

**T.C.
GEBZE YÜKSEK TEKNOLOJİ ENSTİTÜSÜ
MÜHENDİSLİK VE FEN BİLİMLERİ ENSTİTÜSÜ**

**ÇİFTE AĞ METODUYLA
STEREO, ODAK VE BULANIKLIK
BİLGİSİNİ KULLANARAK
RESİMLERDEN DERİNLİK ÇIKARIMI**

**Tarkan AYDIN
DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ
ANABİLİM DALI**

**GEBZE
2011**

**T.C.
GEBZE YÜKSEK TEKNOLOJİ ENSTİTÜSÜ
MÜHENDİSLİK VE FEN BİLİMLERİ ENSTİTÜSÜ**

**ÇİFTE AĞ METODUYLA
STEREO, ODAK VE BULANIKLIK
BİLGİSİNİ KULLANARAK
RESİMLERDEN DERİNLİK ÇIKARIMI**

**Tarkan AYDIN
DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ
ANABİLİM DALI**

**TEZ DANIŞMANI
Yusuf Sinan AKGÜL**

**GEBZE
2011**



DOKTORA TEZİ JÜRİ ONAY SAYFASI

G.Y.T.E. Mühendislik ve Fen Bilimleri Enstitüsü Yönetim Kurulu'nun tarih ve/..... sayılı kararıyla oluşturulan jüri tarafından 18/02/2011 tarihinde tez savunma sınavı yapılan Tarkan AYDIN'ın tez çalışması Bilgisayar Mühendisliği Anabilim Dalında DOKTORA tezi olarak kabul edilmiştir.

JÜRİ

ÜYE

(TEZ DANIŞMANI) : Doç. Dr. Yusuf Sinan AKGÜL

ÜYE

: Yrd. Doç. Dr. Erdoğan SEVİLGİN

ÜYE

: Yrd. Doç. Dr. İlyas KANDEMİR

ÜYE

: Doç. Dr. Erkan ZERGEROĞLU

ÜYE

: Yrd. Doç. Dr. Ersin ÖZÜĞURLU

ONAY

G.Y.T.E. Mühendislik ve Fen Bilimleri Enstitüsü Yönetim Kurulu'nun/...../20... tarih ve/..... sayılı kararı.

İMZA/MÜHÜR

ÖZET

TEZ BAŞLIĞI: ÇİFTE AĞ METODUYLA STEREO, ODAK VE BULANIKLIK BİLGİSİNİ KULLANARAK RESİMLERDEN DERİNLİK ÇIKARIMI
YAZAR : TARKAN AYDIN

Bu tez stereo ve odaktan şekil yöntemlerini, göreceli avantajlarından faydalanan dezavantajlarını gidererek, aynı çayı altında birleştirmeyi amaçlamıştır. İlk olarak, birbirine bağımlı iki eşzamanlı en iyileme kullanan yeni bir stereo eşleşme yöntemi önerilmiştir. Yüzey süreksizlikleri, önerilen regülarizasyon yöntemi ile imge bölütleri ve en iyilemelerin ara adımlarda ürettiği fark haritaları göz önünde bulundurularak ele alınmıştır. Var olan regülarizasyon yöntemlerinden farklı olarak, önerilen regülarizasyon yerel en iyi problemine karşı dirençli olabilmek ve süreksizlikleri koruyabilmek için duruma göre yönbağımsız ve yönbağımlı düzleştirme uygulayabilmektedir.

Bu tezde, ayrıca, odaktan şekil yönteminin problemleri giderilerek bu yöntemin stereo ile birleştirilmesi mümkün kılınmıştır. Bu sebeple, var olan odak ölçütlerinde kapanma durumunda oluşan muğlaklığı giderebilmek için sahnenin net görüntüsünü kullanan yeni bir odak ölçüt operatörü önerilmiştir. Sahnelerdeki yüzey süreksizlikleri uyarlanabilir destek pencereleri kullanılarak etkili bir şekilde ele alınmıştır.

Önerilen odak ölçütü ile elde edilen sonuçlar sahnelerin yaklaşık yapısını bölütlerden daha iyi ifade edebilen katmanlı yapı ile gösterebilmek için kullanılmıştır. Bu yüzden imge bölütü tabanlı regülarizasyon yöntemi katmanlı gösterime dayalı daha güvenilir bir yöntemle değiştirilmiştir.

Ortaya çıkan yöntem başlangıç koşullarına duyarsız ve yerel minimumlara karşı dirençlidir. Bunun yanında, yüzey süreksizlikleriyle başa çıkar ve kapanma bölgeleri için makul sonuçlar üretir. Deneysel sonuçlar önerilen yöntemin keskin süreksizliklerle başa çıkabildiğini ve gelişmiş stereo yöntemleri ile karşılaştırılabilir sonuçlar ürettiğini göstermiştir.

SUMMARY

THESIS TITLE : RECOVERING 3D STRUCTURE FROM IMAGES WITH DUAL MESHES BY USING STEREO, FOCUS AND DEFOCUS INFORMATION

AUTHOR : TARKAN AYDIN

This thesis employs both stereo and shape from focus for depth estimation under the same framework in order to benefit from their relative advantages in depth estimation while addressing their disadvantages. We first introduce a novel stereo correspondence method that uses two synchronous interdependent optimizations. The discontinuities are handled by a novel regularizer which utilizes the image segments and the intermediate disparity maps of the two optimizations. Contrary in the existing regularizers, our regularizer can apply both isotropic and anisotropic smoothing adaptively to be robust against local minima while preserving discontinuities.

The problems of shape from focus are also addressed so that it can be employed in the new framework robustly. Hence, a new focus measurement method is proposed employs an all-focused image of the scene to address the focus measure ambiguity problem of the existing focus measures in the presence of occlusions. Discontinuities in the scene structure are also handled effectively by using adaptively shaped and weighted support windows.

Results of the introduced focus measure are used to represent scenes in a layered form which is a better approximation of the scene structure than the image segments. Therefore segment based regularization is replaced by a more reliable regularizer which is based on layered representation.

The resulting system is initialization insensitive and very robust against local minima. In addition, it accurately handles the depth discontinuities and produce acceptable results at occluded regions. Our experimental results showed that the proposed algorithm can handle sharp discontinuities well and provides disparity maps with accuracy comparable to the state of the art stereo methods.

TEŞEKKÜR

Tez süresince önerileriyle ve yapmış olduğumuz tartışmalarla beni yönlendiren, özellikle yayın çalışmalarımızda detaylı incelemeleri ve azımsanamayacak katkıları bulunan danışmanım Sayın Doç. Dr. Yusuf Sinan AKGÜL'e, beni bilgisayarla görme alanına yönlendiren, yüksek lisans ve doktora eğitimime büyük katkıları bulunan Sayın Yrd. Doç. Dr. Fatih Erdoğan SEVİLGİN'e çok teşekkür ederim.

Bu çalışma "Çifte Ağ Metodu Yardımıyla Cisimlerin 3 Boyutlu Yapılarının Bulunması İçin Bir Bilgisayarla Görme Metodu" adlı TÜBİTAK Kariyer Projesi 105E097 ile desteklenmiştir.

İÇİNDEKİLER DİZİNİ

	<u>Sayfa</u>
ÖZET	iv
SUMMARY	v
TEŞEKKÜR	vi
İÇİNDEKİLER DİZİNİ	vii
SİMGELER VE KISALTMALAR DİZİNİ	x
ŞEKİLLER DİZİNİ	xi
ÇİZELGELER DİZİNİ	xiii
1 Giriş	1
2 Stereo Eşleşme Problemi	5
2.1 Kamera Modeli	5
2.2 Stereo Görüntü Modeli	6
2.3 Stereo Yöntemleri	8
2.3.1 Eşleşme Fonksiyonları	8
2.3.2 Küresel ve Yerel Stereo Yöntemleri	9

	<u>Sayfa</u>
3 Çifte Ağ Yöntemi	13
3.1 Enerji Tabanlı Küresel Stereo Formülü	14
3.2 Çifte Ağ Enerji Formülü	14
3.3 İmge Bölütü Tabanlı Regülerizasyon	17
3.4 Deneyler	22
4 Odaktan Şekil İyileştirmeleri	28
4.1 Görüntü Oluşum Modeli	28
4.2 Odaktan Şekil Elde Etme Yöntemleri	29
4.3 Odak Ölçüt Operatörleri	31
4.4 Uyarlanabilir Odaktan Şekil	39
4.5 Kapanma Problemi	42
4.6 Kapanma-Duyarsız Odak Ölçüt fonksiyonu	45
4.7 Deneyler	46
5 Katmanlı Yapı ile Regülerizasyon	54
5.1 Sahnelerin Katmanlı Gösterilimi	56
5.2 Regülerizasyon	57
5.3 Deneyler	58

Sayfa

6 Sonuç	61
KAYNAKLAR	64
ÖZGEÇMİŞ	78

SİMGELER VE KISALTMALAR DİZİNİ

p	: (x,y) konumundaki piksel
D	: Stereo fark haritası
ϕ	: Eşleşme maliyet fonksiyonu
ψ	: Düzenlileştirme Fonksiyonu
φ	: Ağ çekim fonksiyonu
d_{min}	: En küçük fark değeri
d_{max}	: En büyük fark değeri
Δd	: Ağlar arası uzaklıklar
∇	: Gradyan operatörü
I_b	: Bölütlenmiş imge
∇I_b	: Bölütlenmiş imge gradyanı
f	: Kameranın odak uzaklığı
D	: Kamera diyafram açıklığı
OF	: Odak Ölçüt Operatörü
AKD	: Ayrık Kosinüs Dönüşümü
med	: Ortanca operatörü
ω	: Uyarlanabilir pencere ağırlığı

ŞEKİLLER DİZİNİ

<u>Şekil</u>	<u>Sayfa</u>
2.1 İğne Delikli Kamera Modeli	5
2.2 Stereo sistemler içim görüntü oluşum modeli	7
2.3 Yerel ve küresel yöntemlerin bulduğu örnek fark haritaları	11
3.1 Çifte ağ yöntemiyle ağların eniyileme süreci	13
3.2 Ağların eniyileme süreci	17
3.3 Çifte ağ yöntemi sonuçları	18
3.4 Sınır sızıntısı problemi	20
3.5 İmge bölütü ile düzenleştirilen fark haritaları	25
4.1 Temel imge oluşum geometrisi.	29
4.2 Örnek destek pencereleri.	41
4.3 Kapanma probleminin açıklaması	43
4.4 Odak ölçüt operatörlerinin bulanıklık derecesine göre karşılaştırması.	46
4.5 Kapanma Bölgeleri	47
4.6 Odak ölçüt operatörlerinin karşılaştırması	47
4.7 Venüs sahnesinin deney sonuçları	49
4.8 Cones sahnesinin deney sonuçları	50
4.9 Gerçek bir sahne üzerinde yapılan deneylerin sonuçları	51
4.10 Başka bir gerçek sahne üzerinde yapılan deneylerin sonuçları	52

5.1	İmge bölütleri ve Katmanlı yapı	56
5.2	Sentetik sahnelerde katmanlı yapı ile üretilen fark haritaları	60
5.3	Gerçek sahnelerde katmanlı yapı ile üretilen fark haritaları	60

ÇİZELGELER DİZİNİ

<u>Çizelge</u>	<u>Sayfa</u>
2.1 Yaygın olarak kullanılan eşleşme maliyet fonksiyonları	9
3.1 1 piksel hata eşiği ile hesaplanan hata oranları	24
3.2 0.5 piksel hata eşiği ile hesaplanan hata oranları	26
3.3 Yineleme sayısına göre hata oranları	27
4.1 Venüs resmi için hesaplanan rms hata oranları.	51
4.2 Cones resmi için rms hata oranları.	52
5.1 Katmalı yapı hata oranları	59

1 Giriş

Cisimlerin iki boyutlu görüntülerinden üç boyutlu yapısının elde edilmesi bilgisayarla görme alanının önemli araştırma konularından biridir. Bu amaca ulaşmak için imgelerde bulunan çok çeşitli ipuçlarından yararlanan (stereo eşleşmeler, bulanıklık dereceleri, gölge uzunluğu, vb.) birçok yöntem bulunmaktadır. Stereo bu yöntemlerin en eskilerinden biri olmasına rağmen hala birçok grup tarafından yoğun olarak araştırılmaktadır. Sistemin basitliği ve resim elde etme donanımlarının genişliği, güçlü epipolar ve geometrik kısıtların varlığı, çok geniş kullanım alanı olması gibi unsurlar, problemin popüler olmasındaki sebeplerden bir kaçıdır.

Stereo yöntemleri bir sahne noktasının sahnenin iki farklı açıdan elde edilmiş imgelerindeki izdüşümleri arasındaki farkı bulmaya çalışır. Bu işlem imgelerden birinin tüm pikselleri için yapılır ve ortaya çıkan sonuç fark haritası (disparity map) olarak adlandırılır. Bulunan fark değerleri nesnelerin kamera düzlemine olan uzaklıklarıyla ters orantılıdır. Bu yüzden stereo probleminde çözülmesi gereken ilk ve temel problem aynı sahne noktalarına denk gelen imge piksellerini eşleştirmektir [Horn, 1986; Blake and Zisserman, 1987]. Eşleşme problemini zorlaştıran temel sebepler gürültü, tekrar eden desen ve kapanma problemidir ki bu sorunlar problemin matematiksel ifadesini kötü konumlandırılmış yapar. Kötü konumlandırılmış problemler regülerizasyon işlemi ile iyi konumlandırılmış hale getirilir. En yaygın kullanılan regülerizasyon yöntemi *Tikhonov* regülerizasyonudur [Tikhonov and Arsenin, 1977]. Ancak bu yöntem ile çözülen stereo probleminde, önemli bir bilgi olan yüzey süreksizlikleri aşırı düzeltilir. Yüzey süreksizliklerine izin veren regülerizasyon yöntemlerini kullanmak problemi NP-zor yapar ki bu durumda problemin çokterimli (polinomsal) zamanda çözümünü bulmak mümkün olmamaktadır [Kolmogorov and Zabih, 2004]. Bu yüzden araştırmalar problemin yaklaşık sonucunu bulacak yöntemler üzerinde yoğunlaşmıştır [Boykov et al., 2001; Kolmogorov and Zabih, 2001, 2002; Sun et al., 2003; Saito and Mori, 1995]. Diğer yandan, bazı araştırmacılar da problemin çözümü için diğer derinlik bulma yöntemlerinden faydalanma [Cryer et al., 1993; Bove, 1990], yapılandırılmış ışık [Zhang et al., 2003; Scharstein and Szeliski, 2003a] veya flash [Feris et al., 2008] kullanma gibi alternatif yöntemler denemişlerdir.

Bu tez çalışmasında çiftte ağ yöntemiyle stereo probleminin çözümü üzerinde durulmuştur. Çifte ağ yöntemi, stereo imgelerden sürekli yüzeylerin 3B yapısını bulmak için Akgul and Kambhamettu [1999] tarafından önerilmiştir. Bu çalışmada çiftte ağın geliştirilmesi ve diğer derinlik yöntemleriyle birleştirilmesi amaçlanmıştır.

Yöntemin yüzey süreksizliklerini de doğru bir şekilde çıkarabilmesi için imge bölütü tabanlı yeni bir regülarizasyon yöntemi önerilmiştir [Aydın and Akgul, 2010b]. Önerilen regülarizasyon yöntemi, var olan regülarizasyon yöntemlerinde olmayan bir özellik olarak, imge bölütleriyle beraber ağların birbirlerine olan uzaklık bilgisini de kullanarak yayılım tabanlı düzleştirme işlemini yönlendirir. Bu sayede yöntem, ağların duruma göre uyarlanabilir şekilde yönbağımlı veya yönbağımsız düzleştirme yapma imkanına kavuşmuş olur.

Yayılım tabanlı düzleştir yöntemlerinde, çözümün aşırı düzleştirilmemesi için veya yetersiz regülarizasyona sebep olmamak için, en iyi durdurma şartını tespit etmek oldukça önemli bir problemidir. Önerilen yöntem ağların buldukları ara çözümlerine bağlı oldukları için, iki ağında aynı çözümü bulduğu yerlerde bölütler arası yayılım engellenebilir Bu sayede en iyi sonuca ulaştıktan sonra yapılacak gereksiz yinelemelerde bile çözümün aşırı düzleştirilmesi engellenir. Bu özellik, yöntemi benzer sistemlere göre avantajlı kılmaktadır.

Bu tezde ayrıca, stereo sitemleriyle birleştirilebilmeleri için, odaktan şekil yönteminin sorunları da giderilmeye çalışılmıştır. Odaktan şekil yöntemlerinde göz ardı edilen en önemli problem kapanma problemidir [Marshall et al., 1996; Asada et al., 1995, 1998]. Açık diyaframlı kameralarda görülen kapanma problemi, kapanan nesnenin bir kısmının görünür olması durumudur. Bilindiği gibi böyle bir durumun iğne delikli kameralarda olması mümkün değildir. Kapanmanın olacağı durumlar Schechner and Kiryati [2000] tarafından detaylı bir şekilde incelenmiştir. Odaktan şekil ile kapanma bölgeleri için bulunan derinlik değerleri muğlaklık içerdiğinden güvenilir değildir [Schechner and Kiryati, 2000]. Bu problem çeşitli yöntemler tarafından incelenirse de [Asada et al., 1998; Bhasin and Chaudhuri, 2001; Favaro and Soatto, 2003], odaktan şekil yöntemleri için önerilmiş bir çözüm bulunmamaktadır.

Kapanma problemini çözebilmek için odaklanılmış imge bölgelerini bulmaya

yarayan yeni bir odak ölçüt operatörü önerilmiştir [Aydin and Akgul, 2010a]. Önerilen yöntem, var olan odak ölçüt operatörlerinin aksine imgeler üzerinde yüksek frekanslı içerik bulmak yerine, sahnenin net görüntüsünü kullanarak imgelerin odaklanma derecesini eşleşme işlemi ile hesaplar. Dolayısı ile problem sahnenin net görüntüsünün derinlik değerlerini bulma problemine dönmüş olur ki bu özellik sayesinde yöntemin stereo yöntemleri ile birleştirilmesi için pikseller arası karmaşık kayıtlanma (registration) işlemlerine gerek kalmamış olur.

Odaktan şekil yönteminin çözülmesi gereken bir diğer önemli problemi de yüzey süreksizlikleridir. Çalışmada, bu probleme çözüm olarak odak ölçüt operatörleriyle beraber kullanılacak uyarlanabilir destek pencereleri [Aydin and Akgul, 2008] önerilmiştir. Destek pencereleri sahnenin elimizde var olan net görüntüsü kullanılarak oluşturulur. Uyarlanabilir odaklanma derecesi hesaplaması büyük pencere boyutlarının kullanılmasına izin vermektedir. Bir başka deyişle, pencere boyutunun arttırılması diğer odaktan şekil yöntemlerinin aksine, yöntemin başarımını azaltmamaktadır. Bu sayede ayrıt sızması [Nair and Stewart, 1992] problemine de çözüm getirilmiş olur.

Önerilen odak ölçüt operatörünün en önemli kusuru, sahnelerin net görüntüleri kısmi bulanıklık içerdiği için, bulunan derinlik değerlerinin standart sapmasının yüksek olmasıdır. Bu yüzden bulunan bu derinlik değerleri sahnelerin katmanlı yapısı olarak kullanılmıştır. Katmanlı yapı sahnenin imge bölütlerinde olduğu gibi düzlemlerle ifade edilmiş halidir. Bu yapının imge bölütlerinden farkı katmanların derinliğe bağlı bir sıralama içermesidir. Bu yapının çifte ağla kullanılabilmesi için yeni bir regülarizasyon fonksiyonu tanımlanmıştır. Katmanlı yapı stereo kapanma bölgelerinin de yüzey yapısını içerdiği için, regülarizasyonda bu bölgelere özel bir tasarlama yapılmasına rağmen, yöntemin kapanma bölgelerindeki başarımı artmıştır.

Stereo yöntemi ile odaktan şekil yöntemlerini katmanlı yapıyı kullanarak birleştirmenin en önemli getirisi sistemin herhangi bir kayıtlanma ve ölçümleme işlemine ihtiyacı kalmamasıdır. Kayıtlama yapan alternatif yöntemler mercek sistemlerinde bulunan alan eğriliği, alan derinliği, vb. gibi optik sapmalardan dolayı uygulaması zor çok karmaşık ölçümleme işlemi gerektirmektedir Ahuja and Abbott [1993].

Tezin planlaması şu şekilde yapılmıştır. Bölüm 2 stereo problemini ve öneri-

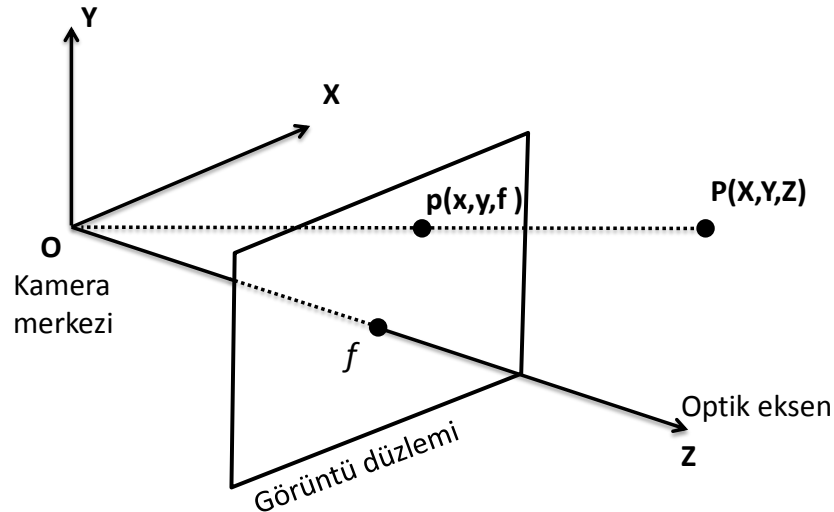
len yaklaşımları özetler. Çifte ağ yöntemi ve önerilen imge bölütü tabanlı regülarizasyon yöntemi Bölüm 3’te anlatılmıştır. Bölüm 4’te odaktan şekil yöntemi ve önerilen kapanma duyarsız odak ölçüt operatörü anlatılmıştır. Bölüm 5’te sahnenin katmanlı yapısı kullanılarak yapılan regülarizasyon yöntemi anlatılmış. Bölüm 6 bulgular yorumlanarak tez sonlandırılmıştır.

2 Stereo Eşleşme Problemi

2.1 Kamera Modeli

Yaygın olarak kullanılan en basit kamera modeli perspektif model olarak da bilinen iğne delikli modeldir. Bu modelde gerçek kameralarda bulunan mercek-diyafram sistemi düşünülmediği için, merceklerin sebep olduğu bulanıklık, geometrik bozulma gibi durumlar açıklanmamıştır. Bu yüzden iğne delikli kamera modeli, kullanılan kameraların lineer yaklaşım modeli olarak bilinir.

İğne delikli kamera modeli O noktasındaki kamera merkezinden f kadar uzağa yerleştirilmiş bir görüntü düzleminde oluşur. Burada f kameranın odak noktası olarak bilinir. Kamera merkezinden çıkan ve görüntü düzlemini dik olarak kesen doğruya optik eksen adı verilir. Görüntü düzlemi ile optik eksenin kesiştiği noktaya imge merkezi denir. Şekil 2.1 anlatılan bu kamera modelini göstermektedir. Modele göre, gerçek dünyadaki herhangi bir P noktasının görüntüsü, P noktası ile kamera merkezi olan O noktası arasında çizilecek doğru parçasının görüntü düzlemini kestiği p noktasında oluşur.



Şekil 2.1 İğne Delikli Kamera Modeli. Gerçek dünyadaki P noktasının görüntü düzlemindeki izdüşümü p noktası olarak verilmiştir.

İğne delikli kamera modeli görüntüsü oluşacak her noktanın görüntü düzlemine perspektif izdüşümünü bulmak için kullanılır. Basit üçgen benzerliklerinden kamera

koordinat sistemindeki sahne noktası $P(X, Y, Z)$ ile bu noktanın görüntüsü $p(x, y, z)$ arasındaki model bağıntıları şu şekilde verilir.

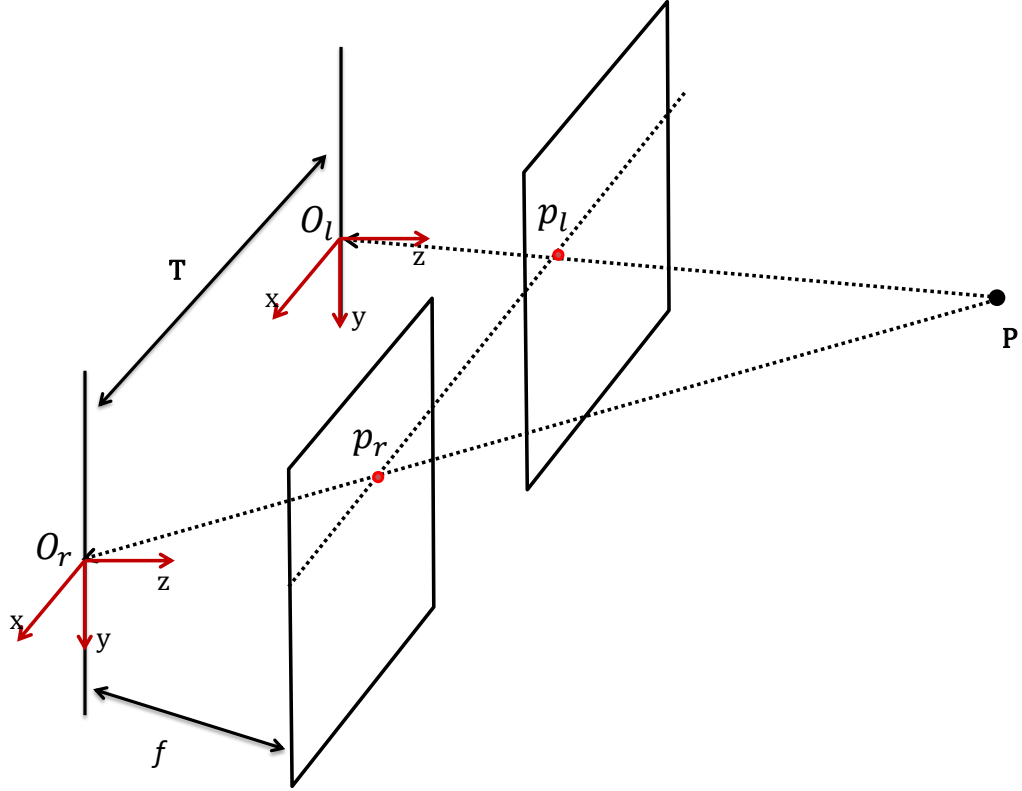
$$\begin{aligned} x &= f \frac{X}{Z} \\ y &= f \frac{Y}{Z} \\ z &= f \end{aligned} \quad (2.1)$$

2.2 Stereo Görüntü Modeli

En basit stereo sistemi iki adet iğne delikli kamerayı aynı yöne bakacak ve görüntü düzlemleri aynı düzlemde olacak şekilde konumlandırarak elde edilir. Şekil 2.2’de bu sistem gösterilmektedir. Şekilde O_l ve O_r sırasıyla sol ve sağ kameranın kamera merkezlerini göstermektedir. I_l ve I_r ise sırasıyla sol ve sağ kameranın görüntü düzlemleridir. Sahnedeki herhangi bir P noktasının görüntü düzlemlerine izdüşümleri sol kamera için p_l ve sağ kamera için p_r olsun. Stereo yöntemlerindeki temel amaç p_l ve p_r noktalarını bulup üçgenleme yöntemi ile P noktasının derinlik değerini (Z) bulmaktır.

Üçgenleme yöntemi kamera merkezleri ve P noktasının oluşturduğu üçgen kullanılarak stereo bağıntılarının çıkarılmasını sağlar. P, O_l ve O_r noktalarını içeren düzlemin kamera görüntü düzlemlerini kestiği doğrulara epipolar doğru adı verilir. Şekilde anlatılan sistem için, kamera görüntü düzlemlerinin aynı düzlemde oldukları göz önüne alınırsa P noktasının p_l ve p_r olan izdüşümlerinin y koordinatlarının aynı değere sahip olduğu kolayca anlaşılır. Kamera merkezleri arasındaki uzaklık $T = |O_l - O_r|$ olsun. Koordinat sistemini kamera merkezlerinin tam ortasına taşırsak P noktasının kamera sistemlerine göre koordinatları şu şekilde değişir.

$$\begin{aligned} (X_l, Y_l, Z_l) &= (X - T/2, Y, Z) \\ (X_r, Y_r, Z_r) &= (X + T/2, Y, Z) \end{aligned} \quad (2.2)$$



Şekil 2.2 İki iğne delikli kamera kullanılarak oluşturulan stereo sisteminde görüntü oluşum modeli

Kameranın P noktasına uyguladığı denklem 2.1'teki perspektif izdüşümü göz önünde bulundurulduğunda p_l ve p_r imge noktalarının x koordinatları şu şekilde olur.

$$x_l = (X - T/2)f/Z \quad (2.3)$$

$$x_r = (X + T/2)f/Z \quad (2.4)$$

Bu iki denklem taraf tarafa çıkarılırsa aşağıdaki bağıntı bulunur.

$$x_l - x_r = Tf/Z. \quad (2.5)$$

Burada $x_l - x_r$ değeri p_l ve p_r için fark değeri (*disparity*) olarak bilinir ve d ile gösterilir. Yukarıdaki denklemler tekrar düzeltildiklerinde P noktasının koordinatları şu şekilde bulunur

$$X = T(x_l + x_r)/2d \quad (2.6)$$

$$Y = Ty/d \quad (2.7)$$

$$Z = Tf/d \quad (2.8)$$

Bulunan bu koordinatlar kamera koordinat sistemindeki koordinatlardır. Nesnelerin gerçek dünya koordinatlarına göre konumları kamere ölçümlemesi yapıldıktan sonra bulunabilir [Grosky and Tamburino, 1990]. Denklem 2.8’de görüldüğü gibi fark değerleri noktaların derinlikleriyle ters orantılıdır. Bu yüzden stereo yöntemlerinin çözünürlüğü kameraya yakın cisimler için daha yüksektir.

Yukarıdaki denklemler Şekil 2.2’de bulunan en basit stereo sistemi için geçerli olsa da, görüntü düzlemleri aynı düzlem üzerinde olmayan sistemler için de uygulanacak düzeltme (rectification) işleminden sonra geçerli olacaktır. Düzeltme işlemi stereo problemini kolaylaştırmak için stereo imgelere uygulanan bir dönüşüm işlemidir [Ayache and Hansen, 1988; Papadimitriou and Dennis, 1996; Fusiello et al., 2000]. Yapılan dönüşümle görüntü düzlemleri aynı düzlem üzerine taşınır.

2.3 Stereo Yöntemleri

Stereo yöntemlerinin sahnedeki nesnelerin derinliklerin bulabilmeleri için ilk önce aynı sahne noktasının imgeler üzerindeki izdüşümlerini eşleştirmesi gerekir. Bu amaçla imge piksellerinin birbirine benzerliğini hesaplayan fonksiyonlar kullanılır.

2.3.1 Eşleşme Fonksiyonları

Literatürde imge parçalarının birbirine benzerliğini ölçmek için çeşitli yöntemler mevcuttur. Kullanılan fonksiyona göre eşleşen imge parçaları, fonksiyonun maksimum ya da minimum değeri üreten parçalar arasından seçilir. Yaygın olarak kullanılan eşleşme fonksiyonları Çizelge 2.3.1’de özetlenmiştir.

Çizelge 2.1 Yaygın olarak kullanılan eşleşme maliyet fonksiyonları. A ve B eşleşme değerinin hesaplanacağı imge parçalarıdır.

Eşleşme Maliyet Fonksiyonu	Formül
Fark Karelerinin Toplamı	$\phi(A, B) = \sum (A(p) - B(p))^2$
Mutlak Farkların Toplamı	$\phi(A, B) = \sum A(p) - B(p) ^2$
Sıfır Ortalamalı Mutlak Fark Toplamı	$\phi(A, B) = \sum (A(p) - \bar{A}) - (B(p) - \bar{B}) ^2$
Çapraz İlinti	$\phi(A, B) = \sum A(p)B(p)$
Normalize edilmiş Çapraz İlinti	$\phi(A, B) = \frac{\sum (A - \bar{A})(B - \bar{B})}{\sqrt{\sum (A - \bar{A})^2 \sum (B - \bar{B})^2}}$
Normalize edilmiş Fark Karelerinin Toplamı	$\phi(A, B) = \sum \left(\frac{(A - \bar{A})}{\sqrt{\sum (A - \bar{A})^2}} - \frac{(B - \bar{B})}{\sqrt{\sum (B - \bar{B})^2}} \right)^2$
Rank	$A'(p) = \sum_{q \in \omega} A(q) < A(p)$ $\phi(A, B) = \sum_{q \in \omega} A'(p) - B'(p)$

2.3.2 Küresel ve Yerel Stereo Yöntemleri

Literatürdeki stereo derinlik bulma yöntemlerinin problemi ele alış bakımından Küresel veya Yerel Yöntem olarak sınıflandırılabilirler. Yerel yöntemler eşleşmeyi her piksel için diğer piksellerden bağımsız olarak yaparlar. Bu sayede toplamda bütün pikseller için minimum eşleşme maliyeti üretecek fark haritası oluşturulur. Küresel yöntemler ise eşleşen pikselleri tespit ederken eşleşen diğer piksellerin durumlarına da bakarlar. Eşleşme problemi kapanmalar, yinelenen desen, imge gürültüsü gibi sebeplerden dolayı kötü konumlandırılmış (ill-posed) bir problemdir.

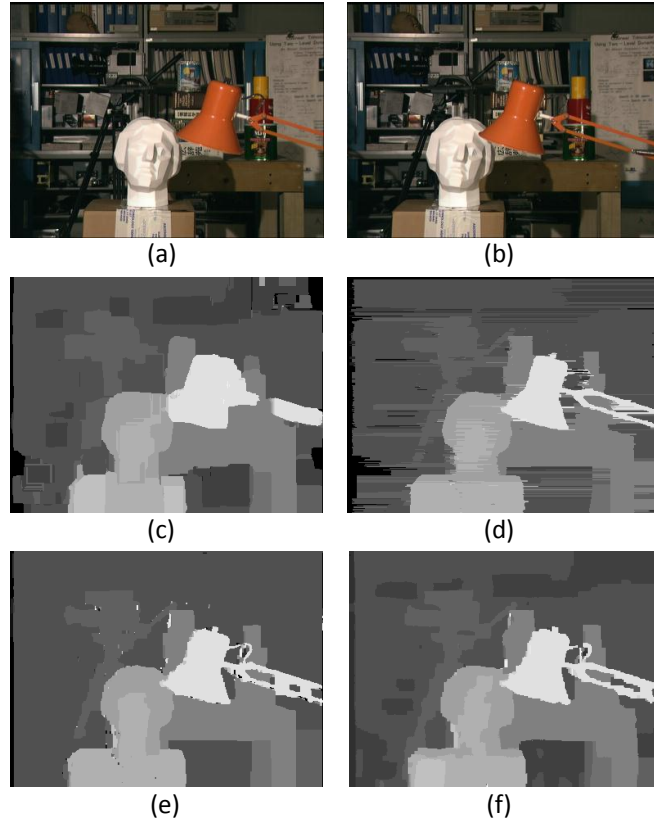
İyi konumlandırılmış (well-posed) problemin şu şartları sağlaması gerekir [Hadamard, 1902];

1. Problemin bir çözümü olmalı (varlık)
2. Problemin tek bir çözümü olmalı (teklik)
3. Çözüm süreklilik arz edecek şekilde veriye bağlı olmalı. Bir başka deyişle, verideki küçük değişiklikler çözümde büyük değişikliklere sebep olmamalı.

İyi konumlandırılmamış problemler Hadamard [1902] tarafından kötü konumlandırılmış olarak adlandırılmıştır. Kötü konumlandırılmış problemler regülarizasyon işlemi ile iyi konumlandırılmış hale getirilirler. Yerel yöntemler regülarizasyonu, eşleşme penceresi kullanma, eşleşmenin yayılımı gibi yöntemlerle eşleşme maliyetleri üzerinde uygulamalar [Intille and Bobick, 1994; Kanade and Okutomi, 1994; Scharstein and Szeliski, 1998; Ben-Ari and Sochen, 2008]. Diğer taraftan, küresel yöntemler problemi ifade eden ve çözüme ulaşmak için eniyilenmesi gereken bir enerji fonksiyonu tanımlar. Regülarizasyon, probleme açık bir şekilde düzlük kısıtı getirilerek fark haritası üzerine uygulanmıştır. Düzlük kısıtını sadece epipolar satırlar üzerinde zorlayarak enerji denkleminin küresel minimum değerine ulaşmak dinamik programlama yöntemiyle mümkündür. Ancak, bu şekilde bulunan fark haritaları epipolar doğrular arası tutarsızlıklardan dolayı düzensiz çizgiler içerirler (Bkz. Şekil 2.3.2.b). Düzlük kısıtı dışbükey bir fonksiyon ile sağlandığında, denklemin minimum değerine ulaşmak çizge kesme yöntemiyle mümkündür [Roy and Cox, 1998; Ishikawa, 2003]. Ancak düzlük fonksiyonu yüzey süreksizliklerinin aşırı düzleştirilmesine yol açar. Potts [Potts, 1952] Modeli gibi yüzey süreksizliklerine izin veren düzleştirici fonksiyonları kullanmak problemi NP-zor yapar ki bu durumdaki problemlerin çokterimli (polynomial) zamanda çözülebilmesi mümkün olmamaktadır [Kolmogorov and Zabih, 2004]. Bu yüzden bu konuda yapılan çalışmalar daha düşük enerji değerlerini üretecek yöntemler geliştirme üzerine yoğunlaşmıştır [Boykov et al., 2001; Kolmogorov and Zabih, 2001, 2002; Sun et al., 2003].

Bazı küresel yöntemler problemi kısmi diferansiyel denklemleri ve değişkenler yöntemlerini kullanırlar. Bu yöntemler tanımlanan enerji fonksiyonunun Euler-Lagrange denklemlerini kullanarak yinelemeli eniyilerler. Shah [1993] doğrusal olmayan yayılım kullanarak fark haritasını ve kapanma haritasını eşzamanlı olarak çıkarır. Benzer bir şekilde Robert and Deriche [1996] fark haritasında oluşan yüzey süreksizliklerini yönbağımlı yayılım kullanarak korurlar. Bazı araştırmacılar stereo problemi Mumford-Shah denklemleriyle ifade edip akış tabanlı düzleştirme uygulamışlardır [Pock et al., 2007]. Bu yöntemlerin en belirgin avantajı üretilen fark haritalarının diğer küresel yöntemlerden farklı olarak sürekli değerler içermesidir.

Geliştirilen yöntemlerin eniyilemedeki başarımları artmasına rağmen, Tappen



Şekil 2.3 Tsukaba sahnensinin (a) Sol ve (b) sağ stereo imgeleri. Bu sahne için (c) Fark Karelerinin Toplamı kullanılarak çıkarılan fark haritası [Scharstein and Szeliski, 2002]. (d) Dinamik programlanma [Intille and Bobick, 1994] (e) çizge kesme [Boykov et al., 2001] ve (f) inanç yayma [Sun et al., 2003] (küresel) yöntemleriyle çıkarılan örnek fark haritaları.

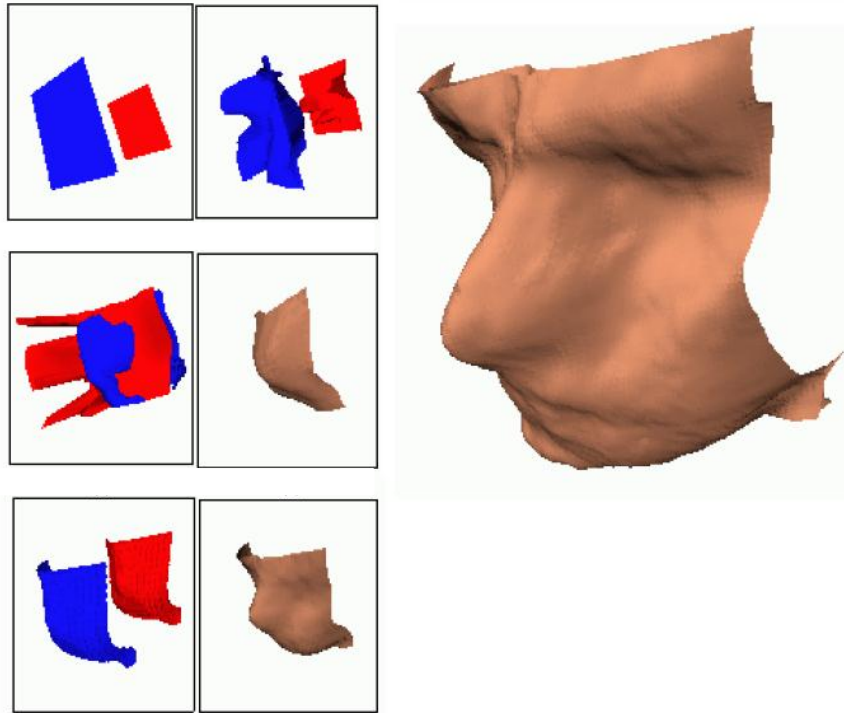
and Freeman [2003] tarafından ortaya konulan çalışmaya göre, daha düşük enerji değerlere ulaşmak, her durumda daha yüksek başarılı sonuçlar ortaya çıkarmamaktadır. Buradan da anlaşılacağı gibi, daha düşük enerji değerlerini üretecek yöntemler bulmak yerine, enerji fonksiyonu üzerinde, sahne geometrisi hakkında daha fazla bilgi içerecek şekilde değişiklikler yapmak yerinde bir yaklaşım olacaktır.

Sahne derinlik bulma yöntemleri için sahne geometrisi hakkındaki en işe yarar bilgi yüzey süreksizlik bölgeleridir. Yüzey süreksizlikleri önceden bilinmediğinden enerji fonksiyonuna eklenmesi yöntemlerin başarımını arttırmaktadır. Bir çok modern stereo yöntemi yüzey süreksizliklerini sahne görüntülerindeki parlaklık süreksizliklerinden tahmin ederler. Buradaki genel varsayım yüzey süreksizliğinin olduğu bölgede parlaklık süreksizliğinin de olacağıdır. Yerel yöntemler bu bilgiyi eşleşme maliyetini iyileştirmek için kullanırlar [Wei and Quan, 2004]. Benzer bir şekilde Yoon and Kweon [2006]'nın önerdiği maliyet iyileştirme yönteminde eşleşme değerleri, sahne görün-

tüsündeki parlaklık değerlerine göre üretilen uyarlanabilir ağırlıklarla iyileştirilir. Bu sayede sadece görüntü bölgelerinin homojen kısımları arasında eşleşmeler hesaplanmış olur. Son zamanlarda, çizge kesme, inanç yayma gibi çeşitli eniyileme yöntemlerini kullanan küresel stereo algoritmaları, yüzey süreksizliklerinin bölümlenmiş görüntülerdeki bölütlerin sınırlarında olacağı kabullenmesiyle, eşleşme işlemini pikseller arasında hesaplamak yerine bölütler arasında üzerinden yapmışlardır [Zhang and Kambhamettu, 2002; Hong and Chen, 2004; Klaus et al., 2006]. Buna karşılık yayılım tabanlı yöntemler imge ayrıt bilgisini kullanarak yönbağımlı yayılım operatörü tanımlamışlardır [Alvarez et al., 2002; Ben-Ari and Sochen, 2008; Aydin and Akgul, 2010b]. Yayılım miktarı ve yönü imge ayrıt büyüklüğüyle orantılanır. Bu sayede tanımlanan operatör ayrıt doğrultusunda yayıma izin verirken ayrıta dik doğrultudaki yayıma izin vermez.

3 Çifte Ağ Yöntemi

Bu tezde önerilen yöntem çifte ağ yapısına dayanmaktadır. Çifte ağ yöntemi yinelemeli küresel stereo yöntemlerinde bulunan başlangıç koşulu problemine çözüm olarak Akgul and Kambhamettu [1999] tarafından önerilmiştir. Yöntemde stereo probleminin çözümü olan fark haritası, şekil değiştirebilen ağ olarak ifade edilmiştir. Ağların şekil değiştirmesi birbirine bağımlı olarak tanımlanan iki stereo enerji fonksiyonunun eşzamanlı olarak eniyilemesi ile sağlanır. İki farklı başlangıç koşuluyla başlatılan iki ağ, enerji fonksiyonuna eklenen çekim enerjisi terimi ile birbirine doğru ilerlemeye zorlanır. Bu yaklaşımın en önemli avantajı sistemin başarımının başlangıç koşullarına olan bağımlılığını ortadan kaldırmasıdır. Eniyilemenin başlatılması tüm problemler için daima aynı şekilde gerçekleştirilir. Yöntemin bir diğer avantajı da hareket takibi problemine de kolayca uygulanabilir olmasıdır [Akgul and Kambhamettu, 1999].



Şekil 3.1 Çifte ağ yöntemiyle ağların şekil değiştirerek eniyileme süreci [Akgul and Kambhamettu, 1999]

Diğer yandan, yöntemin en önemli dezavantajı yüzey süreksizliklerine izin vermemesidir. Ayrıca, ağların etkileşimi sağlayan çekim enerjisi ağlara sabit çekim kuvveti uyguladığı için eniyileme çoğu kez yerel minimumlara takılmaktadır. Çoğu zaman

yöntemin yakınsaması için dışarıdan müdahaleye ihtiyaç vardır. Tezin bu kısmında öncelikle çifte ağın bahsedilen bu problemlerinin çözümü amaçlanmıştır.

3.1 Enerji Tabanlı Küresel Stereo Formülü

Klasik stereo enerji formülü veri terimi ile regülarizasyon teriminin toplamı olarak ifade edilir [Terzopoulos, 1986]. Buna göre, stereo eşleşme problemi aşağıda tanımlanan enerji fonksiyonun eniyileme problemi olarak ifade edilir.

$$E(D) = \int \alpha \phi(D) + \beta \psi(|\nabla D|) dp, \quad (3.1)$$

burada $D(p)$ sol stereo imgesindeki p pikselinin sağ stereo imgesinde eşleştiği q pikseli arasındaki konumsal farkı gösterir. Kolaylık olması için denklemlerde kullanılırken fonksiyon parametresi p gösterilmemiştir. Denklemdaki α ve β ise terimlerin göreceli katkısını ayarlamak için seçilen sabit ağırlıklardır.

Veri terimi ϕ , normalize edilmiş çapraz ilinti, mutlak değerlerin toplamı, ve farkların karelerinin toplamı gibi yaygın olarak bilinen benzerlik ölçütlerinden biri kullanılarak hesaplanan eşleşme maliyetini hesaplayan bir fonksiyondur. Regülarizasyon ya da düzleştirme terimi olarak bilinen ψ ise problemin çözümüne yardımcı olacak kısıtları enerji formülüne eklemek için kullanılır.

3.2 Çifte Ağ Enerji Formülü

Klasik stereo enerji formülüne bağlı olarak eşzamanlı olarak eniyilemesi yapılacak çifte ağ enerji formülleri, formüle ağlar arasında etkileşimi sağlayan çekim enerjisi terimi (φ) eklenerek şu şekilde tanımlanır.

$$E(D_1) = \int \alpha \phi(D_1) + \beta \psi(|\nabla D_1|^2) + \lambda \varphi((D_1 - D_2)^2) dp \quad (3.2)$$

$$E(D_2) = \int \alpha \phi(D_2) + \beta \psi(|\nabla D_2|^2) + \lambda \varphi((D_2 - D_1)^2) dp \quad (3.3)$$

burada D_1 ve D_2 her bir ağın hesapladığı fark haritalarıdır.

Tanımlanan çekim enerjisi terimi φ ağların birbirlerinden etkilenme derecesini ayarlayarak çift ağ yönteminin temel fikrini gösterir. Bu terimin görevi bulunacak fark haritaları D_1 ve D_2 'nin birbirine benzemesini sağlamaktır. Denklem 3.2 ve Denklem 3.3'ten de anlaşılacağı gibi, eniyileme işlemi bu terim olmadan farklı başlangıç koşullarıyla ayrı ayrı yapılırsa herbir ağ değişik fark haritası üretir. Fakat denklemler bu terim ile birlikte eşzamanlı olarak eniyileştirilirse, eniyileme işlemleri aynı sonucu bulmaya zorlanmış olur.

Bilindiği gibi regülarizasyon teriminin görevi komşu piksellere birbirine yakın derinlik değeri atanmasını sağlamaktır. Benzer bir şekilde çekim enerjisi teriminin görevi iki eniyileme işleminin de her piksel için aynı derinlik değerlerini bulmasını sağlamaktır. Dolayısı ile çekim terimi de bir çeşit regülarizasyon yapmaktadır. Yapılan bu ek regülarizasyon işlemi, stereo probleminin kötü konumlandırılmış olmasına ek çözüm getirmiş olur.

Sahnenin derinlik değerlerini hesaplamaya yarayacak fark haritalarını bulmak için tanımlanan enerji fonksiyonlarının en küçük değerini üretecek sonuç aranır. Bu işlem için bir eniyileme yöntemi gerekir ki bu işlem de yokuş inme yöntemi ile gerçekleştirilir. Denklem 3.2 ve Denklem 3.3'ün yokuş inme yöntemi ile eniyilemesi yapay değişim parametresi t tanımlanarak aşağıdaki gibi yapılır.

$$\frac{\partial D_1}{\partial t} = \gamma \left(\alpha \phi'(D_1) + \beta \nabla \cdot (\psi' \nabla D_1) + \lambda \varphi'(D_1 - D_2) \right) \quad (3.4)$$

$$\frac{\partial D_2}{\partial t} = \gamma \left(\alpha \phi'(D_2) + \beta \nabla \cdot (\psi' \nabla D_2) + \lambda \varphi'(D_2 - D_1) \right), \quad (3.5)$$

burada ϕ' , ψ' ve φ' sırasıyla ϕ , ψ ve φ fonksiyonlarının türevleridir.

ψ' yayılım katsayısı [Perona and Malik, 1990] olarak da bilinen düzleştirme kısıtını uygulayan fonksiyondur. Herhangi bir t anında hesaplanan fark haritalarının D_1^t ve D_2^t olduğu kabul edilirse, denklem 3.2 ve 3.3'ün en düşük değeri, denklem 3.4 ve 3.5'nin sonuçur durumunda ($t \rightarrow \infty$) ürettikleri D_1^t ve D_2^t hesaplanarak bulunur.

Çekim terimi φ iki eniyileme işlemini de aynı çözüme yakınsamaya zorlar. Ancak, kontrolsüz olarak iki ağ da birbirlerine doğru hareket etmeye zorlamak ağların

doğru olarak buldukları konumları kaybetmesine yol açar. Bu durum eniyilemelerin yerel minimuma yakınsamasına sebep olabilir. Oluşabilecek bu tip bir problemle karşılaşmamak için çekim katsayısı φ' daha küçük eşleşme değerine (veri terimi) sahip ağı etkilemeyecek şekilde ayarlanmalıdır. Bu yüzden çekim katsayısı birinci eniyileme işlemi için şu şekilde tanımlanmıştır,

$$\varphi' = \begin{cases} 1 - e^{(\frac{\Delta\phi}{\kappa_\varphi})^2} & \Delta\phi \geq 0 \\ 0 & \text{diğer} \end{cases}, \quad (3.6)$$

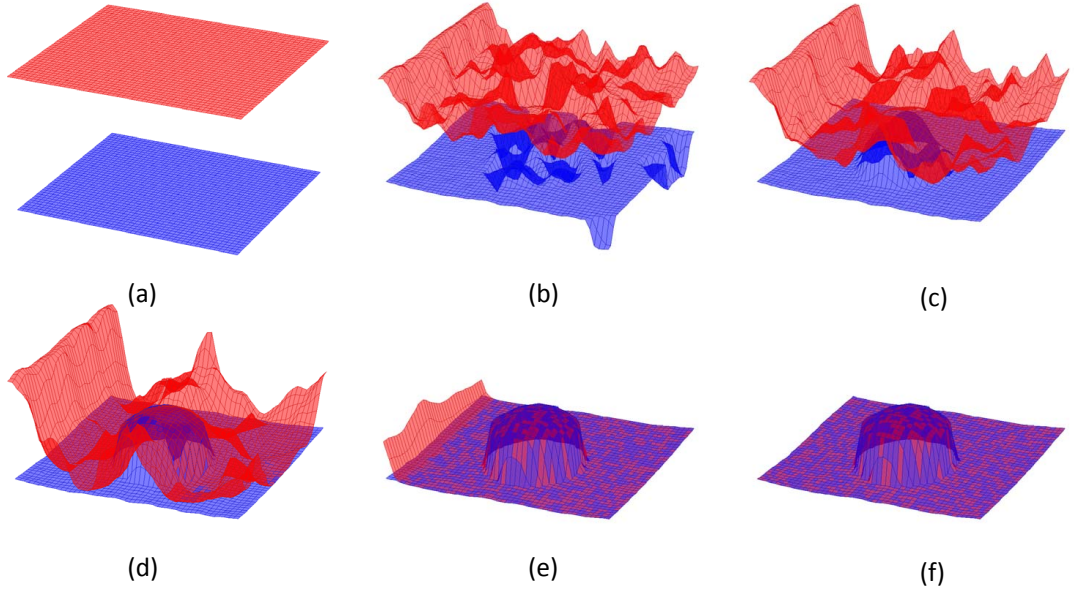
burada κ_φ sabit katsayı değerini, $\Delta\phi$ ise iki ağı arasındaki eşleşme değerlerinin farkını gösterir ve şu şekilde hesaplanır,

$$\Delta\phi = \phi(D_1^t) - \phi(D_2^t) \quad (3.7)$$

Çekim katsayısı diğer eniyileme işlemi için de benzer şekilde hesaplanır.

Formülden de anlaşılacağı üzere çekim enerjisi simetrik değildir. Ayrıca ağların bulunduğu konuma göre değişmektedir. Diğer ağı göre daha düşük benzerlik değerleri olan ağı noktaları için bu değer sıfır olacağından, bu durumdaki ağı diğer ağıdan etkilenmeden eniyilemeye devam eder. Buna karşılık diğer ağıdaki aynı pozisyondaki noktalar bu ağı doğru hareket eder. Bu yaklaşım yerel minimumlara takılma problemini önemli ölçüde gidermiş olur, çünkü yerel minimuma takılan ağı diğer ağı tarafından regülasyon işlemine maruz kalır. Sonuç olarak eş zamanlı yapılan eniyileme, birbirinden bağımsız yapılacak eniyilemelerden daha iyi sonuçlar üretir.

Yöntemlerin istenilen sonuca yakınsaması için döngülü eniyileme işlemlerinin başlangıç koşullarının belirlenmesi çok önemli adımdır. Küresel minimumdan uzak başlangıç koşullarıyla yapılan eniyileme çoğu durumda yerel minimuma yakınsar. Önerilen yöntemde başlangıç koşulları küresel minimum değerini veren çözüm hakkında herhangi bir varsayıma bağlı değildir. Tüm örnek problemler için başlangıç koşulları aynı şekilde verilmektedir. Eniyileme işlemlerinden biri minimum fark (d_{min}) değerleriyle başlatılırken diğeri de maksimum (d_{max}) fark değerleriyle başlatılır.



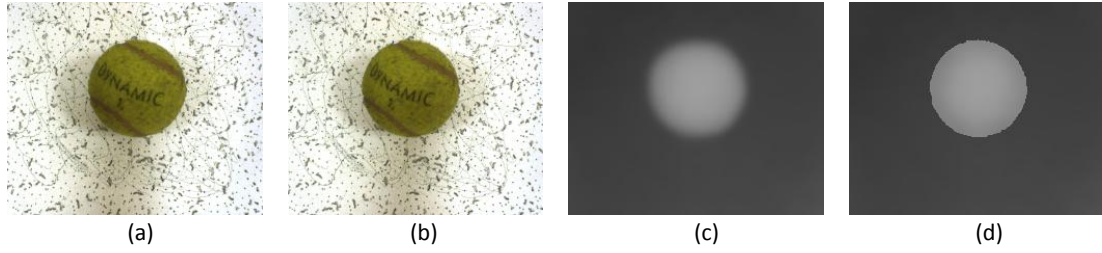
Şekil 3.2 Şekil 3.3’de gösterilen sahne için çalıştırılan eniyileme işleminin ara adımlarında üretilen fark haritaları. (a) Ağların başlangıç durumları. (b-e) eniyileme boyunca ağlar (D_1^t ve D_2^t) aynı sonucu bulmak için birbirine doğru hareketi. (f) Eniyileme sonucunda üretilen fark haritaları.

$$\begin{aligned} D_1^0 &= d_{min} \\ D_2^0 &= d_{max}. \end{aligned} \quad (3.8)$$

Örnek bir sahnenin (Şekil 3.3 a) eniyileme süreci Şekil 3.2’de görülmektedir. Görüldüğü gibi ağlardan biri minimum fark değerlerinden başlarken diğeri de maksimum fark değerlerinden başlamıştır (Şekil 3.2 a). Eniyileme işlemi devam ederken, ağlar çekim enerjisinin de etkisiyle birbirlerine doğru ilerlerler (Şekil 3.2 b-e). Eniyileme işlemi iki ağ da aynı fark haritasını bulana kadar devam eder (Figure 3.2 f). Sahnenin çıkarılan fark haritası Şekil 3.3.d’de görülmektedir. Şekilde görüldüğü gibi bir sonraki bölümde anlatılacak çifte ağ yapısının avantajlarından faydalanan imge bölütü tabanlı regülarizasyon sayesinde yüzey süreksizlikleri de korunmuştur.

3.3 İmge Bölütü Tabanlı Regülarizasyon

Regülarizasyon kötü konumlandırılmış ters problemlerin çözülebilmesi için problemin ilk formülüne bir dengeleme terimi eklenerek iyi konumlandırılmış hale getiril-



Şekil 3.3 Örnek bir sahnenin sol (a) ve sağ (b) kamera imgeleri. Tikhonov regülarizasyonu ile (c) [Aydın and Akgül, 2006] ve imge bölütü tabanlı regülarizasyon (d) [Aydın and Akgül, 2010b] ile çalıştırılan çiftte ağ yönteminin ürettiği fark haritaları.

mesidir. Dengeleme terimi çözüm hakkında bilinen öncül bilgileri kısıt (örneğin düzlük kısıtı) olarak problem formülüne eklemiş olur.

En iyi bilinen ve en sık kullanılan regülarizasyon yöntemi *Tikhonov* regülarizasyonudur [Tikhonov and Arsenin, 1977]. Bu yöntem, yönbağımsız yayınlımda olduğu gibi, çözümün gradyanını karesel formül ile $(|\nabla D|^2)$ cezalandırmaya dayanmaktadır. Bu yöntem ile yapılan eniyileme yerel minimum problemlerinden etkilenmez Ancak, fark haritalarında olabilecek süreksizlikler tamamen göz ardı edilmiştir. Doğal olarak bu yöntem ile çözülecek stereo problemiyle elde edilecek fark haritaları aşırı düzleştirilirler. Sonuçta sahnelerin yalnız kaba fark haritaları çıkarılabilir. Yüzey süreksizliklerinin de doğru olarak çıkarılabilmesi için stereo problemine uygun bir regülarizasyon yöntemine ihtiyaç vardır.

Yüzey süreksizliklerinin fark haritalarındaki yerleri önceden bilinmiş olsaydı, süreksizlikleri koruyan regülarizasyon yöntemi basit bir şekilde geliştirilebilirdi. Ancak pratikte böyle bir bilgi başlangıç verisi olarak bulunmamaktadır. Zaten bu bilgi bulmaya çalıştığımız çözümün bir parçasıdır. Bu konuda yapılması gereken yüzey süreksizlik bölgelerinin yerlerini makul bir kaynaktan tahmin etmek ya da çıkarmaktır.

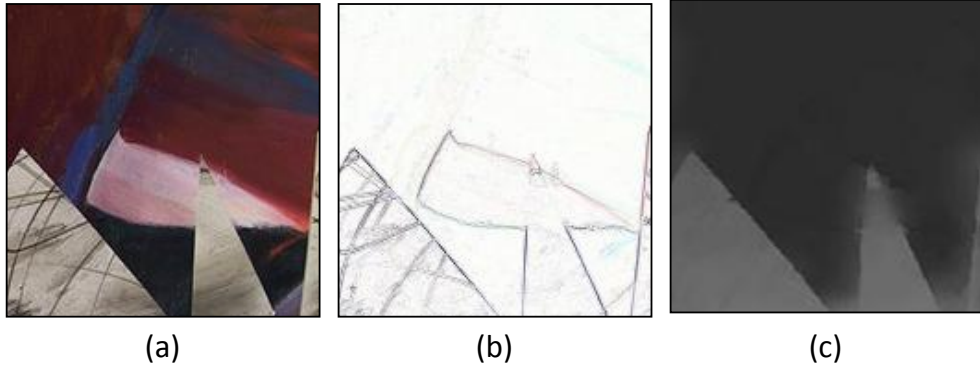
Yüzey süreksizliklerini koruyan regülarizasyon yöntemi geliştirmede verilmesi gereken en önemli karar süreksizliklerini tespit edecek fonksiyondur. Bu fonksiyon fark haritalarındaki yüzey süreksizliklerinin muhtemel yerlerini belirlemeye yarar. Süreksizliklerin yerleri belirlendikten sonra düzleştirmeyi sadece sürekli bölgelerde yapmak regülarizasyon işlemi için yeterli olabilmektedir.

Regülarizasyon işlemleri süreksizlikleri tahmin ettikleri kaynaklara göre fark haritası tabanlı ve imge tabanlı olmak üzere iki gruba ayrılabilirler. Fark haritası ta-

banlı (ya da çözüm tabanlı) regülarizasyon yöntemleri yüzey süreksizlik bölgelerini eniyileme işlemi sırasında bulunan ara çözümlerden çıkarırlar [Robert and Deriche, 1996]. Bu yaklaşımın başarılı olabilmesi için başlangıç fark haritasının aranan çözüme oldukça yakın olması gerekmektedir. Ayrıca ara adımlarda ortaya çıkan süreksizliklerin yüzey süreksizliklerinden mi yoksa gürültü, kapanma, tekrar eden desen gibi sorunlardan dolayı mı ortaya çıktığını tespit etmek oldukça zordur.

Diğer taraftan imge tabanlı regülarizasyon yöntemleri yüzey süreksizliklerini sahnenin imgelerinden tahmin ederler [Alvarez et al., 2002; Kim et al., 2004]. Bu yöntemlerin temel varsayımı yüzey süreksizliklerinin olduğu imge bölgelerinde parlaklık süreksizliklerinin de olacağıdır. Regülarizasyon işlemi imge gradyan büyüklüklerine bakarak türdeş imge bölgelerinde yönbağımsız düzleştirme, ayrışık imge bölgelerinde ise yönbağımlı düzleştirme uygulanır. Bu yüzden imge ayrıt bölgelerinde fark haritası tabanlı regülarizasyon yöntemleri daha yüksek başarımlar sağlarlar. Fakat, imge tabanlı yüzey süreksizlik tahmini her durumda tutarlı sonuçlar üretemez. Çünkü sahnede farklı derinlikte bulunan yüzeylerin sınırlarını da gösteren yüzey süreksizliklerini imgelerden güvenilir bir şekilde çıkarmak mümkün değildir. Farklı derinlikteki sahne nesnelerinin yüzeylerini ayırması beklenen imge ayrıtları nesnenin yüzey sınırı boyunca sürekli olmayabilir. Bu durum sınır sızıntısı olarak bilinen farklı derinlikteki yüzeylerin sınırlarında karşılıklı yayınma yol açar (bkz. 3.4 b ve c). Benzer şekilde, aynı derinlikte bulunan yüzeyler arasında imge gradyanına göre yayınım engellenebilir. Özellikle aşırı desenli veya gürültülü imgeler için bu durum geçerlidir. Son olarak, düzleştirme veya yayınım derecesi imge gradyanına bağlı olduğu için aynı büyüklükteki yüzey süreksizlikleri farklı derecelerde düzleştirmeye maruz kalır. Gerçekte imge gradyanlarının büyüklüğü yüzey süreksizliklerinin büyüklüğüyle ilintili değildir.

İmgelerin gradyanına bağlı regülarizasyon yöntemlerindeki problemler imge bölütleri kullanılarak çözülebilir. İmge bölütlemesi türdeş ve ayrışık imge bölgerindeki bilgilerden yararlanılarak gerçekleştirilir. Bu yüzden imge bölütleri sahnedeki nesneler hakkında imge gradyanlarına nazaran daha genel bilgiler içerir. Ayrıca, imge bölütleriyle yapılacak regülarizasyon ile piksel seviyesinde yapılan regülarizasyon işlemi bölüt seviyesine taşınmış olur. Sonuç olarak, regülarizasyon işlemi sahne hakkında daha genel bilgiye sahip olacağından yüzey süreksizlikleri daha tutarlı bir model ile ele alın-



Şekil 3.4 Sınır sızıntısını gösteren bir örnek. (a) Middlebury imge veritabanında bulunan Sawtooth imgesinden kırılmış bir bölüm, (b) imge gradyan büyüklükleri, (c) imge tabanlı regülarizasyon ile elde edilen fark haritası [Kim et al., 2004]. Fark haritasında da görüldüğü gibi nesne sınırının belirgin olmadığı yüzey süreksizlikleri aşırı regülarizasyona maruz kalmıştır.

miş olur.

İmge bölütü tabanlı regülarizasyon yöntemleri yüzey süreksizliklerinin bölüt sınırlarında olacağını ve bölüt içlerinde yüzey süreksizliklerinin olmayacağını kabul ederler. Bu sebeple aynı bölüt içinde kalan pikseller arasındaki yayınıma herhangi bir sınırlandırma getirmezler. Bu yaklaşımda dikkat edilmesi gereken asıl konu bölütler arası yayınının ayarlanmasıdır. Bölütler arası yayınının tamamen engellenmesi her bölütün birbirinden bağımsız eniyilendiği bir sistem ortaya çıkarır ki böyle bir sistemde küçük boyuttaki ve kapanma bölgelerindeki bölütlerin yerel minimumlara takılması kaçınılmaz olur. Bölütler arası yayınıma tamamen izin verildiği durumlarda ise regülarizasyon birebir yönbağımsız yayınıma dönmüş olur. Burumda da yerel minimuma takılmayan fakat yüzey süreksizliklerinin aşırı düzeltilildiği bir regülarizasyon yapılmış olur. Sonuç olarak, bahsi geçen problemlerden etkilenmeden bölütler arası yayının miktarını belirleyecek bir yöntem ihtiyacı vardır.

Önerilen imge tabanlı regülarizasyonda, çifte ağ yöntemin avantajlarından faydalanılarak, bölütler arası yayının ağların arasındaki uzaklığa göre ayarlanmıştır. Böylece, hem yönbağımlı hem de yönbağımsız yayınının avantajlarından faydalanan bir regülarizasyon işlemi gerçekleştirmek mümkün olmuştur. Bu operatör çifte ağın geçici durum bilgilerini kullanmaktadır. İşlemler arasındaki geçici uzaklıklar aşağıda verilmiştir: Eniyileme süresince değişen ağlar arası uzaklık t anında şu şekilde hesaplanır,

$$\Delta d^t(x, y) = D_1^t(x, y) - D_2^t(x, y). \quad (3.9)$$

Önceki bölümde de anlatıldığı gibi ağlardan biri en küçük fark değerleriyle başlatılırken diğeri de en büyük fark değerleriyle başlatılır. Dolayısı ile Δd başlangıçta en büyük değerini alır. Bu anda eniyilemenin yerel minimumlardan etkilenmemesi için yayınının yönbağımsız olması gerekir. Bu yaklaşım yüzey süreksizliklerine izin vermesi de, eniyilemenin ara adımlarda sahnenin kaba fark haritasını çıkarmasına yardımcı olur. Eniyileme süreci devam ettikçe ağlar birbirine doğru yaklaştıklarından Δd değerleri gittikçe küçülmeye başlar ve fark haritalarında sahne yapısı belirmeye başlar. Yüzey süreksizliklerinin de doğru olarak çıkarılması için ağlar birbirine yaklaştıkça regülarizasyonun yavaş yavaş yönbağımlı davranması gerekir. İki ağın da aynı fark değerini bulduğu yerlerde regülarizasyon tamamen yönbağımlı olmalı ve bölütler arası yayınıma izin verilmemelidir.

Bu prensipler ışığında ağlar arası farkı gösteren Δd denklem 3.4 ve 3.5'teki yayılım fonksiyonuna (ψ') değişken olarak eklenmiştir. Tanımlanan yayılım fonksiyonu bölüt içi ve bölütler arası yayılım miktarını ağlar arasındaki uzaklığa göre ayarlar. Yayılım fonksiyonu ψ' şu şekilde tanımlanmıştır.

$$\psi' = g(\nabla I_b(x, y), \Delta d(x, y)). \quad (3.10)$$

burada g süreksizlik tespit fonksiyonu olarak bilinir ve şu şekilde tanımlanmıştır.

$$g(\nabla I_b, \Delta d) = 1 - \|\nabla I_b\| e^{-(\Delta d/\kappa_\psi)^2}, \quad (3.11)$$

burada I_b bölütlenmiş imgeyi κ_ψ ise yayınının yönbağımlılığını veya yönbağımsızlığını ayarlayan sabit bir katsayıdır. Bölütlenmiş imgenin gradyan büyüklüğü ($\|\nabla I_b\|$) şu şekilde tanımlanmıştır.

$$\|\nabla I_b(x, y)\| = \begin{cases} 1 & , \quad (x, y) \text{ bölüt sınırında ise} \\ 0 & , \quad \text{diğer} \end{cases} \quad (3.12)$$

Yayılım fonksiyonun açıklaması şu şekilde yapılabilir. Eğer Δd değeri büyükse yayılım fonksiyonun değeri 1 olarak hesaplanır. Bu durumda düzleştirme işlemi

Tikhonov regülarizasyonunda olduğu gibi yönbağımsız olarak çalışır ve eniyileme yerel minimumlardan etkilenmeden devam eder. Eniyileme devam ederken Δd gittikçe küçülür. Buna bağlı olarak, yayılım fonksiyonu yönbağımlı davranmaya başlar. Başka bir ifadeyle, bölütler arası yayılım azalmaya başlar. Nihayetinde, ağlar aynı fark haritasını bulduklarında ($\Delta d = 0$) bölüt sınırlarında yayılım fonksiyonunun değeri 0 olarak hesaplanır. Yani bölütler arası yayılım tamamen engellenir. Bu aşamada düzeltme işlemi, süreksizlikleri koruyan regülarizasyonda olması gerektiği gibidir; düzeltme sadece türdeş bölgeler içinde yapılır. Bu yaklaşımın getirdiği bir ek avantaj da şudur ki en iyi sonuç elde edildikten sonra bile devam edilen gereksiz eniyileme aşırı düzeltmeye sebep olamaz. Bu özellik diğer imge tabanlı regülarizasyon yöntemlerinde bulunmamaktadır çünkü küçük de olsa, bu tip regülarizasyon yöntemlerinde ayrışık bölgeler arasında yayıma izin verilir.

3.4 Deneyler

Önerilen yöntemin başarımı gerçek fark haritaları bilinen sahnelerin olduğu Middlebury imge setleri [Scharstein and Szeliski, 2003b] kullanılarak ölçülmüştür. Yayılım fonksiyonunda kullanılan imge bölütleri sahnenin sol stereo imgesine ortalama kayması [Comaniciu and Meer, 2002] yöntemi uygulanarak elde edilmiştir.

İmge pikselleri arasındaki eşleşme derecesi için Normalize edilmiş Çapraz İlinti fonksiyonu kullanılmıştır. Yöntemin performansını arttırmak ve yakınsama süresini kısaltmak için eşleşme değerlerine ön düzeltme işlemi uygulanmıştır. En basit ön düzeltme işlemi eşleşme ölçümü için kullanılan pencere boyutunu arttırmaktır. Ancak, pencere boyutunu arttırmak yüzey süreksizliklerinin konumlarında kaymaya sebep olacaktır [Scharstein and Szeliski, 2002]. Yüzey süreksizliklerinin konumlarını değiştirmeden ön düzeltme yapmak için stereo imgelerinden oluşturulan çift yönlü süzgeçlerden [Tomasi and Manduchi, 1998] yararlanılmıştır. Bu yöntem sol ve sağ stereo resimlerinden üretilen süzgeçlerin çarpımıyla yapılan ön düzeltmeye benzerdir [Yoon and Kweon, 2006].

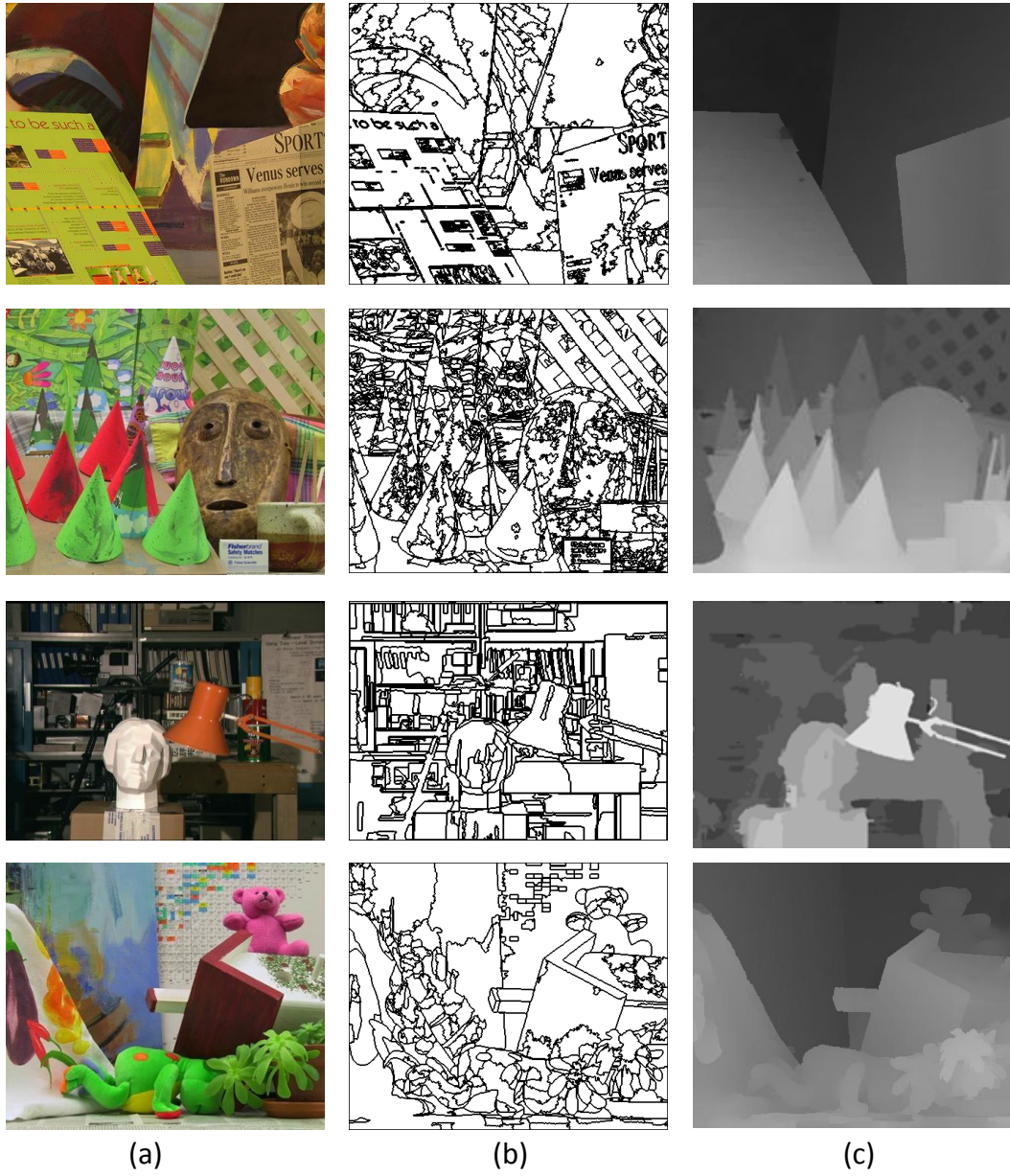
Deneyler yüksek oranda yüzey süreksizlikleri içeren Venus, Teddy, Tsukuba ve Cones sahnelerinin imge setleri kullanılarak gerçekleştirilmiştir. Önerilen yöntemin

parametreleri tüm setler için sabit olarak alınmıştır ($\alpha = 0.1$, $\beta = 0.1$, $\lambda = 0.15$, $\gamma = 0.1$, and $\kappa_\varphi = 0.01$). Yayınım fonksiyonunun katsayısı sahnelerde olabilecek maksimum ve minimum fark değerleriyle ilişkilendirilmiş ve $\kappa_\psi = d_{max} - d_{min}$ olacak şekilde alınmıştır. Önerilen yöntemin hata oranları sahnenin tüm yüzeyi, sadece görünen yüzeyi ve sadece yüzey süreksizliklerin olduğu bölgelerde ayrı ayrı hesaplanmıştır. Çizelge 3.1 önerilen yöntem ve gelişmiş stereo yöntemleri için hesaplanan hata oranlarını göstermektedir. Bu çizelgedeki hata oranları hatalı eşik değeri 1 piksel olacak şekilde hesaplanmıştır. Şekil 3.5 deneylerde kullanılan sahnelerin sol stereo imgesini, bu imgeler için çıkarılan bölütlerin çeperlerini ve elde edilen fark haritalarını göstermektedir. Görsel ve sayısal değerlerden de anlaşıldığı gibi önerilen yöntemin yüzey süreksizliklerini bulmadaki başarımı yüksektir.

Önerilen yöntem sürekli fark değerleri ürettiği için, hata oranları 0.5 piksel hata eşiği için de hesaplanmıştır. Çizelge 3.2 hesaplanan bu hata oranlarını göstermektedir. Çizelgede görüldüğü gibi tüm sahneler için en iyi sonucu üreten bir yöntem bulunmamaktadır. Hata eşiği düşürüldükçe yöntemlerin başarımları da değişmektedir. Önerilen yöntem diğer gelişmiş stereo yöntemleri ile karşılaştırılabilir sonuçlar üretmektedir. Ayrıca, karşılaştırılan diğer yöntemlerin düzlem uydurma, birden çok imge bölütü kullanma ya da elde edilen fark haritasını işleme gibi yöntemlere başvurduğu göz önüne alınırsa, önerilen yöntemin başarımı daha iyi anlaşılır.

Çizelge 3.1 Middlebury stereo sahneleri kullanılarak 1 piksellik hata eşiği ile hesaplanan hata oranları. Hata oranları sahnelerin görünen yüzeyi (gör), tüm yüzeyi (tüm) ve süresiz yüzeylerde (srksz) ayrı ayrı hesaplanmıştır. Çizelgedeki bazı değerler ilgili makalelerde de bulunmadığından yazılmamıştır.

	Tsukaba			Venus			Teddy			Cones		
	gör	tüm	srksz	gör	tüm	srksz	gör	tüm	srksz	gör	tüm	srksz
Yöntem												
Woodford et al. [2008]	2.91	3.56	0.24	0.24	0.49	2.76	10.9	15.4	20.6	5.42	10.8	12.5
Yang et al. [2007]	1.24	1.76	5.98	0.12	0.46	1.74	3.45	8.38	10.0	2.93	8.73	7.91
Yoon and Kweon [2006]	1.38	1.85	6.90	0.71	1.19	6.13	7.88	13.3	18.6	3.97	9.79	8.26
Ben-Ari and Sochen [2010]	3.97	5.23	14.9	0.28	0.76	3.78	9.34	14.3	20.0	4.14	9.91	11.4
Pock et al. [2008]	3.61	5.72	18.0	1.16	2.50	12.4	6.10	15.7	16.8	3.88	14.4	11.5
Pock et al. [2007]	-	-	-	-	1.10	-	-	6.63	-	-	3.67	-
Önerilen Yöntem	2.51	2.71	7.13	0.32	0.40	3.45	8.08	12.3	21.2	4.08	11.5	13.7



Şekil 3.5 (a) Middlebury sitesinden alınan imge setleri (Venus, Cones, Tsukaba, ve Teddy seti). (b) İmge bölütlerinin sınırlar. (c) Segment tabanlı çiftte ağ yöntemiyle elde edilen fark haritaları.

Çizelge 3.2 0.5 piksel hata eşiği ile hesaplanan hata oranları

Yöntem	Tsukaba			Venus			Teddy			Cones		
	gör	tüm	srksz	gör	tüm	srksz	gör	tüm	srksz	gör	tüm	srksz
Woodford et al. [2008]	7.10	7.70	0.56	0.56	0.83	4.21	17.5	22.7	31.5	11.6	17.1	20.4
Yang et al. [2007]	8.78	9.45	14.9	0.72	1.12	5.24	10.1	16.4	21.3	8.49	14.7	16.5
Yoon and Kweon [2006]	18.1	18.8	18.6	7.77	8.40	15.83	17.6	23.9	34.0	14.0	19.7	20.6
Ben-Ari and Sochen [2010]	7.18	8.56	20.1	1.46	2.12	7.87	12.9	19.4	27.5	6.22	12.6	15.8
Pock et al. [2008]	11.1	13.3	27.2	5.99	7.40	22.3	10.5	19.9	25.8	5.99	16.5	16.7
Pock et al. [2007]	-	-	-	-	3.45	-	-	11.2	-	-	7.52	-
Önerilen Yöntem	8.53	9.29	16.98	1.23	1.52	9.55	10.2	16.1	25.4	7.34	14.2	14.4

Çizelge 3.3 Farklı yineleme değerleriyle elde edilen fark haritalarının hata oranları. Hata oranları 1 ve 0.5 piksellik eşik değerleri için hesaplanmıştır.

Yineleme sayısı	hata eşiği 1 piksel			hata eşiği 0.5 piksel		
	gör	tüm	srksz	gör	tüm	srksz
1000	27.6	28.3	35.4	31.0	31.9	47.0
1400	12.4	12.9	26.0	16.6	17.2	33.2
1800	1.91	2.29	12.0	3.17	3.77	19.6
2000	0.95	1.35	9.38	1.61	2.04	10.9
2500	0.32	0.40	3.45	1.23	1.52	9.55
5000	0.32	0.40	3.45	1.23	1.52	9.55

Son deney olarak yöntemin farklı yineleme sayısına göre ürettiği fark haritaları çıkarılmıştır. Bu sayede yöntemin yineleme sayısına bağımlılığı görülmüştür. Çizelge 3.3 yineleme sayısına göre hesaplanan hata oranlarını göstermektedir. Çizelgede görüldüğü hata oranları, en iyi sonuca ulaştıktan sonra yapılan gereksiz yinelemelere rağmen değişmemektedir.

Önerilen yöntem çoğunlukla küçük boyutlu bölütlerde ve bölütlemesi zor olan imge bölgelerinde başarısız olmuştur. Ayrıca yöntemin başarımı imge bölütlemenin başarısına bağlıdır. Bölütlerin içerisinde kalan yüzey süreksizlikleri üretilen fark haritalarında aşırı düleştirmeye maruz kalacaklardır. Bu problem farklı bölütleme yöntemlerini aynı anda kullanarak aşılabılır [Woodford et al., 2008]. Bir diğer başarısız olunan konu da yetersiz desen bulunan bölgelerdir. Regülerizasyon bu bölgelerin büyük bir kısmının üstesinden gelse de tamamında başarılı olamamaktadır. Son olarak, yöntem kapanma problemine karşı herhangi bir çözüm üretmediğinden bu bölgelerin hata oranları yüksek çıkmaktadır.

4 Odaktan Şekil İyileştirmeleri

Sahnelerin kamera düzlemine olan uzaklıklarını hesaplamada kullanılabilecek bir diğer ipucu da sahnelerin açık diyaframlı merceklerle çekilmiş 2B imgelerinde bulunan göreceli netlik veya bulanıklık miktarıdır. Kameralarda kullanılan merceklerin alan derinliği kısıtlı olduğu için, bu tip kameralarla elde edilen imgeler, odaklanılmış bölgeler hariç, uzaklığa bağlı bulanıklık içerir. İmgelerdeki bu özellik, çeşitli yöntemler tarafından sahne derinliğini hesaplamada etkin bir şekilde kullanılmıştır. Bu yöntemler Odaktan Şekil [Nayar and Nakagawa, 1994; Nair and Stewart, 1992; Krotkov, 1987; Jarvis, 1983; Helmlı and Scherer, 2001; V. Michael Bove, 1993; Bhasin and Chaudhuri, 2001; Aydın and Akgul, 2008] ve Bulanıktan Şekil [Krotkov, 1987; M.Subbarao and A.Nikzad, 1993; Pentland et al., 1994; Subbarao and Surya, 1994; Nayar et al., 1996; Rajagopalan and Chaudhuri, 1997; Favaro and Soatto, 2007; Favaro et al., 2008; Subbarao et al., 1992] olarak adlandırılır.

4.1 Görüntü Oluşum Modeli

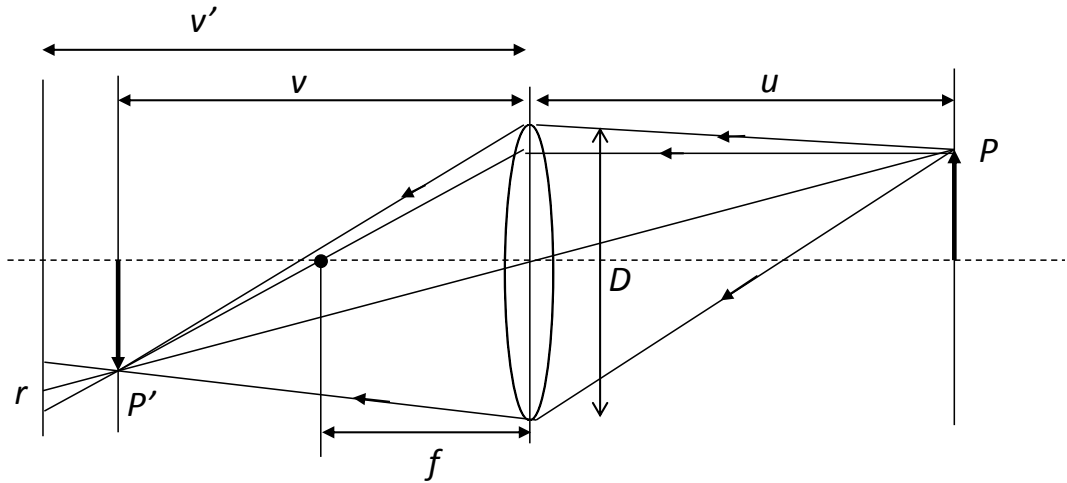
Bu bölümde ideal mercekler kullanılarak elde edilen imgelerin oluşum modelini açıklayacağız. Şekil 4.1 de en basit kamera modeli görülmektedir. Sistem birbirlerine paralel olarak yerleştirilmiş bir mercek ve görüntünün oluştuğu düzlemde oluşmaktadır. Görüntü düzlemi üzerlerine düşen ışık miktarını ölçebilen alıcılardan oluşmaktadır. Sistemde kullanılan merceğin görevi üzerine gelen ışığı belirli bir doğrultuda yönlendirmektir. Herkesçe bilinen ideal mercek kuralına göre mercekten u kadar uzakta bulunan bir nokta kaynaktan çıkan bütün ışınlar mercekten v kadar uzaktaki görüntü düzlemi üzerine toplanması için şu şartı sağlaması gerekmektedir [Born and Wolf, 1965],

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f} \quad (4.1)$$

Burada f merceğin odak uzaklığı olarak bilinir. İdeal mercekler için tam olarak odakta olan bölge görüntü düzlemine paralel bir düzlemdir. Bir başka deyişle odaklanılmış

düzlem üzerindeki herhangi bir sahne noktasının görüntü düzleminde oluşacak görüntüsü yine bir noktadır. Bu noktaların bulunduğu bölgeler için odaklanılmış tabiri kullanılır ve bu noktalara karşılık gelen görüntü bölgeleri de sahne noktalarının en net görüntüsünü içerir. Ayrıca, Şekil 4.1'te de görüleceği gibi görüntü düzlemi v' konumuna kaydırıldığında değişen odaklanılmış düzleminde olmayan her hangi bir noktadan çıkan ışınlar görüntü düzleminde bir daire içine düşer. Bu daireye bulanıklık dairesi denir. Bulanıklık dairesinin yarıçapı r basit geometrik işlemler kullanılarak kolayca çıkarılabilir. Bu dairenin yarıçapı u, v, f ve D diyafram açıklık miktarına bağlıdır ve şu bağıntı ile bulunur;

$$r = \frac{(v' - v)D}{v} \quad (4.2)$$



Şekil 4.1 Temel imge oluşum geometrisi.

4.2 Odaktan Şekil Elde Etme Yöntemleri

Odaktan Şekil, aşamalı olarak değiştirilen odak ayarlarıyla elde edilen N farklı sahne imgesi kullanılarak gerçekleştirilen bir derinlik hesaplama yöntemidir. Standart Odaktan Şekilde elimizde her bir sahne noktasının N farklı imgesi mevcuttur. Sahnedeki herhangi bir noktanın elimizdeki N farklı imgede oluşan görüntülerinin her imgede aynı piksel pozisyonuna denk geldiğini kabul edersek, Odaktan Şekil yöntemlerindeki asıl amaç sahne noktalarının tam olarak odaklanılmış imgelerini, dolayısı

ile bu imgelerin bilinen odak ayarlarını bulmaktır. Denklem 4.1'den de anlaşılacağı üzere odak ayarları bilinen odaklanılmış imge noktasına karşılık gelen sahne noktasının uzaklığı hesaplanabilir.

Sahne noktalarının en iyi odaklanılmış resmini bulabilmek için imge piksellerinin odaklanma miktarını ölçen fonksiyonlar kullanılmaktadır. Bu fonksiyonların genel adı Odak Ölçüt Operatörüdür. Odaklanma dereceleri hesaplandıktan sonra, her bir piksel pozisyonu için uzaklık hesaplanmasında kullanılacak odak ayarı f şu şekilde seçilir,

$$f_i(x, y) = \arg \max_i (OF_i(x, y)), \quad (4.3)$$

burada $i = 1..N$ ve N kullanılan imge sayısını gösterir. $OF_i(x, y)$ ise i . resmin (x, y) konumundaki pikseli için hesaplanan Odak Ölçüt değeridir. Her piksel için bulunan f_i değerleri Denklem 4.1'de yerine konularak sahne noktalarının uzaklık değerleri hesaplanır.

Odaktan Şekil ile bulunan sonuçları iyileştirmek için Subbarao and Choi [1995] üç boyutlu görüntü uzayındaki nesnenin kısmi düzlemsel yaklaşımı olan Odaklanılmış Görüntü Yüzeyi (Focused Image Surface) fikrini önerdiler. Odaktan Şekil için elde edilen imgelerin arka arkaya dizilmesiyle oluşturulan küçük kübik hacimde olası düzlemsel yüzey parçaları göz önünde bulundurularak en yüksek Odak Ölçüt değerini üretecek yüzey aranır. Olası tüm yüzeyler için arama yapmak çok yüksek hesaplama maliyeti gerektirdiğinden, arama yalnızca standart Odaktan Şekil ile bulunan ilk sonucun etrafındaki yüzeyler için yapılır. Hesaplama Maliyetini düşürmek için Ahmad and Choi [2005] dinamik programlamayı kullanmışlardır. Asif and Choi [2001] yapay sinir ağları kullanarak odak ölçütünü en iyileştirecek Odaklanılmış Görüntü Yüzeyini (OGY) öğrenmeye çalışmışlardır. Bu çalışmada yapay sinir ağlarındaki ağırlıklar odak ölçütünü arttıracak şekilde geri bildirim mekanizmasıyla güncellenir.

OGY düzlemsel olmayan daha karmaşık geometrilere sahip nesneler için iyi sonuçlar vermediği için Yun and Choi [1999] kıvrık arama pencerelerini kullanan Eğri Pencere Odak Ölçütünü önerdiler. Bunun için Dokuz Kontrol Noktası adını verdikleri dokuz nokta ile Lagrange polinomları ile yüzey denklemlerini çıkarmaya çalıştılar. Bunların dışında, bazı araştırmacılar OGY'yi birleşimsel eniyileme Shim and Choi

[2010] ve Beizer yüzeyi yaklaşımıyla [Muhammad and Choi, 2010] bulmayı önermişlerdir.

Sahnelerdeki cisimlerin yüzeylerini OGY'deki gibi düzlemsel parçalarla veya eğri yüzey parçalarıyla modellenmesi yüzey süreksizliklerini göz ardı ettiği için bu bölgeler için hatalı sonuçlar hesaplanır. Bölüm 4.4'te bu probleme çözüm olarak önerdiğimiz yüzey süreksizliklerini de göz önünde bulunduran uyarlanabilir pencereler modelini açıklanmaktadır.

4.3 Odak Ölçüt Operatörleri

Nayar and Nakagawa [1994]'nın görüntü oluşumunun geometrik analizine göre bulanıklaşma bir alçak geçirgen filtreleme işlemidir. Bir başka deyişle bulanıklaşma derecesi görüntülerdeki yüksek frekanslı içerik miktarıyla ters orantılıdır. Dolayısı ile görüntülerde odaklanılmış bölgeleri bulmak için kullanılacak Odak Ölçütleri yüksek frekanslı imge içeriğine duyarlı olması ve tam odaklanılmış bölgeler için en yüksek değeri üretmesi gerekmektedir. Bu yüzden çoğu Odak Ölçüt Operatörlerinde görüntülerdeki yüksek frekanslı içerik miktarını ölçebilen fonksiyonlardan yararlanılır.

İdeal Odak Ölçüt operatörlerinin şu özelliklere sahip olması gerekir;

1. İmge parlaklık değerlerindeki yüksek frekanslı değişkenliklere karşı yüksek değerler üretmeli. Belirli bir sahne noktasının tam odaklanılmış görüntüsü için en büyük değeri üretmesi gerekir.
2. İmge içeriğinden bağımsız olmalı.
3. Tek modlu olmalı. Yani operatör sadece bir maksimum değere sahip olmalı ki bu da tam odaklanılmış duruma karşılık gelmelidir.
4. Tekdüze bir fonksiyon olmalı. Fonksiyonun maksimum noktasının her iki tarafı da monotonik olmalı. Yani bulanıklık miktarı arttıkça odak ölçütünün azalması gerekir.

Literatürde odak ölçüt operatörü olarak uzaysal tabanda veya frekans tabanında tanımlı çeşitli fonksiyonlar kullanılmıştır. Uzaysal tabanda çalışan operatörler genellikle yüksek geçirgen filtrelerden yararlanır. Frekans tabanlı operatörler ise genellikle frekans dağılımındaki toplam enerjiyi ölçerler ya da yüksek frekanslı bileşenlerin düşük frekanslı bileşenlere oranını hesaplar.

Bu bölümde literatürde odak ölçütü olarak kullanılan operatörleri açıklayacağız. Verilen I imgesinin x_0, y_0 konumundaki pikselinin değeri $I(x_0, y_0)$ ve bu piksel için hesaplanacak odak ölçüt değeri OF olsun. $\Omega(x_0, y_0)$ da x_0, y_0 konumundaki pikselin komşuluğunda olan pikseller kümesi olsun.

Varyans

En basit odak ölçüt operatörü olarak imgenin belirli bir bölge içindeki varyasını kullanılabilir [M.Subbarao and A.Nikzad, 1993]. $N \times N$ boyutunda bir imge bloğu için varyans şu şekilde hesaplanır

$$OF_{VAR} = \frac{1}{N^2} \sum_{x,y \in \Omega} (I(x, y) - \bar{I})^2 \quad (4.4)$$

burada $\bar{I}$ seçilen bölge içerisinde kalan piksel değerlerinin ortalamasıdır ve şu formülle hesaplanır,

$$\bar{I} = \sum_{x,y \in \Omega} I(x, y) / N \quad (4.5)$$

Ortalama Oran

Odak ölçütü hesaplanacak pikselin komşuluğundaki piksellerin ortalama değerine olan oranı da odak ölçüt değeri olarak kullanılabilir [Helmlı and Scherer, 2001]. Odak ölçüt değerinin 1 olarak hesaplanması bu bölgedeki piksellerin aynı değeri taşıdığını gösterir. Eğer sahnede yeterince desen mevcut ise tam odaklanılmış pozisyonlara denk gelen imge bölgeleri için yüksek değerler üretilecektir.

$$OF_{ORT} = \frac{I(x, y)}{\bar{I}(x, y)} \quad (4.6)$$

Diğer tüm odak ölçütü hesaplama yöntemlerinde olduğu gibi hesaplanan değerler yerel bir pencere etrafında toplanır.

Kıvrım Odak Ölçütü

İmgelerdeki piksel değerlerinin 3B uzayda bir yüzey oluşturduğunu farz edersek, nesnelerin tam olarak odaklanılmış imgelerine denk gelen yüzey bölgelerinde çokça kıvrım olacağını söyleyebiliriz. Helmlı and Scherer [2001] yüzey parçacıklarını ikinci dereceden polinomlara benzeterek odaklanma derecesini ölçmeyi önermişlerdir. Bu yöntemde ilk olarak yüzeylerin $S(x, y) = ax + by + cx^2 + dy^2$ denkleminde uyacağı varsayılarak en küçük kareler yöntemiyle katsayılar hesaplanır. Daha sonra katsayıların mutlak değerleri toplanarak odak ölçütü hesaplanır.

$$OF_{KVR} = |a| + |b| + |c| + |d| \quad (4.7)$$

Gradyan Enerjisi

Bu odak ölçütünde imgelerin her iki yöndeki gradyanlarının karelerinin toplamının kareleri toplanır M.Subbarao and A.Nikzad [1993].

$$OF_{GR} = \sum_{x,y \in \Omega} \left(\left(\frac{\partial I(x, y)}{\partial x} \right)^2 + \left(\frac{\partial I(x, y)}{\partial y} \right)^2 \right)^2 \quad (4.8)$$

Tenengrad

Tenengrad odak ölçütü de Gradyan enerjisine benzer bir yüksek geçirgen filterinin imgeye uygulanması ile hesaplanır Tenenbaum [1971].

$$OF_{GR} = \sum_{x,y \in \Omega} I_x(x, y)^2 + I_y(x, y)^2 \quad (4.9)$$

burada $I_x(x, y)$ ve $I_y(x, y)$ imgeye Sobel operatorünün x ve y bileşenlerinin uygulanması ile elde edilir.

Laplace

Yüksek geçişli filtre yapmanın bir diğer yolu imgelerin ikinci türevini almaktır. İki boyutlu imgeler için Laplace operatörü yaygın olarak kullanılır Krotkov [1987].

$$OF_{LE} = \sum_{x,y \in \Omega} \frac{\partial^2 I(x,y)}{\partial x^2} + \frac{\partial^2 I(x,y)}{\partial y^2} \quad (4.10)$$

Laplace Enerjisi

Bu odak ölçütünde imgelerin ikinci türevlerinin kareleri toplanır.

$$OF_{LE} = \sum_{x,y \in \Omega} \left(\frac{\partial^2 I(x,y)}{\partial x^2} + \frac{\partial^2 I(x,y)}{\partial y^2} \right)^2 \quad (4.11)$$

Değiştirilmiş Laplace Toplamı

Laplace operatörünün yatay ve dikey yöndeki bileşenleri birbirlerini yok edebilirler. Bu yüzden Nayar and Nakagawa [1994] Değiştirilmiş Laplace Toplamı, klasik Laplacian'dan farklı olarak kısmi ikinci türevlerin kendileri toplanmak yerine mutlak değerlerini alarak toplar:

$$OF_{DLT} = \sum_{x,y \in \Omega} \nabla^2 I(x,y), \quad \text{ve } \nabla^2 I(x,y) > T, \quad (4.12)$$

burada

$$\nabla^2 I(x,y) = \left| \frac{\partial^2 I(x,y)}{\partial x^2} \right| + \left| \frac{\partial^2 I(x,y)}{\partial y^2} \right| \quad (4.13)$$

olarak hesaplanır.

İmge Momenti

[Zhang et al., 2000; Xiong and Shafer, 1997, 1995] yüksek geçişli filtre kullanmayan bir odak ölçüt operatörü geliştirmişlerdir. Geliştirilen bu odak ölçüt operatörü Flusser and Suk [1998]'un imge momenti konusunda ulaştığı teorik sonuçları kullanır. Flusser ve Suk çift dereceli momentlerin bulanıklaşmaya duyarlı olduğunu ispatlamıştır. Bu bulgudan yola çıkarak Zhang et al. [2000] ikinci ve dördüncü derece momentleri kullanarak odak ölçüt operatörü geliştirmiş ve bu operatörün monoton olduğunu ispatlamıştır. Moment tabanlı odak ölçütlerinin dezavantajı ise imge sınırlarına duyarlı

olmasıdır. Bu yüzden bu operatör sadece ön plandaki nesneler için etkili sonuçlar üretir.

SUSAN Operatörü

Bu odak ölçütünde SUSAN [Smith and Brady, 1997] operatörü kullanılarak çıkarılan imge özellik haritalarından yararlanılır. Piksellerin odaklanma miktarı, çıkarılan özellik noktalarına olan uzaklıklarına göre üstel olarak azalan fonksiyon tarafından belirlenir. İmgelerdeki herhangi bir x, y piksel pozisyonu için odak ölçütü şu şekilde hesaplanır [Mendapara et al., 2009]:

$$OF_{SUSAN} = e^{-(d(x,y)/2)} I(x, y) + (1 - e^{-(d(x,y)/2)}) \quad (4.14)$$

burada $d(x, y)$ odak ölçütü hesaplanacak pikselin en yakın özellik noktasına olan uzaklığıdır ve SUSAN operatörü uygulanarak elde edilen özellik haritasına uzaklık dönüşümü uygulanarak hesaplanır.

Ayrık Dalgacık Dönüşümü

Bu odak ölçütünde imgelere ayrık dalgacık dönüşümü uygulanarak elde edilen yüksek frekans katsayılarının düşük frekans katsayılarına oranı kullanılmıştır. Yöntemin normal imgelerde performansı iyi olmasına rağmen imgelerdeki gürültü oranı arttıkça performansı düşer.

Verilen bir imgenin dalgacık dönüşümü ile elde edilen yüksek frekanslı katsayıları $h_W(f)$ ve düşük frekanslı katsayıları $l_W(f)$ olsun. Bu durumda odak ölçütü şu şekilde hesaplanır [Kautsky et al., 2002; Huang et al., 2005]:

$$OF_{WL} = \frac{\|h_W(f)\|}{\|l_W(f)\|} \quad (4.15)$$

Benzer bir şekilde, Yang and Nelson [2003]'ün geliştirdiği dalgacık dönüşümü tabanlı yöntemde *Daubechies D6* dalgacık filtresi kullanılarak imgeler (W) yüksek geçişli (H) ve alçak geçişli (L) filtrelerden geçirilir. Bu işlem sonucu ortaya çıkan yeni imgeler W_{LL1} , W_{LH1} , W_{HL1} , ve W_{HH1} olacak şekilde dört alt parçaya bölünür. Bu parçalar kullanılarak odak ölçütü aşağıdaki gibi hesaplanır.

$$\begin{aligned}
OF_{WL2} = & \sum_{x,y \in \Omega} (W_{LH1}(x,y) - \mu_{LH1})^2 + \sum_{x,y \in \Omega} (W_{HL1}(x,y) - \mu_{HL1})^2 \\
& + \sum_{x,y \in \Omega} (W_{HH1}(x,y) - \mu_{HH1})^2
\end{aligned} \quad (4.16)$$

burada N pencere boyutunu, μ ise pencere içindeki imgenin ortalama piksel değerini gösterir.

Ayrık Kosinüs Dönüşümü

Literatürde Ayrık Kosinüs Dönüşümü (AKD) tabanında tanımlı odak ölçüt operatörleri de mevcuttur [Charfi et al., 1991; Baina and Dublet, 1995; Kristan et al., 2006]. .AKD piksel değerleri ile ifade edilen imgeleri uzaysal frekanslar cinsinden ifade edebilmek için bu frekanslara karşılık gelen katsayılarını hesaplar. Hesaplanan katsayılar, sıfır frekanslı katsayılar DC, diğerleri de AC olacak şekilde ikiye ayrılır. AKD tabanlı operatörler imgelerin kosinüs dönüşümünü sonucu elde edilen katsayılarından yararlanarak odaklanma miktarını hesaplarlar. Baina and Dublet [1995] AKD'nin DC katsayılarının enerjisinin odak ölçütü olarak kullanılabileceğini göstermişlerdir. Önerdikleri odak ölçütü, $F(u,v)$ 'nin imgenin AKD uygulanmış hali olduğunu düşünürsek şu şekilde hesaplanır;

$$OF_{DCT_1} = \sum_{u,v \in \Omega} F(u,v)^2 \quad (4.17)$$

Shen and Chen [2006] düşük kontrastlı imgeler için AC ve DC katsayılarının enerjilerinin oranının daha iyi bir odak ölçütü olduğunu göstermişlerdir.

$$OF_{DCT_2} = \frac{E_{AC}}{E_{DC}} \quad (4.18)$$

burada E_{AC} ve E_{DC} sırasıyla AKD'nin AC ve DC bileşenlerinin enerjisidir. Lee et al. [2006] AKD'ne tabi tutulan imge bloğunun merkeze yakın noktadaki katsayıların hesaplanan odak ölçütüne etkisinin kenarlara yakın noktadaki katsayılarından daha fazla olduğu göstermişlerdir. Dolayısı ile önerdikleri AKD tabanlı odak ölçütü sadece her imge bloğu için merkezde bulunan katsayıyı hesaplarlar. Hesaplama maliyeti diğer

odak ölçüt operatörlerine göre çok daha az olduğu için cep telefonu gibi kısıtlı enerjiye sahip cihazlarda kullanılmaktadır.

Bir başka AKD tabanlı odak ölçütü [Kristan et al., 2006] tarafından önerilmiştir. Bu yöntemde imgelerdeki odaklanma miktarı normalize edilmiş AKD katsayılarının entropisi hesaplanarak bulunur. Bayes-dağılım entropi fonksiyonu uygulanan AKD katsayılarına $\tilde{F}(u, v)$ dersek, odak ölçütü şu şekilde hesaplanır;

$$OF_{DCT_3} = 1 - \sum_{u,v \in \Omega} \tilde{F}(u, v)^2 \quad (4.19)$$

S Dönüşümü Enerjisi

S-Dönüşümü [Stockwell et al., 1996] dalgacık dönüşümü ve fourier dönüşümü gibi frekans analiz araçları karşısında sahip olduğu avantajlardan dolayı görüntü ve sinyal işlemede sıklıkla kullanılmaktadır. Mahmood and Choi [2010] S dönüşümünü odak ölçüt operatörü olarak kullanmışlardır. I_{SD} 'nin imgenin S dönüşümü uygulanmış hali olduğunu kabul edersek, odak ölçütü şu şekilde hesaplanır,

$$OF_{SD} = \sum_{u,v \in \Omega} I_{SD}(u, v)^2 \quad (4.20)$$

Frekans Seçici Ağırlık Ortanca Filtresi

İmgelerin yüksek frekanslı bileşenleri bulmak için sıklıkla başvurulanan yöntemlerden birisi de ortanca filtresi kullanmaktır. Bu filtre imgelerdeki kenar bulmada kullanılan doğrusal olmayan bir filtredir. FSAO süzgeci sadece imgedeki yüksek frekans bileşenlerine tepki göstermek ile kalmaz, aynı zamanda dürtü gürültüsünden de etkilenmez [Choi et al., 1999]. İmgelerdeki her iki doğrultu için ağırlıklı ortanca filtresi şu şekilde hesaplanır.

$$\begin{aligned} H_x = & \text{med}\{I(x-1, y), I(x, y), I(x+1, y)\} \\ & - \frac{1}{2} \text{med}\{I(x-1, y), I(x-2, y), I(x-3, y)\} \\ & - \frac{1}{2} \text{med}\{I(x+1, y), I(x+2, y), I(x+3, y)\} \end{aligned} \quad (4.21)$$

$$\begin{aligned}
H_y = & \text{med}\{I(x, y-1), I(x, y), I(x, y+1)\} \\
& - \frac{1}{2} \text{med}\{I(x, y-1), I(x, y-2), I(x, y-3)\} \\
& - \frac{1}{2} \text{med}\{I(x, y+1), I(x, y+2), I(x, y+3)\}
\end{aligned} \quad (4.22)$$

H_x ve H_y değerleri hesaplandıktan sonra odak ölçütü şu şekilde hesaplanır,

$$OF_{FSAO} = \sum_{x,y \in \Omega} (H_x^2 + H_y^2) \quad (4.23)$$

İki Kutuplu Görüntü Filtresi

Malik and Choi [2008] iki kutuplu uyumsuz görüntü analizinden [Poon, 1985; Indebetouw and Poon, 1986] yararlanarak görüntüdeki gürültüye duyarsız odak ölçütü geliştirmiştir. Bu odak ölçütünde görüntüler iki kutuplu filtreden (h) geçirilerek I_c imgeleri elde edilir.

$$I_c = I * h(\sigma_1, \sigma_2) \quad (4.24)$$

burada σ_1 ve σ_2 iki kutuplu filtre parametrelerini, $*$ ise evrişim işlemini gösterir. Elde edilen imgeler kullanılarak odak ölçütü şu şekilde hesaplanır.

$$OF_{PLR} = \sum_{x,y \in \Omega} I_c(x, y) \quad (4.25)$$

Bunların dışında literatürde, imge entropisi [V. Michael Bove, 1993], güç dağılım analizi [Jutamulia et al., 1994], gabor filtresi [Xiong and Shafer, 1994], yönelim kod eşleşmesi [Li et al., 2007], şekilcik ayrışımı [Meneses et al., 2008], AKD tabanında [Mahmood et al., 2008] ve dalgacık tabanında [Mahmood et al., 2009] ana bileşen analizi gibi değişik fonksiyonlardan yararlanan odak ölçüt operatörleri mevcuttur. Değişik odak ölçüt operatörlerinin farklı durumlar için performans karşılaştırması için literatürde çeşitli çalışmalar mevcuttur [Groen et al., 1985; Subbarao and Tyan, 1998; Huang and Jing, 2007; Tian et al., 2007; Aslantas and Kurban, 2009].

4.4 Uyarlanabilir Odaktan Şekil

Klasik OŞ metotları sahne bir sahne noktasına denk gelen piksellerin arasından tam olarak odaklanılan pikseli tespit edebilmek için her piksel için odaklanma miktarını hesaplarlar. Ancak her hangi bir pikselin odaklanma miktarı sadece bu pikselin parlaklık değerine bakarak hesaplamak mümkün değildir. Bu yüzden bir piksel için odak ölçüt değerleri hesaplanırken bu pikselin komşuluğunda bulunan piksellerin değerleri de kullanılır. Bunun yanında, hesaplanan odak ölçüt değerleri, gürültü, yanlış hizalama, vb. gibi problemlerin etkisini azaltmak için sabit boyutlu yerel pencere içinde toplanır [Jarvis, 1983; Krotkov, 1987]. Odaklanma miktarının kenar yayılması probleminden etkilenmemesi için pencere boyutunun en az kameranın imgelerde oluşturabileceği bulanıklık dairesinin boyutu kadar olmalıdır [Nair and Stewart, 1992; Stewart and Nair, 1989]. Odak ölçütü hesaplamasında yaygın olarak kabul görmüş bu yaklaşım, birbirine yakın piksellere karşılık gelen derinlik değerlerinin de birbirine yakın olacağı varsayımına dayanmaktadır. Bu varsayım sahnenin sürekli yüzeyleri için gerçekçi olsa da yüzey süreksizliklerinin olduğu bölgelerde tamamıyla ihlal edilmiş olur. Dolayısı ile bu bölgeler için hesaplanacak odak ölçüt değerleri ve derinlik değerleri hatalı olur.

Yüzey süreksizliklerinin bulunduğu bölgeler için de doğru derinlik tespiti yapabilmek için odak ölçütlerini toplarken sabit boyut ve şekilli pencere kullanmak yerine, sahnenin yüzey yapısına göre uyarlanabilir pencereler kullanmak gerekir [Aydın and Akgül, 2008]. Bu sayede seçilecek pencerelerin yüzey süreksizliklerini içermemesini sağlamış oluruz. Ancak başlangıçta sahne yapısı dolayısı ile yüzey süreksizliklerinin bulunduğu bölgeler elimizde veri olarak bulunmamaktadır. Zaten bu bilgiler çözmeye çalıştığımız problemin sonucunda ortaya çıkacaktır. Bu yüzden sahnenin yapısı hakkında elimizdeki verilerden gerçekçi çıkarımlarda bulunmamız gerekir.

Sahnenin fotometrik özelliklerinin sahne yapısı hakkında ipuçları içerdiği varsayılırsa, olası sahne yapısı ve yüzey süreksizlikleri sahne imgelerinden çıkarılabilir. Bu şekilde yerel imge özelliklerine dayalı her piksel için uyarlanabilir destek penceresi üretmek mümkündür. Bu yaklaşım insanların görsel algılama sisteminden esinlenmiştir. İnsan algı sistemi *Gestalt* kuralları uyarınca açıklanır. Buna göre insanlar etraftan

gelen uyarıcıları birbirinden bağımsız işlemek yerine bunları gruplandırarak bir bütün (gestalt) olarak değerlendirirler. Gruplandırma işlemi algılanan nesnelerin ya da özelliklerin birbirlerine yakınlık ve benzerlik derecesine göre yapılır. Örneğin bir insan resmine bakarken teker teker el, kol, gövde, baş gibi parçalar değil, insanın tamamı algılanır.

Sahne yapısına göre uyarlanabilir destek penceresi oluşturmanın en basit yolu pencere içerisindeki her piksele bir ağırlık katsayısı atamaktır. Bu şekilde imgedeki her piksel için farklı ağırlıklara sahip pencereler oluşturulur. Ağırlıklar hesaplanırken *Gestalt* kurallarından yararlanılarak pencerelerin sahne yapısına göre uyarlanması sağlanır. Birbirine konum ve renk değeri bakımından yakın olan pikseller aynı yüzeyin imgeleri olarak gruplandırılır. Buna göre pencerenin hesaplanacağı piksel ile pencere içerisinde kalan piksellerin yakınlık ve benzerlik derecesi hesaplanarak, yakın ve benzer pikseller için daha yüksek ağırlık değerleri verilir.

Buradaki en büyük problem odaktan şekil çıkarımı için elimizde bulunan sahne imgelerinin tümünde derinliğe bağlı kısmi bulanıklık bulunmasıdır. Bir başka deyişle, elimizdeki imgelerden hiç biri sahnenin fotometrik özelliklerini tam olarak içermez. Ayrıca, bu imgeler üzerinde yapılacak tümleştirme işlemiyle, sonraki bölümde açıklanan kapanma probleminden dolayı, sahnenin net görüntüsünü elde etmek her durumda mümkün olamamaktadır. Dolayısı ile elimizdeki kısmi bulanıklık içeren imgelerden uyarlanabilir destek pencereleri oluşturmak randımanlı olacaktır. Önerilen sistem sahnenin net görüntüsüne ihtiyaç duymaktadır ki bu da ancak iğne delikli kameralarda oluşmaktadır. Odaktan şekil için kullanılan kamera sistemini iğne delikli kameraya benzetmek için kameranın diyafram aralığını en düşük seviyeye getirmek yeterli olacaktır. Bu şekilde sahnenin kısmi bulanıklık içeren imgelerine ek olarak sahnenin fotometrik özelliklerini içeren net resmi de elde edilebilir.

Verilen (x_0, y_0) pozisyonunda bulunan piksel için ω_{x_0, y_0} destek penceresinin katsayıları sahnenin net görüntüsü I_f olsun. Görüntü pikselleri arasındaki yakınlık derecesi Δd , benzerlik derecesi de ΔI_f olarak alırsak destek penceresindeki (x_0, y_0) konumundaki ağırlık katsayısı şu şekilde hesaplanır;



Şekil 4.2 Örnek destek pencereleri.

$$\omega_{x_0 y_0}(x, y) = e^{-(\Delta d / \gamma_1 + \Delta I_f / \gamma_2)}, \quad (4.26)$$

burada γ_1 ve γ_2 ise bu iki uzaklık değerinin göreceli katkısını ayarlamak kullanılan pozitif sabit katsayılarıdır. Formül uzaysal düzlemde ve renk düzleminde çalışan iki gauss denkleminin birleşimidir. Destek penceresinin hesaplandığı piksel ile pencere içerisindeki herhangi bir pikselin uzaysal düzlemdeki öklid uzaklığı Δd ve renk düzlemindeki öklid uzaklığı da ΔI_f şu şekilde hesaplanır;

$$\Delta d = \sqrt{(x - x_0)^2 + (y - y_0)^2}, \quad (x, y) \in \Omega_{x_0, y_0} \quad (4.27)$$

$$\Delta I_f = \|I_f(x, y) - I_f(x_0, y_0)\|. \quad (4.28)$$

Şekil 4.2’de örnek bir imge üzerinde seçilen noktalar için oluşturulan destek pencereleri görülmektedir. Şekilde görülen destek pencerelerinde parlak noktalar yüksek ağırlık değerlerine denk gelmektedir. Pencere görüntülerinden de anlaşılacağı gibi seçilen noktalar birbirlerine çok yakın dahi olsalar oluşturulan destek pencereleri birbirinden çok farklı olabilmektedir. Pencere şekilleri seçilen noktanın dahil olduğu gruba, bir başka deyişle ait olduğu sahne parçasının yapısına göre değişmektedir.

Sahnenin net görüntüsündeki her piksel için ayrı ayrı oluşturulan destek pencereleri ve Bölüm 4.3’te anlatılan odak ölçüt operatörlerinden birisi kullanılarak Uyarla-

nabilir Odak Ölçüt (OF_{UYR}) değerleri aşağıdaki formüle göre hesaplanır.

$$OF_{UYR}(x_0, y_0) = \sum \omega_{x_0 y_0}(x, y) OF(x, y) \quad (4.29)$$

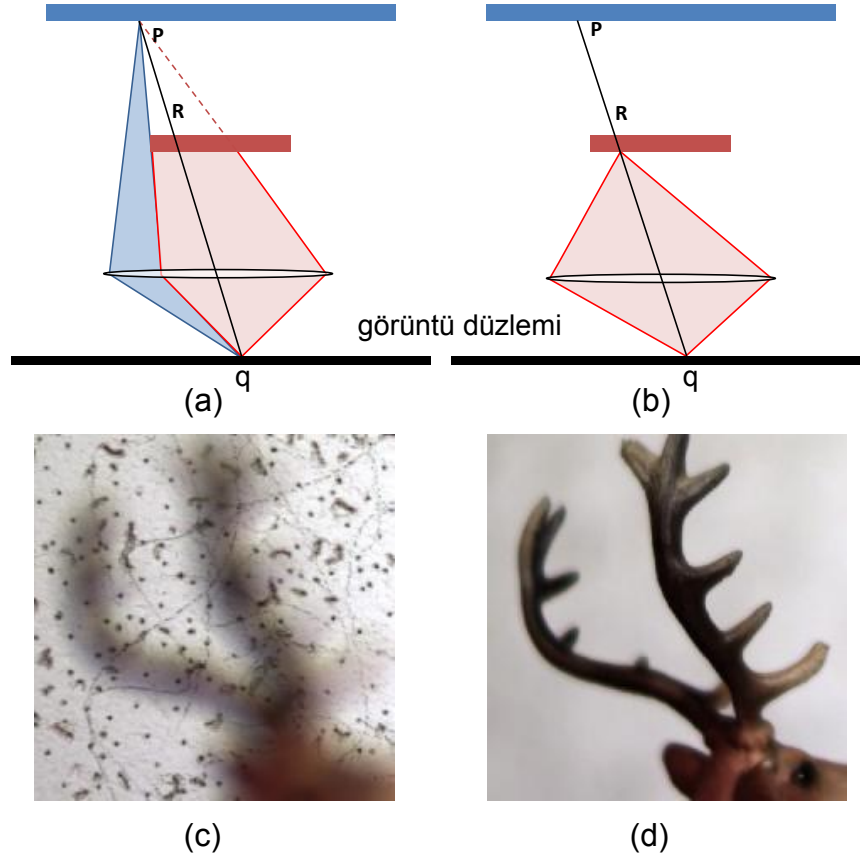
burada (x_0, y_0) odak ölçüt değeri hesaplanan pikselin pozisyonudur.

Aynı sahne noktasının elimizdeki farklı odak ayarlarıyla oluşturulmuş N farklı imgesi için uyarlanabilir odak ölçütü hesaplandıktan sonra, bu noktanın en net görüntüsünü ve bu görüntüyü oluşturan odak ayarını elde edilir. Bulunan odak ayar değeri Denklem 4.1'te yerine koyularak noktanın derinlik değeri hesaplanır.

4.5 Kapanma Problemi

Literatürdeki odak ölçüt operatörleri imgelerin odaklanma derecesini ölçmek için yerel imge bölgelerindeki parlaklık değişkenliklerini incelerler. Bu yaklaşım, açık diyaframlı kameraların tam olarak odaklanılmamış sahne noktaları için oluşturduğu imgelerin kısmi bulanıklık içereceği varsayımına dayanmaktadır. Görüntülerdeki kısmi bulanıklık, kamera diyafram-mercek düzeneğinin bulanık bölgelerde denk gelen yüksek frekanslı içeriği filtrelemesi sonucu ortaya çıkar. Bulanıklık miktarı odaklanma derecesine bağlı olduğundan, bir sahne noktasının tam odaklanılmış durumdaki görüntüsü o sahne noktası için gözlenebilecek en yüksek frekanslı içeriğe sahip olacaktır. Bu varsayım düzgün ve sürekli yüzeyler için gerçeğe uygundur. Ancak yüzey süreksizliklerinin olduğu durumlarda, sahnelerin açık diyaframlı kameralar ile oluşturulan imgelerinde iğne delikli kameralarda görülmeyen ve yukarıdaki varsayımı geçersiz kılan durumlar ortaya çıkar. Bunlardan en belirgin olanı, kameranin arka plana odaklandığı durumlarda, yüzey süreksizliklerine sebep olan ön plandaki cisimlerin tıpkı saydam cisimlerdeki gibi arka planı görünür kılmasıdır. Bunun sonucu olarak görüntü düzleminin ön plan cisimlerine denk gelen bölgelerinde, odak ayarı arka plana ayarlandığı durumlarda bile yüksek frekanslı içerik oluşabilmektedir. Bir başka deyişle, ön planda bulunan bir cismin en net görüntüsü olarak kameranin arka plandaki cisme odaklandığı durumda oluşturduğu bir imge seçilebilir. Böyle bir durumda ön plandaki cismin için

yapılacak derinlik hesaplaması hatalı olacaktır [Schechner and Kiryati, 2000]. Literatürde bu problem kapanma problemi olarak bilinmektedir. Çoğu Odaktan Şekil yöntemi yüzey süreksizliklerini göz ardı ettiği için bu problemi de buna bağlı olarak göz ardı etmiştir.



Şekil 4.3 Görüntü düzlemine paralel iki düzlemin bulunduğu bir sahne için görüntü oluşum geometrisi. (a) Kamera arka plandaki (kapanan) düzleme odaklanmış. (b) Kamera ön plandaki (kapatan) düzleme odaklanmış. İmge noktası q her iki durumda da odaklanılmış olarak algılanabilir. Bu durumu gösteren gerçek sahnenin odak ayarının arka plana ayarlandığında (c) ve önplana ayarlandığında (d) elde edilen imgeleri. (c)'de geyiğin boynuzları odak düzlemi dışında olmasına rağmen çekilen imgede geyiğin boynuzlarına odaklanılarak elde edilen imgeye oranla daha fazla yüksek frekanslı bileşenler vardır. (c) ve (d) imgeleri Şekil 4.9'deki sahne imgesinin kırpılmış imgeleridir.

Yüzey süreksizliğinin olduğu durumlarda açık diyaframlı kamera sistemlerinde görüntü oluşumu Şekil 4.3'te açıklanmıştır. Şekilde aynı sahnenin, kameranın ön plandaki cisme odaklandığı ve arka plandaki cisme odaklandığı durumlar için iki farklı görüntüsü oluşturulmuştur. Kamera merceği arka plandaki nesneye odaklandığı zaman, yüzey süreksizliğine yakın bir noktaya (Şekilde R noktası) denk gelen imge noktası (şekilde q noktası) odaklanılan arka planın ve bulanık ön planın görüntüsünü içerir.

Benzer şekilde, kamera ön plandaki nesneye odaklandığı zaman, aynı noktanın görüntüsü sadece odaklanılan ön planın görüntüsünü içerir. Sonuç olarak aynı noktanın iki farklı odak ayarıyla oluşturulan imgelerin hangisinin daha yüksek frekanslı içeriğe sahip olacağı arka plandaki nesnenin parlaklık desenine bağlıdır. Şekil 4.3 c’de kamera arka plana odaklanmıştır. Şekilden de anlaşılacağı üzere, geyiğin boynuzlarını gösteren imge bölgeleri boynuzun bulanık görüntüsünü ve duvarın net görüntüsünü içermektedir. Boynuzun Şekil 4.3 d’deki odaklanılmış imgesine bakılarak, bulanık görüntünün net görüntüden daha yüksek frekanslı içeriğe sahip olduğu durumlar görülebilir.

Kamera düzleminde oluşacak görüntü sahnenin parlaklık değerlerinin kamera diyafram-mercek sisteminin nokta yayılım fonksiyonu ile evrişimi olarak modellenebilir [Horn, 1986]. Nokta yayılım fonksiyonu tam odaklanma durumu için *dirac delta* veya birim dürtü δ fonksiyonu şeklindedir. Odaklanma düzlemi dışındaki bölgeler için bulanıklık dairesi büyüklüğünde 2B gauss fonksiyonu olarak alınabilir. Şekil 4.3.a’daki q noktası için görüntü oluşumu şu şekilde yazılır,

$$I_q = L_{\text{ön-plan}} * \delta(R), \quad (4.30)$$

burada $L_{\text{ön-plan}}$ kapanmaya sebep olan ön plandaki nesnenin parlaklık desenini, $*$ ise evrişim (convolution) operatörünü gösterir. Benzer biçimde Şekil 4.3.b’deki durum için aynı q noktasının görüntü oluşum denklemi şu şekilde yazılır,

$$I_q = \alpha L_{\text{arka-plan}} * \delta(P) + (1 - \alpha) L_{\text{ön-plan}} * H_R, \quad (4.31)$$

burada $L_{\text{arka-plan}}$ arka plandaki nesnenin parlaklık deseni, $0 < \alpha < 1$ olmak üzere arka plandan gelen ışık oranı ve H_R R noktası için geçerli olan 2B gauss fonksiyonunu gösterir. Denklemlerden de kolayca anlaşılacağı üzere, q noktası, α ve $L_{\text{arka-plan}}$ değerlerine bağlı olarak iki durumda da odaklanılmış olarak algılanabilir. Buna bağlı olarak, bu nokta için R noktasının derinlik değeri atanması gerekirken, bazı durumlar için P noktasının derinlik değeri atanabilir. Dolayısı ile imgelerdeki odaklanma derecesini bulmak için imgeler üzerinde parlaklık değişkenliklerini incelemek her durumda doğru sonucu vermeyebilir. Odaklanma derecesini ölçmek için kapanmaya duyarsız bir yöntem gerekmektedir.

4.6 Kapanma-Duyarsız Odak Ölçüt fonksiyonu

Literatürdeki mevcut odak ölçüt operatörlerinin tamamı imgelerdeki parlaklık değişkenliklerine dayalı olduğu için kapanma probleminin bulunduğu durumları çözmekte yetersiz kalmaktadır. Sahnenin görünür yüzeyinin fotometrik özelliklerinden yararlanılarak kapanma probleminden etkilenmeyen odak ölçüt operatörü geliştirilebilir. Odak ölçütü hesaplamasına dâhil edilecek bu ek bilgiyle, kameranın sadece arka plana odaklanması sonucu görünür olan sahne kısımları tespit edilebilir.

Sahnenin fotometrik özellikleri Odaktan Şekil için elde edilen sahne imgelerine tümleştirme işlemi uygulanarak çıkarılabilir [Subbarao et al., 1995]. Fakat tümleştirme yöntemleri de mevcut odak ölçüt operatörlerini kullandığı için kapanma probleminden etkilenirler. Sahnenin fotometrik özelliklerini en basit şekilde sahnenin net görüntüsünden elde edebiliriz ki zaten bu görüntü uyarlanabilir pencere oluşturmak için mevcut yöntemde ek bilgi olarak kullanılmaktadır. Odaktan şekil yöntemlerinde sahnenin net görüntüsü kenar bölgelerinin odaklanma derecesini [Nair and Stewart, 1992] ve bulanıklık derecesini ölçmek için [V. Michael Bove, 1993] kullanılmıştır

Sahnenin belirli bölgelerine odaklanılarak elde edilen imgelerdeki odaklanılan bölgelerin görüntüleri, sahnenin net görüntüsündeki bölgelere benzerler. Buradan yola çıkarak sahnenin net görüntüsü ile kısmi bulanıklık içeren görüntülerin benzerliği odak ölçütü olarak kullanılmıştır. Bu yöntem tek başına imgelerdeki yüksek frekanslı içeriklere bağlı olmadığı için diğer odak ölçütlerinde bulunan kapanma probleminden etkilenmez. Ancak görüntüler arasındaki benzerliği hesaplamak için kullanılacak yöntemin görüntülerdeki parlaklık farklılıklarından etkilenmemelidir. Çünkü karşılaştırma için kullanılan görüntüler farklı diyafram ve odak ayarlarıyla elde edildiği için parlaklık farklılığı içerir. Bu yüzden benzerlik ölçüsü olarak Normalize edilmiş Çapraz İlinti (Normalized Cross Correlation) ölçütü kullanılmıştır. Aynı büyüklükteki A ve B imge parçaları için Normalize edilmiş Çapraz İlinti değeri şu şekilde hesaplanır,

$$N(A, B) = \frac{\sum_{\Omega} (A - \bar{A})(B - \bar{B})}{\sqrt{\sum_{\Omega} (A - \bar{A})^2 \sum_{\Omega} (B - \bar{B})^2}}, \quad (4.32)$$

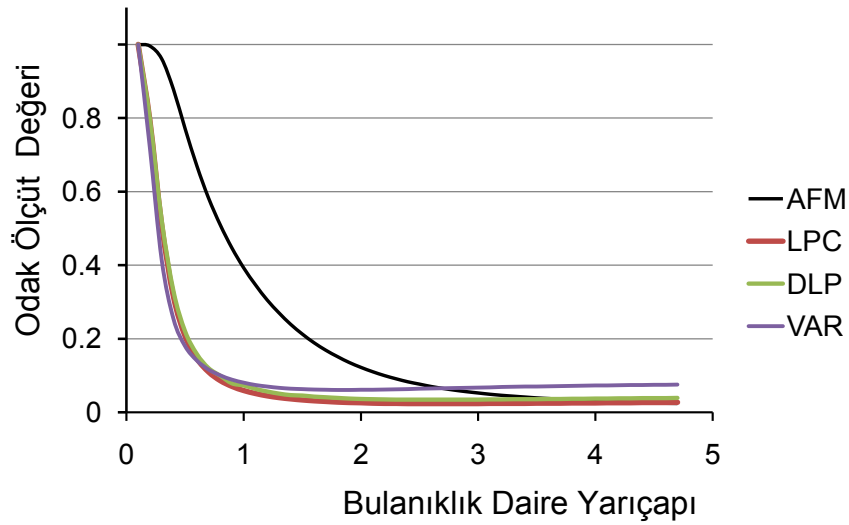
burada Ω hesaplamının yapıldığı görüntü bölgesinin büyüklüğü olarak alınmıştır. NÇİ denklemi kullanılarak kapanmaya duyarsız odak ölçüt operatörü şu şekilde tanımlanmıştır,

$$OF_i(x, y) = N_{\Omega}(I_f(x, y), I_i(x, y)), \quad (4.33)$$

burada I_f sahnenin net görüntüsü, I_i ise odaktan şekil imge dizisindeki i . imge olarak alınmıştır.

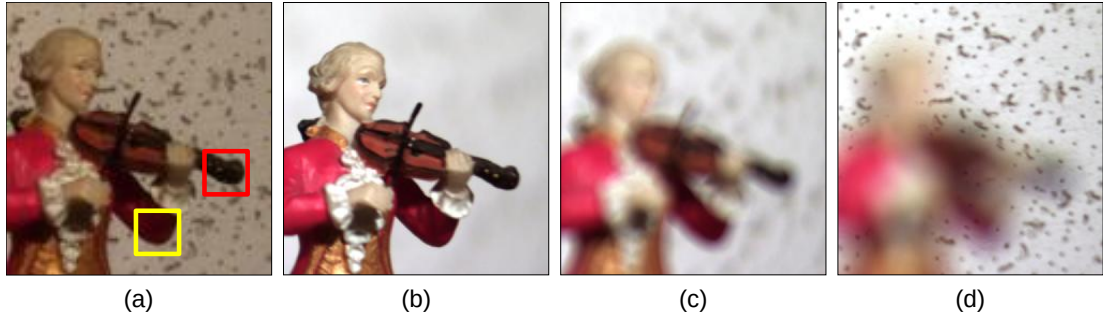
Burada sahnenin net görüntüsünü elde ederken kullanılan farklı diyafram ayarı imgeler arası hizalama problemi ortaya çıkaracağından, odaktan şekil imge dizisindeki bütün imgeler sahnenin net görüntüsüne göre hizalanmalıdır. Bu işlem imgeler arasındaki büyütme faktörü hesaplanarak basitçe yapılabilir [Favaro and Soatto, 2007].

4.7 Deneyler



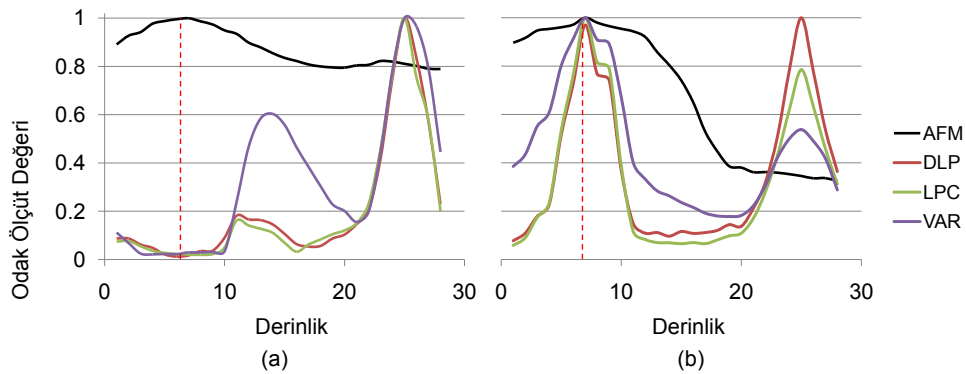
Şekil 4.4 Odak ölçüt operatörlerinin bulanıklık derecesine göre karşılaştırması.

Önerilen odaktan şekil metodunun doğruluğunu ve başarımını göstermek için, kapanma probleminin görüldüğü yapay ve gerçek sahneler üzerinde çeşitli deneyler yapılmıştır. Deneyler boyunca uyarlanabilir pencereler oluşturmada kullanılan γ_1 parametresi 5 olarak, γ_2 parametresi ise seçilen pencere boyutunun yarısı olarak alınmıştır.



Şekil 4.5 Örnek bir sahnenin net görüntüsü (a) ve sahnenin farklı odak ayarlarıyla elde edilmiş imgeleri (b-d). Odak ölçütlerinde hataya sebep olan örnek kapanma bölgeleri kırmızı ve sarı renkli çerçevelerle işaretlenmiştir (a).

İlk deney olarak, önerilen odak ölçüt operatörünün odaklanma derecesine karşı davranışını görebilmek için, yüksek frekanslı içeriğe sahip yapay bir imge ve bu imgenin yapay olarak bulanıklaştırılmış imgeleri oluşturulmuştur. Bulanıklaştırma işlemi imgelerin gauss fonksiyonu ile evrişimi ile gerçekleştirilmiştir. Şekil 4.4 farklı odak ölçüt operatörlerinin artan bulanıklık derecesine göre ürettikleri değerlerin grafiğini göstermektedir. Grafik önerilen kapanmaya duyarlı odak ölçütünün (AFM) Laplace (LPC), Değiştirilmiş Laplace Toplamı (DLP), ve Varyans (VAR) odak ölçütleri ile olan karşılaştırmasını göstermektedir. Grafikten de anlaşılacağı gibi, önerilen odak ölçütü ideal odak ölçütlerinde olması gereken bir özellik olan tekdüzeliği taşımaktadır. Bir başka deyişle odaklanma derecesi arttıkça ki bu bulanıklığın azalması demektir, odak ölçüt değeri artmaktadır.



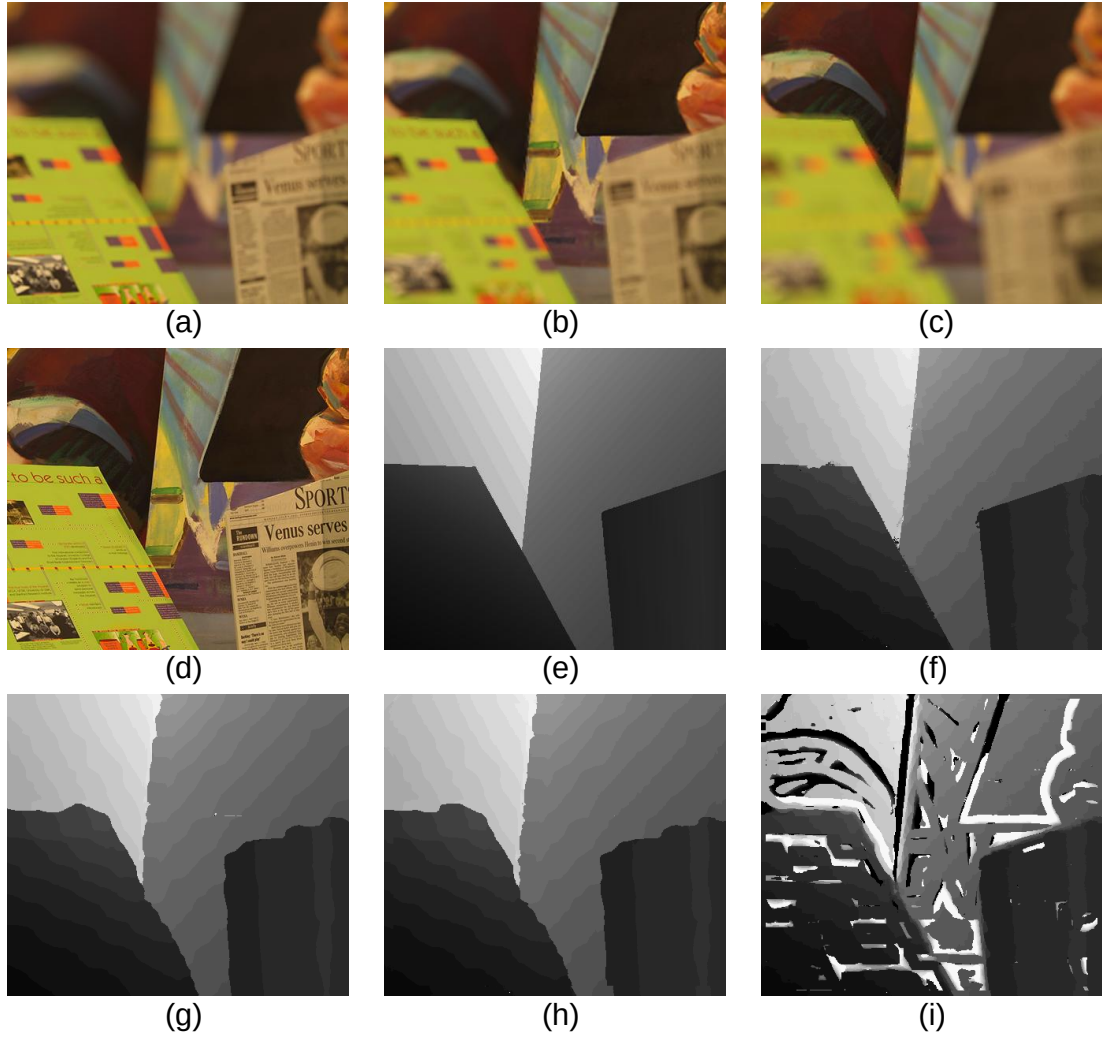
Şekil 4.6 Şekil 4.5.a 'da gösterilen pencereler için odak ölçüt operatörlerinin karşılaştırması.

İkinci deneyde önerilen odak ölçütünün kapanmaya karşı duyarlılığının gösterilmesi amaçlanmıştır. Bu yüzden bu deney için kapanmanın görüldüğü gerçek bir

sahnenin gerçek imgeleri kullanılmıştır. Şekil 4.5 kullanılan sahnenin örnek imgele-
rini göstermektedir. Kapanma probleminin görüldüğü iki farklı nokta Sekil 4.5.a 'da
kırmızı ve sarı renkli pencerelerle gösterilmiştir. Bu iki noktaya ve arka plana odak-
lanıldığı durumlarda üretilen sahne imgeleri Şekil 4.5.b-d' de gösterilmiştir. Bu iki
nokta için farklı odak ölçütlerinin odaklanılan sahne derinliğine bağlı olarak ürettikleri
değerlerin grafikleri Şekil 4.6'de gösterilmiştir. Grafiklerdeki kesikli kırmızı çizgilerle
bu noktaların gerçek derinlik değerleri gösterilmiştir. Grafiklerde de görüldüğü gibi var
olan odak ölçüt operatörleri kapanma probleminin olduğu durumlarda ya yanlış derin-
lik değerlerinde en yüksek tepkiyi verir ya da iki modlu sonuçlar üretir. Diğer yandan
önerilen odak ölçüt operatörü en yüksek değeri doğru derinlik ayarında verdiği gibi tek
modlu sonuçlar üretir. Bu özellikler de ideal odak ölçüt operatörlerinde olması gereken
özelliklerdendir.

Gerçek sahnelerin derinlik değerleri bilinmediği için, Odaktan Şekil yöntem-
lerinin bu sahnelerdeki başarımını sayısal olarak göstermek mümkün olmamaktadır.
Bu yüzden bu yöntemlerin sayısal doğrulukları, net görüntüsü ve derinlik değerleri
bilinen bir sahnenin ardışık derinlik düzlemlerine odaklanılmış sentetik imgeleri üreti-
lerek hesaplanmıştır. Sahnelerin sentetik imgeleri Iris filtresi [Sakurai, 2004] kullanı-
larak oluşturulmuştur. Iris filtresi sahnelerin farklı derinliklere odaklanılmış imgelerini
oluşturmak için sahnenin net görüntüsüne ve sahnenin derinlik bilgisine ihtiyaç du-
yar. Derinlik değerleri bilinen sahne imgeler Middlebury veri tabanından [Scharstein
and Szeliski, 2002] elde edilmiştir. Bu şekilde Middlebury veri tabanındaki Venus ve
Cones sahneleri için 30'ar adet yapay imge oluşturulmuştur.

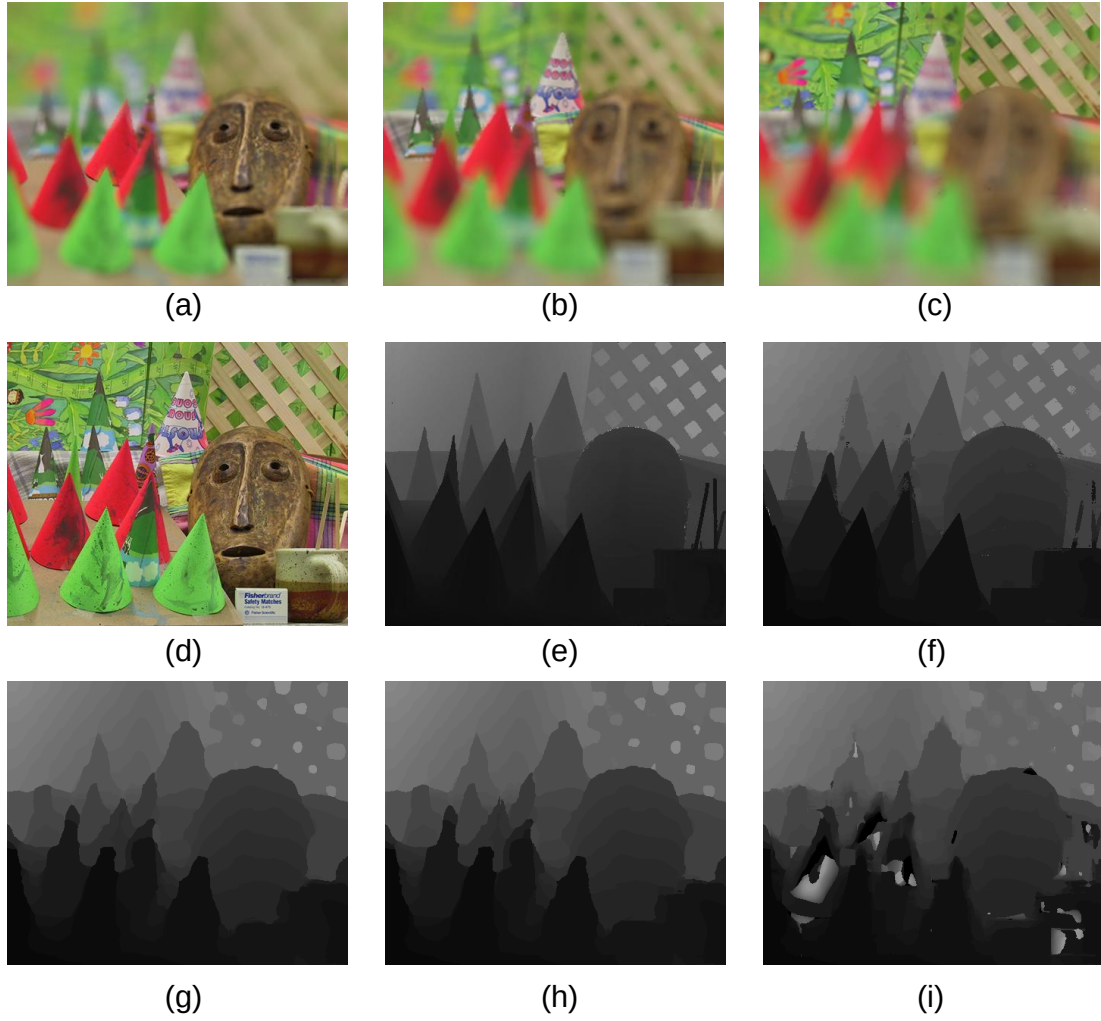
Seçilen örnek sahneler yüksek miktarda yüzey süreksizlikleri içermesi öneri-
len odak ölçüt operatörünün bu bölgelerdeki başarımını da sayısal olarak gösterme
imkânı sağlar. Şekil 4.7 sahnenin net görüntüsünü, gerçek derinlik haritasını, önerilen
yöntemin ürettiği derinlik haritasını ve literatürde bulunan bazı odak ölçüt operatörleri
ile bulunan derinlik haritasını göstermektedir. Buna benzer olarak şekil 4.8'de Cones
sahnesi için elde edilen imgeleri gösterir. Şekil 4.7 ve 4.8'te de görüleceği gibi yüzey
sürekliliklerinden dolayı var olan odak ölçüt operatörleri bu bölgelerde beklendiği gibi
hatalı sonuçlar üretirler. Buna karşılık önerilen odak ölçütü kullanılarak bulunan derin-
lik haritalarında yüzey süreksizlikleri içeren bölgelerin derinlik değerleri doğru olarak



Şekil 4.7 (a-c)Venus sahnesi için Iris filtresi ile üretilen sentetik odaklanma görüntüleri. (d) Sahnenin net görüntüsü. (e) Sahnenin gerçek derinlik haritası. (f) önerilen yöntemin bulduğu derinlik haritası. (g) DLP, (h) LPC ve (i) VAR odak ölçütleri ile bulunan derinlik haritaları.

hesaplanmıştır.

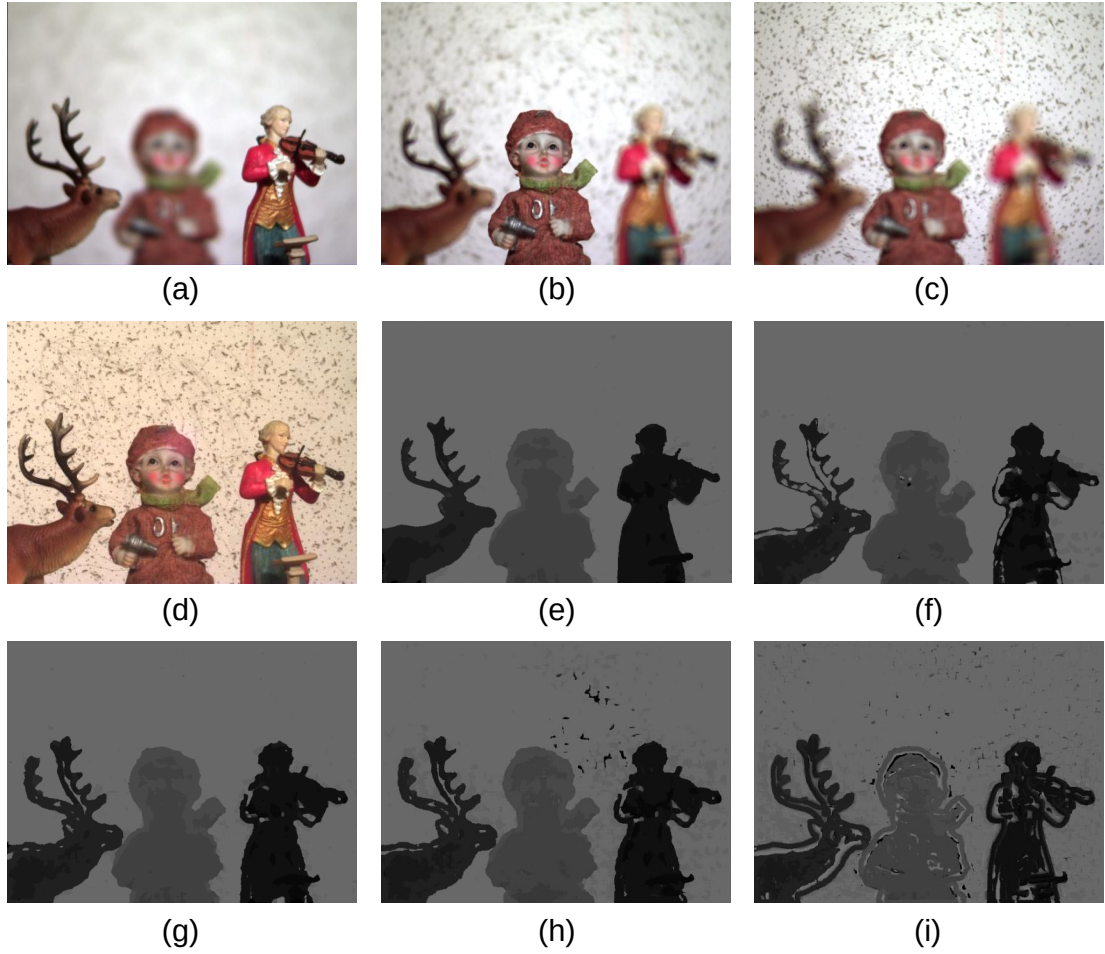
Venus ve Cones sahneleri farklı odak ölçütleri kullanılarak elde edilen derinlik haritalarının sahnenin gerçek derinlik haritasıyla karşılaştırarak odak ölçütlerinin başarımı karşılaştırılmıştır. Karşılaştırma ölçütü olarak Ortalama Karekök (rms) hataları kullanılmıştır. Çizelge 4.1 ve Çizelge 4.2 kullanılan odak ölçütlerinin farklı pencere boyutlarında ürettiği hata oranlarını göstermektedir. Sonuçlara göre küçük boyutlu pencereler için önerilen odak ölçütünün başarımı var olan odak ölçütleriyle karşılaştırılabilir seviyededir. Düşük pencere boyutları bütün odak ölçütleri için yüksek hata oranlarına sebep olmaktadır. Pencere boyutu arttıkça hata oranları kademeli olarak düşmektedir. Ancak, pencere boyutu arttıkça önerilen odak ölçütü kullanılarak elde edilen sonuç-



Şekil 4.8 (a-c)Cones sahnesi için üretilen sentetik odakta şekil imgeleri. (d) Sahnenin net görüntüsü. (e) Sahnenin gerçek derinlik haritası. (f) önerilen yöntemin bulduğu derinlik haritası. (g) DLP, (h) LPC ve (i) VAR odak ölçütleri ile bulunan derinlik haritaları.

ların hata oranı diğer yöntemlere göre çok daha hızlı bir şekilde düşmektedir. Ayrıca pencere boyutunu aşırı derecede arttırmak önerilen yöntem için problem teşkil etmesine rağmen diğer yöntemler için belirli bir dereceden sonra problem oluşturmaktadır. Özellikle yüzey süreksizliklerinin olduğu bölgelerde önerilen yöntemin başarımı çok daha yüksektir çünkü önerilen yöntemin uyarlanabilir özelliği, onun gelişi güzel biçimdeki yüzey süreksizliklerinden etkilenmemesini sağlar.

Sonuncu deney olarak, bariz yüzey süreksizlikleri ve kapanma bölgeleri içeren zorlayıcı iki farklı gerçek sahne seçilmiştir. Sahne imgeleri laboratuvar ortamında gerçek diyaframlı kameralar kullanılarak elde edilmiştir. Her sahne için farklı odak ayarlarında 30 farklı sahne imgesi oluşturulmuştur. Sahnelerin net görüntüsü kameranın

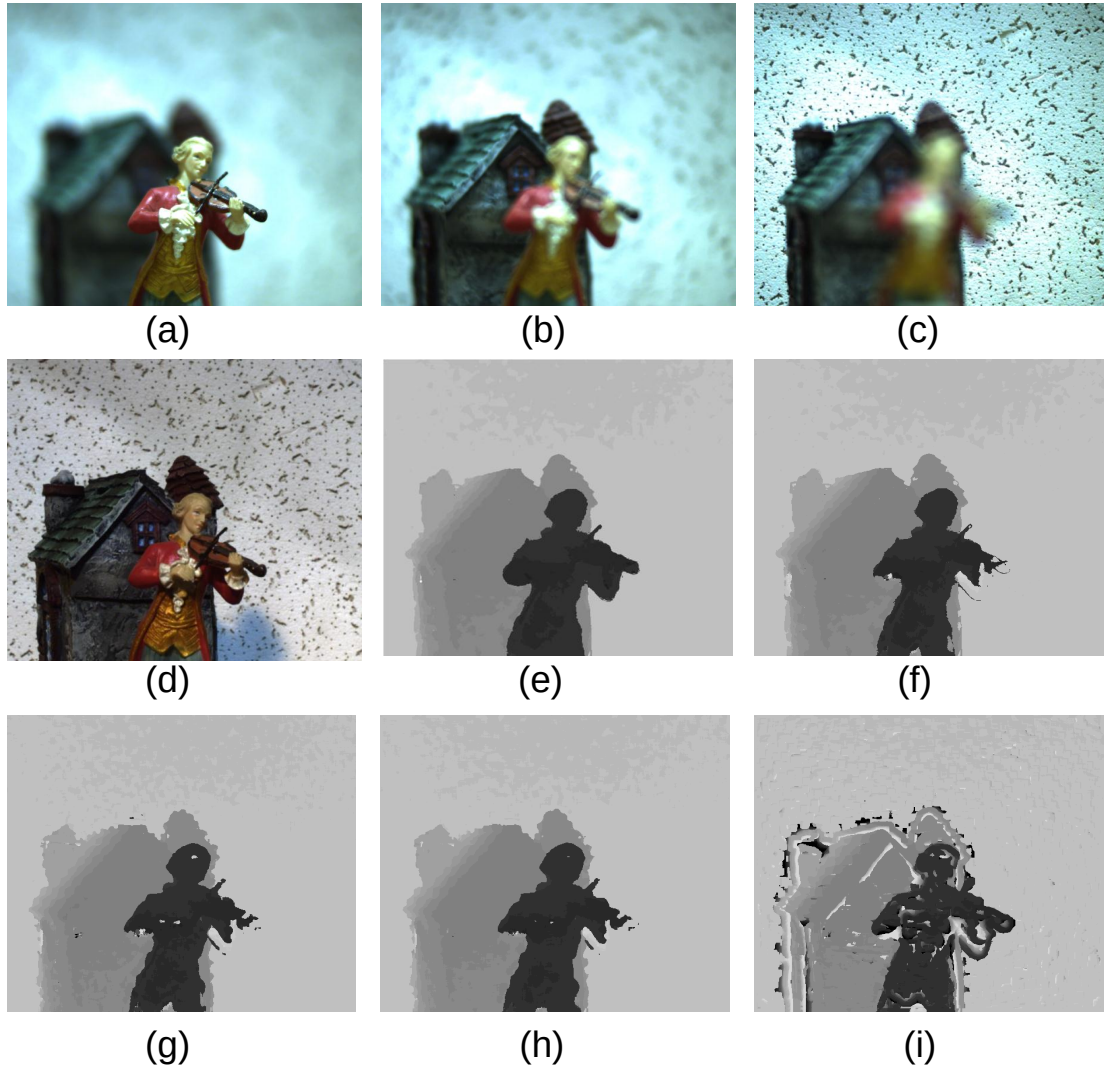


Şekil 4.9 (a) Gerçek bir sahnenin net görüntüsü. (b-d) Odaktan şekil setindeki örnek imgeler ve (e) önerilen yöntemin bulduğu derinlik haritası. (f) uyarlanabilir pencere-lerle kullanarak DLP, (g) DLP, (h) LPC ve (i) VAR odak ölçütleri ile bulunan derinlik haritaları.

Çizelge 4.1 Venüs resmi için hesaplanan rms hata oranları.

Odak Ölçüt Operatörü	Pencere Boyutu				
	25x25	21x21	15x15	11x11	9x9
SML	14	12	12	13	14
LPC	14	13	13	13	15
VAR	59	56	55	55	56
AFM	5	6	7	13	15

en düşük diyafram açıklığı ayarıyla elde edilmiştir. Şekil 4.9 ve Şekil 4.10 sahnelerin farklı odak ayarlarıyla elde edilen örnek imgelerini, net görüntüsünü ve farklı odak ölçütü kullanılarak bulunan derinlik haritalarını göstermektedir. Sahnelerin gerçek derinlik haritaları bilinmemektedir. Ancak, görsel olarak anlaşılacağı üzere önerilen odak ölçüt operatörü yüzey süreksizliklerinin ve kapanma bölgelerinin üstesinden gelebilmiştir. Diğer odak ölçüt operatörleri ise bu bölgelerde hatalı sonuçlar üretmişlerdir.



Şekil 4.10 (a) Başka bir sahnenin net görüntüsü. (b-d) Odaktan şekil setindeki örnek imgeler ve (e) önerilen yöntemin bulduğu derinlik haritası. (f) uyarlanabilir pencere-lerle kullanarak DLP, (g) DLP, (h) LPC ve (i) VAR odak ölçütleri ile bulunan derinlik haritaları.

Çizelge 4.2 Cones resmi için rms hata oranları.

Odak Ölçüt Operatörü	Pencere Boyutu				
	25x25	21x21	15x15	11x11	9x9
SML	12	11	10	9	9
LPC	12	11	10	9	9
VAR	21	20	19	19	22
AFM	6	6	6	7	8

Uyarlanabilir pencere ile Değiştirilmiş Laplace operatörü kullanan Aydın and Akgul [2008]'ün yöntemi yüzey süreksizliklerini doğru olarak bulmasına rağmen kapanma bölgelerinde diğer yöntemler gibi başarısız olmuştur.

Çeşitli pencere büyüklükleri ile denenen önerilen odak ölçütünün ve diğer odak

ölçütlerinin sonuçlarından da anlaşıldığı gibi önerilen yöntem yüzey süreksizliklerini başarılı bir şekilde çıkarır. Özellikle, klasik yöntemlerin başarımı pencere boyutu artarken azalmasına rağmen, önerilen yöntemin başarımı artmaktadır. Bunun yanında, önerilen yöntem kapanma bölgelerindeki derinlik değerlerini de doğru bir şekilde bulmaktadır ki bu durum diğer odak ölçütlerinde mümkün olmamaktadır.

5 Katmanlı Yapı ile Regülerizasyon

İmge tabanlı regülerizasyon yöntemlerinin en belirgin dezavantajı imgede zayıf ayrıtlarla ifade edilen yüzey süreksizliklerine aşırı düzleştirme uygulamasıdır. Ayrıca düzleştirme oranını imge ayrıt büyüklüğüyle ilişkilendirmek tutarlı bir yaklaşım değildir, çünkü yüzey süreksizliklerinin derecesi ayrıt büyüklükleri ile ilintili değildir. Bir başka deyişle komşu iki piksel arasındaki parlaklık farkının büyüklüğü aynı piksellerin derinlik farklarının büyüklüğünden bağımsızdır.

Bahsedilen yöntemlerdeki yaklaşım sonuçlarda gözle görülür bir iyileştirmeye sebep olsa da, aşırı desenli imge bölgelerinde başarısız olur. Ayrıca bu yöntemlerin stereo problemi için bir diğer ciddi problem olan kapanma probleminin çözümüne katkısı yoktur. Kapanma bölgelerindeki piksellerin derinlik değerlerini, diğer resimdeki eşleşeceği pikseller görünür olmadığı için, tespit etmek oldukça zordur. Bu problem için sunulan en etkin çözüm aktif ışıklandırma kullanmaktır [Raskar et al., 2004; Zickler et al., 2002]. Ancak tüm sahneler için bu çözüm uygulanabilir değildir.

İmge parlaklığına dayalı süreksizlik tahmin etme yöntemlerinin kısıtlamalarını aşmaya yardımcı olacak ve kapanma bölgeleri için de çözüm olabilecek daha güvenilir ipuçlarına ihtiyaç vardır.

Bu konuda önerilebilecek en basit yaklaşım farklı yöntemler ile üretilecek derinlik haritaları üzerinde tümleştirme işlemi gerçekleştirmektir. Sahnelerin derinliği odaktan şekil [Krotkov, 1987; Nayar and Nakagawa, 1994], bulanıklıktan şekil [Subbarao et al., 1995; Subbarao and Surya, 1994; Nayar et al., 1996], fotometrik stereo [Woodham, 1980], hareketten şekil [Waldmann and Merhav, 1992], vb. gibi yöntemlerle de bulunabilir. Ancak bu yöntemlerinde kendine has problemleri olduğu için tümleştirme işlemleri dikkatlice yapılmalıdır. Literatürde, bu yöntemlerle stereo yöntemini birleştiren gölgeden şekil ile stereo [Cryer et al., 1993; Lange, 1999; Karaman and Bilge, 2009], odaktan şekil ile stereo [Bove, 1990; Frese and Gheeta, 2006], bulanıklıktan şekil ile stereo [Rajagopalan and Mudenagudi, 2004; Kim and Sikora, 2007], hareketten şekil ile stereo [Waldmann and Merhav, 1992; Yuk-Kwan Ho and Pong, 1994;

Kambhamettu et al., 1995], optik akış ile stereo [Braillon et al., 2006] gibi tümleştirme yöntemleri mevcuttur.

Tezin bu kısmında odaktan şekil ile stereo yardımlaşması ile fark haritası bulma amaçlanmıştır. Yaklaşımımız stereo ve odak bilgisini kullanarak derinlik tespiti yapan insan görme sisteminden [Pentland, 1987] esinlenmiştir. Odaktan şekil ile stereo tümleştirmesi üzerine yapılan en eski çalışma, iki yöntem ile bulunan derinlik haritalarını ağırlıklı ortalama yöntemi ile birleştirme yoluna gitmiştir [Bove, 1990]. Benzer bir şekilde, Krotkov and Bajcsy [1993] edle edilen derinlik haritalarını karşılıklı doğrulama yapmak için kullanmışlardır. Bazı araştırmacılar odaktan şekilden elde ettikleri derinlik haritasını stereo yöntemlerinde başlangıç değeri olarak almışlar [Ahuja and Abbott, 1993] ya da arama uzayını daraltmak için kullanmışlardır [Subbarao et al., 1997]. Bahsedilen çalışmaların hepsi odaktan şekil ve stereo derinlik haritaları arasında kayıtlama işlemini yapmak zorundadırlar. Bu işlem gerçek diyaframlı lens sistemlerinde bulunan büyütme, alan eğriliği, alan derinliği gibi problemlerden dolayı oldukça zor ve hassas ölçümleme gerektirmektedir. Ayrıca bilinen odaktan şekil yöntemlerinde göz ardı edilen yüzey süreksizlikleri ve kapanma problemi, yöntemleri birleştirme işleminden önce çözülmesi gerekmektedir.

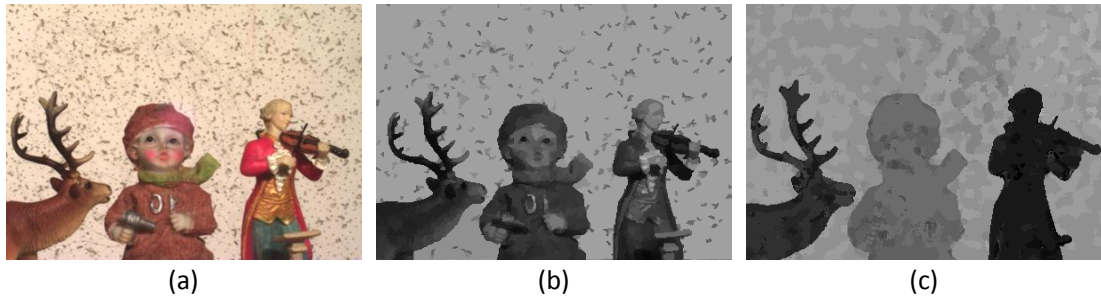
Çalışmanın bu kısmında önerilen yöntemde görünür sahne yüzeyi kameraya olan uzaklıklarına göre sıralanmış katmanlı bir yapıyla modellenmiştir. Sahnenin katmanlı yapısı odaktan şekil yöntemindeki gibi sahne üzerindeki kamera düzlemine paralel sanal katmalara odaklanılarak elde edilen imgelerden çıkarılmıştır. Katmanlı yapıyı çıkarırken bölüm 4'te önerilen odak ölçüt operatörü kullanarak klasik odaktan şekil yöntemlerinde bulunan kapanma ve yüzey süreksizliği problemi çözülmüş olur. Sahnenin katmanlı yapısı önerdiğimiz stereo yönteminde yüzey süreksizliklerini belirlemede ve kapanma bölgelerinin yüzey yapısını tahmin etmede kullanılmıştır.

Bunun yanında, sahnenin katmanlı yapısından ve ağların zamansal konumlarından etkin bir şekilde faydalanabilecek yayılım operatörü tasarlanmıştır. Operatör daha önce önerilen yayılım yöntemlerinden farklı olarak, ağların konumsal ve zamansal durumlarına göre yönbağımlı ya da yönbağımsız düzleştirme işlemi uygulayabilmektedir. Ayrıca, sahnenin katmanlı yapısını sadece yayılım operatörü içinde kullanarak, diğer

odaktan şekil ile stereo birleştirme yöntemlerinde yapılması gereken ancak gerçek sahneler için uygulaması pratik olmayan [Ahuja and Abbott, 1993] iki sistem arasındaki kayıtlama işlemine gerek kalmamıştır.

5.1 Sahnelerin Katmanlı Gösterilimi

Bir sahnenin sabit bir açıdan görülebilen yüzeyi, birbirleriyle çakışmayan ve görüntü düzlemine paralel yüzey katmanlarıyla ifade edilebilir. Bu şekilde ifade edilen sahne yapısı, stereo yöntemlerinde sahne yüzeyindeki süreksizliklerinin muhtemel yerlerini tahmin etmede kullanılabilir. Birçok modern stereo yöntemi basitçe görüntüler üzerinde imge bölütleme algoritması uygulayarak bu yaklaşımın avantajını kullanırlar [Zhang and Kambhamettu, 2002; Hong and Chen, 2004; Bleyer and Gelautz, 2004, 2005b,a; Evans and Ferryman, 2006; Gerrits and Bekaert, 2006; Bleyer and Gelautz, 2007]. Ancak, aşırı desenli yüzeyler için normalden fazla katman üretilebilmektedir. Buna ek olarak, resim bölütleme ile oluşturulan katmanlarda derinlik ile ilgili bilgi bulunmadığı için, düzleştirmede katmanlar arası ilişkiyi ayarlayacak belirgin bir yöntem yoktur. Bunun sonucunda, kapanmanın olduğu bölgelerin içinde kalan, dolayısı ile tutarsız eşleşme değerleri olan katmanların derinlik değerleri düzleştirme ile düzeltilmeyecektir.



Şekil 5.1 (a) Örnek bir sahnenin net görüntüsü, (b) imge bölütleri, (c) sahnenin katmanlı gösterimi.

Önceki bölümde önerilen odak ölçüt operatörünü kullanarak sahnenin katmanlı yapısı oluşturulabilir. Önerilen odak ölçüt operatörü imgelerin en net olduğu bölgeleri bulmak için bu imgeler ile sahnenin net görüntüsü arasında eşleşme işlemi gerçekleştirir. Sahnenin net görüntüsü gerçek diyaframlı kamerayı, diyaframı en küçük açıklık değerine getirmek suretiyle, iğne delikli kameraya benzeterek elde ediliyordu. Ancak,

kamera ideal iğne delikli kamera gibi davranamayacağından, bu yöntem ile elde edilecek sahne görüntüleri kısmi bulanıklık içerecektir. Bu yüzden eşleşmeler en net imge bölgeleri yerine ona yakın netlikteki imge bölgeleriyle olacaktır. Dolayısı ile katmanlı yapıda yüzey süreksizliklerinin olduğu bölgeler ve kapanma bölgelerinin yüzey yapısı doğru olarak gösterilmesine rağmen, gösterilen derinlik değerlerinin standart sapmasının yüksek çıkmaktadır. Bu yüzden bulunan bu derinlik değerlerini stereo sistemine taşımak yerine, bu yapıyı regülarizasyon işleminde kullanmak önerilmiştir.

Şekil 5.1 örnek bir sahnenin net görüntüsünü, imge bölütlerini ve katmanlı yapıyı göstermektedir. Şekildeki katmanlı yapının gösterdiği derinlik değerleri yüksek standart sapmaya sahip olsa da sahne yapısı hakkında imge bölütlerinden çok daha fazla bilgi içermektedir. Katmanlı yapı yüzey süreksizliklerini ve kapanma bölgelerindeki noktaların yüzey şekillerini barındırdığından ve katmanlar arasındaki göreceli uzaklıklar bilindiğinden, regülarizasyon işleminde sistemin başarımını arttıracak şekilde kullanılabilir.

5.2 Regülarizasyon

Katmanlı sahne yapısının imge bölütlerine karşı sahip olduğu avantajlardan faydalanabilmek için bu yapı regülarizasyon işleminde kullanılabilir. Bu amaçla, denklem 3.10'teki yayılım fonksiyonu g , katmanlı imge I_l kullanılarak şu şekilde tanımlanmıştır

$$g(\nabla I_l) = e^{-(|\nabla I_l|/\kappa_1)^2}, \quad (5.1)$$

burada κ_1 süreksizlik sabitidir ve hangi derinlik farklarının yüzey süreksizliği olarak algılanacağını belirler. Fakat, sadece denklem 5.1 kullanılarak tanımlanacak yönbağımlı yayılım fonksiyonu eniyileme işlemini yerel minimumlara karşı duyarlı yapar. Özellikle gürültülü ve küçük yüzey parçaları regülarizasyondan yeterince faydalanamazlar. Bilindiği gibi yayılım katsayısının sabit olduğu yönbağımsız yayılım işlemi yerel minimumlara karşı duyarsız olmasına rağmen yüzey süreksizliklerine aşırı düzeltme uygularlar.

Bu tip problemleri çözebilmek için katmanlı sahne yapısını ve ağların birbirine olan uzaklıklarını kullanarak bölüm 3.3’de önerilen regülarizasyon yöntemine benzer bir yol izlenmiştir. Dolayısı ile eniyileme sürecinde regülarizasyon yönbağımsız yayı-
nımdan yönbağımlı yayınıma doğru evrilecek şekilde ayarlanmıştır. İstenilen özellik-
leri sağlaması için yayının fonksiyonu şu şekilde tanımlamıştır.

$$\psi' = (1 - h(\Delta d)) + h(\Delta d) \cdot g(\nabla I_l), \quad (5.2)$$

burada

$$h(\Delta d) = e^{-(\Delta d/\kappa_2)^2}, \quad (5.3)$$

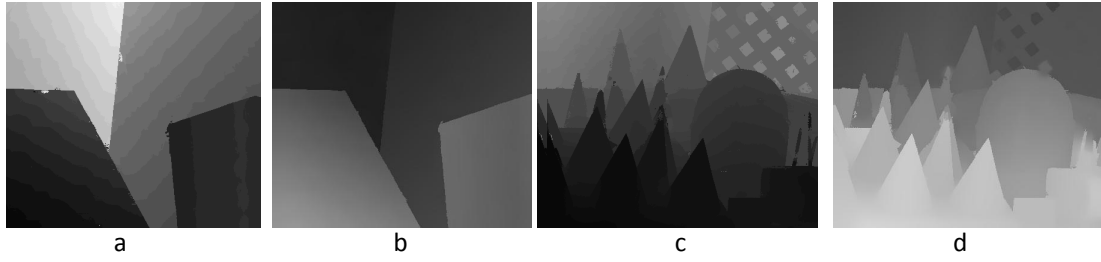
olarak tanımlanmıştır. κ_2 yayının yönbağımlılığını veya yönbağımsızlığı ayarlar. Başlangıçta Δd yüksek bir değer alacağından ($\Delta d \gg \kappa_2$) h fonksiyonunun değeri 0 olarak hesaplanır. Dolayısı ile yayının fonksiyonunun değeri de 1 olarak hesaplanır ki bu da yayının yön bağımsız olduğunu gösterir. Eniyileme ilerledikçe ağlar birbi-
rine doğru ilerlerler. Eniyilemenin sonlarına doğru ağlar aynı fark haritasını üretmeye
başlar ($\Delta d/\kappa_2 \approx 0$). Bu anda h fonksiyonun değeri 1 olur. Buna bağlı olarak yayının
fonksiyonu da yönbağımlı yayımında olduğu gibi $g(\nabla D_f)$ fonksiyonuna dönmüş olur.

5.3 Deneyler

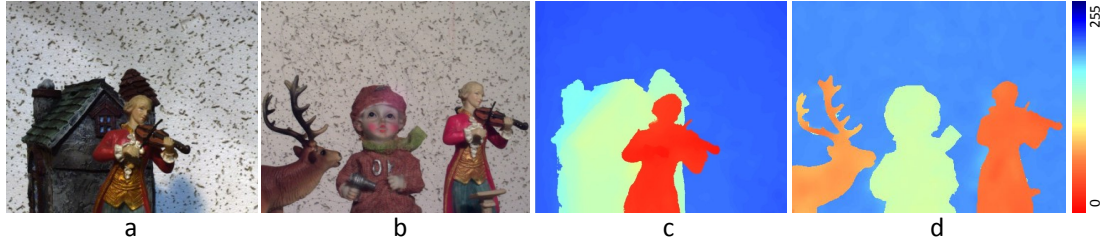
Önerilen yöntem gerçek sahne imgeleriyle ve derinlik değerleri bilinen gerçek sahnelerin Iris filtresi [Sakurai, 2004] ile oluşturulan sentetik imgelerle test edilmiştir. Yöntemlerin performansını sayısal olarak göstermeyi mümkün kılan derinlik değerleri bilinen gerçek sahne imgeleri Middlebury [Scharstein and Szeliski, 2003b] stereo imge veri tabanından alınmıştır. Bu veri tabanında bulunan Venus ve Cones imge setleri için farklı odak ayarları ile 16 sentetik imge üretilmiştir. Önerilen yöntem ile bulunan sahne derinlik değerlerinin hata oranları sahnenin tüm yüzeyi, sadece görünen yüzeyi ve sadece yüzey süreksizliklerin olduğu bölgelerde ayrı ayrı hesaplanmıştır. Çizelge 5.1 sahnenin katmanlı yapısı ve imge bölütü kullanılarak bulunan sonuçların hata oranlarını göstermektedir. Şekil 5.2 Venus ve Cones sahneleri için katmanlı gösterimi ve katmanlı yapı kullanılarak elde edilen fark haritalarını göstermektedir. Sonuçlardan da anlaşılacağı gibi katmanlı yapı kullanımı sonuçları oldukça iyileştirmektedir.

Çizelge 5.1 Sahnenin katmanlı yapısı ve imge bölütleri kullanılarak elde edilen sonuçların hata oranları. Hata oranları sahnenin tüm yüzeyi, görünür yüzeyi ve yüzey süreksizlik bölgeleri için hata eşiği 1 piksel ve 0.5 piksel seçilerek ayrı ayrı hesaplanmıştır.

hata eşiği	Cones											
	Venus						Cones					
	1 piksel			0.5 piksel			1 piksel			0.5 piksel		
	gör	tüm	srksz	gör	tüm	srksz	gör	tüm	srksz	gör	tüm	srksz
katmanlı sahne yapısı ile	0.10	0.25	1.41	0.5	0.89	5.30	5.74	12.0	16.4	8.24	15.0	21.3
imge bölütü ile	0.32	0.40	3.45	1.23	1.52	9.55	7.08	14.5	19.7	8.34	16.2	22.4



Şekil 5.2 Venus (a) ve Cones (a) sahnelerinin katmanlı gösterilimi. Katmanlı yapı kullanılarak bulunan fark haritaları (b,d).



Şekil 5.3 Sahnelerin net görüntüleri içeren imgeler (a,b), ve önerilen yöntem ile bulunan fark haritaları (c,d).

Gerçek sahneler üzerindeki deneyler 25mm'lik lens kullanan kamera sistemleri ile gerçekleştirilmiştir. Sahnenin katmanlı yapısını çıkarmak için kameranın odak ayarını aşamalı olarak değiştirerek 10 farklı imge elde edilmiştir. Odak ölçüt operatörü ve stereo sistemi için gerekli olan sahnelerin net görüntüsü kameranın en düşük diyafram açıklığıyla elde edilmiştir. Önerilen yöntem ile çıkarılan sahne fark haritaları Şekil 5.3'te gösterilmiştir.

6 Sonuç

Stereo eşleşme problemi ters bir problem olduğu için, çoğu ters problemde olduğu gibi, problemin matematiksel ifadesi kötü konumlandırılmıştır. Bu tip problemlerin çözümünde başvurulmuş en yaygın yöntem çözüm üzerine kısıt eklemek veya çözüm hakkında ek bilgi kullanmak suretiyle regülarizasyon uygulamaktır.

Bu tezde, çifte ağ olarak adlandırılan, birbirine bağımlı iki stereo enerji fonksiyonunun eşzamanlı eniyilemesine dayalı yeni bir yöntem geliştirilmiştir. Sistemle beraber önerilen regülarizasyon yöntemi imge bölütlerini kullanarak yüzey süreksizliklerinin başarılı bir şekilde çıkarılmasını sağlar. Ayrıca, regülarizasyon yöntemi ağların bulundukları konumları göz önünde bulundurarak uyarlanabilir düzleştirme uygulamaktadır. Sonuç olarak yönbağımsız olarak başlayan yayılım, eniyilemenin ileri safhalarında yönbağımlı yayılım olarak devam eder.

Önerilen yöntem yaklaşık eniyileme yöntemlerinde bulunan başlangıç koşullarına bağımlılık, yerel minimumlara takılma gibi problemlere çözüm getirmiştir. Bu yüzden elde edilen fark haritaları tek bir eniyileme yöntemi ile elde edilen fark haritalarından daha yüksek başarıma sahiptir. Bunun yanında, önerilen yöntem yinelemeli eniyileme yöntemlerinde önemli bir problem olan eniyilemeyi durdurma şartı seçimi problemine de çözüm getirmiştir. Eniyileme sürecince ağların aynı fark haritalarını üreten bölgelerindeki imge bölütleri arasındaki yayılım tamamen kesilir. Bu sayede, en iyi çözüm bulunduğundan sonra devam edilen gereksiz eniyileme yüzey süreksizliklerinin aşırı düzleştirilmesine neden olmaz.

Önerilen yöntemin başarımı Middlebury sitesindeki gerçek fark haritaları bilinen sahneler üzerinde test edilmiştir. Çıkan sonuçlar, yöntemin gelişmiş stereo yöntemleriyle karşılaştırılabilir olduğunu göstermiştir.

Bu tezde asıl amaçlanan hedef stereo ile odaktan şekil yöntemlerini birleştirmektir. Bu amaca ulaşabilmek için odaktan şekil yönteminde de bulunan kapanma ve yüzey süreksizlikleri problemlerinin çözülmesi gerekir. Bu yüzden bu çalışmada yeni bir odak ölçüt operatörü geliştirilmiştir. Geliştirilen bu operatör, var olan operatörlerden farklı olarak sahnenin net görüntüsünden faydalanır. Bu yaklaşım açık diyaframlı

kameralara has kapanma problemine etkili bir çözüm getirir. Ayrıca geliştirilen uyarlanabilir destek pencereleri kullanılarak hesaplanan odak ölçüt değerleri üzerinde düzeltme işlemi uygulanır. Bu sayede diğer odaktan şekil yöntemlerinin göz ardı ettiği yüzey süreksizlikleri de doğru olarak bulunur.

Gerçek ve sentetik sahneler üzerinde yapılan deneyler önerilen odak ölçüt operatörünün yüzey süreksizliklerini ve kapanma bölgelerini diğer odaktan şekil yöntemlerinden çok daha iyi bulduğunu göstermektedir. Bunun yanında uyarlanabilir destek pencerelerinin boyutunun artırılması, diğer odaktan şekil yöntemlerinin tersine, yöntemin başarımını arttırmaktadır.

Geliştirilen odaktan şekil yönteminin bir diğer avantajı da şudur. Yöntem sahnenin net görüntüsü için derinlik haritası ürettiği için yöntemin stereo yöntemleriyle birleştirilmesi için herhangi bir piksel kayıtlanma işlemine ihtiyaç duyulmaz.

Yöntemin en önemli kusuru olarak, her piksel için farklı uyarlanabilir pencere oluşturulduğundan, hesaplama maliyeti gösterilebilir. Ancak hesaplama süreci çift taraflı ızgara [Chen et al., 2007] gibi probleme uygun veri yapısı kullanılarak hızlandırılabilir. Yöntemle ilgili bir diğer problem de sahnenin net görüntüsünün açık diyaframlı kameralarla verimli bir şekilde elde edilememesidir. Bu yüzden üretilen derinlik değerleri yüksek standart sapmaya sahiptir. Bu durum, bulunan derinlik değerleri civarında frekans tabanlı odak ölçüt operatörü değerleri üzerinden arama yapılarak çözülebilse de, stereo yönteminde kullanılabilmesi için çözülmesi gereken bir problem değildir. Üretilen derinlik haritası stereo yönteminde sahnenin katmanlı yapısı olarak kullanılmıştır.

Önerilen odak ölçüt operatörüyle çıkarılan sahnenin katmanlı yapısı imge bölütlerine nazaran sahne hakkında daha fazla bilgiye sahiptir. Çıkarılan katmanlı yapıda, imge bölütlerinde bulunmayan, katmanlar arası derinliğe bağlı bir sıralama vardır. Katmanlı yapının bu özelliğinden yararlanarak farklı derinlikteki yüzeyler arası yayılım daha tutarlı bir şekilde ayarlanabilir. Bu amaçla sahnenin katmanlı yapısını ve çifte ağ yönteminin avantajlarını kullanan regülerizasyon yöntemi önerilmiştir. Önerilen yöntem yüzey süreksizliklerinin olduğu bölgelerdeki yayılım miktarını katmanlar arasındaki derinlik farkına göre ayarlar.

Katmanlı yapının imge bölütlerine karşı sahip olduğu bir diğer önemli avantajı da stereo imgelerde bulunan kapanma bölgelerinde sahnenin yüzey yapısı hakkında bilgi içermesidir. Önerilen yöntemde kapanma bölgelerine açık bir çözüm getirilmesine rağmen, katmanlı yapının bu özelliğinden dolayı, yöntemin kapanma bölgelerindeki başarımı kabul edilebilir düzeydedir. Bu durumun bölüt tabanlı regülarizasyon yöntemlerinde oluşması mümkün değildir.

Son olarak, önerilen yöntemin eşzamanlı çalışan iki eniyileme işlemini içermesi hesaplama maliyetini diğer stereo yöntemlerine nazaran iki kat arttırmaktadır. Sistemin yakınsama süresini çok ölçekli yaklaşımlarla azaltmak mümkün olsa da, bu yaklaşımların önerilen yönteme uyarlaması aşikâr değildir. Çifte ağ için uyarlanmış örnek çok ölçekli eniyileme yöntemi Akgul and Kambhamettu [1999]'nun çalışmasında bulunmaktadır. Sistemi hızlandırmak için getirilebilecek bir diğer çözüm de yöntemi GPU mimarileri üzerinde uygulamaktır çünkü KDD tabanlı çözümler paraleleleştirilmeye oldukça uygundur.

KAYNAKLAR

Ahmad, M. and Choi, T.-S. (2005). A heuristic approach for finding best focused shape. *Circuits and Systems for Video Technology, IEEE Transactions on*, 15(4):566 – 574.

Ahuja, N. and Abbott, A. (1993). Active stereo: Integrating disparity, vergence, focus, aperture and calibration for surface estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(10):1007–1029.

Akgul, Y. S. and Kambhamettu, C. (1999). Recovery and tracking of continuous 3d surfaces from stereo data using a deformable dual-mesh. In *International Conference on Computer Vision*, pages 765–772.

Alvarez, L., Deriche, R., Sanchez, J., and Weickert, J. (2002). Dense disparity map estimation respecting image discontinuities: A pde and scale-space based approach. *Journal of Visual Communication and Image Representation*, 13(1/2):3–21.

Asada, N., Fujiwara, H., and Matsuyama, T. (1995). Seeing behind the scene: analysis of photometric properties of occluding edges by the reversed projection blurring model. *Computer Vision, IEEE International Conference on*, 0:150.

Asada, N., Fujiwara, H., and Matsuyama, T. (1998). Seeing behind the scene: Analysis of photometric properties of occluding edges by the reversed projection blurring model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:155–167.

Asif, M. and Choi, T.-S. (2001). Shape from focus using multilayer feedforward neural networks. *Image Processing, IEEE Transactions on*, 10(11):1670 – 1675.

Aslantas, V. and Kurban, R. (2009). A comparison of criterion functions for fusion of multi-focus noisy images. *Optics Communications*, 282(16):3231 – 3242.

Ayache, N. and Hansen, C. (1988). Rectification of images for binocular and trinocular stereovision. In *Pattern Recognition, 1988., 9th International Conference on*, pages 11 –16 vol.1.

Aydin, T. and Akgul, Y. (2006). 3d structure recovery from stereo using synchronous optimization processes. In *BMVC06*, page III:1179.

Aydin, T. and Akgul, Y. (2008). A new adaptive focus measure for shape from focus. In *British Machine Vision Conference*.

Aydin, T. and Akgul, Y. S. (2010a). An occlusion insensitive adaptive focus measurement method. *Opt. Express*, 18(13):14212–14224.

Aydin, T. and Akgul, Y. S. (2010b). Stereo depth estimation using synchronous optimization with segment based regularization. *Pattern Recognition Letters*, 31(15):2389 – 2396.

Baina, J. and Dublet, J. (1995). Automatic focus and iris control for video cameras. pages 232 –235.

Ben-Ari, R. and Sochen, N. (2008). A geometric approach for regularization of the data term in stereo-vision. *J. Math. Imaging Vis.*, 31(1):17–33.

Ben-Ari, R. and Sochen, N. (2010). Stereo matching with mumford-shah regularization and occlusion handling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 99(Preliminary).

Bhasin, S. S. and Chaudhuri, S. (2001). Depth from defocus in presence of partial self occlusion. *Computer Vision, IEEE International Conference on*, 1:488.

Blake, A. and Zisserman, A. (1987). *Visual Reconstruction*. MIT Press.

Bleyer, M. and Gelautz, M. (2005a). Graph-based surface reconstruction from stereo pairs using image segmentation. In *SPIE Symposium on Electronic Imaging (Videometrics VIII)*, volume 5665, pages 288–299.

Bleyer, M. and Gelautz, M. (2005b). A layered stereo matching algorithm using image segmentation and global visibility constraints. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(3):128–150.

Bleyer, M. and Gelautz, M. (2007). Graph-cut-based stereo matching using image segmentation with symmetrical treatment of occlusions. *Image Commun.*, 22(2):127–143.

Bleyer, M. and Gelautz, M. (24-27 Oct. 2004). A layered stereo algorithm using image segmentation and global visibility constraints. *Image Processing, 2004. ICIP '04. 2004 International Conference on*, 5:2997–3000 Vol. 5.

Born, M. and Wolf, E. (1965). *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Pergamon, NY, third edition.

Bove, V. M. J. (1990). Probabilistic method for integrating multiple sources of range data. *Journal of the Optical Society of America A*, 7:2193–2198.

Boykov, Y., Veksler, O., and Zabih, R. (2001). Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222 – 1239.

Braillon, C., Usher, K., Pradalier, C., Crowley, J., and Laugier, C. (2006). Fusion of stereo and optical flow data using occupancy grids. In *Intelligent Transportation Systems Conference, 2006. ITSC '06. IEEE*, pages 1240 –1245.

Charfi, M., Nyeck, A., and Tosser, A. (1991). Focusing criterion. *Electronics Letters*, 27(14):1233 –1235.

Chen, J., Paris, S., and Durand, F. (2007). Real-time edge-aware image processing with the bilateral grid. In *SIGGRAPH 07*, page 103, New York, NY, USA. ACM.

Choi, K.-S., Lee, J.-S., and Ko, S.-J. (1999). New autofocusing technique using the frequency selective weighted median filter for video cameras. *Consumer Electronics, IEEE Transactions on*, 45(3):820 –827.

Comaniciu, D. and Meer, P. (2002). Mean shift: a robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(5):603–619.

Cryer, J., Tsai, P.-S., and Shah, M. (1993). Integration of shape from x modules: combining stereo and shading. In *Computer Vision and Pattern Recognition, 1993. Proceedings CVPR '93., 1993 IEEE Computer Society Conference on*, pages 720–721.

Evans, M. and Ferryman, J. (2006). Cross validation and segment support for stereo belief propagation. *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, 1:115–118.

Favaro, P. and Soatto, S. (2003). Seeing beyond occlusions (and other marvels of a finite lens aperture). *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 2:579.

Favaro, P. and Soatto, S. (2007). *3-D Shape Estimation and Image Restoration: Exploiting Defocus and Motion-Blur*. Springer London.

Favaro, P., Soatto, S., Burger, M., and Osher, S. J. (2008). Shape from defocus via diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(3):518–531.

Feris, R., Raskar, R., Chen, L., Tan, K.-H., and Turk, M. (2008). Multiflash stereopsis: Depth-edge-preserving stereo with small baseline illumination. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(1):147–159.

Flusser, J. and Suk, T. (1998). Degraded image analysis: an invariant approach. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(6):590–603.

Frese, C. and Gheeta, I. (2006). Robust depth estimation by fusion of stereo and focus series acquired with a camera array. In *Multisensor Fusion and Integration for Intelligent Systems, 2006 IEEE International Conference on*, pages 243–248.

Fusiello, A., Trucco, E., and Verri, A. (2000). A compact algorithm for rectification of stereo pairs. *Machine Vision and Applications*, 12:16–22. 10.1007/s001380050120.

Gerrits, M. and Bekaert, P. (2006). Local stereo matching with segmentation-based outlier rejection. In *CRV '06: Proceedings of the The 3rd Canadian Conference on Computer and Robot Vision (CRV'06)*, page 66.

Groen, F. C., Young, I. T., and Ligthart, G. (1985). A comparison of different focus functions for use in autofocus algorithms. *Cytometry*, 6(2):81–91.

Grosky, W. and Tamburino, L. (1990). A unified approach to the linear camera calibration problem. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(7):663 –671.

Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 13:40–52.

Helmli, F. and Scherer, S. (2001). Adaptive shape from focus with an error estimation in light microscopy. *Image and Signal Processing and Analysis*, pages 188–193.

Hong, L. and Chen, G. (2004). Segment-based stereo matching using graph cuts. In *IEEE Computer Vision and Pattern Recognition or CVPR*, pages I: 74–81.

Horn, B. K. P. (1986). *Robot Vision (MIT Electrical Engineering and Computer Science)*. The MIT Press, mit press ed edition.

Huang, J.-T., Shen, C.-H., Phoong, S.-M., and Chen, H. (2005). Robust measure of image focus in the wavelet domain. pages 157 – 160.

Huang, W. and Jing, Z. (2007). Evaluation of focus measures in multi-focus image fusion. *Pattern Recognition Letters*, 28(4):493 – 500.

Indebetouw, G. and Poon, T.-C. (1986). Parallel synthesis of bipolar point spread functions in a scanning heterodyne optical system. *Optica Acta: International Journal of Optics*, 33(7):827–834.

Intille, S. S. and Bobick, A. F. (1994). Disparity-space images and large occlusion stereo. In *ECCV '94: Proceedings of the third European conference on Computer Vision (Vol. II)*, pages 179–186, Secaucus, NJ, USA. Springer-Verlag New York, Inc.

Ishikawa, H. (2003). Exact optimization for markov random fields with convex priors. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(10):1333–1336.

Jarvis, R. (1983). A perspective on range-finding techniques for computer vision. *Pattern Analysis and Machine Intelligence*, 3:122–139.

Jutamulia, S., Asakura, T., Bahuguna, R. D., and DeGuzman, P. C. (1994). Autofocusing based on power-spectra analysis. *Appl. Opt.*, 33(26):6210–6212.

Kambhamettu, C., Palaniappan, K., and Hasler, A. (1995). Coupled, multi-resolution stereo and motion analysis. In *Symposium on Computer Vision*, pages 43–48.

Kanade, T. and Okutomi, M. (1994). A stereo matching algorithm with an adaptive window: Theory and experiment. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16(9):920–932.

Karaman, M. and Bilge, H. (2009). Adaptive fusion of shape from shading and photometric stereo. In *Signal Processing and Communications Applications Conference, 2009. SIU 2009. IEEE 17th*, pages 892–895.

Kautsky, J., Flusser, J., Zitová, B., and Simberová, S. (2002). A new wavelet-based measure of image focus. *Pattern Recognition Letters*, 23(14):1785–1794.

Kim, H., Choe, Y., and Sohn, K. (2004). Disparity estimation using a region-dividing technique and energy-based regularization. *Optical Engineering*, 43(8):1882–1890.

Kim, J. and Sikora, T. (2007). Confocal disparity estimation and recovery of pinhole image for real-aperture stereo camera systems. In *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, volume 5, pages V–229–V–232.

Klaus, A., Sormann, M., and Karner, K. (2006). Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. In *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition*, pages 15–18.

Kolmogorov, V. and Zabih, R. (2001). Computing visual correspondence with occlusions via graph cuts. In *International Conference on Computer Vision*, pages II: 508–515.

Kolmogorov, V. and Zabih, R. (2002). Multi-camera scene reconstruction via graph cuts. In *European Conference on Computer Vision*, page III: 82 ff.

Kolmogorov, V. and Zabih, R. (2004). What energy functions can be minimized via graph cuts? *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(2):147–159.

Kristan, M., Pers, J., Perse, M., and Kovacic, S. (2006). A bayes-spectral-entropy-based measure of camera focus using a discrete cosine transform. *Pattern Recognition Letters*, 27(13):1431 – 1439.

Krotkov, E. (1987). Focusing. *Int. J. Comput. Vision*, 1(3):223–237.

Krotkov, E. and Bajcsy, R. (1993). Active vision for reliable ranging: Co-operating focus, stereo, and vergence. *International Journal of Computer Vision*, 11(2):187–203.

Lange, H. (1999). Advances in the cooperation of shape from shading and stereo vision. In *3-D Digital Imaging and Modeling, 1999. Proceedings. Second International Conference on*, pages 46 –58.

Lee, S., Park, S., Kim, C., Kumar, Y., and Kim, S. (2006). Low-power auto focus algorithm using modified dct for the mobile phones. pages 67 – 68.

Li, Y., Ohmura, I., Takauji, H., Kaneko, S., and Tanaka, T. (2007). Robust focusing by orientation code matching. pages 752 –758.

Mahmood, M. and Choi, T.-S. (2010). Focus measure based on the energy of high-frequency components in the s transform. *Opt. Lett.*, 35(8):1272–1274.

Mahmood, M., Choi, W.-J., and Choi, T.-S. (2008). Dct and pca based method for shape from focus. In *Computational Science and Its Applications ICCSA 2008*, volume 5073 of *Lecture Notes in Computer Science*, pages 1025–1034. Springer Berlin / Heidelberg.

Mahmood, M. T., Shim, S.-O., and Choi, T.-S. (2009). Shape from focus using principal component analysis in discrete wavelet transform. *Optical Engineering*, 48(5):057203.

Malik, A. S. and Choi, T.-S. (2008). A novel algorithm for estimation of depth map using image focus for 3d shape recovery in the presence of noise. *Pattern Recognition*, 41(7):2200 – 2225.

Marshall, J. A., Burbeck, C. A., Ariely, D., Rolland, J. P., and Martin, K. E. (1996). Occlusion edge blur: a cue to relative visual depth. *J. Opt. Soc. Am. A*, 13(4):681–688.

Mendapara, P., Minhas, R., and Wu, Q. (2009). Depth map estimation using exponentially decaying focus measure based on susan operator. pages 3705 –3708.

Meneses, J., Suarez, M. A., Braga, J., and Gharbi, T. (2008). Extended depth of field using shapelet-based image analysis. *Appl. Opt.*, 47(2):169–178.

M.Subbarao, T. and A.Nikzad (1993). Focusing technique. *Image and Signal Processing and Analysis*, 32(11):2824–2836.

Muhammad, M. S. and Choi, T.-S. (2010). A novel method for shape from focus in microscopy using bezier surface approximation. *MICROSCOPY RESEARCH AND TECHNIQUE*, 73(2):140–151.

Nair, H. and Stewart, C. (1992). Robust focus ranging. In *CVPR92*, pages 309–314.

Nayar, S. and Nakagawa, Y. (1994). Shape from Focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(8):824–831.

Nayar, S. K., Watanabe, M., and Noguchi, M. (1996). Real-time focus range sensor. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:1186–1198.

Papadimitriou, D. and Dennis, T. (1996). Epipolar line estimation and rectification for stereo image pairs. *Image Processing, IEEE Transactions on*, 5(4):672–676.

Pentland, A., Scherrock, S., Darrell, T., and Girod, B. (1994). Simple range cameras based on focal error. *J. Opt. Soc. Am. A*, 11(11):2925–2934.

Pentland, A. P. (1987). A new sense for depth of field. *IEEE Trans. Pattern Anal. Mach. Intell.*, 9(4):523–531.

Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-12(7):629–639.

Pock, T., Schoenemann, T., Graber, G., Bischof, H., and Cremers, D. (2008). A convex formulation of continuous multi-label problems. In *European Conference on Computer Vision (ECCV)*, Marseille, France.

Pock, T., Zach, C., and Bischof, H. (2007). Mumford-shah meets stereo: Integration of weak depth hypotheses. *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pages 1–8.

Poon, T.-C. (1985). Method of two-dimensional bipolar incoherent image processing by acousto-optic two-pupil synthesis. *Opt. Lett.*, 10(5):197–199.

Potts, R. (1952). Some generalized order-disorder transformation. In *Proceedings of the Cambridge Philosophical Society*, volume 48, pages 106–109.

Rajagopalan, A. N. and Chaudhuri, S. (1997). A variational approach to recovering depth from defocused images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(10):1158–1164.

Rajagopalan, A. N. and Mudenagudi, U. (2004). Depth estimation and image restoration using defocused stereo pairs. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(11):1521–1525. Senior Member-S. Chaudhuri.

Raskar, R., Tan, K.-H., Feris, R., Yu, J., and Turk, M. (2004). Non-photorealistic camera: depth edge detection and stylized rendering using multi-flash imaging. *ACM Trans. Graph.*, 23(3):679–688.

Robert, L. and Deriche, R. (1996). Dense depth map reconstruction: A minimization and regularization approach which preserves discontinuities. In *ECCV '96*:

Proceedings of the 4th European Conference on Computer Vision-Volume I, pages 439–451, London, UK. Springer-Verlag.

Roy, S. and Cox, I. J. (1998). A maximum-flow formulation of the n-camera stereo correspondence problem. In *ICCV*, pages 492–502.

Saito, H. and Mori, M. (1995). Application of genetic algorithms to stereo matching of images. *Pattern Recogn. Lett.*, 16(8):815–821.

Sakurai, R. (2004). Irisfilter, <http://www.reiji.net/>.

Scharstein, D. and Szeliski, R. (1998). Stereo matching with nonlinear diffusion. *Int. J. Comput. Vision*, 28(2):155–174.

Scharstein, D. and Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47(1-3):7–42.

Scharstein, D. and Szeliski, R. (2003a). High-accuracy stereo depth maps using structured light. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 1, pages I–195 – I–202 vol.1.

Scharstein, D. and Szeliski, R. (2003b). High-accuracy stereo depth maps using structured light. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 1, pages I–195–I–202 vol.1.

Schechner, Y. and Kiryati, N. (2000). Depth from defocus vs. stereo: How different really are they? *International Journal of Computer Vision*, 39(2):141–162.

Shah, J. (1993). A nonlinear diffusion model for discontinuous disparity and half-occlusions in stereo. *Computer Vision and Pattern Recognition, 1993. Proceedings CVPR '93., 1993 IEEE Computer Society Conference on*, pages 34–40.

Shen, C.-H. and Chen, H. (2006). Robust focus measure for low-contrast images. pages 69 – 70.

Shim, S.-O. and Choi, T.-S. (2010). A novel iterative shape from focus algorithm based on combinatorial optimization. *Pattern Recognition*, 43(10):3338 – 3347.

Smith, S. M. and Brady, J. M. (1997). Susan a new approach to low level image processing. *International Journal of Computer Vision*, 23:45–78. 10.1023/A:1007963824710.

Stewart, C. and Nair, H. (1989). New results in automatic focusing and a new method for combining focus and stereo. In *In Proc. SPIE Sensor Fusion II: Human and Machine Strategies*, volume 1198, page 102i₆¹/₂113.

Stockwell, R., Mansinha, L., and Lowe, R. (1996). Localization of the complex spectrum: the s transform. *Signal Processing, IEEE Transactions on*, 44(4):998 –1001.

Subbarao, M. and Choi, T. (Mar 1995). Accurate recovery of three-dimensional shape from image focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(3):266–274.

Subbarao, M., Choi, T., and Nikzad, A. (1992). Focusing techniques. *Journal of Optical Engineering*, 32:2824–2836.

Subbarao, M. and Surya, G. (1994). Depth from defocus: A spatial domain approach. *International Journal of Computer Vision*, 13(3):271–294.

Subbarao, M. and Tyan, J.-K. (1998). Selecting the optimal focus measure for autofocus and depth-from-focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8):864 –870.

Subbarao, M., Wei, T.-C., and Surya, G. (1995). Focused image recovery from two defocused images recorded with different camera settings. *Image Processing, IEEE Transactions on*, 4(12):1613 –1628.

Subbarao, M., Yuan, T., and kwei Tyan, J. (1997). Integration of defocus and focus analysis with stereo for 3d shape recovery. In *Shape Recovery, Three-Dimensional Imaging and Laser-based Systems for Metrology and Inspection III*, pages 14–15.

Sun, J., Zheng, N.-N., and Shum, H.-Y. (2003). Stereo matching using belief propagation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(7):787–800.

Tappen, M. F. and Freeman, W. T. (2003). Comparison of graph cuts with belief propagation for stereo, using identical mrf parameters. In *ICCV '03: Proceedings of the Ninth IEEE International Conference on Computer Vision*, page 900.

Tenenbaum, J. M. (1971). *Accommodation in computer vision*. PhD thesis, Stanford, CA, USA.

Terzopoulos, D. (1986). Regularization of inverse visual problems involving discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(4):413–424.

Tian, Y., Shieh, K., and Wildsoet, C. F. (2007). Performance of focus measures in the presence of nondefocus aberrations. *J. Opt. Soc. Am. A*, 24(12):B165–B173.

Tikhonov, A. N. and Arsenin, V. Y. (1977). *Solutions of Ill-Posed Problems*. V. H. Winston & Sons, Washington, D.C.: John Wiley & Sons, New York,.

Tomasi, C. and Manduchi, R. (1998). Bilateral filtering for gray and color images. In *ICCV '98: Proceedings of the Sixth International Conference on Computer Vision*, page 839, Washington, DC, USA. IEEE Computer Society.

V. Michael Bove, J. (1993). Entropy-based depth from focus. *J. Opt. Soc. Am. A*, 10(4):561–566.

Waldmann, J. and Merhav, S. (1992). Fusion of stereo and motion vision for 3-d reconstruction. In *Pattern Recognition, 1992. Vol.I. Conference A: Computer Vision and Applications, Proceedings., 11th IAPR International Conference on*, pages 5–8.

Wei, Y. and Quan, L. (2004). Region-based progressive stereo matching. *CVPR 2004: Proceedings of the 2004 IEEE Computer Society Conference on*, 1:106–113.

Woodford, O., Torr, P., Reid, I., and Fitzgibbon, A. (2008). Global stereo reconstruction under second order smoothness priors. *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8.

Woodham, R. J. (1980). Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144.

Xiong, Y. and Shafer, S. (1995). Moment filters for high precision computation of focus and stereo. *Intelligent Robots and Systems, IEEE/RSJ International Conference on*, 3:3108.

Xiong, Y. and Shafer, S. (1997). Moment and hypergeometric filters for high precision computation of focus, stereo and optical flow. *International Journal of Computer Vision*, 22(1):25–59.

Xiong, Y. and Shafer, S. A. (1994). Variable window gabor filters and their use in focus and correspondence. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pages 668–671.

Yang, G. and Nelson, B. (2003). Wavelet-based autofocusing and unsupervised segmentation of microscopic images. volume 3, pages 2143 – 2148 vol.3.

Yang, Q., Yang, R., Davis, J., and Nister, D. (2007). Spatial-depth super resolution for range images. *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pages 1–8.

Yoon, K.-J. and Kweon, I. S. (2006). Adaptive support-weight approach for correspondence search. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(4):650–656.

Yuk-Kwan Ho, A. and Pong, T.-C. (1994). Cooperative fusion of stereo and motion. In *Speech, Image Processing and Neural Networks, 1994. Proceedings, IS-SIPNN '94., 1994 International Symposium on*, pages 292 –295 vol.1.

Yun, J. and Choi, T. (1999). Accurate 3-d shape recovery using curved window focus measure. *ICIP 99*, 3:910–914 vol.3.

Zhang, L., Curless, B., and Seitz, S. M. (2003). Spacetime stereo: Shape recovery for dynamic scenes. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 367–374.

Zhang, Y. and Kambhamettu, C. (2002). Stereo matching with segmentation-based cooperation. In *ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part II*, pages 556–571.

Zhang, Y., Zhang, Y., and Wen, C. (2000). A new focus measure method using moments. *Image and Vision Computing*, 18(12):959 – 965.

Zickler, T. E., Belhumeur, P. N., and Kriegman, D. J. (2002). Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. *Int. J. Comput. Vision*, 49(2-3):215–227.

ÖZGEÇMİŞ

1978 yılında İstanbul’da doğdu. İlk ve orta öğretimini 1984–1992 yılları arasında sırasıyla Hüvviyet Bekir İlkokulu, Cihangir Ortaokulunda tamamladı. 1992 yılında girdiği Mersin Fenlisesini bitirdikten sonra girdiği Boğaziçi Üniversitesi Fizik Bölümünden 2000 yılında mezun oldu. 2004 yılında Gebze Yüksek Teknoloji Enstitüsü Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü’nde doktora eğitimine başladı. Halen Bahçeşehir Üniversitesinde öğretim görevlisi olarak çalışmaktadır.