemlenmiştir.

Literatürde çok sayıda stereo görme tabanlı sistem önerilmiş olsa da bazı problemler hala açık araştırma konusu olarak kalmıştır. Zayıf dokulu yüzeylerde stereo eşleme aşamasındaki hatalar nedeniyle oluşturulan derinlik haritasının yetersiz kaldığı gözlemlenmiştir. Ayrıca gerçekleştirilen çalışmalarda tahmin edilen aykırılık haritaları

sahnelerdeki aydınlanma değişimlerinden oldukça etkilenmektedir. Yapılan bazı çalışmalarda ise önerilen yöntemlerin hesaplama maliyetlerinin fazlalığı veya başarım seviyesinin düşüklüğü sebebiyle görme engelli navigasyonunda kullanımı oldukça güçtür. Şekil 2.13'de literatürde öne çıkan bazı stereo kamera tabanlı görme engelli navigasyon sistemleri sunulmuştur.



(a) Kawai v.d. [122]



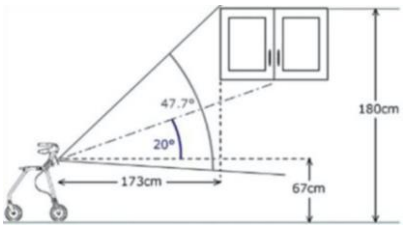
(b) Balakrishnan v.d. [123, 124]



(c) Martinez v.d. [125]



(d) Rodrigez v.d. [126]



(e) Smart Walker [127]



(f) Robot Vision [128]



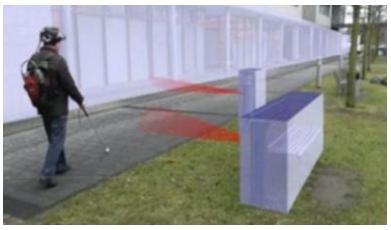
(g) Everding v.d. [129]



(h) Leung v.d. [130]



(i) Lin v.d. [131]



(j) Schwarze v.d. [132]

Şekil 2.13. Stereo kamera tabanlı görme engelli navigasyon sistemleri

2.3.2.3. Derinlik (RGB-D) Kamerası Tabanlı Sistemler

RGB-D kameraların ortaya çıkması ile son zamanlarda bu teknolojilerden yararlanan yeni bir görme engelli navigasyon sistemleri ailesini ortaya çıkarmıştır. Literatürdeki çeşitli çalışmalar [135-138], RGB verisiyle birlikte derinlik bilgisini algılama özelliğine de sahip olması sebebiyle görme engelli navigasyon cihazının ana donanım bileşeni olarak RGB-D sensörlerini kullanmanın uygulanabilirliğini göstermiştir. Ancak belirtilen bu çalışmaların hiçbirinde gerçekleştirilen engel tespit görevi başarım seviyesi ve hassasiyet açısından değerlendirilmemiştir.

Bing ve diğerleri tarafından Tango cihazına üzerinde çalışan ve engel tespit özelliklerine sahip ISANA isminde bir iç mekân navigasyon sistemi önerilmiştir [139]. Yaklaşım özellikle iç mekanlar için tasarlanmıştır ve bina mimarisi modelinden bilgi ayrıştırarak çevrenin anlamsal bir haritasını çıkarmaktadır. Engel algılama modülü derinlik sensörünün önünde bulunan nesneleri tanımlamaktadır. Yol gösterici mesajlar görme engelli kullanıcıya gerçek zamanlı olarak akustik yolla iletilir. Çalışmada görme engelli kullanıcılar ile yapılan deneysel sonuçlar ile yöntem değerlendirilmiştir. Sistemin dezavantajı sadece bina yapısı hakkında daha önceden bilgi sahibi olunan senaryolarda çalışabilmesidir. Tango tabletlere entegre edilmiş RGB-D kamera sensörü ile çalışan benzer bir yöntem Munoz ve diğerleri tarafından önerilmiştir [140]. Bu çalışmada ise iç mekanlardaki merdivenlerin tespiti amaçlanmıştır. Çalışmada içerisinde merdiven barındırma ihtimali olan RGB görüntüler Hough dönüşümü kullanılarak bir dizi eşzamanlı paralel çizgi çıkarımı yapılarak tespit edilmektedir. Daha sonra derinlik görüntüleri üst kat, alt kat ve koridor gibi birden çok kategoriyle eğitilmiş bir Destek Vektör Makinesi (Support Vector Machines, SVM) sınıflandırıcısı kullanarak analiz edilmiştir.

Blessenohl ve diğerleri görme engelli bireylerin iç mekanlarda güvenli bir şekilde dolaşabilmesi için baş bölgesine takılan ve derinlik kamerası kullanılan bir sistem dizayn etmişlerdir [141]. Sistem sahnedeki zemin, duvar veya engelleri tespit etmek ve algılama modülü tarafından verilen bilgileri 3B bir ses haritasına dönüştürmek için tasarlanmıştır. Yöntem gerçek zamanlı olarak çalışmaktadır ve zemin üzerinde bulunan küçük ve büyük nesneleri tespit etmede başarılı sonuçlar göstermiştir.

Brilhault ve diğerleri [142] coğrafi bilgi sistemleri, bilgisayarlı görme ve GPS tabanlı çalışan NAVIG isimli görme engelli navigasyon sistemini önermişlerdir. Sistem RGB-D kamera, mikrofon, kulaklık ve cep bilgisayarından oluşmaktadır. Diğer sistemlerin aksine nesne konumlandırma fonksiyonu tehlikeli durumlardan kaçınmak için değil, görme engelli kullanıcıyı ilgilenilen noktaya ulaşmalarında yönlendirmek için kullanılır. Deneysel sonuçlar incelendiğinde sistemin işleme hızının nesne boyutlarından ve veri tabanına dahil edilen maksimum yer işareti sayısından önemli ölçüde etkilendiği gözlemlenmektedir. Ayrıca kullanıcı için kapsamlı bir eğitim aşaması gerektirmektedir.

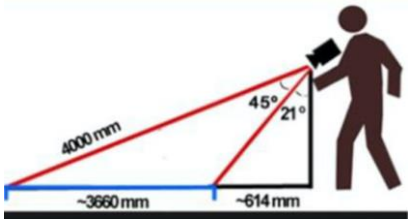
Caraiman ve diğerleri tarafından RGB-D kamera tabanlı çalışan Sound of Vision isimli görme engelli navigasyon sistemi önerilmiştir [143, 144]. Sistem çevredeki ortamdan elde edilen bilgileri görme engelli kullanıcıya ses ve dokunsal yolla iletmektedir. Önerilen sistemde iç ve dış mekân sahnelerin 3B yeniden yapılandırması gerçekleştirilmektedir. İlgi alanındaki nesnelerin görüntülerden çıkarımı için görüntü ayrıştırması teknikleri uygulanmaktadır. Sistem oldukça karmaşıktır ve görme engelli kullanıcının uyarı mesajlarını anlaması için bir eğitim aşaması gerekmektedir. Şekil 2.14'de literatürde öne çıkan bazı RGB-D kamera tabanlı çalışmalar sunulmuştur.



(a) Object Sonif. [135]



(b) KinDectect [136]



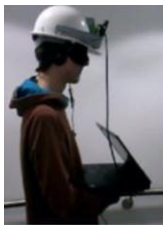
(c) Filipe v.d. [137]



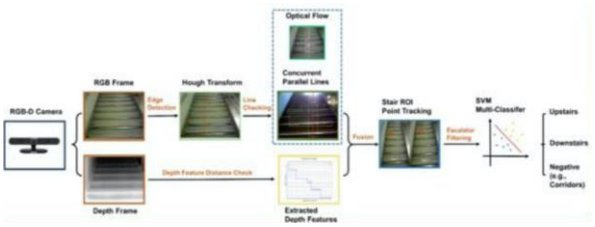
(d) Kinect Cane [138]



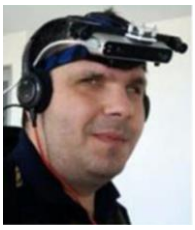
(e) ISANA [139]



(f) Blessenohl v.d. [141]



(g) Munoz v.d. [140]



(h) Sound of Vision [143, 144]

Şekil 2.14. Derinlik kamerası tabanlı görme engelli navigasyon sistemleri

BÖLÜM 3. MATERYAL VE YÖNTEM

Bu bölümde tez çalışması sırasında kullanılan materyallerin, önerilen yöntemin ve geliştirilen duyusal ikame cihazı prototipinin detayları verilmektedir.

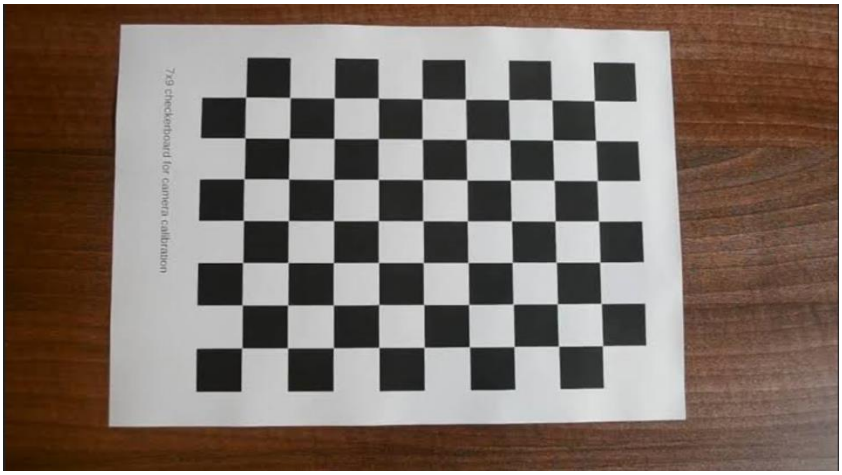
3.1. Kullanılan Materyal

Çalışma kapsamında görüntülerin elde edilmesi için Fijutsu tarafından üretilen FinePix REAL 3D W3 model dijital stereo kamera ekipmanı kullanılmıştır. Kullanılan kamera ekipmanı ve görme engelli bireyin kamerayı taşıyabilmesi için kullanılan askı aparatının görüntüsü Şekil 3.1'de gösterilmektedir.



Şekil 3.1. Deneyde kullanılan stereo kamera ekipmanı ve askı aparatı

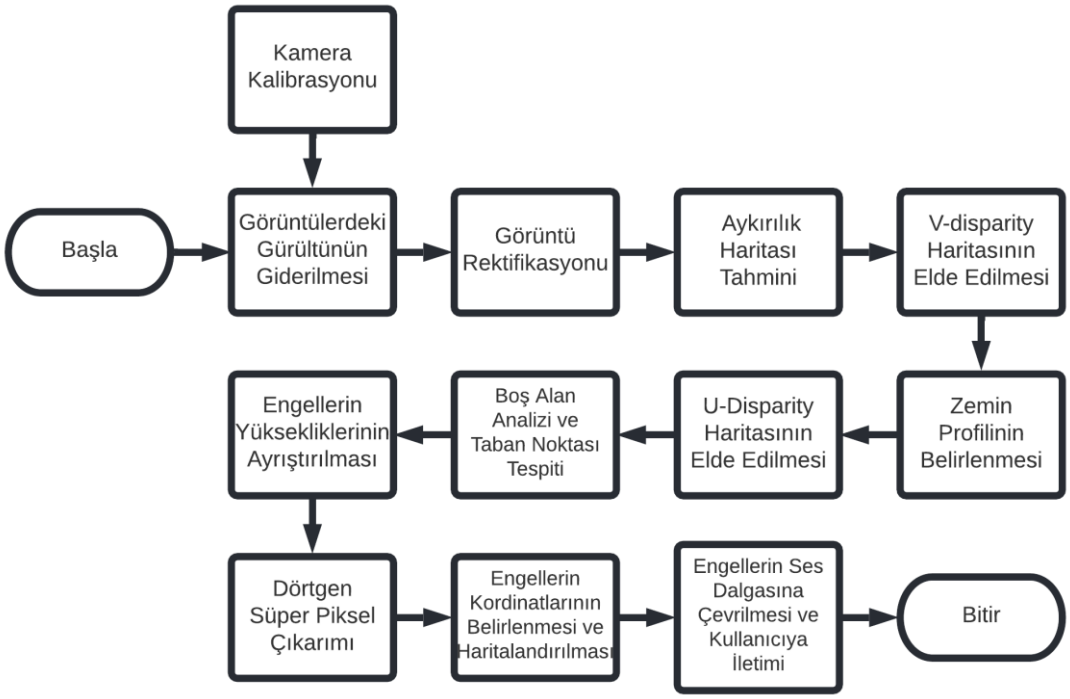
Deneyde kullanılan prototipin yazılım modülü MATLAB geliştirme ortamı kullanılarak geliştirilmiştir. Geliştirilen prototip, i7 işlemcili 16 GB bellek kapasitesine sahip bir bilgisayar üzerinde koşturulmuştur. Kamera kalibrasyonu aşamasında Şekil 3.2'deki satranç tahtası deseni kullanılmıştır.



Şekil 3.2. Kamera kalibrasyonu işleminde kullanılan satranç tahtası deseni

3.2. Önerilen Yöntem

Bu tez çalışmasında, pasif stereo kamera ekipmanı ile elde edilen görüntüler işlenerek görme engelli bireylerin yürüyüş yolları üzerindeki engellerden kaçınmalarını sağlamayı hedefleyen gerçek zamanlı bir sistem geliştirilmiştir. Geliştirilen sistem üç aşamadan oluşmaktadır. İlk aşamada kullanılan stereo kameranın dahili ve harici parametrelerinin belirlenebilmesi için kamera kalibrasyonu işlemi uygulanmaktadır. Eğer kullanılan stereo kameranın piksel alanı değeri biliniyorsa ilk aşamanın uygulanmasına gerek olmadan ikinci aşama uygulanmaktadır. İkinci aşama ise görüntülerdeki engellerin tespitinin yapılmasını sağlayan gerçek zamanlı adımlardan oluşmaktadır. Son aşamada ise sahnede tespit edilen engeller akustik yolla kullanıcıya iletilmektedir. Geliştirilen prototipin akış şeması Şekil 3.3'de gösterilmiştir.



Şekil 3.3. Duyusal ikame cihazının akış şeması

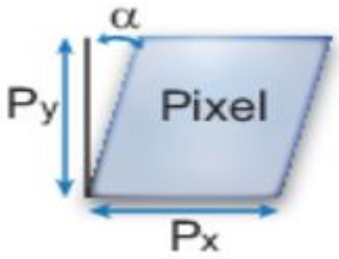
3.2.1. Stereo Kamera Kalibrasyonu

Güvenilir engel tespiti sonucu elde etmek için stereo kamera kalibrasyonu işlemi gereklidir. Stereo kamera kalibrasyon problemi kameraların iç (intrinsic), dış (extrinsic) ve gürültü (distortion) parametrelerinin tahminini içermektedir. Bu çalışma kapsamında Matlab Camera Calibration Toolbox [145] programı kullanılarak her iki kamera için kalibrasyon işlemi uygulanmıştır. Böylelikle her kamera için dahili kamera matrisi (intrinsic camera matrix) K elde edilmiştir.

$$K = \begin{bmatrix} f_x & 0 & 0 \\ s & f_y & 0 \\ c_x & c_y & 1 \end{bmatrix} \tag{23}$$

Eşitlik (23)'de f_x ve f_y terimleri kameranın odak uzaklığının (focal length) piksel cinsinden değerini, c_x ve c_y terimleri kamera optik merkezinin (optical center) piksel

koordinatını ve s terimi ise görüntü eksenlerinin dörtgen olmaması durumunda ortaya çıkan eğrilik katsayısını ifade etmektedir. Şekil 3.4'de piksel eğriliği gösterilmektedir.



Şekil 3.4. Piksel eğriliği

Piksel eğrilik katsayısı Eşitlik (24) ile elde edilmektedir.

$$s = f_x \tan \alpha \tag{24}$$

Stereo kamera kalibrasyon parametrelerinin elde edilebilmesi için satranç tahtası deseni düz bir yüzeye yapıştırılmıştır. Satranç tahtası deseninin bulunduğu çerçevenin çeşitli mesafelerden ve açılardan görüntüsü alınmıştır. Kalibrasyon işlemi için kullanılan bazı örnekler görüntüler Şekil 3.5'de paylaşılmıştır.



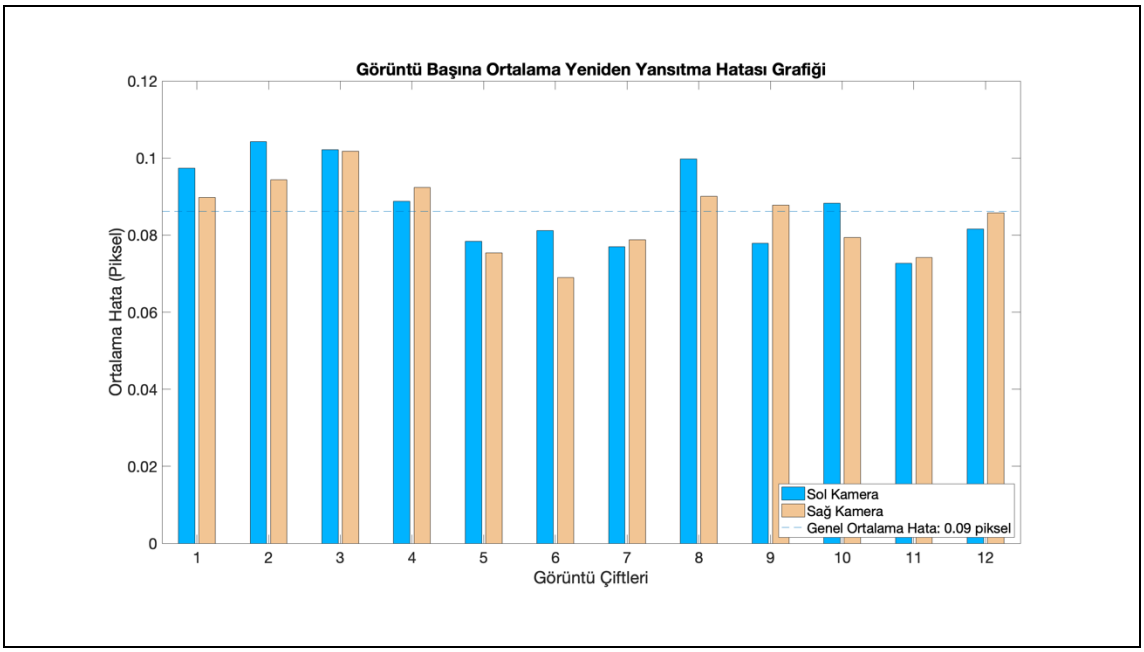
Şekil 3.5. Kalibrasyon işleminde kullanılan görüntülerden örnekler

Kameraların her birinin kalibrasyonundan sonra, stereo kamera kurulumunun dışsal parametreleri tahmin edilir. Harici parametreler (extrinsic parameters), stereo kamera kurulumundaki sol ve sağ kameralar arasındaki ilişkiyi tanımlayan bir rotasyon matrisi (rotation matrix) R ve bir kayma vektörü (translation vector) t içermektedir. Kameraların dahili ve harici parametreleri elde edildikten sonra her bir kameranın kamera matrisi P aşağıdaki eşitlik ile elde edilmektedir.

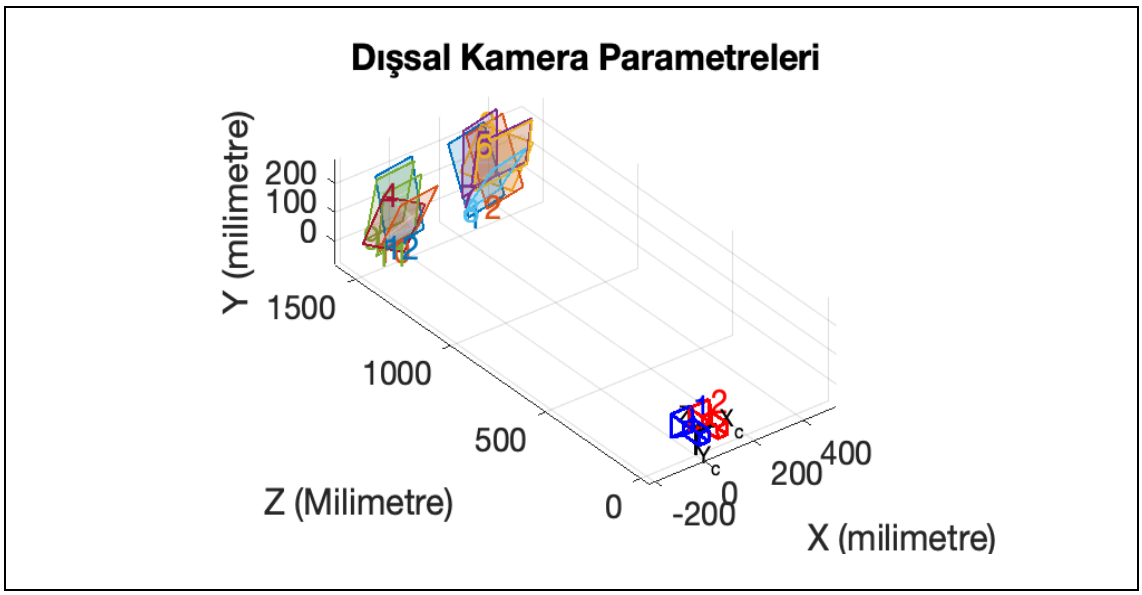
$$P = \begin{bmatrix} R \\ t \end{bmatrix} K \tag{25}$$

Stereo kamera kalibrasyon işleminin doğruluğunu gösteren sütun grafiği Şekil 3.6'da gösterilmektedir. Grafikte her bir stereo görüntü çifti için hesaplanan ortalama yansıtma hatası paylaşılmaktadır. Kalibrasyon işlemi sonucunda genel ortalama hata ise 0.09 piksel olarak hesaplanmıştır. Stereo kamera kalibrasyon işleminde kullanılan satranç

tahtası desenlerinin stereo kamera ekipmanına göre konumlarını gösteren grafik Şekil 3.7'de paylaşılmaktadır.



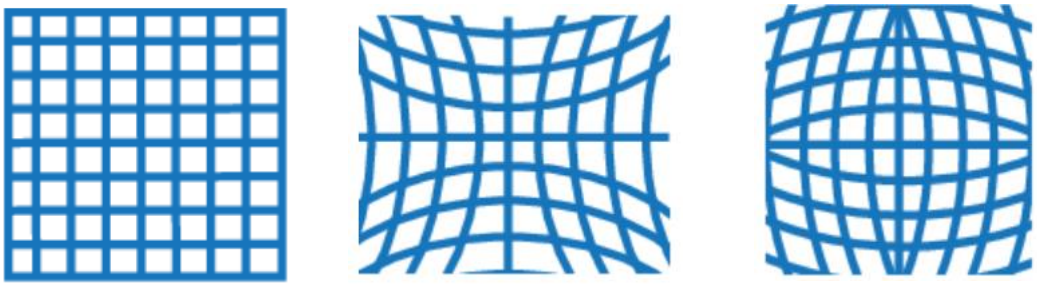
Şekil 3.6. Kalibrasyon işlemi sonucunda görüntü çiftleri için hesaplanan ortalama yeniden yansıtma hatası grafiği



Şekil 3.7. Kullanılan kalibrasyon desenlerinin kameraya göre pozisyonlarının 3 boyutlu görünümü

3.2.2. Görüntülerdeki Geometrik Bozulmaların Giderilmesi

Kamera matrisi lens gürültüsünü hesaba katmamaktadır çünkü ideal bir iğne deliği modeli kameranın (Pinhole Camera Model) lens gürültüsü bulunmamaktadır. Gerçek bir kamerayı doğru şekilde temsil etmek için kamera modeli radyal ve teğetsel lens bozulmalarını içermektedir. Radyal bozulma, kameraya ulaşan ışınların lensin kenarlarında optik merkezde olduğundan daha fazla büküldüğünde meydana gelmektedir. Şekil 3.8'de negatif ve pozitif radyal bozulmalar gösterilmektedir.



a) Bozulma yok b) Negatif radyal bozulma c) Pozitif radyal bozulma

Şekil 3.8. Radyal bozulmalar

Mercek ve görüntü düzlemi paralel olmadığında ise teğetsel bozulma meydana gelmektedir. Teğetsel bozulma katsayıları bu tür bozulmaları modellemektedir. Radyal (k_1, k_2, k_3) ve teğetsel (p_1, p_2) bozulma parametreleri polinom yaklaşımları kullanılarak modellenmiştir [146]. Şekil 3.9-a'da kamera sensörü ile lens birbirine paralel olduğu için teğetsel bozulma olmamaktadır. Şekil 3.9-b'de ise kamera sensörü ve lensin birbirine paralel olmaması durumu gösterilmektedir. Bu durumda teğetsel bozulma gerçekleşmektedir.



(a) Lens ve sensör paralel (b) Lens ve sensör paralel değil

Şekil 3.9. Teğetsel bozulma

3.2.3. Görüntülere Rektifikasyon Uygulanması

Görüntüler içerisinde birbirine karşılık gelen stereo eşleşmeleri tahmin etmeden önce görüntülere geometrik düzeltme işlemi uygulamak yaygın bir yöntemdir. Geometrik düzeltme işlemi Bölüm 2.1.1.3'de bahsedildiği gibi görüntü çifti içerisindeki her bir görüntüye izdüşüm dönüşümü uygulanarak gerçekleştirilmektedir. Rektifikasyon işlemi sayesinde her iki görüntüdeki eşleşen noktaların konumları arasında sadece yatay yönde farklılık olması sağlanmaktadır. Bu sayede stereo eşleme aşamasında arama uzayı 3 boyuttan 2 boyuta düşmektedir. Şekil 3.10'da görüntü rektifikasyonu uygulanmış bir stereo görüntü çiftinin kırmızı-mavi anaglifi gösterilmektedir.



Şekil 3.10. Görüntü rektifikasyonu uygulanmış bir stereo görüntü çiftinin anaglifi

3.2.4. Aykırılık Haritası Tahmini

Aykırılık haritası tahmininin sonucu Bölüm 2.1.2'de detayları verilen stereo eşleme yöntemlerine göre farklılık göstermektedir. Yoğun (dense) eşleme yaklaşımı ile aykırılık haritası hesaplama maliyeti daha yüksektir ancak görüntülerdeki tüm pikseller için aykırılık değeri tahmini yapılmaktadır. Bu nedenle yoğun aykırılık haritası üretilmektedir. Seyrek (sparse) eşleme yaklaşımında ise hesaplama maliyeti düşüktür ama sadece görüntüdeki özellik noktaları için seyrek aykırılık haritası oluşturulmaktadır. Seyrek aykırılık haritasında görüntülerdeki sahneye ait 3 boyutlu bilgilerde kayıp yaşanması nedeniyle bu çalışma kapsamında yoğun aykırılık haritası tahmini yapılmıştır. Aykırılık haritası oluşturmak için sağlamlığı ve hesaplama performansı nedeniyle Bölüm 2.1.2.5'de detayları verilen Hirschmuller'in [46] yarı global eşleme yöntemi kullanılmıştır. Şekil 3.11'de deneyler sırasında kullanılan veri setindeki örnek bir görüntü ve o görüntü için tahmin edilen yoğun aykırılık haritası gösterilmektedir.



(a) Orijinal sol görüntü (b) Stereo aykırılık haritası

Şekil 3.11. Deneylerde kullanılan bir görüntü ve stereo aykırılık haritası

Stereo görüntü çiftleri arasında eşleme yapılırken Bölüm 2.1.2.3'de bahsedilen benzerlik ölçüleri kullanılabilir. Mutlak farkların toplamı (SAD) veya kare farkların toplamı (SSD) gibi benzerlik ölçüleri çalışma zamanı bakımından daha iyi performans göstermektedir. Normalleştirilmiş çapraz korelasyon (NCC) benzerlik ölçüsü, görüntü çiftleri arasındaki aydınlanma değişimlerine daha az duyarlı olmakla beraber iç ve dış alan fark etmeksizin çalışacak bir sistem için daha iyi sonuçlar vermektedir. Bu nedenle bu tez çalışması kapsamında normalleştirilmiş çapraz korelasyon (Eşitlik (26)) benzerlik ölçüsü olarak kullanılarak stereo eşleme işlemi gerçekleştirilmiştir [44].

$$\gamma(u,v) = \frac{\sum_{x,y}\left[f(x,y) - \bar{f}_{u,v}\right]\left[t(x-u,y-v) - \bar{t}\right]}{\sqrt{\left(\sum_{x,y}\left[f(x,y) - \bar{f}_{u,v}\right]^2 \sum_{x,y}\left[t(x-u,y-v) - \bar{t}\right]^2\right)}} \tag{26}$$

Eşitlik (26)'da γ terimi korelasyon çıktısını, f terimi korelasyon bulunmaya çalışılan büyük görüntüyü, t terimi karşılaştırma yapılacak pencereyi, $\bar{t}$ pencere içerisindeki

değerlerin ortalamasını ve $\bar{f}_{u,v}$ ise pencere ile karşılaştırma yapılan $f(x,y)$ bölgesinin içerisindeki değerlerin ortalamasını vermektedir.

3.2.5. Zemin Profilinin Belirlenmesi

Sahnedeki zemin profilinin belirlenmesi v-disparity haritasının hesaplanması ve sahnedeki zemin profilinin tespiti aşamalarından oluşmaktadır.

3.2.5.1. V-Disparity Haritasının Hesaplanması

V-Disparity haritası [57], bir sahnenin geometrisinin iyi bir şekilde yorumlanmasına olanak sağlamaktadır. V-Disparity haritası, aykırılık haritasının her satırında ortaya çıkan çeşitli aykırılık değerlerinin histogramı olarak açıklanmaktadır. V-Disparity haritasının en önemli özelliklerinden biri sahnedeki ana düzlemsel yüzeylere karşılık gelen çizgisel temsillerin v-disparty görüntüsünde bulunmasıdır. Dikey yüzeyler, v-disparity görüntüsündeki dikey çizgi parçalarına karşılık gelirken, sahnenin zemini eğik bir çizgi parçasına karşılık gelmektedir. Bu çalışma kapsamında v-disparity haritası, stereo görüntü çiftleri için oluşturulan aykırılık haritası (Şekil 3.11) kullanılarak oluşturulmaktadır. Matematiksel olarak Eşitlik (27)'deki gibi ifade edilmektedir. Eşitlik (27)'de F terimi aykırılık haritasının herhangi bir satırındaki aynı aykırılık değerine sahip piksellerin sayısını toplayan fonksiyonu, $D_v(x,d)$ terimi v-disparity haritasını ve $D(x,y)$ terimi ise aykırılık haritasını temsil etmektedir.

$$F_v\big(D(x,y)\big) = D_v(x,d) \tag{27}$$

Algoritma 1'de v-disparity haritasını oluşturma adımları sunulmuştur. Kullanılan veri seti içerisindeki örnek bir görüntü için v-disparity oluşturma algoritmasının adımlarının uygulanması sonucu elde edilen v-disparity görüntüsü Şekil 3.12'de gösterilmektedir.

Algoritma 1: V-Disparity Haritası Hesaplaması

Girdi: Aykırılık Haritası $D(x, y)$

Çıktı: V-Disparity Haritası $D_v(x, d)$

for *each* i[th] column in D **do**

  for *each* j[th] row in D **do**

    if $D(i, j) > 0$ **then**

      $D_v\big(j, D(i, j)\big) + +$



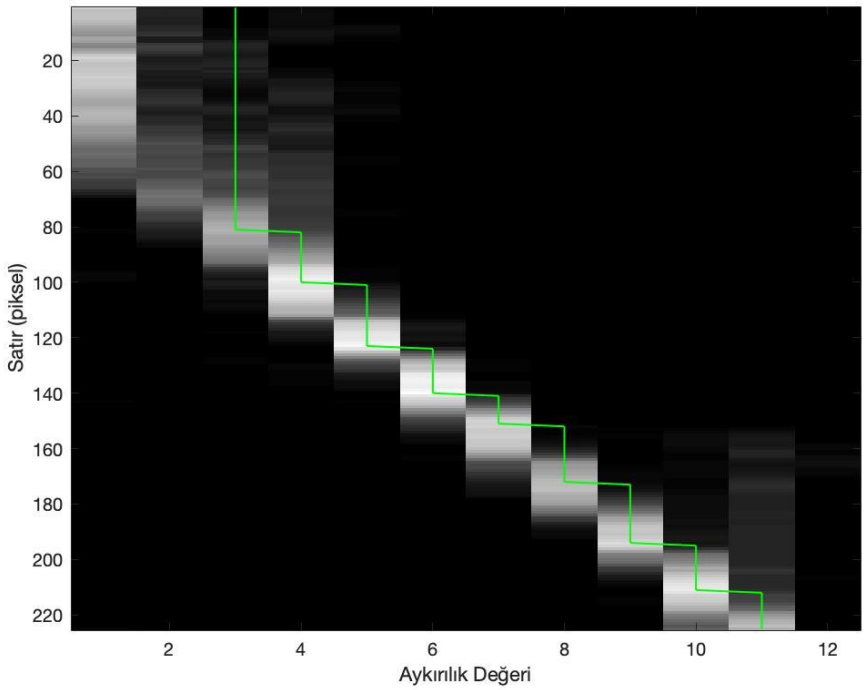
(a) (b)

Şekil 3.12. V-Disparity haritası ve orijinal sol görüntü

3.2.5.2. Sahnedeki Zemin Profilinin Tespiti

Sahnedeki zeminin profilini çıkarabilmek için bir önceki aşamada elde edilen v-disparity haritasına dinamik programlama tekniği ile çizge kesme işlemi uygulanmıştır. Çizge kesme işlemi sonucu elde edilen en küçük kesim, sahnedeki zemin profilini ortaya çıkarmaktadır. Sahnedeki zemin profilini daha hassas bir şekilde ortaya çıkarabilmek için zeminin düzlemsel olduğunu varsaymak yerine v-disparity haritasına B-Splayn (B-

Spline) eğrisi uyarlayarak tahmin edilmiştir. Çizge kesme işlemi uygulanırken maliyet terimi olarak v-disparity haritasındaki dikey yönlü gradyent değerleri kullanılmıştır. Şekil 3.13'de v-disparity görüntüsüne uygulanan çizge kesme işlemi sonrası elde edilen en küçük kesi yeşil çizgi ile gösterilmektedir. Elde edilen en küçük kesim, sahnedeki zeminin profilini temsil etmektedir.



Şekil 3.13. V-Disparity haritasında zemin profilini gösteren en küçük kesi (yeşil çizgi)

3.2.6. Engellerin Taban Noktalarının Tespiti

Engellerin taban noktalarının tespiti, u-disparity haritasının ve görüntüdeki boş alanların hesaplanması aşamalarından oluşmaktadır.

3.2.6.1. U-Disparity Haritasının Hesaplanması

U-Disparity haritası [147], v-disparity haritası ile aynı konsepte sahiptir. Başlıca ve en önemli fark ise u-disparity haritasının, aykırılık haritasındaki aykırılık değerlerinin sütun şeklinde temsili olmasıdır (Şekil 3.14). U-Disparity haritası, sahnede bulunan

engelleri birden çok yatay çizgiyle temsil ederek bilgi sağlamaktadır. Matematiksel olarak Eşitlik (28)'deki gibi ifade edilmektedir. Eşitlik (28)'de F terimi aykırılık haritasının herhangi bir kolonundaki aynı aykırılık değerine sahip piksellerin sayısını toplayan fonksiyonu, $D_u(d,y)$ u-disparity haritasını ve $D(x,y)$ ise aykırılık haritasını temsil etmektedir.

$$F_u\big(D(x,y)\big) = D_u(d,y) \tag{28}$$

Algoritma 2'de u-disparity haritasını oluşturma adımları sunulmuştur. Kullanılan veri seti içerisindeki örnek bir görüntü için u-disparity oluşturma algoritmasının adımlarının uygulanması sonucu elde edilen u-disparity görüntüsü Şekil 3.14'de gösterilmektedir.

Algoritma 2: U-Disparity Haritası Hesaplaması

Girdi: Aykırılık Haritası $D(x,y)$

Çıktı: U-Disparity Haritası $D_u(d,y)$

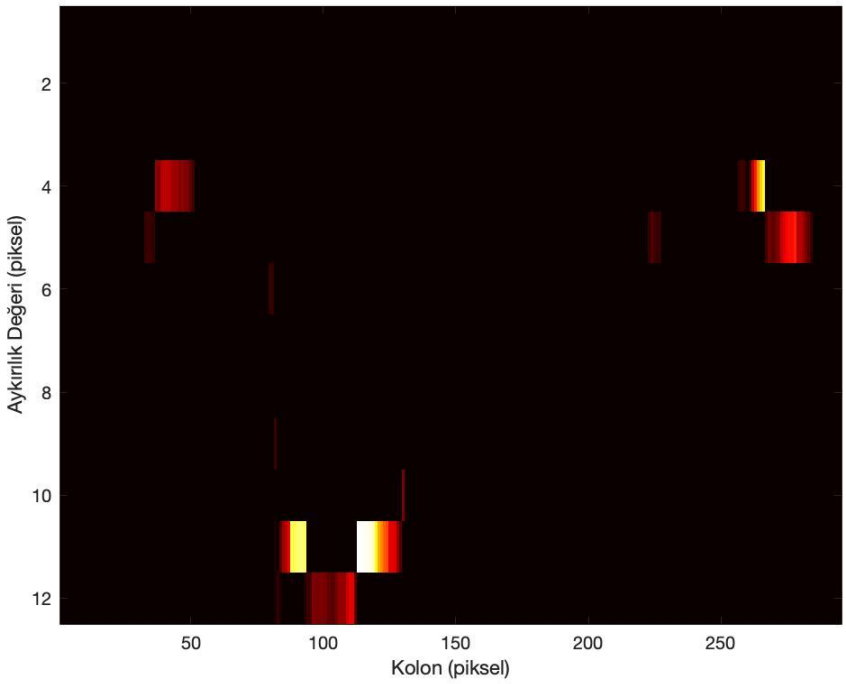
```
for each i^th column in D do
    for each j^th row in D do
        if D(i,j) > 0 then
            D_u(D(i,j), i) ++
```



Şekil 3.14. Orijinal sol görüntü ve görüntü için oluşturulan U-Disparity haritası

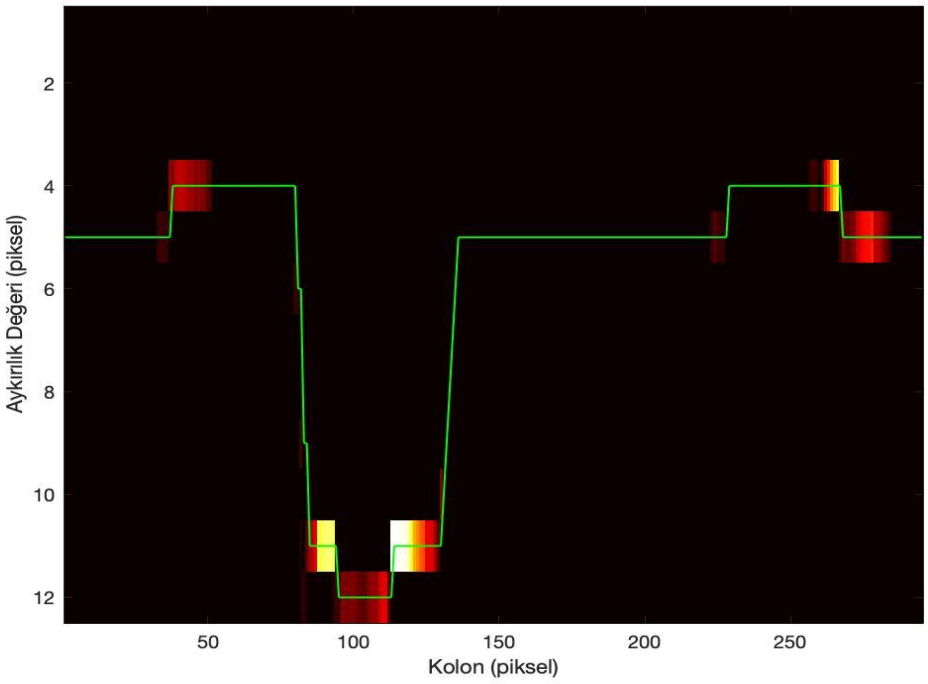
3.2.6.2. Boş Alanların Hesaplanması

Boş alan analizindeki görev pozitif derinlik yönündeki ilk engelin bulunmasıdır. Şekil 3.15'de görülebileceği üzere bunun anlamı arama işleminin dikey yönde görüntünün en altından başlaması ve ilk işgal edilen pikselin bulunmasıdır. Bu pikselin altındaki boşluk, sahnedeki boş alan olarak kabul edilmektedir.

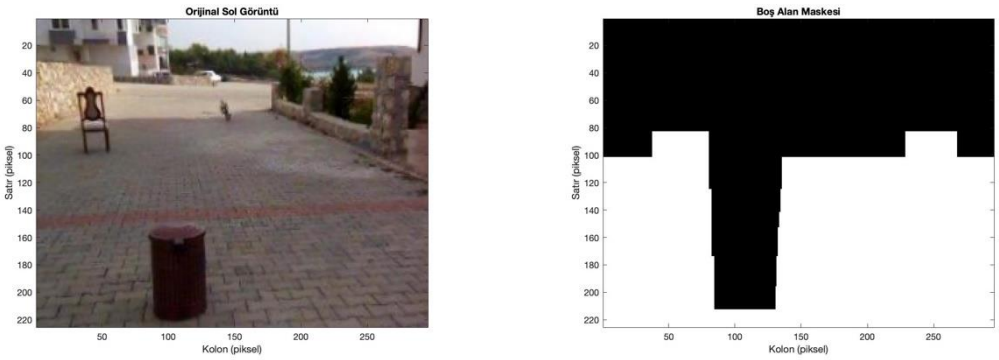


Şekil 3.15. U-Disparity haritasının işgal gridi olarak gösterimi

U-Disparity haritasındaki her bir kolona eşikleme işlemi uygulamak yerine dinamik programlama kullanılarak çizge kesme işlemi uygulanmıştır. Çizge kesme işlemi sonucunda elde edilen en küçük kesi, sahnedeki boş alanlarını bitimini göstermekle birlikte o pozisyonda bulunan potansiyel engellerin taban noktalarını da vermektedir. Şekil 3.16'da dinamik programlama kullanılarak elde edilen en küçük kesi yeşil çizgi ile gösterilmektedir. Elde edilen en küçük kesi (u, d_u) vektör koordinatları kümesidir. u değeri görüntüdeki kolonları, d_u değeri ise boş alanın mevcut olduğu mesafeye karşılık gelen aykırılık değerini temsil etmektedir. Şekil 3.17'de boş alan analizi sonucu elde edilen en küçük kesinin noktaları kullanılarak oluşturulan boş alan maskesi gösterilmektedir.



Şekil 3.16. U-Disparity haritasına uygulanan çizge kesme işlemi sorası ortaya çıkan en küçük kesi (yeşil çizgi)



Şekil 3.17. Orijinal sol görüntü ve boş alan analizi sonucu edilen boş alan maskesi (beyaz bölge)

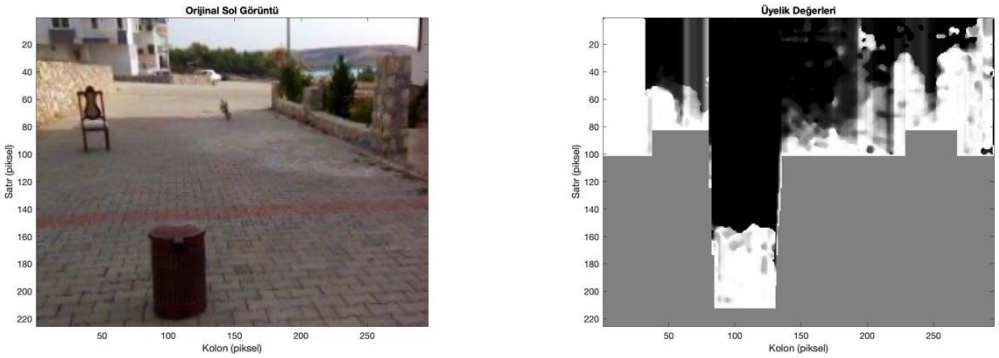
3.2.7. Engellerin Yüksekliklerinin Tespiti

Sahnede bulunan engellerin yükseklikleri, ön plan ve arka plan aykırılık değerleri arasındaki en uygun ayrıştırmayı bularak elde edilmektedir. Bu ayrıştırma için öncelikle bir maliyet görüntüsü hesaplanmaktadır. Daha sonra nesnelerin üst sınırını bulmak için hesaplanan maliyet görüntüsüne dinamik programlama ile çizge kesme işlemi uygulanmaktadır. Çizge kesme işlemindeki görev, boş alan analizinden elde edilen (u, d_u) vektör koordinatları kümesi içerisindeki her bir nokta için o konumdaki nesnenin üst sınırını gösteren en iyi satır pozisyonu olan v_t değerini elde etmektir.

Önerdiğimiz yaklaşımda aykırılık görüntüsünde bulunan her bir aykırılık değeri $d(u, v)$ (sütun u ve satır v'deki aykırılık değeri), ön plan nesnesi olup olmadığına dair oy vermektedir. Daha açık bir ifadeyle, herhangi bir aykırılık değeri, taban noktası tespit edilen nesnenin beklenen aykırılık değerinden maksimum uzaklıktan fazla sapmaması durumunda ön plan nesnesi olduğuna dair pozitif oy aksi durumda ise negatif oy kullanmaktadır. Mantıksal (boolean) atamalar maksimum uzaklık için belirlenen eşik değerini çok hassas hale getirmektedir. Belirlenen eşik değeri çok büyükse tüm noktalar ön plan üyeliği için oy verebilmektedir veya eşik değeri çok küçükse tüm noktalar arka plan üyeliği için oy verebilmektedir. Bu nedenle noktaların üyelik oylaması Eşitlik (29)'daki gibi sürekli üstel bir fonksiyon kullanarak yapılmaktadır.

$$M_{u,v}(d) = 2^{(1-\,|d-\,d_u|)} - 1 \tag{29}$$

Eşitlik (29)'da d_u terimi boş alan analizinden elde edilen vektördeki aykırılık değerini temsil etmektedir. Diğer bir ifadeyle d_u terimi u kolonundaki ön plan nesnesinin beklenen başlangıç aykırılık değeridir. Şekil 3.18'de deneyler sırasında kullanılan bir görüntü için elde edilen üyelik değerleri gösterilmektedir.



Şekil 3.18. Örnek bir görüntü ve bu görüntü için elde edilen üyelik değerleri

Elde edilen üyelik değerleri kullanılarak maliyet görüntüsü Badino ve diğerlerinin çalışmasında [2] önerildiği şekilde Eşitlik (30)'daki gibi hesaplanmaktadır. Eşitlikteki v_f terimi boş alan analizinden elde dilen nesnelerin taban noktalarının satır pozisyonunu temsil etmektedir.

$$C(u,v) = \sum_{i=0}^{i=v-1} M_{u,v}\big(d(u,i)\big) - \sum_{i=v}^{i=v_f} M_{u,v}\big(d(u,i)\big) \tag{30}$$

Oluşturulan maliyet görüntüsü üzerinde çizge kesme işlemi sonucunda elde edilen en küçük kesi sahnedeki engellerin yüksekliklerini göstermektedir. Şekil 3.18'deki görüntüye uygulanan yükseklik ayrıştırması sonucu Şekil 3.19'da gösterilmektedir.



Şekil 3.19. Yükseklik ayrıştırması sonucu (sarı çizgi)

3.2.8. Sahnedeki Engellerin Ayrıştırılması

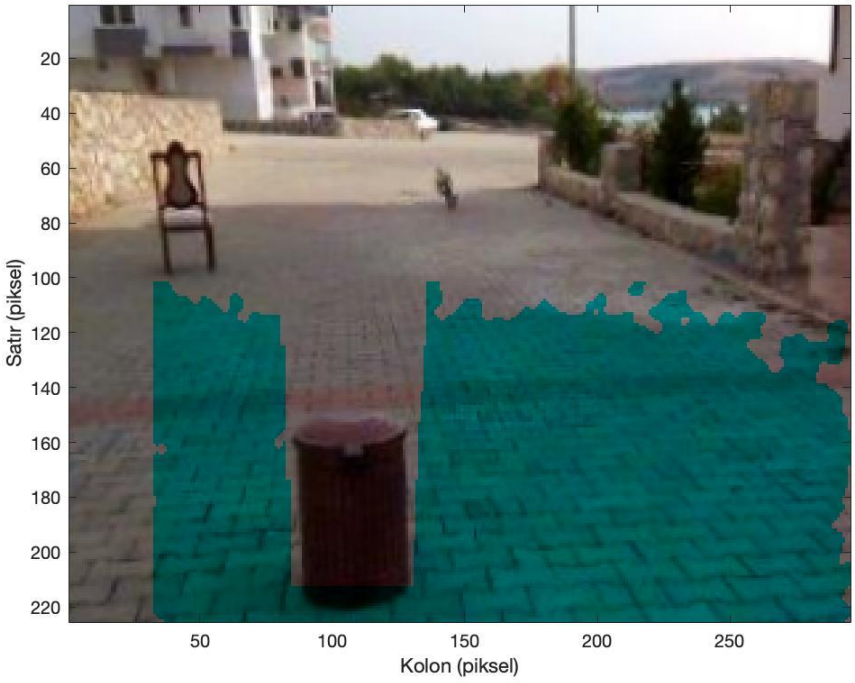
Görüntüdeki her kolon için boş alanlar ve engellerin yükseklikleri hesaplandıktan sonra dörtgen süper piksel çıkarımı yaklaşımı ile sahnede bulunan engeller görüntüden ayrıştırılmaktadır. Bu işlemi gerçekleştirebilmek için öncelikle görüntünün üzerinde arama yapılacak olan dörtgen süper piksellerin genişlikleri atanmaktadır. Bu genişliğin az olması görüntü üzerinde arama yapılacak alanı arttırmaktadır. Dolayısıyla ayrıştırmanın daha detaylı yapılmasına olanak vermektedir. Ancak arama yapılacak alanın artması sebebiyle işlem zamanı da artmaktadır. Örneğin dörtgen süper piksel genişliği 1 piksel belirlenirse 320 piksel genişliğinde bir görüntü için 320 kolonda arama yapılması anlamına gelmektedir. Bu bakımdan ihtiyaç duyulan performans ve istenilen ayrıştırma detayı dengesini koruyacak ideal bir genişlik seçilmesi faydalıdır. Gerçekleştirilen çalışmada 320 piksel genişliğindeki görüntüler için dörtgen süper piksel

genişliği 4 piksel olarak atanmıştır. Yani görüntü üzerinde 80 kolonda arama yapılmaktadır.

Taban noktası v_b, üst nokta v_t ve dörtgen süper piksel genişliği kullanılarak bir dörtgen süper piksel çerçevesi elde edilebilmektedir. Şekil 3.20'de test veri setindeki bir görüntü için çıkarımı yapılan dörtgen süper pikseller gösterilmektedir. Dörtgen süper piksellerin renkleri, içerisinde bulunan aykırılık değerlerini temsil etmektedir. Görme engelli bir bireyin gezinti yolu üzerindeki sadece belirli bir mesafede arama yapılması yeterli olduğu için belirli bir aykırılık değerinden daha küçük aykırılık değerine sahip olan dörtgen süper piksel çıkarımları elenmiştir. Şekil 3.21'de, Şekil 3.20'deki sahne için tespit edilen ve kullanıcıya belirli bir mesafe uzaklıktaki boş alanlar gösterilmektedir. Geliştiren duyusal ikame sistemi prototipinde görme engelli kullanıcı bu boş alanlara yönlendirilmektedir.



Şekil 3.20. Görüntüde tespit edilen dörtgen süper pikseller



Şekil 3.21. Görme engelli kullanıcının yönelebileceği sahnedeki boş alanlar

3.2.9. Derinlik Çıkarımı

Görme engelli bir bireyin güvenli bir şekilde mobilize olabilmesi için sahnedeki engellerin bireye ulan uzaklıklarının tespit edilmesi gerekmektedir. Engellere dair derinlik bilgisinin çıkarılabilmesi için engellerin ağırlık merkezi noktasına üçgenleştirme işlemi uygulanmaktadır. Matematiksel ifadesi Eşitlik (31) ve (32)'de gösterilmektedir.

$$Z = \frac{b\alpha_x}{d} \tag{31}$$

$$\alpha_x = f / s_x \tag{32}$$

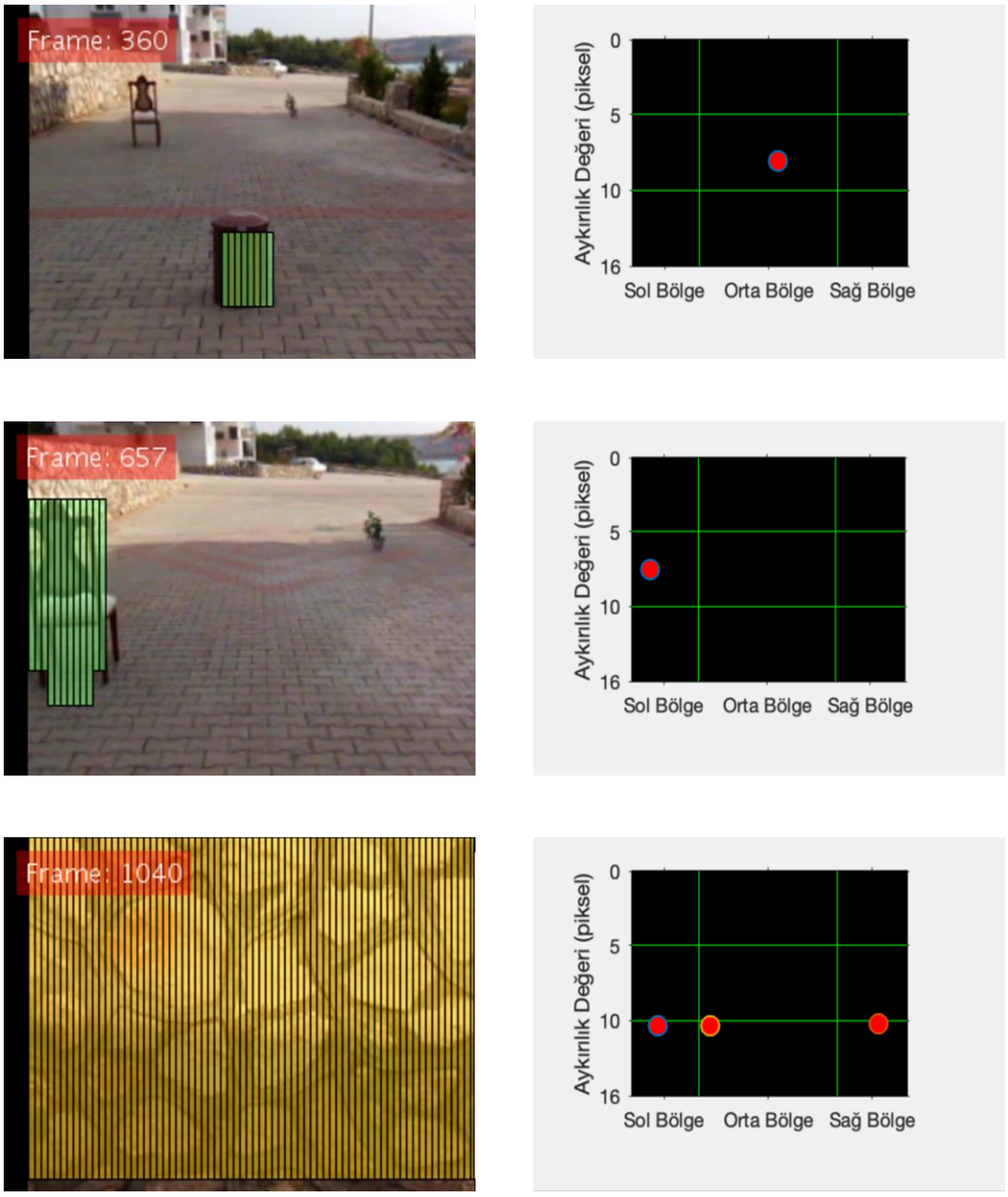
Eşitliklerde kullanılan Z terimi kameraya olan uzaklığı temsil etmektedir. f terimi kameranın odak uzaklığını, d terimi aykırılık değerini, b terimi kameralar arasındaki

mesafeyi ve s_x terimi piksel alanını temsil etmektedir. Verilen bu eşitlikler kullanılarak görüntülerden 3 boyutlu bilgi çıkarımı yapılmaktadır.

3.2.10. Sahnedeki Engel Haritasının Oluşturulması

Dörtgen süper piksel ayrıştırması sonucu her bir görüntüde bulunan engeller pozisyonlarına göre bir ızgara (grid) üzerinde konumlandırılmıştır. Görüntünün birinci kolonu ile toplam kolon sayısının çeyreğine kadar olan bölge ızgaradaki "sol bölge", toplam kolon sayısının çeyreğinden dörtte üçüne kadar olan bölge ızgaradaki "ön bölge" ve toplam kolon sayısının dörtte üçünden sonuncu kolona kadar olan bölge ise "sağ bölge" olarak belirlenmiştir. Oluşturulan sahne haritasının satır pozisyonları ise aykırılık değerlerine göre belirlenmiştir ve üç eşit parçaya ayrılmıştır.

Deneyin uygulandığı veri setindeki bazı örnek görüntüler ve bu örnek görüntüler için oluşturulan sahne haritası temsili Şekil 3.22'de paylaşılmıştır. Engellerin elde edilen sahne haritaları üzerindeki pozisyonlarına göre görme engelli kullanıcıya akustik geri bildirim sağlanmaktadır.



Şekil 3.22. Deneyde kullanılan bazı örnek görüntüler ve bunlar için oluşturulan sahne haritaları

3.2.11. Akustik Geri Bildirim Sağlanması

Görme engelli kullanıcının sahnede bulunan engellerden kaçınabilmesi amacıyla görüntülerde tespit edilen engeller ses yoluyla görme engelli kullanıcıya iletilmektedir. Bu işlevi sağlayabilmek için dörtgen süper pikselleri ses sinyallerine dönüştüren bir şema tasarlanmıştır. Sahnedeki engelleri ses sinyali ile eşleştirmek ve akustik geri bildirim üretmek için belirlenen kural kümesi Tablo 1'de paylaşılmıştır.

Tablo 1. Akustik geri bildirim oluşturma kuralları

Dörtgen Süper Piksel	**Stereo Ses**	
Derinlik (metre)	**Ses Yüksekliği**	
>10	0	
<2	1	
>2 ve <10	Ses yüksekliği mesafe arttıkça doğrusal olarak azalır.	
Kolon Koordinatı (Piksel)	**Frekans (Hz)**	**Ses Kanalı**
$< 0.25 * Toplam\ Kolon\ Sayısı$	340	Sol
$> 0.25 * Toplam\ Kolon\ Sayısı$ ve $< 0.75 * Toplam\ Kolon\ Sayısı$	440	Sol ve Sağ
$> 0.75 * Toplam\ Kolon\ Sayısı$	540	Sağ

Tasarlanan prototipte 10 metre içerisindeki tüm engeller için üretilen dörtgen süper pikseller "bip" sesine dönüştürülmüştür. Bip sesini çıkaracak dalga formunu (waveform) üretmek için Eşitlik (33) kullanılmıştır. Eşitlik (33)'de $y(t)$ terimi dalga formunu, A terimi ses dalgasının genliğini, f terimi ses dalgasının Hz (Hertz) cinsinden frekansını ve t terimi ise saniye cinsinden zaman değerini temsil etmektedir.

$$y(t) = A\sin(2\pi f t) \tag{33}$$

Üretilen bip sesinin yüksekliğine ve frekans farkına görüntüde tespit edilen engelin uzaklığına ve yönüne göre karar verilmiştir. Akustik geri bildirim sesi olarak bip sesini kullanmamızın sebebi arabaların park sensörleri gibi cihazlarda da genellikle bip sesinin tercih ediliyor olması ve bu nedenle insanların bu sese aşina olmasıdır. 10 metrelik mesafe içerisinde çeşitli uzaklıklarda birden fazla engelin olması durumunda ise kullanıcıya en yakın olan engeller için akustik geri bildirim üretilmiştir.

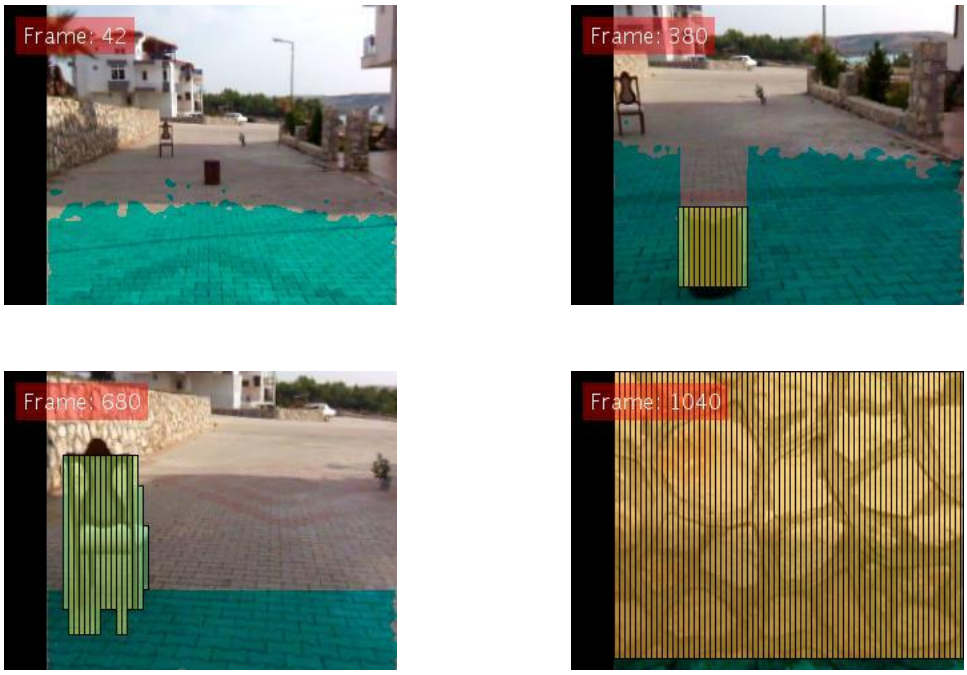
Akustik geri bildirim için stereo ses tercih edilmiştir. Böylelikle sahne haritasının sol bölgesinde bulunan engeller kullanıcının sol kulaklığından, sağ bölgede bulunan engeller kullanıcının sağ kulaklığından ve ön bölgede bulunan engeller ise her iki kulaklıktan verilmektedir. Böylelikle kullanıcı sahnedeki engellerin hangi bölgede olduğunu ayırt edebilmektedir.

BÖLÜM 4. DENEYSEL SONUÇLAR

Bu çalışma kapsamında görme engelli bireylerin mobilize olmasını sağlayacak pasif stereo görme tabanlı duyusal ikame prototipi geliştirilmiştir. Deneysel kurulumumuzda FUJIFILM'in FinePix Real 3D kamera ekipmanı kullanılmıştır. FinePix Real 3D, 6.3 mm odak uzaklığına, yaklaşık 75 mm taban uzunluğuna ve geniş açılı lenslere sahip bir stereo kameradır. Deneylerde kullanılan kamera, satranç tahtası deseni kullanılarak kalibre edilmiş ve kamera parametreleri elde edilmiştir. Veri setindeki videolar kent ortamında, 320x240 çözünürlükte ve 30 fps olarak yakalanmıştır. Prototipimiz 16 GB RAM ve Intel Core i7 2.70 GHz CPU'ya sahip bir bilgisayar üzerinde koşturulmuştur.

Kaydedilen videolar oldukça zorlayıcı sahneler içermektedir. Videolar yürüyerek kaydedildiği için test yapılan görüntülerde titremeler ve açısal bozulmalar mevcuttur. Ayrıca videolarda iki kamera arasında aydınlanma farklılığının bulunduğu sahneler de bulunmaktadır. Veri setindeki görüntülerde statik engeller bulunmaktadır.

UV-Disparity haritalarına çizge kesme işlemi uygulanarak ve dörtgen süper piksel yaklaşımı kullanılarak sahnelerde bulunan zemin profili ve engellerin gerçek zamanlı olarak tespitlerinin yapılması sağlanmıştır. Sahnelerdeki engellerin yüksekliklerinin hesaplanması adımında yeni bir üyelik oylaması yaklaşımı önerilmiştir. Geliştirilen görme engelli duyusal ikame sistemi prototipi umut verici sonuçlar vermektedir. Geliştirilen sistem, tamamen pasif stereo görme tekniklerine dayanması sebebiyle günlük hayatta sıkça kullanılan telefon, bilgisayar ve tablet gibi birçok cihazda da ek bir ekipman kullanımına ihtiyaç duyulmadan çalışabilecektir. Yapılan deneylerde görme engelli bireyin yürüyüş yolu üzerindeki engellerin yüksek başarım ile tespit edilmiştir. Şekil 4.1'de geliştirilen sistemin testler sırasında ürettiği örnek sonuçlar gösterilmektedir.



Şekil 4.1. Gerçekleştirilen deneyden örnek görüntüler

Geliştirdiğimiz prototipin engel tespitindeki başarısını ölçümleyebilmek için öncelikle her bir video karesindeki engeller tarafımızdan etiketlenerek kesin referans değerlerini içeren video kareleri oluşturulmuştur. Sistemin engel tespiti performansını ölçümleyebilmek için "yanlış tespit edilen engel sayısı", "tespit edilmesi gerekirken tespit edilemeyen engel sayısı", "doğru tespit edilen engel sayısı" ve "sahnedeki ilgi alanı içerisinde herhangi bir engelin var olmadığı durumda prototipin alarm üretmeme sayısı" gözlem parametreleri belirlenmiştir. Bu parametrelerin yardımıyla doğruluk (accuracy), duyarlılık (recall), kesinlik (precision) ve F1 skoru metrikleriyle prototip objektif olarak değerlendirilmiştir [148].

Oluşturduğumuz veri seti içerisindeki 1080 görüntü karesi tek tek tarafımızdan incelenerek her bir görüntü karesi içerisindeki doğru olarak tespit edilen engel (gerçek pozitif, true positive), yanlış olarak tespit edilen engel (yanlış pozitif, false positive), tespit edilmesi gerekirken tespit edilemeyen engel (yanlış negatif, false negative) ve sahnede herhangi bir engelin var olmadığı durumda prototipin alarm üretmeme (doğru negatif, true negative) sayıları hesaplanmıştır. Hata matrisi Tablo 2'de paylaşılmıştır.

Tablo 2. Hata Matrisi

		Gerçek	
		Pozitif	**Negatif**
Tahmin	**Pozitif**	TP	FP
	Negatif	FN	TN

Elde edilen hata matrisi ışığında $Doğruluk$ $(accuracy)$, $Duyarlılık$ $(recall)$, $Kesinlik$ $(precision)$ ve $F1$ skoru hesaplamaları sırasıyla Eşitlik (34), Eşitlik (35), Eşitlik (36) ve Eşitlik (37) kullanılarak yapılmıştır. Eşitliklerde gerçek pozitif değerleri TP, hatalı pozitif değerleri FP, gerçek negatif değerleri TN ve hatalı negatif değerler ise FN temsil etmektedir.

$$Doğruluk = \frac{TP + TN}{TP + TN + FP + FN} \tag{34}$$

$$Duyarlılık = \frac{TP}{TP + FN} \tag{35}$$

$$Kesinlik = \frac{TP}{TP + FP} \tag{36}$$

$$F1 = 2 \times \frac{Kesinlik \times Duyarlılık}{Kesinlik + Duyarlılık} \tag{37}$$

Önerdiğimiz prototipin engel tespit performansı Tablo 3'de gösterilmektedir. Doğruluk oranı %86,4, duyarlılık oranı %73,3, kesinlik oranı %97,7 ve F1 skoru %83,6 olarak hesaplanmıştır.

Tablo 3. Geliştirilen prototipin engel tespit performansı

Metrik	Oran
Doğruluk	%86,4
Duyarlılık	%73,3
Kesinlik	%97,7
F1-skor	%83,6

Önerdiğimiz görüntü işleme algoritması, zemin görüntüsünün kaybolduğu ve duvar gibi büyük ve düz engellerin bulunduğu bazı görüntülerde engelleri tespit edememiştir. Bunun sebebi ise zemin profilinin kaybolmasından dolayı dörtgen süper piksel ayrıştırmasın düzgün bir şekilde yapılamamasıdır. Dörtgen süper piksel ayrıştırması aşamasında karşılaşılan hata bu tip sahnelerde ön plan ve arka plan ayrıştırmasını etkilemekte ve hatalı negatif sonuçlara sebebiyet vermektedir.

Görme engelli navigasyon sistemleri veya duyusal ikame cihazlarıyla alakalı araştırmalarda kullanılabilecek, geliştirilen sistemlerin bir bütün olarak değerlendirilip kıyaslanabileceği standartlaşmış bir değerlendirme metodolojisi henüz oluşmamıştır. Ancak Tapu ve diğerleri gerçekleştirdikleri tarama çalışmasında [3] görme engelli bireyler için geliştirilen sistemlerin performansının değerlendirilebilmesi için bir yöntem önermişlerdir. Böylelikle literatürdeki prototiplerin karşılaştırılabilmesinin imkânı olmuştur. Yaptıkları çalışmada görme engelli kullanıcılar, araştırmacılar ve yazılım geliştiricilerden oluşan çeşitli gruplar ve derneklerle kapsamlı danışma ve tartışma görüşmeleri yaparak prototipler için ihtiyaç duyulan gereksinimleri analiz etmişlerdir. Böylelikle görme engelli topluluğu tarafından kabul görecek bir prototipin sunması gereken işlevsel özellikleri belirlemişlerdir. Yaptıkları analiz sonucunda belirlenen işlevsel özellikler Tablo 4'de paylaşılmaktadır.

Tablo 4. Prototiplerin sunması gereken işlevsel özellikler

Kısaltma	Özellik	Açıklama
F1	İşleme Hızı	Sistem gerçek zamanlı olarak çalışmalı ve kullanıcıyı 1,5 metreden daha yakın mesafede bulunan bir engel hakkında anında uyarmalıdır.
F2	Kullanılabilirlik	Sistem hem iç hem de dış mekân sahnelerinde çalışmalıdır.
F3	Sağlamlık	Sistem sahne dinamikleri veya aydınlatma koşullarından etkilenmemelidir.
F4	Kapsama Mesafesi	Kullanıcı ile tespit edilen nesne arasındaki maksimum mesafe.
F5	Engel Tespiti	Sistem konumu, şekli, boyutu veya nesne dinamikleri fark etmeksizin her tür nesneyi algılayabilmelidir.
F6	Taşınabilirlik	Sistem hafif, ergonomik ve kolay giyilebilir olmalıdır.
F7	Kullanım Kolaylığı	Sistemin kullanımı için orta veya uzun vadeli özel bir eğitime ihtiyaç duyulmamalıdır ve öğrenmesi kolay olmalıdır.

Tablo 4'de paylaşılan kriterler, görme engelli kullanıcılara bilinmeyen ortamlarda otonom navigasyon sağlanabilmesi için geliştirilen herhangi bir sistemin içermesi gereken temel işlevsel özellikler olarak kabul edilebilir. Bu nedenle yaptığımız çalışmada geliştirdiğimiz prototipin performans değerlendirmesini yapabilmek ve diğer stereo kamera tabanlı sistemlerle karşılaştırabilmek için Tapu ve diğerlerinin tarama çalışmasında önerdikleri değerlendirme yöntemini [3] çalışmamızda kullandık.

Tapu ve diğerlerinin gerçekleştirdikleri tarama çalışmasında [3] görme engelli bireyler için geliştirilen yardımcı sistemlerin sistematik bir değerlendirilmesi yapılmıştır. Araştırmacılar sistemlerin etkinliğini global olarak değerlendirebilmek için Tablo 4'de paylaşılan her bir işlevsel eleman (F_i) için 0 ile 10 arasında değişen nitel puan vermişlerdir. Araştırmacıların önerdikleri yönteme göre belirli bir özellik tamamen karşılanıyorsa maksimum değer olan 10 atanmıştır. Bununla birlikte dikkate alınan işlevsel özelliği kısmi bir şekilde gerçekleştiren sistemler için ise ara puanlar kullanılmıştır.

İşleme hızı kriteri için kullanıcının normal bir şekilde yürümesini sağlamaya yetecek hızda hesaplama yapabilen sistemlere 10 puan vermişlerdir. Kullanılabilirlik özelliği için yalnızca iç veya dış ortamlarda çalışacak şekilde tasarlanmış bir sisteme 5, her ikisinde de çalışan sistemlere 10 puan vermişlerdir.

Sağlamlık parametresi için doğruluk (accuracy) ve duyarlılık (recall) oranlarının her ikisi de birden %99'dan yüksekse araştırmacılar bu kritere 10 puan atamışlardır. Aksi durumda atadıkları puan yanlış pozitif (false positive) ve yanlış negatiflerin (false negative) sayısı ile ilişkili olarak atanmıştır. Tam puan alabilen sistemlerin dışındakilerden kesinlik (precision) ve duyarlılık (recall) oranları %50'nin altında olan sistemler oldukça hassas olarak kabul edilmiş ve bu nedenle sağlamlık parametresi için "*Düşük"* kategorisinde değerlendirilmiştir. Doğruluk (accuracy) puanları %75'in altında olan sistemler "*Orta*" kategorisinde değerlendirilmiştir. F1-skoru %75'in üzerinde olan sistemler "*Yüksek*" kategorisinde değerlendirilmiştir.

Kapsama mesafesi kriteri için ise görme engelli kullanıcıyı 5 metre ve daha yüksek bir mesafede bulunan her türlü statik veya dinamik engel hakkında bilgilendiren

sistemlere 10 puan verilmiştir. Daha düşük kapsama mesafesine sahip sistemler için ise araştırmacılar oransal olarak puanı azaltmışlardır.

Araştırmacılar değerlendirme kriterleri arasındaki engel tespit işlevi için ise herhangi bir ön bilgi olmadan engel tespiti yapabilen sistemlere maksimum puan vermişlerdir. Taşınabilirlik işlevi için ise giyilebilir ve hafif sistemler için 10 puan vermişlerdir.

Son olarak, kullanım kolaylığı kriteri için herhangi bir ek eğitim olmaksızın doğrudan kullanılabilen prototipler maksimum puan almışlardır. Bu kategoride 4 ve altında puan alan sistemeler özel bir personel ile kapsamlı eğitim gerektirdiği için "*Düşük*" olarak etiketlenmiştir. Herhangi bir eğitim gerektirmeden doğrudan görme engelli kullanıcılar tarafından kullanılabilen sistemler "*Yüksek*" olarak etiketlenmiştir ve bu prototiplere araştırmacılar tarafından 8 ve üzeri puan verilmiştir.

Tablo 4'de verilen işlevlere göre kendi geliştirdiğimiz prototip de dahil olmak üzere literatürdeki en güncel diğer stereo kamera tabanlı prototiplerin karşılaştırmalı genel görünümü Tablo 5'de sunulmaktadır.

Tablo 5. Stereo kamera tabanlı sistemlerin karşılaştırması

Sistem Adı	**F1**	**F2** *(İç mekân / Dış mekân)*	**F3**	**F4**	**F5** *(Şekil - Boyut - Statik/Dinamik - Pozisyon)*	**F6**	**F7**
Robot Vision	10 fps	İç ve Dış Mekân	Orta	3 m	Hepsi - Hepsi - Statik - Hepsi	Hayır	Yok
SmartVision	10 fps	İç ve Dış Mekân	Düşük	2 m	Geometrik kısıt - Küçük - Statik - Yer seviyesi	Hayır	Yok
Lin v.d.	0.5 fps	İç ve Dış Mekân	Düşük	3 m	Hepsi - Hepsi - Hepsi - Hepsi	Evet	Düşük
Smart Walker	4 fps	İç ve Dış Mekân	Orta	2 m	Spesifik - Hepsi - Statik - Hepsi	Hayır	Yok
Leung v.d.	1 fps	Dış Mekân	Orta	3,5 m	Yok	Hayır	Yok
Schwarze v.d.	15 fps	Dış Mekân	Yüksek	10 m	Hepsi - Hepsi - Hepsi - Hepsi	Evet	Orta
Everding v.d.	20 fps	İç Mekân	Orta	6 m	Hepsi - Hepsi - Hepsi - Hepsi	Evet	Orta
Bizim prototipimiz	***18 fps***	***Dış Mekân***	***Yüksek***	***10 m***	***Hepsi - Hepsi - Statik - Yer seviyesi***	***Evet***	***Yüksek***

Literatürde öne çıkan duyusal ikame cihazlarına global bir skor belirleyebilmek amacıyla araştırmacılar 42 tane görme engelli kullanıcıdan oluşan bir gruba Tablo 4'de paylaşılan 7 tane özelliği azalan önem sırasına göre sıralamaları istemişlerdir. Ardından araştırmacılar puanları toplayıp görme engelli kullanıcı sayısını ve toplam özellik sayısını kullanarak ağırlık değerlerini normalleştirmişlerdir. Böylelikle her bir işlev için global skor hesabında kullanılacak ağırlık değerlerini belirlemişlerdir. İşlevler için hesaplanan ağırlık değerleri Tablo 6'da verilmiştir.

Tablo 6. Özellikler için belirlenen ağırlık değerleri

Kısaltma	**Özellik**	**Belirlenen Ağırlık**
F1	İşleme Hızı	0.18
F2	Kullanılabilirlik	0.12
F3	Sağlamlık	0.21
F4	Kapsama Mesafesi	0.16
F5	Engel Tespiti	0.14
F6	Taşınabilirlik	0.10
F7	Kullanım Kolaylığı	0.09

Tablo 6'daki özellikler ve ağırlık değerleri için Tablo 5'deki kriterlere göre atanan puanlar kullanılarak her bir prototipin global skor hesabının yapılması için araştırmacılar tarafından Eşitlik (38) önerilmiştir [3]. Eşitlikte F_i özellik i için atanan puanı, N değerlendirmede kullanılan toplam özellik sayısını ve w_i ise her özellik için atanan ağırlık değerini temsil etmektedir.

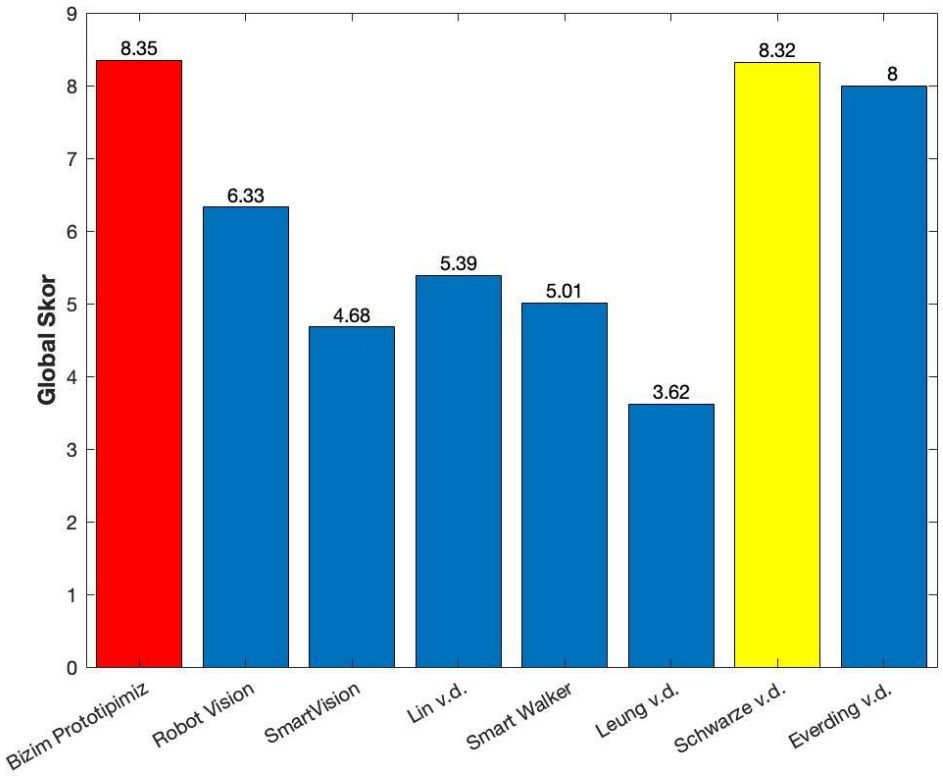
$$Global\ Skor = \frac{\sum_{i=1}^{N} w_i \times F_i}{N} \tag{38}$$

Tablo 7'de stereo kamera tabanlı sistemler için Tablo 5'deki kriterlere göre atanan puanlar ve global skorlar paylaşılmıştır. Tablo 7'de karşılaştırılan literatürdeki diğer stereo kamera tabanlı çalışmaların puanları ve global skoru için Tapu ve diğerlerinin çalışmasında [3] verilen puanlar direkt olarak alınmıştır. Önerdiğimiz prototipin puanları da aynı çalışmada belirlenen kriterlere göre verilmiştir ve bu puanlara göre global skor hesabı yapılmıştır. Önerdiğimiz prototip, bu sistematik değerlendirme yöntemine göre karşılaştırdığımız diğer stereo kamera tabanlı sistemlerden daha yüksek global skor elde etmiştir.

Tablo 7. Stereo kamera tabanlı sistemlerin performans değerlendirmesi

Sistem Adı	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**	**F7**	**Global Skor**
Robot Vision	8	10	6	7	8	1	1	6.33
SmartVision	8	10	3	5	3	1	1	4.68
Lin v.d.	1	10	3	7	10	5	4	5.39
Smart Walker	4	10	6	5	6	1	1	5.01
Leung v.d.	2	5	5	8	1	1	1	3.62
Schwarze v.d.	10	5	8	10	10	7	6	8.32
Everding v.d.	10	5	6	10	10	8	6	8
Bizim prototipimiz	***10***	***5***	***7***	***10***	***7***	***10***	***10***	***8.35***

Çalışmamız kapsamında performans değerlendirmesinde dikkate alınan stereo kamera tabanlı sistemlerin ve önerdiğimiz prototipin global skorlarının karşılaştırmasını gösteren sütun grafiği Şekil 4.2'de sunulmuştur. Karşılaştırmada 8,35 puan ile en yüksek global skoru (kırmızı sütun) önerdiğimiz prototip elde etmiştir. İkinci en yüksek global skoru ise (sarı sütun) 8,32 puan ile Schwarze ve diğerlerinin çalışması elde etmiştir.



Şekil 4.2. Performans değerlendirmesinde elde edilen global skorlar

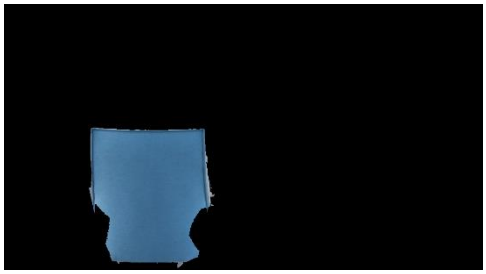
Şekil 4.3'de deneyler sırasında kullanılan örnek bir sahnede tespit edilen engellerin kameraya olan uzaklıklarının üçgenleme yöntemiyle tahmin edilmesinin sonucu gösterilmektedir. Şekil 4.4'de görme engelli kullanıcının yürüyüş patikasındaki engellerinin tespitinin ve zemin profilinin ayrıştırılmasının sonuçları gösterilmektedir.



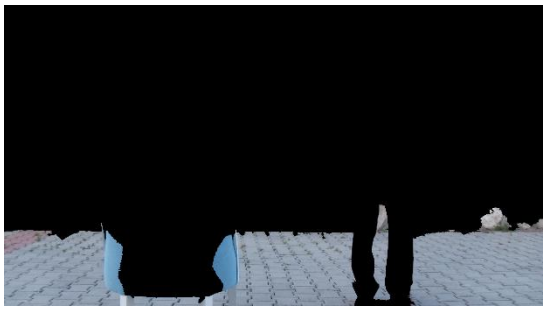
Şekil 4.3. Tespit edilen engellerin kameraya olan uzaklıkları



(a) Patika üzerindeki engel



(b) Patika üzerindeki engel



(c) Zemin profili

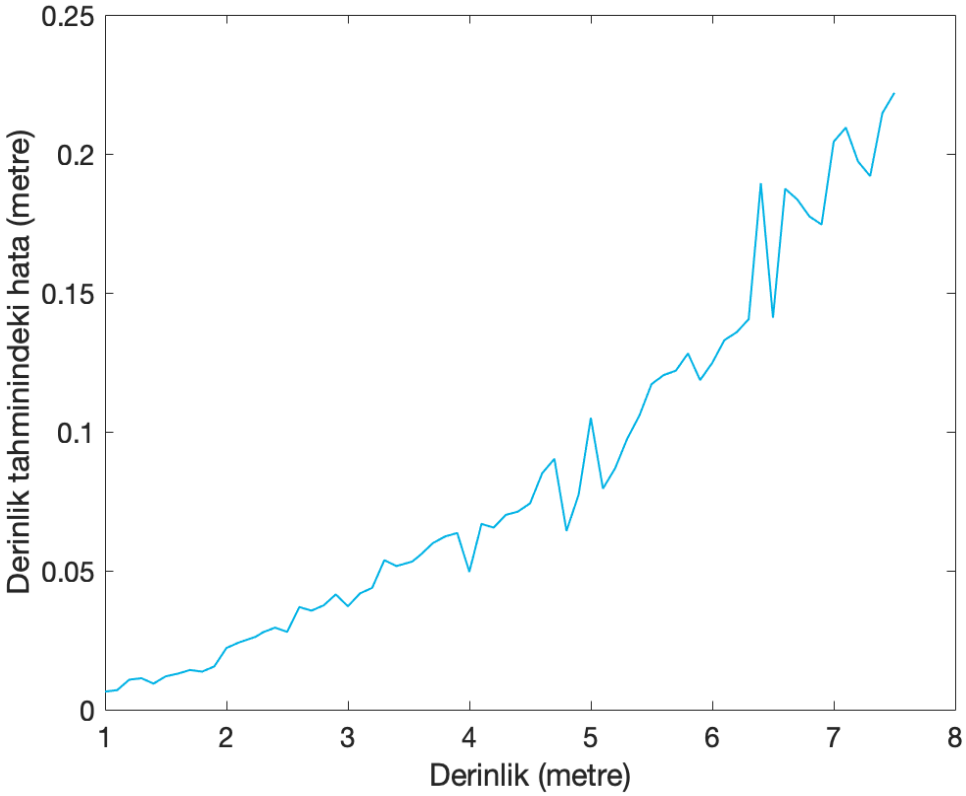
Şekil 4.4. Yürüyüş patikası üzerinde bulunan belirli uzaklıktaki engellerin görüntüden çıkarımı ve tespit edilen zemin profili.

Tablo 8'de görüntülerde tespiti yapılan engeller için hesaplanan uzaklık değerleri ile gerçek uzaklık değerlerinin karşılaştırılması verilmiştir. Sonuçlar incelendiğinde başarı seviyesinin görme engelli bir bireyin güvenli bir şekilde gezintisini sağlamak için yeterli düzeyde olduğu görülmektedir. Görüntülerdeki engeller kameradan uzaklaştıkça hesaplanan uzaklıklarındaki hata oranının arttığı gözlemlenmiştir.

Tablo 8. Tespit edilen engeller için uzaklık tahminleri ve hata oranları

	Gerçek Uzaklık (cm)	**Hesaplanan Uzaklık (cm)**	**Hata Oranı**
Koltuk	224	221.4	%1,17
İnsan	353	347.7	%1,51
Duvar	750	727.8	%2,96

Deneyler sırasında hesaplanan uzaklık değerleri ile gerçek uzaklık değerleri 10'ar santimetrelik mesafe ile 7,5 metre boyunca karşılaştırılmıştır. Şekil 4.5'de gerçek uzaklık ile tahmin edilen uzaklıktaki hata miktarının karşılaştırmasının grafiği paylaşılmaktadır. 1 ile 7,5 metre arasındaki bir mesafede tahmin edilen uzaklıklardaki hata miktarı 0,65 ile 22,2 santimetre arasında değiştiği gözlenmiştir.



Şekil 4.5. Derinlik tahmini hatası grafiği

Tablo 9'da geliştirilen sistemin ana aşamalarının ortalama işlem zamanları gösterilmektedir. Sistemin, saniyede yaklaşık 18 görüntüyü işleyebildiği gözlemlenmiştir. Bu açıdan bakıldığında görme engelli bir bireyin mobilizasyonu için gerekli olan gerçek zamanlı akış sağlanabilmektedir.

Tablo 9. Geliştirilen sistemdeki ana aşamaların çalışma süreleri

Aşama	**İşlem Süresi (saniye)**
Görüntülerdeki Gürültünün Temizlenmesi	0,0046
Görüntü Rektifikasyonu	0,0016
Aykırılık Haritası Tahmini	0,0091
Zemin Profilinin Belirlenmesi	0,0021
Engellerin Taban Noktalarının Tespiti	0,0033
Engellerin Yüksekliklerinin Tespiti	0,0271
Dörtgen Süper Piksellerin Oluşturulması	0,0008
Derinlik Çıkarımı	0,0050
Toplam	0,0536 (~ 18 fps)

BÖLÜM 5. SONUÇ

Bu çalışma kapsamında stereo kamera ve kulaklığın kullanıldığı giyilebilir bir görme engelli navigasyon sistemi prototipi geliştirilmiştir. Önerdiğimiz prototipin temel amacı cep telefonu ve kulaklık gibi insanların gündelik hayatında sıklıkla kullandığı standart ekipmanlarla kullanıcının mobilizasyonunun sağlanmasıdır. Prototipin engel tespit performansının nicel değerlendirmesi sonucunda doğruluk oranı %86,4, duyarlılık oranı %73,3, kesinlik oranı %97,7 ve F1 skoru ise %83,6 olarak hesaplanmıştır. Engel tespiti için önerilen görüntü işleme algoritması saniyede ortalama 18 görüntü karesi işleyebilmektedir. Prototipin nitel değerlendirmesi ise Tapu ve diğerlerinin [3] önerdikleri yöntem ile yapılmıştır. Bu değerlendirme sonucunda 8,35 global skor puanı ile literatürde öne çıkan stereo kamera tabanlı çalışmalardan daha yüksek puan elde etmiştir. Elde edilen sonuçlar, büyük çoğunluğu düşük gelir grubuna mensup görme engelli bireyler için geliştirilecek düşük maliyetli, kolay taşınabilir ve ekstra bir eğitime ihtiyaç olmadan kullanılabilecek bir görme engelli duyusal ikame sistemi için umut vericidir.

Geliştirilen prototip kapsamında önerdiğimiz görüntü işleme algoritması otonom sürüş teknolojilerinde sıklıkla kullanılan dörtgen süper piksel (stixel) ayrıştırması ve uv-disparity haritası analizine dayanmaktadır. Çalışmamız dörtgen süper piksel yaklaşımının görme engelli bireylerin duyusal ikamesi için kullanıldığı ender çalışmalardan biridir. Çalışmamızda Badino, Franke ve Pfeiffer tarafından önerilen dörtgen süper piksel tekniğinin [2] görme engelli bireyin mobilizasyonu için uygun olmayan kısımları çıkarılmış ve görme engelli bireylerin ihtiyaçları göz önünde bulundurularak bazı iyileştirmeler yapılmıştır. Algoritmanın gerçek zamanlı çalışabilmesi için Badino ve diğerlerinin [2] çalışmasındaki yükseklik tespiti yönteminden farklı olarak arka plan ayrıştırması için yeni bir maliyet fonksiyonu önerilmiştir. Böylelikle orijinal yöntemdeki 3 boyutlu yapılandırma adımına gerek olmaksızın kullanıcıların gerçek zamanlı mobilizasyonuna olanak verecek zemin profili ve engel tespiti yapılmıştır.

Önerdiğimiz görüntü işleme algoritmasının engel tespit ve sınıflandırma başarısının objektif olarak ölçümlenebilmesi için algoritmanın ürettiği sonuçlar doğruluk, duyarlılık, kesinlik ve F1-skoru metrikleri ile değerlendirilmiştir. Önerdiğimiz görüntü işleme algoritmasının sonuçları kameranın hareket etmesine, sahnelerdeki ışıklandırma

koşullarının farklılaşmasına ve engellerin şekillerindeki değişimlere rağmen sağlamlık ve tutarlılık göstermiştir.

Geliştirdiğimiz prototip pasif stereo görme tekniklerine dayanmaktadır ve tamamen yazılım tabanlıdır. Yazılım sistemi modüler ve entegre edilebilir bir yapıdadır. Bu nedenle sistemin görüntü işleme modülü herhangi bir görme engelli navigasyon sistemi ile kolaylıkla entegre edilebilir.

Sistemin görme engelli bireyler açısından sistemsel başarısının ölçümlenebilmesi için ise Tapu ve diğerlerinin görme engelli navigasyon sistemleri ile ilgili yaptıkları tarama çalışmasında önerdikleri nitel değerlendirme yöntemi [3] kullanılmıştır. Yapılan değerlendirmeler sonucunda önerdiğimiz prototipin literatürde öne çıkan diğer pasif stereo görme tabanlı sistemlerle karşılaştırıldığında daha yüksek başarı seviyesine ulaştığı gözlemlenmiştir.

Her ne kadar prototipin global skoru benzer çalışmalardan yüksek olsa da hala yapılabilecek birtakım iyileştirmeler bulunmaktadır. Sistem sadece dış mekân görüntüleri üzerinde test edilmiştir. Bu nedenle yapılan sistemsel değerlendirmede kullanılabilirlik açısından literatürdeki benzer sistemlere göre daha az puan almıştır.

Sistemsel değerlendirmede kullanılan ve prototipimizin tam puandan az değerleme aldığı bir diğer parametre ise şekil, boyut, statik/dinamik ve pozisyon ölçütleri açısından tespit edilen engel tipleridir. Geliştirdiğimiz prototipte sadece statik ve yer seviyesindeki engellerin tespiti yapılmaktadır. Ayrıca çukur gibi negatif engellerin tespiti, kamera açısındaki bozulmaların tahmini ve dörtgen süper piksel ayrıştırmasının başarısının arttırılması geliştirdiğimiz sistemin ve önerdiğimiz görüntü işleme algoritmasının sağlamlığını arttıracak unsurlardır. Gelecek çalışmalarda sistemin kullanılabilirlik, engel tespiti ve sağlamlık performansını yükseltecek bu konular hakkında iyileştirmeler yapılması planlanmaktadır.

Sonuç olarak, otonom sürüş teknolojisi araştırmaları tüm hızıyla devam etmesine rağmen görme engelli bireylerin mobilizasyonuna yardım sağlamayı amaçlayan araştırmalar daha nadir olarak gerçekleştirilmektedir. Yaptığımız çalışma otonom sürüş için kullanılan yaklaşımların çeşitli modifikasyon ve iyileştirmelerle görme engelli

bireylerin duyusal ikamesi için de kullanılabileceğini göstermiştir. Geliştirdiğimiz prototip, görme engelli bireyler için düşük maliyetli ve mobil telefon ve kulaklık gibi insanların sıklıkla sahip olduğu ekipmanlarla birlikte çalışabilecek yazılım tabanlı bir duyusal ikame çözümünün geliştirilebileceğine dair umut verici sonuçlar vermiştir.

KAYNAKÇA

[1] R. R. A. Bourne *et al.*, "Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis," *Lancet Glob Health*, vol. 5, no. 9, pp. e888-e897, Sep 2017, doi: 10.1016/S2214-109X(17)30293-0.

[2] H. Badino, U. Franke, and D. Pfeiffer, "The Stixel World - A Compact Medium Level Representation of the 3D-World," in *Pattern Recognition*, Berlin, Heidelberg, J. Denzler, G. Notni, and H. Süße, Eds., 2009// 2009: Springer Berlin Heidelberg, pp. 51-60.

[3] R. Tapu, B. Mocanu, and T. Zaharia, "Wearable assistive devices for visually impaired: A state of the art survey," *Pattern Recognition Letters*, vol. 137, pp. 37-52, 2020/09/01/ 2020, doi: https://doi.org/10.1016/j.patrec.2018.10.031.

[4] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003, p. 700.

[5] Y. Li, S. Velipasalar, and M. C. Gursoy, "An improved evolutionary algorithm for fundamental matrix estimation," in *2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance*, 27-30 Aug. 2013 2013, pp. 226-231, doi: 10.1109/AVSS.2013.6636644.

[6] R. I. Hartley, "In defense of the eight-point algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 6, pp. 580-593, 1997, doi: 10.1109/34.601246.

[7] R. I. Hartley, "Theory and Practice of Projective Rectification," *International Journal of Computer Vision*, journal article vol. 35, no. 2, pp. 115-127, November 01 1999, doi: 10.1023/a:1008115206617.

[8] M. J. Hannah, "Computer matching of areas in stereo images," Stanford University, 1974.

[9] Y. C. Hsieh, D. M. McKeown, and F. P. Perlant, "Performance evaluation of scene registration and stereo matching for cartographic feature extraction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 214-238, 1992, doi: 10.1109/34.121790.

[10] D. Marr and T. Poggio, "Cooperative computation of stereo disparity," *Science*, vol. 194, no. 4262, pp. 283-287, 1976, doi: 10.1126/science.968482.

[11] S. T. Barnard and M. A. Fischler, "Computational Stereo," *ACM Comput. Surv.*, vol. 14, no. 4, pp. 553-572, 1982, doi: 10.1145/356893.356896.

[12] U. R. Dhond and J. K. Aggarwal, "Structure from stereo-a review," *IEEE Transactions on Systems, Man, and Cybernetics,* vol. 19, no. 6, pp. 1489-1510, 1989, doi: 10.1109/21.44067.

[13] D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms," *International Journal of Computer Vision,* journal article vol. 47, no. 1, pp. 7-42, April 01 2002, doi: 10.1023/a:1014573219977.

[14] M. Z. Brown, D. Burschka, and G. D. Hager, "Advances in computational stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 25, no. 8, pp. 993-1008, 2003, doi: 10.1109/TPAMI.2003.1217603.

[15] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, 17-22 June 2006 2006, vol. 1, pp. 519-528, doi: 10.1109/CVPR.2006.19.

[16] D. Marr, *Vision : a computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman, 1982, pp. xvii, 397 p.

[17] H. P. Moravec, "The Stanford Cart and the CMU Rover," *Proceedings of the IEEE,* vol. 71, no. 7, pp. 872-884, 1983, doi: 10.1109/PROC.1983.12684.

[18] K. Konolige, "Small Vision Systems: Hardware and Implementation," London, 1998: Springer London, in Robotics Research, pp. 203-212.

[19] S. Thrun *et al.*, "Stanley: The robot that won the DARPA Grand Challenge," *Journal of Field Robotics,* vol. 23, no. 9, pp. 661-692, 2006, doi: 10.1002/rob.20147.

[20] D. Marr and T. Poggio, "A Theory of Human Stereo Vision," Massachusetts Institute of Technology, 1977.

[21] E. L. W. Grimson, *From Images to Surfaces: A Computational Study of the Human Early Visual System*. MIT Press, 1981, p. 274.

[22] H. H. Baker and T. O. Binford, "Depth from Edge and Intensity Based Stereo," in *IJCAI*, 1981.

[23] S. T. Barnard and W. B. Thompson, "Disparity Analysis of Images," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 2, no. 4, pp. 333-340, 1980, doi: 10.1109/tpami.1980.4767032.

[24] D. Marr, T. Poggio, and S. Brenner, "A computational theory of human stereo vision," *Proceedings of the Royal Society of London. Series B. Biological Sciences,* vol. 204, no. 1156, pp. 301-328, 1979, doi: doi:10.1098/rspb.1979.0029.

[25] J. E. W. Mayhew and J. P. Frisby, "The Computation of Binocular Edges," *Perception,* vol. 9, no. 1, pp. 69-86, 1980, doi: 10.1068/p090069.

[26] R. D. Arnold, "Automated stereo perception," Stanford University, 1983.

[27] W. E. L. Grimson, "Computational Experiments with a Feature Based Stereo Algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. PAMI-7, no. 1, pp. 17-34, 1985, doi: 10.1109/TPAMI.1985.4767615.

[28] Y. Ohta and T. Kanade, "Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. PAMI-7, no. 2, pp. 139-154, 1985, doi: 10.1109/TPAMI.1985.4767639.

[29] R. C. Bolles, H. H. Baker, and D. H. Marimont, "Epipolar-plane image analysis: An approach to determining structure from motion," *International Journal of Computer Vision,* journal article vol. 1, no. 1, pp. 7-55, March 01 1987, doi: 10.1007/bf00128525.

[30] L. Matthies, T. Kanade, and R. Szeliski, "Kalman filter-based algorithms for estimating depth from image sequences," *International Journal of Computer Vision,* journal article vol. 3, no. 3, pp. 209-238, September 01 1989, doi: 10.1007/bf00133032.

[31] R. C. Bolles, H. H. Baker, and M. J. Hannah, "The JISCT stereo evaluation," 1993.

[32] R. T. Collins, "A space-sweep approach to true multi-image matching," in *Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* 18-20 June 1996 1996, pp. 358-363, doi: 10.1109/CVPR.1996.517097.

[33] Z. Zhang and Y. Shan, "A Progressive Scheme for Stereo Matching," Berlin, Heidelberg, 2001: Springer Berlin Heidelberg, in 3D Structure from Images — SMILE 2000, pp. 68-85.

[34] M. Lhuillier and L. Quan, "Match Propogation for Image-Based Modeling and Rendering," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 24, no. 8, pp. 1140-1146, 2002, doi: 10.1109/tpami.2002.1023810.

[35] M. Lhuillier and L. Quan, "A quasi-dense approach to surface reconstruction from uncalibrated images," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 27, no. 3, pp. 418-433, 2005, doi: 10.1109/TPAMI.2005.44.

[36] C. Strecha, T. Tuytelaars, and L. V. Gool, "Dense Matching of Multiple Wide-baseline Views," presented at the Proceedings of the Ninth IEEE International Conference on Computer Vision - Volume 2, 2003.

[37] M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. M. Seitz, "Multi-View Stereo for Community Photo Collections," in *2007 IEEE 11th International*

Conference on Computer Vision, 14-21 Oct. 2007 2007, pp. 1-8, doi: 10.1109/ICCV.2007.4408933.

[38] P. Giblin and R. Weiss, "Reconstruction of Surfaces from Profiles," University of Massachusetts, 1987.

[39] R. Cipolla and A. Blake, "Surface shape from the deformation of apparent contours," *International Journal of Computer Vision*, journal article vol. 9, no. 2, pp. 83-112, November 01 1992, doi: 10.1007/bf00129682.

[40] R. Vaillant and O. D. Faugeras, "Using Extremal Boundaries for 3-D Object Modeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 157-173, 1992, doi: 10.1109/34.121787.

[41] J. Y. Zheng, "Acquiring 3-D Models from Sequences of Contours," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 2, pp. 163-178, 1994, doi: 10.1109/34.273734.

[42] R. Szeliski and R. Weiss, "Robust Shape Recovery from Occluding Contours Using a Linear Smoother," *International Journal of Computer Vision*, journal article vol. 28, no. 1, pp. 27-44, June 01 1998, doi: 10.1023/a:1008050630628.

[43] H. Rostam Affendi, R. Rosman Abd, and N. Zarina Mohd, "Sum of Absolute Differences algorithm in stereo correspondence problem for stereo matching in computer vision application," in *2010 3rd International Conference on Computer Science and Information Technology*, 9-11 July 2010 2010, vol. 1, pp. 652-657, doi: 10.1109/ICCSIT.2010.5565062.

[44] J. P. Lewis, "Fast Normalized Cross-Correlation," 1995.

[45] M. Okutomi and T. Kanade, "A multiple-baseline stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 4, pp. 353-363, 1993, doi: 10.1109/34.206955.

[46] H. Hirschmuller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 20-25 June 2005 2005, vol. 2, pp. 807-814 vol. 2, doi: 10.1109/CVPR.2005.56.

[47] S. T. Barnard, "A stochastic approach to stereo vision," presented at the Proceedings of the Fifth AAAI National Conference on Artificial Intelligence, Philadelphia, Pennsylvania, 1986.

[48] E. Horowitz, S. Sahni, and S. Rajasckaran, *Computer Algorithms: C++*. W. H. Freeman Co., 1996, p. 769.

[49] P. N. Belhumeur and D. Mumford, "A Bayesian treatment of the stereo correspondence problem using half-occluded regions," in *Proceedings 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 15-18 June 1992 1992, pp. 506-512, doi: 10.1109/CVPR.1992.223143.

[50] I. J. Cox, S. L. Hingorani, S. B. Rao, and B. M. Maggs, "A Maximum Likelihood Stereo Algorithm," *Comput. Vis. Image Underst.*, vol. 63, no. 3, pp. 542-567, 1996, doi: 10.1006/cviu.1996.0040.

[51] S. Birchfield and C. Tomasi, "Depth discontinuities by pixel-to-pixel stereo," in *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, 7-7 Jan. 1998 1998, pp. 1073-1080, doi: 10.1109/ICCV.1998.710850.

[52] G. Van Meerbergen, M. Vergauwen, M. Pollefeys, and L. Van Gool, "A Hierarchical Symmetric Stereo Algorithm Using Dynamic Programming," *International Journal of Computer Vision*, journal article vol. 47, no. 1, pp. 275-285, April 01 2002, doi: 10.1023/a:1014562312225.

[53] A. Gonçalves, A. Godinho, and J. Sequeira, "Low cost sensing for autonomous car driving in highways," in *ICINCO-RA*, 2007.

[54] T. Pollard and J. L. Mundy, "Change Detection in a 3-d World," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, 17-22 June 2007 2007, pp. 1-6, doi: 10.1109/CVPR.2007.383073.

[55] K. Sebesta and J. Baillieul, "Animal-inspired agile flight using optical flow sensing," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, 10-13 Dec. 2012 2012, pp. 3727-3734, doi: 10.1109/CDC.2012.6426163.

[56] Y. Lin, C. Lin, W. Liu, and L. Chen, "A vision-based obstacle detection system for parking assistance," in *2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA)*, 19-21 June 2013 2013, pp. 1627-1630, doi: 10.1109/ICIEA.2013.6566629.

[57] R. Labayrade, D. Aubert, and J. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through "v-disparity" representation," in *Intelligent Vehicle Symposium, 2002. IEEE*, 17-21 June 2002 2002, vol. 2, pp. 646-651 vol.2, doi: 10.1109/IVS.2002.1188024.

[58] Z. Hu and K. Uchimura, "U-V-disparity: an efficient algorithm for stereovision based scene analysis," in *IEEE Proceedings. Intelligent Vehicles Symposium, 2005.*, 6-8 June 2005 2005, pp. 48-54, doi: 10.1109/IVS.2005.1505076.

[59] A. Broggi, C. Caraffi, R. I. Fedriga, and P. Grisleri, "Obstacle Detection with Stereo Vision for Off-Road Vehicle Navigation," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, 21-23 Sept. 2005 2005, pp. 65-65, doi: 10.1109/CVPR.2005.503.

[60] N. Soquet, M. Perrollaz, R. Labayrade, and D. Aubert, "Free Space Estimation for Autonomous Navigation," 03/21 2007.

[61] M. Perrollaz, J. Yoder, A. Spalanzani, and C. Laugier, "Using the disparity space to compute occupancy grids from stereo-vision," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 18-22 Oct. 2010 2010, pp. 2721-2726, doi: 10.1109/IROS.2010.5649690.

[62] M. Perrollaz, J. Yoder, A. Nègre, A. Spalanzani, and C. Laugier, "A Visibility-Based Approach for Occupancy Grid Computation in Disparity Space," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 3, pp. 1383-1393, 2012, doi: 10.1109/TITS.2012.2188393.

[63] N. Fakhfakh, D. Gruyer, and D. Aubert, *Weighted V-disparity Approach for Obstacles Localization in Highway Environments*. 2013, pp. 1271-1278.

[64] A. Iloie, I. Giosan, and S. Nedevschi, "UV disparity based obstacle detection and pedestrian classification in urban traffic scenarios," in *2014 IEEE 10th International Conference on Intelligent Computer Communication and Processing (ICCP)*, 4-6 Sept. 2014 2014, pp. 119-125, doi: 10.1109/ICCP.2014.6936963.

[65] A. Elfes, "Sonar-based real-world mapping and navigation," *IEEE Journal on Robotics and Automation*, vol. 3, no. 3, pp. 249-265, 1987, doi: 10.1109/JRA.1987.1087096.

[66] M. Yguel, O. Aycard, and C. Laugier, *Efficient GPU-based Construction of Occupancy Grids Using several Laser Range-finders*. 2006, pp. 105-110.

[67] L. Matthies and A. Elfes, "Integration of sonar and stereo range data using a grid-based representation," in *Proceedings. 1988 IEEE International Conference on Robotics and Automation*, 24-29 April 1988 1988, pp. 727-733 vol.2, doi: 10.1109/ROBOT.1988.12145.

[68] M. Brandão, R. Ferreira, K. Hashimoto, J. Santos-Victor, and A. Takanishi, *Integrating the whole cost-curve of stereo into occupancy grids*. 2013, pp. 4681-4686.

[69] M. Perrollaz, A. Spalanzani, and D. Aubert, "Probabilistic representation of the uncertainty of stereo-vision and application to obstacle detection," in *2010 IEEE Intelligent Vehicles Symposium*, 21-24 June 2010 2010, pp. 313-318, doi: 10.1109/IVS.2010.5548010.

[70] H. Badino, U. Franke, and R. Mester, "Free Space Computation Using Stochastic Occupancy Grids and Dynamic Programming," 03/19 2012.

[71] Z. Zhang, "A stereovision system for a planetary rover: calibration, correlation, registration, and fusion," *Machine Vision and Applications*, vol. 10, no. 1, pp. 27-34, 1997/05/01 1997, doi: 10.1007/s001380050056.

[72] F. Oniga and S. Nedevschi, "Processing Dense Stereo Data Using Elevation Maps: Road Surface, Traffic Isle, and Obstacle Detection," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 3, pp. 1172-1182, 2010, doi: 10.1109/TVT.2009.2039718.

[73] M. Vergauwen, M. Pollefeys, and L. Van Gool, "A stereo-vision system for support of planetary surface exploration," *Machine Vision and Applications*, vol. 14, no. 1, pp. 5-14, 2003/04/01 2003, doi: 10.1007/s00138-002-0097-7.

[74] J. Rebut, G. Toulminet, and A. Bensrhair, "Road obstacles detection using a self-adaptive stereo vision sensor: a contribution to the ARCOS French project," in *IEEE Intelligent Vehicles Symposium, 2004*, 14-17 June 2004 2004, pp. 738-743, doi: 10.1109/IVS.2004.1336476.

[75] S. Kubota, T. Nakano, and Y. Okamoto, "A Global Optimization Algorithm for Real-Time On-Board Stereo Obstacle Detection Systems," in *2007 IEEE Intelligent Vehicles Symposium*, 13-15 June 2007 2007, pp. 7-12, doi: 10.1109/IVS.2007.4290083.

[76] R. Benenson, R. Timofte, and L. V. Gool, "Stixels estimation without depth map computation," in *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, 6-13 Nov. 2011 2011, pp. 2010-2017, doi: 10.1109/ICCVW.2011.6130495.

[77] R. Benenson, M. Mathias, R. Timofte, and L. Van Gool, "Fast Stixel Computation for Fast Pedestrian Detection," Berlin, Heidelberg, 2012: Springer Berlin Heidelberg, in Computer Vision – ECCV 2012. Workshops and Demonstrations, pp. 11-20.

[78] D. A. P. Carrillo and A. Sutherland, "Fast Obstacle Detection Using Sparse Edge-Based Disparity Maps," in *2016 Fourth International Conference on 3D Vision (3DV)*, 25-28 Oct. 2016 2016, pp. 66-72, doi: 10.1109/3DV.2016.80.

[79] T. Williamson and C. Thorpe, "A specialized multibaseline stereo technique for obstacle detection," in *Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No.98CB36231)*, 25-25 June 1998 1998, pp. 238-244, doi: 10.1109/CVPR.1998.698615.

[80] R. Shade and P. Newman, "Discovering and mapping complete surfaces with stereo," in *2010 IEEE International Conference on Robotics and Automation*, 3-7 May 2010 2010, pp. 3910-3915, doi: 10.1109/ROBOT.2010.5509337.

[81] M. Bleyer, C. Rhemann, and C. Rother, "Extracting 3D Scene-Consistent Object Proposals and Depth from Stereo Images," Berlin, Heidelberg, 2012: Springer Berlin Heidelberg, in Computer Vision – ECCV 2012, pp. 467-481.

[82] S. Kramm and A. Bensrhair, "Obstacle detection using sparse stereovision and clustering techniques," in *2012 IEEE Intelligent Vehicles Symposium*, 3-7 June 2012 2012, pp. 760-765, doi: 10.1109/IVS.2012.6232283.

[83] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," presented at the Proceedings of the 7th international joint conference on Artificial intelligence - Volume 2, Vancouver, BC, Canada, 1981.

[84] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, no. 1–3, pp. 185–203, 1981, doi: 10.1016/0004-3702(81)90024-2.

[85] S. Vedula, P. Rander, R. Collins, and T. Kanade, "Three-dimensional scene flow," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 27, no. 3, pp. 475-480, 2005, doi: 10.1109/TPAMI.2005.63.

[86] K. Yamaguchi, D. McAllester, and R. Urtasun, "Efficient Joint Segmentation, Occlusion Labeling, Stereo and Flow Estimation," Cham, 2014: Springer International Publishing, in Computer Vision – ECCV 2014, pp. 756-771.

[87] M. Menze and A. Geiger, "Object scene flow for autonomous vehicles," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7-12 June 2015 2015, pp. 3061-3070, doi: 10.1109/CVPR.2015.7298925.

[88] O. Lahav, D. Schloerb, S. Kumar, and M. Srinivasan, "A virtual environment for people who are blind – a usability study," *Journal of Assistive Technologies,* vol. 6, no. 1, pp. 38-52, 2012, doi: 10.1108/17549451211214346.

[89] M. Leo, G. Medioni, M. Trivedi, T. Kanade, and G. M. Farinella, "Computer vision for assistive technologies," *Computer Vision and Image Understanding,* vol. 154, pp. 1-15, 2017/01/01/ 2017, doi: https://doi.org/10.1016/j.cviu.2016.09.001.

[90] W. Elmannai and K. Elleithy, "Sensor-Based Assistive Devices for Visually-Impaired People: Current Status, Challenges, and Future Directions," (in eng), *Sensors (Basel),* vol. 17, no. 3, Mar 10 2017, doi: 10.3390/s17030565.

[91] L. Kay, "Air Sonars with Acoustical Display of Spatial Information," in *Animal Sonar Systems*, R.-G. Busnel and J. F. Fish Eds. Boston, MA: Springer US, 1980, pp. 769-816.

[92] N. Pressey, "Mowat sensor," vol. 3, no. 3, pp. 35-39, 1977.

[93] R. Kuc, "Binaural sonar electronic travel aid provides vibrotactile cues for landmark, reflector motion and surface texture classification," *IEEE Transactions on Biomedical Engineering,* Article vol. 49, no. 10, pp. 1173-1180, 2002, doi: 10.1109/TBME.2002.803561.

[94] I. Ulrich and J. Borenstein, "The GuideCane-applying mobile robot technologies to assist the visually impaired," *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans.*, Conference Paper vol. 31, no. 2, pp. 131-136, 2001, doi: 10.1109/3468.911370.

[95] S. Cardin, D. Thalmann, and F. Vexo, "A wearable system for mobility improvement of visually impaired people," *The Visual Computer,* vol. 23, no. 2, pp. 109-118, 2007/02/01 2007, doi: 10.1007/s00371-006-0032-4.

[96] J. Akita, T. komatsu, K. Ito, T. Ono, and M. Okamoto, "CyARM: Haptic Sensing Device for Spatial Localization on Basis of Exploration by Arms," *Advances in Human-Computer Interaction,* vol. 2009, p. 901707, 2010/01/12 2009, doi: 10.1155/2009/901707.

[97] G.-Y. Jeong and K.-H. Yu, "Multi-Section Sensing and Vibrotactile Perception for Walking Guide of Visually Impaired Person," (in eng), *Sensors (Basel)*, vol. 16, no. 7, p. 1070, 2016, doi: 10.3390/s16071070.

[98] D. A. Ross and A. Lightman, "Talking braille: a wireless ubiquitous computing network for orientation and wayfinding," presented at the Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility, Baltimore, MD, USA, 2005. [Online]. Available: https://doi.org/10.1145/1090785.1090805.

[99] K. Yelamarthi, D. Haas, D. Nielsen, and S. Mothersell, "RFID and GPS integrated navigation system for the visually impaired," in *2010 53rd IEEE International Midwest Symposium on Circuits and Systems*, 1-4 Aug. 2010 2010, pp. 1149-1152, doi: 10.1109/MWSCAS.2010.5548863.

[100] N. Márkus, A. Arató, Z. Juhász, G. Bognár, and L. Késmárki, "MOST-NNG: An Accessible GPS Navigation Application Integrated into the MObile Slate Talker (MOST) for the Blind," in *Computers Helping People with Special Needs*, Berlin, Heidelberg, K. Miesenberger, J. Klaus, W. Zagler, and A. Karshmer, Eds., 2010// 2010: Springer Berlin Heidelberg, pp. 247-254.

[101] M. Bessho, S. Kobayashi, N. Koshizuka, and K. Sakamura, "Assisting mobility of the disabled using space-identifying ubiquitous infrastructure," presented at the Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility, Halifax, Nova Scotia, Canada, 2008. [Online]. Available: https://doi.org/10.1145/1414471.1414539.

[102] A. S. Martinez-Sala, F. Losilla, J. C. Sánchez-Aarnoutse, and J. García-Haro, "Design, Implementation and Evaluation of an Indoor Navigation System for Visually Impaired People," (in eng), *Sensors (Basel)*, vol. 15, no. 12, pp. 32168-32187, 2015, doi: 10.3390/s151229912.

[103] M. H. A. Wahab *et al.*, "Smart cane: Assistive cane for visually-impaired people," *International Journal of Computer Science*, vol. 8, no. 4, pp. 21-27, 2011.

[104] W. Penrod, M. D. Corbett, and B. Blasch, "Practice Report: A Master Trainer Class for Professionals in Teaching the UltraCane Electronic Travel Device," *Journal of Visual Impairment & Blindness*, vol. 99, no. 11, pp. 711-715, 2005/11/01 2005, doi: 10.1177/0145482X0509901110.

[105] B. Hoyle and D. Waters, "Mobility AT: The Batcane (UltraCane)," in *Assistive Technology for Visually Impaired and Blind People*, M. A. Hersh and M. A. Johnson Eds. London: Springer London, 2008, pp. 209-229.

[106] L. H. Villamizar, M. Gualdron, F. Gonzalez, J. Aceros, and C. V. Rizzo-Sierra, "A necklace sonar with adjustable scope range for assisting the visually impaired," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 3-7 July 2013 2013, pp. 1450-1453, doi: 10.1109/EMBC.2013.6609784.

[107] V. Kulyukin, C. Gharpure, and J. Nicholson, "RoboCart: toward robot-assisted navigation of grocery stores by the visually impaired," in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2-6 Aug. 2005 2005, pp. 2845-2850, doi: 10.1109/IROS.2005.1545107.

[108] D. López-de-Ipiña, T. Lorido, and U. López, "Indoor Navigation and Product Recognition for Blind People Assisted Shopping," in *Ambient Assisted Living*, Berlin, Heidelberg, J. Bravo, R. Hervás, and V. Villarreal, Eds., 2011// 2011: Springer Berlin Heidelberg, pp. 33-40.

[109] S. Chumkamon, P. Tuvaphanthaphiphat, and P. Keeratiwintakorn, "A blind navigation system using RFID for indoor environments," in *5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2008*, 2008, vol. 2, pp. 765-768, doi: 10.1109/ECTICON.2008.4600543. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-52949137024&doi=10.1109%2fECTICON.2008.4600543&partnerID=40&md5=943d3c74eceaee0449470a750ba759c5

[110] A. Meliones and D. Sampson, "Blind MuseumTourer: A System for Self-Guided Tours in Museums and Blind Indoor Navigation," *Technologies*, vol. 6, p. 4, 01/04 2018, doi: 10.3390/technologies6010004.

[111] P. B. L. Meijer, "An experimental system for auditory image representations," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 2, pp. 112-121, 1992, doi: 10.1109/10.121642.

[112] G.Sainarayanan, "On intelligent image processing methodologies applied to navigation assistance for visually impaired," 2002.

[113] D. Croce, L. Giarré, F. G. L. Rosa, E. Montana, and I. Tinnirello, "Enhancing tracking performance in a smartphone-based navigation system for visually impaired people," in *2016 24th Mediterranean Conference on Control and Automation (MED)*, 21-24 June 2016 2016, pp. 1355-1360, doi: 10.1109/MED.2016.7535871.

[114] R. Manduchi, "Mobile Vision as Assistive Technology for the Blind: An Experimental Study," Berlin, Heidelberg, 2012: Springer Berlin Heidelberg, in Computers Helping People with Special Needs, pp. 9-16.

[115] J. du Buf *et al.*, "The SmartVision Navigation Prototype for Blind Users," *International Journal of Digital Content Technology and its Applications*, vol. 12, pp. 351-361, 05/31 2011, doi: 10.4156/jdcta.vol5.issue5.39.

[116] S. Kammoun, M. J.-M. Macé, B. Oriola, and C. Jouffrais, "Towards a Geographic Information System Facilitating Navigation of Visually Impaired Users," Berlin, Heidelberg, 2012: Springer Berlin Heidelberg, in Computers Helping People with Special Needs, pp. 521-528.

[117] R. Tapu, B. Mocanu, A. Bursuc, and T. Zaharia, "A Smartphone-Based Obstacle Detection and Classification System for Assisting Visually Impaired People," presented at the Proceedings of the 2013 IEEE International Conference on Computer Vision Workshops, 2013.

[118] B. D. Lucas and T. Kanade, "ITERATIVE IMAGE REGISTRATION TECHNIQUE WITH AN APPLICATION TO STEREO VISION," 1981, vol. 2, pp. 674-679. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-0019647180&partnerID=40&md5=19abd273b2b030069091e2afabd0f3ff. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-0019647180&partnerID=40&md5=19abd273b2b030069091e2afabd0f3ff

[119] J. j. Lee and G. Kim, "Robust Estimation of Camera Homography Using Fuzzy RANSAC," Berlin, Heidelberg, 2007: Springer Berlin Heidelberg, in Computational Science and Its Applications – ICCSA 2007, pp. 992-1002.

[120] M. Capp and P. Picton, "The optophone: an electronic blind aid," *Engineering Science and Education Journal*, vol. 9, no. 3, pp. 137-143, 2000, doi: 10.1049/esej:20000306.

[121] J. S. Zelek, S. Bromley, D. Asmar, and D. Thompson, "A Haptic Glove as a Tactile-Vision Sensory Substitution for Wayfinding," *Journal of Visual Impairment & Blindness*, vol. 97, no. 10, pp. 621-632, 2003, doi: 10.1177/0145482x0309701007.

[122] Y. Kawai and F. Tomita, "A support system for visually impaired persons to understand three-dimensional visual information using acoustic interface," in *Object recognition supported by user interaction for service robots*, 11-15 Aug. 2002 2002, vol. 3, pp. 974-977 vol.3, doi: 10.1109/ICPR.2002.1048200.

[123] G. Balakrishnan, G. Sainarayanan, R. Nagarajan, and S. Yaacob, "Wearable real-time stereo vision for the visually impaired," *Engineering Letters*, vol. 14, no. 2, pp. 6-14, 2007.

[124] G. Balakrishnan, G. Sainarayanan, R. Nagarajan, and S. Yaacob, "A Stereo Image Processing System for Visually Impaired," *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering*, vol. 2, pp. 2794-2803, 2008.

[125] J. M. Saez Martinez and F. Escolano Ruiz, "Stereo-based Aerial Obstacle Detection for the Visually Impaired," presented at the Workshop on Computer Vision Applications for the Visually Impaired, Marseille, France, 2008-10, 2008. [Online]. Available: https://hal.inria.fr/inria-00325455.

[126] A. Rodríguez, J. J. Yebes, P. F. Alcantarilla, L. M. Bergasa, J. Almazán, and A. Cela, "Assisting the visually impaired: obstacle detection and warning system by acoustic feedback," (in eng), *Sensors (Basel)*, vol. 12, no. 12, pp. 17476-17496, 2012, doi: 10.3390/s121217476.

[127] S. Cloix, V. Weiss, G. Bologna, T. Pun, and D. Hasler, "Obstacle and planar object detection using sparse 3D information for a smart walker," in *VISAPP 2014 - Proceedings of the 9th International Conference on Computer Vision Theory and Applications*, 2014, vol. 2, pp. 292-298, doi: 10.5220/0004661602920298. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84906895182&doi=10.5220%2f0004661602920298&partnerID=40&md5=8cabb65e6c6022bf53ffa9afaa96ef97

[128] V. Pradeep, G. Medioni, and J. Weiland, "Robot vision for the visually impaired," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, 13-18 June 2010 2010, pp. 15-22, doi: 10.1109/CVPRW.2010.5543579.

[129] L. Everding, L. Walger, V. S. Ghaderi, and J. Conradt, "A mobility device for the blind with improved vertical resolution using dynamic vision sensors," in *2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)*, 14-16 Sept. 2016 2016, pp. 1-5, doi: 10.1109/HealthCom.2016.7749459.

[130] T. Leung and G. Medioni, "Visual Navigation Aid for the Blind in Dynamic Environments," in *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 23-28 June 2014 2014, pp. 579-586, doi: 10.1109/CVPRW.2014.89.

[131] K. W. Lin, T. K. Lau, C. M. Cheuk, and Y. Liu, "A wearable stereo vision system for visually impaired," in *2012 IEEE International Conference on Mechatronics and Automation*, 5-8 Aug. 2012 2012, pp. 1423-1428, doi: 10.1109/ICMA.2012.6284345.

[132] T. Schwarze, M. Lauer, M. Schwaab, M. Romanovas, S. Böhm, and T. Jürgensohn, "An Intuitive Mobility Aid for Visually Impaired People Based on Stereo Vision," in *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, 7-13 Dec. 2015 2015, pp. 409-417, doi: 10.1109/ICCVW.2015.61.

[133] B. Deville, G. Bologna, M. Vinckenbosch, and T. Pun, "Guiding the focus of attention of blind people with visual saliency," 10/01 2008.

[134] L. Chen, B.-L. Guo, and W. Sun, "Obstacle Detection System for Visually Impaired People Based on Stereo Vision," *2010 Fourth International Conference on Genetic and Evolutionary Computing*, pp. 723-726, 2010.

[135] F. Ribeiro, D. Florêncio, P. A. Chou, and Z. Zhang, "Auditory augmented reality: Object sonification for the visually impaired," in *2012 IEEE 14th International Workshop on Multimedia Signal Processing (MMSP)*, 17-19 Sept. 2012 2012, pp. 319-324, doi: 10.1109/MMSP.2012.6343462.

[136] A. Khan, F. Moideen, J. Lopez, W. L. Khoo, and Z. Zhu, "KinDectect: Kinect Detecting Objects," Berlin, Heidelberg, 2012: Springer Berlin Heidelberg, in Computers Helping People with Special Needs, pp. 588-595.

[137] V. Filipe, F. Fernandes, H. Fernandes, A. Sousa, H. Paredes, and J. Barroso, "Blind Navigation Support System based on Microsoft Kinect," *Procedia Computer Science*, vol. 14, pp. 94-101, 2012/01/01/ 2012, doi: https://doi.org/10.1016/j.procs.2012.10.011.

[138] H. Takizawa, S. Yamaguchi, M. Aoyagi, N. Ezaki, and S. Mizuno, "Kinect cane: Object recognition aids for the visually impaired," in *2013 6th International Conference on Human System Interactions (HSI)*, 6-8 June 2013 2013, pp. 473-478, doi: 10.1109/HSI.2013.6577867.

[139] B. Li, J. P. Muñoz, X. Rong, J. Xiao, Y. Tian, and A. Arditi, "ISANA: Wearable Context-Aware Indoor Assistive Navigation with Obstacle Avoidance for the Blind," Cham, 2016: Springer International Publishing, in Computer Vision – ECCV 2016 Workshops, pp. 448-462.

[140] R. Munoz, R. Xuejian, and Y. Tian, "Depth-aware indoor staircase detection and recognition for the visually impaired," in *2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, 11-15 July 2016 2016, pp. 1-6, doi: 10.1109/ICMEW.2016.7574706.

[141] S. Blessenohl, C. Morrison, A. Criminisi, and J. Shotton, "Improving Indoor Mobility of the Visually Impaired with Depth-Based Spatial Sound," in *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, 7-13 Dec. 2015 2015, pp. 418-426, doi: 10.1109/ICCVW.2015.62.

[142] A. Brilhault, S. Kammoun, O. Gutierrez, P. Truillet, and C. Jouffrais, "Fusion of Artificial Vision and GPS to Improve Blind Pedestrian Positioning," *2011 4th IFIP International Conference on New Technologies, Mobility and Security*, pp. 1-5, 2011.

[143] S. Caraiman *et al.*, "Computer Vision for the Visually Impaired: the Sound of Vision System," in *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, 22-29 Oct. 2017 2017, pp. 1480-1489, doi: 10.1109/ICCVW.2017.175.

[144] S. Caraiman, O. Zvoristeanu, A. Burlacu, and P. Herghelegiu, "Stereo Vision Based Sensory Substitution for the Visually Impaired," *Sensors (Basel)*, vol. 19, no. 12, doi: 10.3390/s19122771.

[145] J. Y. Bouguet, "Camera calibration toolbox for matlab," 2001.

[146] J. Heikkila and O. Silven, "A four-step camera calibration procedure with implicit image correction," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 17-19 June 1997 1997, pp. 1106-1112, doi: 10.1109/CVPR.1997.609468.

[147] H. Zhencheng, F. Lamosa, and K. Uchimura, "A complete U-V-disparity study for stereovision based 3D driving environment analysis," in *Fifth International Conference on 3-D Digital Imaging and Modeling (3DIM'05)*, 13-16 June 2005 2005, pp. 204-211, doi: 10.1109/3DIM.2005.6.

[148] D. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *ArXiv*, vol. abs/2010.16061, 2020.