

**ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

**UYDU GÖRÜNTÜLERİ ÜZERİNDE ROTASYON, ÖLÇEKLEME VE
ÖTELEME DEĞİŞMEZLİKLİ NESNE TANIMA**

Yusuf SOYMAN

ELEKTRİK - ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

ANKARA

2013

Her hakkı saklıdır.

TEZ ONAYI

Yusuf SOYMAN tarafından hazırlanan “**Uydu Görüntüleri Üzerinde Rotasyon, Ölçekleme ve Öteleme Değişmezlikli Nesne Tanıma**” adlı tez çalışması 16.09.2013 tarihinde aşağıdaki jüri tarafından oy birliği ile Ankara Üniversitesi Elektrik – Elektronik Mühendisliği Anabilim Dalı’nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman : Yrd. Doç. Dr. Hakkı Alparslan ILGIN
Ankara Üniversitesi, Elektrik – Elektronik Mühendisliği Anabilim Dalı

Jüri Üyeleri :

Başkan : Doç. Dr. Süleyman TOSUN
Ankara Üniversitesi, Bilgisayar Mühendisliği Anabilim Dalı

Üye : Prof. Dr. H. Gökhan İLK
Ankara Üniversitesi, Elektrik – Elektronik Mühendisliği Anabilim Dalı

Üye : Yrd. Doç. Dr. Hakkı Alparslan ILGIN
Ankara Üniversitesi, Elektrik – Elektronik Mühendisliği Anabilim Dalı

Yukarıdaki sonucu onaylarım.

Prof. Dr. İbrahim DEMİR

Enstitü Müdürü

ÖZET

Yüksek Lisans Tezi

UYDU GÖRÜNTÜLERİ ÜZERİNDE ROTASYON, ÖLÇEKLEME VE ÖTELEME DEĞİŞMEZLİKLİ NESNE TANIMA

Yusuf SOYMAN

Ankara Üniversitesi

Fen Bilimleri Enstitüsü

Elektrik – Elektronik Mühendisliği Anabilim Dalı

Danışman: Yrd. Doç. Dr. Hakkı Alparslan ILGIN

Nesne tanıma üzerine literatürde bütünsel yöntemlerden dönüşüm – arama tabanlı yöntemlere çok sayıda çalışma bulunmaktadır. Bu yöntemler rotasyon, ölçekleme, öteleme, kapatma ve arka plan gürültüsü gibi etkiler karşısında yüksek yanlış kabul oranıyla nesne tanımayı gerçekleştirmektedir. Düşük yanlış kabul oranıyla nesne tanımayı gerçekleştirmek için rotasyon, ölçekleme ve öteleme etkilerinden bağımsız, kapatmaya ve arka plan gürültüsüne dayanıklı bir öznitelik tabanlı nesne tanıma yöntemi kullanılmıştır. Kullanılan yöntemi oluşturan öznitelik çıkartımı ve sınıflandırma gibi işlemler parametrik olarak incelenmiştir. Farklı parametreler kullanılarak gerçekleştirilen deneysel çalışmalar sonucunda performans başarımları elde edilmiştir. Karşılaştırmalı çalışma sonucunda bu parametrelerin nesne tanıma performansına etkileri belirtilmiştir.

Eylül 2013, 99 sayfa

Anahtar Kelimeler: Nesne tanıma, Rotasyon-Ölçekleme-Öteleme Değişmezlik, öznitelik çıkartımı, sınıflandırma, kümeleme, görsel kelimeler sözlüğü, uydu görüntüleme

ABSTRACT

Master Thesis

ROTATION, SCALING AND TRANSLATION INVARIANT OBJECT RECOGNITION ON SATELLITE IMAGERY

Yusuf SOYMAN

Ankara University

Graduate School of Natural and Applied Sciences

Department of Electrical – Electronic Engineering

Supervisor: Asst. Prof. Dr. Hakkı Alparslan ILGIN

In the literature of the object recognition, there have been various proposals ranging from global methods to transformational-search based methods; yet these methods do not promise object recognition with low false alarm rate under the impacts of rotation, scaling and translation (RST), occlusion and background clutter effects. In order to recognize the objects with low false alarm rate a feature-based object recognition method, which is RST invariance and robust to occlusion and background clutter, has been utilized. The components of the utilized method such as feature extraction and classification have been checked out parametrically. As a result of experimental studies with the usage of different parameters successful performances have been attained. In consequence of contrasting studies the effects of these parameters upon object recognition have been emphasized.

September 2013, 99 pages

Key Words: Object recognition, Rotation-Scale-Translation (RST) invariance, feature extraction, classification, clustering, Bag of Visual Word (BOVW), satellite imagery

Aileme,

TEŞEKKÜR

Çalışmalarım boyunca çalışmalarımı yönlendiren, pozitif tavsiyeleriyle motive olmamı sağlayan, araştırmalarımın her aşamasında bilgi, öneri ve yardımlarını esirgemeyerek akademik ortamda olduğu kadar beşeri ilişkilerde de engin fikirleriyle yetişme ve gelişmeye katkıda bulunan danışman hocam sayın Yrd.Doç.Dr. Hakkı Alparslan ILGIN'a (Ankara Üniversitesi Elektrik – Elektronik Mühendisliği) en içten duygularla teşekkür ederim.

Hayatımın bütün aşamalarında beni her zaman destekleyen, hiç bir fedakarlıktan kaçınmayan, karşılıksız ve sonsuz sevgisini esirgemeyen aileme, anneme, babama ve ablama tüm sevgilerimi sunarım.

Yusuf SOYMAN

Ankara, Eylül 2013

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT	ii
TEŞEKKÜR	iv
KISALTMALAR DİZİNİ	vii
ŞEKİLLER DİZİNİ	viii
ÇİZELGELER DİZİNİ	x
1. GİRİŞ	1
2. UYDU GÖRÜNTÜLEME	3
2.1 Ikonos Uydu Görüntüleri	4
2.2 Quickbird Uydu Görüntüleri	6
3. NESNE TANIMA.....	9
3.1 Evrensel Yöntemler.....	9
3.1.1 İki boyutlu ilinti.....	10
3.1.1.1 Ön işleme.....	11
3.1.1.2 Alt örnekleme veya görüntü piramitleme	12
3.1.1.3 Faz ilinitisi.....	13
3.1.2 Bütünsel öznitelik vektörleri	15
3.2 Dönüşüm - Arama Tabanlı Yöntemler	16
3.3 Öznitelik Tabanlı Yöntemler	17
3.3.1 Öznitelik algılayıcılar	19
3.3.1.1 Öznitelik özellikleri	30
3.3.1.1.1 Tekrarlanabilirlik miktarı	30
3.3.1.1.2 Ölçekten bağımsızlık.....	31
3.3.1.1.3 Rotasyon değişmezlik ve yönelim kestirimi	34
3.3.1.1.4 İlgin dönüşümünden bağımsızlık:	36
3.3.1.2 Öznitelik tanımlayıcılar	38

3.3.1.2.1 Yanlılık ve kazanç normalizasyonu	40
3.3.1.2.2 SIFT	41
3.3.1.2.3 PCA – SIFT.....	42
3.3.1.2.4 SURF	42
3.3.1.2.5 OpponentSIFT - OpponentSURF	43
3.3.1.2.6 Gradyan konum – yönelim histogramı	43
3.3.1.2.7 Yönlendirilebilir filtreler:.....	44
3.3.1.2.8 Yerel tanımlayıcıların performansı:.....	44
3.3.1.3 Öznitelik eşleştirme	46
3.3.1.3.1 Eşleştirme stratejisi ve hata oranı	47
3.4 Kullanılan Yöntem	53
3.4.1 Öznitelik çıkartımı	55
3.4.2 Görsel kelime sözlüğü	56
3.4.2.1 K-ortalama kümelemesi.....	57
3.4.2.2 K-ortalama++ kümeleme.....	60
3.4.3 Sınıflandırma	60
3.4.3.1 Bayes.....	61
3.4.3.2 Destek Vektör Makinaları.....	62
3.4.3.2.1 Doğrusal-olmayan sınıflandırma	67
4. DENEYSEL SONUÇLAR.....	69
5. SONUÇLAR	76
KAYNAKLAR	79
ÖZGEÇMİŞ.....	87

KISALTMALAR DİZİNİ

ACC	Accuracy
AUC	Area Under Curve
BOW	Bag of Word
BOVW	Bag of Visual Word
DoG	Difference of Gaussian
FPR	False Positive Rate
LoG	Laplacian of Gaussian
MOPS	Multi-Scale Oriented Patches
MSER	Maximally Stable Extremal Region
NNDR	Nearest Neighbor Distance Rate
PPV	Positive Predictive Value
RANSAC	Random Sample Consensus
ROC	Receiver Operating Characteristic
SIFT	Scale Invariant Feature Transform
SVM	Support Vector Machine
SURF	Speeded-Up Robust Features
TPR	True Positive Rate

ŞEKİLLER DİZİNİ

Şekil 3.1 Şablon görüntüsü	10
Şekil 3.2 Şablonun aranacağı görüntü.....	11
Şekil 3.3 Önişleme kullanarak iki boyutlu ilintileme.....	12
Şekil 3.4 Görüntü seyrek örnekleme	13
Şekil 3.5 Şablonun uzaysal ve frekans domaindeki görünümü.....	14
Şekil 3.6 Görüntünün uzaysal ve frekans domaindeki görünümü	14
Şekil 3.7 Faz ilintisi sonucu tespit edilen nesne.....	15
Şekil 3.8 Eşleştirecek iki görüntü çifti	17
Şekil 3.9 Görüntülerin analiz, tanımlama ve eşleştirmelerinde kullanılabilecek.....	19
Şekil 3.10 Görüntü çiftleri ve görüntülerden çıkartılan parçalar	21
Şekil 3.11 Farklı görüntü parçaları için açıklık problemi	22
Şekil 3.12 Gri seviye görüntü ve düzlem çizimi olarak gösterilen	24
Şekil 3.13 Oto-korelasyon matrisinin bir öz değeri analizine karşılık gelen	26
Şekil 3.14 Popüler anahtar nokta tespit etme fonksiyonların eş çevreleri	28
Şekil 3.15 İlgili operatörlerinin yanıtları	29
Şekil 3.16 Uyarlamalı Maksimum Olmayan Noktaların Bastırımı (ANMS)	30
Şekil 3.17 Çoklu ölçek yönelimli görüntü parçaları	32
Şekil 3.18 Gauss farklarının (DoG) alt oktav piramidi kullanılarak ölçek – uzam öz nitelik tespiti	33
Şekil 3.19 Baskın yönelim kestirimi	35
Şekil 3.20 İlgin bölge algılayıcı kullanımı	37
Şekil 3.21 İlgin normalizasyonu	37
Şekil 3.22 MSER kullanılarak çıkartılan ve eşleştirilen bölgeler	38
Şekil 3.23 MOPS tanımlayıcılar	39
Şekil 3.24 Öz nitelik eşleştirme	40
Şekil 3.25 SIFT tanımlayıcı gösterimi	41
Şekil 3.26 Gradyan konum – yönelim histogram tanımlayıcısı	44
Şekil 3.27 Öz nitelik tanımlayıcı performans değerlendirilmesi.....	45
Şekil 3.28 SIFT ve GLOH yöntemleri için uzamsal toplama blokları.....	46
Şekil 3.29 Gürültülü ortamda nesne tanıma	47

Şekil 3.30 Yanlış kabuller ve redler	48
Şekil 3.31 ROC eğrisi ve ilgili oranlar	51
Şekil 3.32 Sabit eşik değeri, en yakın komşu ve en yakın komşu	52
Şekil 3.33 Öznitelik tanımlayıcıların eşleştirme stratejisine göre.....	53
Şekil 3.34 Voronoi hücreleri	57
Şekil 3.35 k-ortalama gösterimi	59
Şekil 3.36 Sınıflar arası ayrım sağlayan doğrular	63
Şekil 3.37 Maksimum-sınır hiperdüzlemi.....	65
Şekil 3.38 Kernel makinası	68
Şekil 4.1 Eğitim veri kümesi.....	69
Şekil 4.2 Değerlendirme veri kümesi.....	70
Şekil 4.3 Görüntülerde öznitelik gösterimi	71
Şekil 4.4 Aynı ölçeklemedeki nesne tanıma hata oranları	72
Şekil 4.5 İlgin-değişimsiz öznitelikle aynı ölçeklemede elde edilen	74

ÇİZELGELER DİZİNİ

Çizelge 2.1 Uydulardaki görüntüleyicilerin spektral bantları	3
Çizelge 2.2 Bazı ülkelerin sahip oldukları uydular ve özellikleri	4
Çizelge 3.1 Optik karakter tanıma işlemi için moment özniteliği	16
Çizelge 3.2 Temel öznitelik tespit etme algoritması ana hatları	28
Çizelge 3.3 Doğru ve yanlış eşleştirme sayıları	49
Çizelge 3.4 Kernel fonksiyonları	68
Çizelge 4.1 Aynı ölçekleme seviyesindeki deneysel sonuçlar	72
Çizelge 4.2 İlgin-değişimsiz öznitelik bulucuyla aynı ölçekleme seviyesindeki deneysel sonuçlar	73
Çizelge 4.3 Farklı ölçekleme seviyesindeki deneysel sonuçlar	74

1. GİRİŞ

Uydu görüntüleme tarım, bölge ve şehir planlaması, ormancılık, istihbarat/haber alma, ulusal güvenlik gibi farklı alanlarda kullanılmaktadır. Örneğin, yerleşimin yoğun olduğu bölgeler tespit edilerek baz istasyonları daha uygun olarak konumlandırılmakta; bir ülkede tarımın nerelerde yapıldığı ve hangi ürünlerin ekildiği tespit edilmekte; sahip olunan orman türleri ve bunların bulunduğu bölgelerin dağılımı gösterilmekte; yerleşim bölgelerindeki kaçak yapılaşma tespit edilerek karar mekanizmalarına gerekli bilgiler sağlanmaktadır. Ayrıca istihbarat amacıyla da kullanılan uydu görüntüleme sayesinde füze ateşlenmesi, nükleer tesis, havaalanı ve uçak tespiti gibi askeri öneme sahip uygulamalar da yapılabilmektedir. Bu kullanım amaçları sebebiyle uydu görüntüleme büyük öneme sahiptir.

Uydulardan elde edilen görüntüler, grinin tonlarında olabileceği gibi kırmızı, mavi, yeşil renk bileşenlerinden de oluşabilmektedir. Ayrıca hiperspektral görüntüler sayesinde daha farklı spektral bantlarda da görüntüler alınabilmektedir. Tespit edilmek istenen nesnenin farklı spektral bantlardaki davranışı incelenerek uygun bant seçilebilir. Böylece yüksek nesne tanıma başarımı elde edilebilir. Bu çalışma kapsamında, kırmızı, yeşil ve mavi renk bileşenlerinden oluşan 3 kanallı görüntüler üzerinde çalışılmıştır.

Uydulardan farklı ölçeklemelerde, rotasyonlarda ve değişik bakış açılarında görüntüler alınabilmektedir. Nesne tanıma için geliştirilen algoritmanın bu değişimlerden bağımsız olarak nesneyi bulması gerekmektedir. Uydulardan çeşitli bakış açılarında görüntü temin edilebilmesine rağmen genellikle 45° bakış açısından düşük görüntüler kullanılmaktadır. Bu yüzden, geliştirilecek algoritmanın 45 derecelik bakış açısı karşısında ilginç değişmezlikli olarak çalışması yeterli görülmüştür. Bu çalışmada, ölçekleme ve rotasyon bozucu etkilerine karşı yüksek başarımda nesne tanıma konusuna yoğunlaşmıştır. Bu anlamda nesne tanımak için bölütleme gerektirmeyen öznitelik tabanlı bir algoritma kullanılmıştır. Bu algoritmayı oluşturan öznitelik çıkartımı, kümeleme ve sınıflandırma yöntemleri parametrik olarak incelenmiştir ve karşılaştırmalı bir çalışma yapılmıştır.

Bu alıřma kapsamında ncelikle Blm 2’de uydu grntlemesinden ve uydulardan elde edilen grntlerin zelliklerinden bahsedilecektir. Blm 3’te nesne tanıma iin kullanılan yntemler incelenecektir. alıřma kapsamında uydu grntlerinde nesne tanıma iin kullanılan yntem Blm 3.4’te detaylandırılacaktır. Kullanılan yntem farklı parametrelerle test edilmiřtir. Yntemin bařarısı, Blm 4’te deneysel sonular blmnde ele alınacaktır. Son olarak, performansı etkileyen parametrelerden bahsedilecektir.

2. UYDU GÖRÜNTÜLEME

Son yıllarda uydu sistemlerinde meydana gelen çok önemli teknolojik gelişmeler sonucunda, uydulardan yüksek çözünürlüklü kaliteli görüntüler elde edilmesi olanaklı hale gelmiştir. Başlangıçta yalnızca çok gelişmiş ülkeler tarafından gerçekleştirilen uydu görüntüleme sistemleri, günümüzde çok sayıda ülke ve ticari kuruluş tarafından kullanıma sunulmaya başlamış, bunun sonucunda kaliteli uydu görüntüsü temini konusunda bir rekabet ortamı doğmuştur (Özbalımcı 1999).

Uydulardan farklı dalga boylarında görüntüler alınabilmektedir. Bunlar, pankromatik olarak adlandırılan siyah-beyaz görüntüleme, multispektral olarak adlandırılan mavi-yeşil-kırmızı-NIR(yakın kızılötesi) görüntüleme, görüntü demeti olarak adlandırılan bu iki görüntüleme yeteneğini birlikte sunan görüntüleme ve çok sayıda spektral bant görüntüsü sağlayan hiperspektral görüntülemedir. Bu görüntülerin elde edildiği bazı spektral bantlar Çizelge 2.1’de gösterilmektedir.

Çizelge 2.1 Uydulardaki görüntüleyicilerin spektral bantları

	Spektral Bantlar
Pankromatik	470 - 830 nm
Mavi	430 - 550 nm
Yeşil	500 - 620 nm
Kırmızı	590 - 710 nm
Yakın Kızılötesi (NIR)	740 - 940 nm

Günümüzde uydu görüntülerinin piksel boyutlarında meydana gelen küçülme nedeniyle konumsal ayırma gücü artmış olup, görüntülerden, 0,50 cm ve daha küçük boyuttaki nesneleri ayırt etmek olanaklı hale gelmiştir. Uydu görüntülerinde meydana gelen en önemli gelişme, arazideki ayırma gücünün siyah-beyaz görüntülerde 0,50 cm, renkli görüntülerde ise 2 metreye ulaşmasıdır.

Dünyada uydu geliştirilmesinde öne çıkan üç firma bulunmaktadır. Bunlar ABD’nin DigitalGlobe ve GeoEye firmaları ve Fransa’nın Spot firmasıdır. Yeni uydu sistemleri,

yüksek maliyetlerin düşürülmesi için genellikle belli bir amaca yönelik küçük sistemler şeklinde tasarlanmaktadır. Bu tip uydu görüntüleme sistemlerinin geliştirilmesinde, genellikle iki temel yaklaşımdan birisi uygulanmaktadır. Birinci yaklaşım, mevcut uydu sistemlerinin geliştirilerek kullanımına devam edilmesi; ikinci yaklaşım ise, tamamen yeni ticari uydu sistemlerinin geliştirilmesidir (Özbalmumcu 2001).

İlk yaklaşıma örnek olarak Amerika Birleşik Devletleri'nin DigitalGlobe firmasının WorldView-1 ve WorldView-2 uyduları; ABD'nin GeoEye firmasının GeoEye-1 ve GeoEye-2 uyduları; Fransa'nın SPOT-1, SPOT-2, SPOT-3 ve SPOT-4 uyduları örnek verilebilir. Bazı ülkelerin sahip oldukları uydular ve özellikleri Çizelge 2.2'de verilmiştir (İşcan 2004).

Çizelge 2.2 Bazı ülkelerin sahip oldukları uydular ve özellikleri

Uydu Adı	Ülke	Fırlatma Tarihi	Algılayıcı Tipi	Konumsal Ayırma Gücü
QuickBird	Amerika	2003	Pankromatik + Multispektral	60 cm 2,4 m
WorldView I	Amerika	2000	Pankromatik	50 cm
WorldView II	Amerika	2004	Pankromatik + Multispektral	46 cm 1,84 m
Ikonos	Amerika	1999	Pankromatik + Multispektral	82 cm 3,2 m
GeoEye-1	Amerika	2001	Pankromatik + Multispektral	41 cm 1,65 m
GeoEye-2	Amerika	2007	Pankromatik	25 cm

2.1 Ikonos Uydu Görüntüleri

Yüksek çözünürlüklü uydu görüntülerinin ilk başarılı denemesi olan IKONOS uydu görüntüleri, pek çok ülkede kamu, özel, sivil ve askeri kuruluşlar tarafından yaygın

şekilde kullanılmaya başlanmıştır. Ayrıca uydu görüntülerine olan ilginin artmasına yol açmış ve büyük çaplı harita üretim projeleri / planlamalarda uydu görüntülerinin de göz önüne alınması ile yeni bir dönemin başlamasına neden olmuştur.



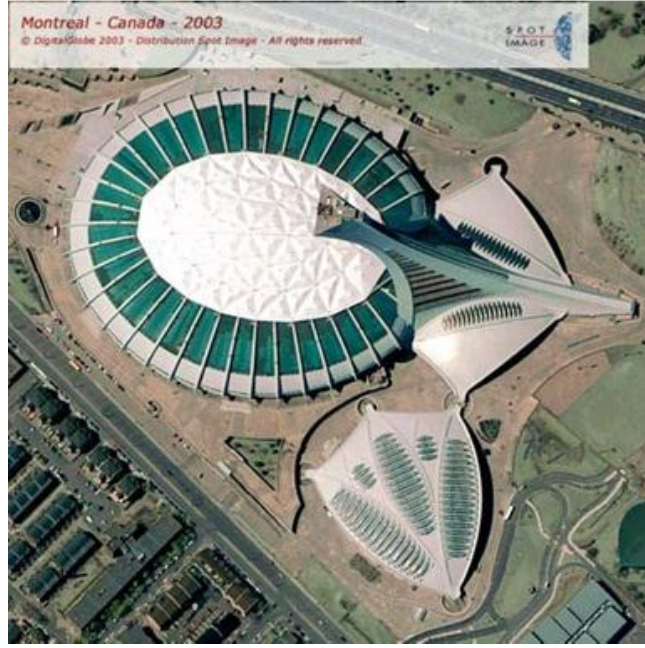
Şekil 2.1 IKONOS uydu görüntüsü (<http://glcf.umd.edu/data/ikonos/>)

Uydunun, eğik alımda görüntü yinleme aralığı 1-3 gün, normal (dik, düşey, nadir gözlem) alımda görüntü yinleme aralığı ise 14 gündür. Normal alımda 700 km enindeki şerit içinde 1 m, eğik alımda 1450 km enindeki şerit içinde 1,5 m çözünürlükte görüntüler alabilmektedir. Bir görüntünün arazide kapladığı alan ortalama 11 km x 11 km olup, dünya çevresinden bir kez geçişte 10.000 km²'lik bir alanın görüntüsü alınabilmektedir. IKONOS uydusu ile uydu yörüngesi boyunca veya yan yana iki yörünge geçişi ile stereo görüntü alımı olanaklıdır. İki temel ürün yanında, IKONOS uydusunun pankromatik (0,82 m) ve multispektral (3,2 m) görüntüleri birleştirilerek, 1 m çözünürlüğe sahip renklendirilmiş görüntüler elde edilebilmektedir (Şekil 2.1).

2.2 Quickbird Uydu Görüntüleri

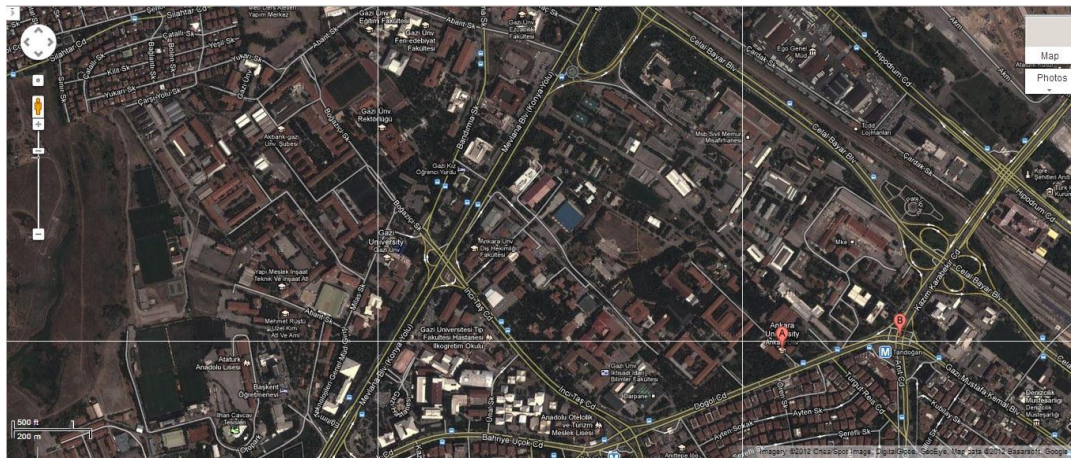
QUICKBIRD uydusu, Digital Globe firması tarafından 18 Ekim 2001'de, Kaliforniya'daki Vandenberg Hava Kuvvetleri'nde bulunan Boeing Delta II fırlatma aracından fırlatılmıştır. Pankromatik görüntülerin çözünürlüğü normal alımda 0,60 m, nadirden 25° eğim açısı ile eğik alımda 0,72 metredir. Multispektral görüntülerin çözünürlüğü ise, normal alımda 2,40 m, 25° eğim açısında 2,88 metredir. Bir QUICKBIRD uydu görüntüsünün arazide kapladığı alan (tarama alanı) 16,5 km x 16,5 km'dir (Alexandrov vd. 2004).

QUICKBIRD bir metrenin altında konumsal çözünürlükte veri toplayabilen ilk ticari uydudur. Günün herhangi bir saatinde dünyanın herhangi bir yerine ait görüntülerin alımı olanaklıdır. Bir günde yaklaşık 900 adet görüntü alımı yapılabilmekte ve yaklaşık 137 GByte veri toplayabilmektedir. QUICKBIRD uydusundan elde edilen görüntü türleri; pankromatik, multispektral ve zenginleştirilmiş (pan-sharpened) görüntülerdir. QUICKBIRD uydusunun multispektral algılayıcısı dört ayrı renkte görüntü almaktadır. Zenginleştirilmiş görüntüler olarak doğal renkli (RGB; red, green, blue; kırmızı, mavi, yeşil) veya renkli kızılötesi (NIRR-G; near infrared-red-green; yakın kızılötesi-kırmızı-yeşil) görüntüler alınabilmektedir (Eisenbeiss vd. 2004). QUICKBIRD uydusuna ait örnek görüntüler şekil 2.2'de gösterilmiştir.

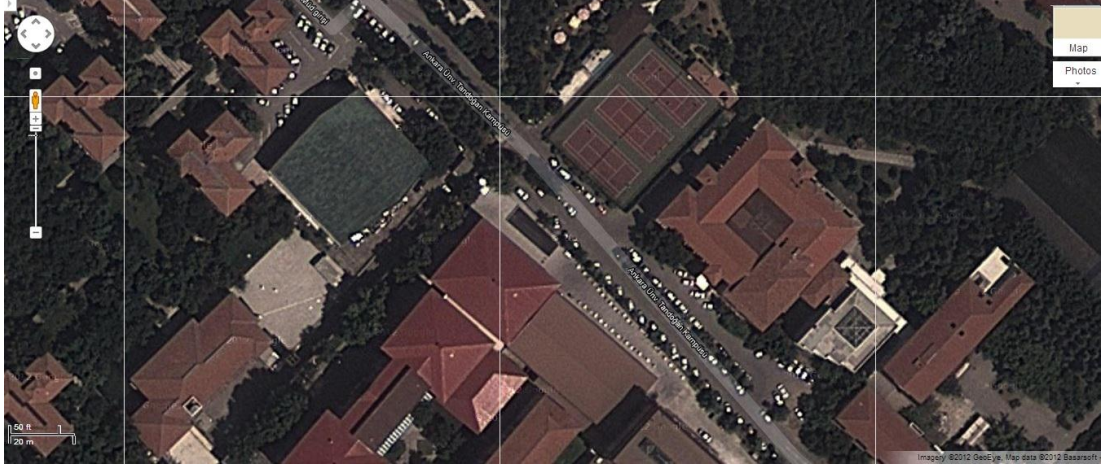


Şekil 2.2 QUICKBIRD uydu görüntüleri (<http://glcf.umd.edu/data/quickbird/>)

Uydu görüntüleri, ticari uydu firmalarından ücreti mukabilinde veya Google Earth programı veya Google Maps web uygulaması kullanılarak ücretsiz olarak temin edilebilir. Google Maps ve Google Earth görüntülerinin, DigitalGlobe ve GeoEye uyduları tarafından sağlandığı bilinmektedir. Farklı çözünürlüklerle elde edilebilen uydu görüntüleri için örnek görüntüler şekil 2.3 - 2.4'te verilmiştir.



Şekil 2.3 200 metre çözünürlüklü Google Maps görüntüsü



Şekil 2.4 20 metre çözünürlüklü Google Maps görüntüsü

3. NESNE TANIMA

Nesne tanıma, uzun bir süredir üzerinde yoğun olarak araştırma yapılan bir alandır. Özellikle son 10 yılda çok fazla sayıda çalışma yapılmış ve birçok çözüm yöntemi önerilmiştir. Bunun sebebine daha yakından bakılacak olursa; nesne tanıma, her biri kendine has gereksinimlere ve kısıtlamalara sahip olan geniş bir uygulama alanına sahiptir. Nesne tanıma da bu alanlarda önerilen tüm algoritmaları bir şemsiye altında toplayan terimdir.

Bilgisayar donanımlarının hızlı gelişimi, otomatik nesne tanıma kullanımını endüstriyel görüntü işlemeden medikal uygulamalara, internette fazla miktarda sorgulanan görüntülere benzer görüntülerin internetten elde edilmesi gibi görevlere varıncaya kadar kullanım alanları yaygınlaşmaktadır. İç mekan (in-door) ve dış mekan (out-door) gibi farklı ortamlarındaki, esnek olmayan (rigid) ve esnek (non-rigid) gibi farklı özelliklerdeki nesnelerin tanınması her bir uygulamanın kendine özgü gereksinimleri ve kısıtları olduğunu göstermektedir. Sonuç olarak tek bir genel amaçlı algoritma ile tüm bu alanlara çözüm önerilemez. Bu bölümde nesne tanıma için çözüm oluşturulabilecek bir kaç yöntem üzerinde durulacaktır.

Bu bölümde, nesne tanıma üzerine literatürde bulunan yöntemler anlatılmıştır. Bunlar evrensel yöntemler, dönüşüm-arama tabanlı ve öznitelik-tabanlı yöntemlerdir.

3.1 Evrensel Yöntemler

Evrensel yöntemler, tanımlanacak nesneyi genellikle eğitim aşamasındaki örnek resimlerden öğrenen ve bir bütün olarak sunan veriler üzerinde çalışır. Model, tanımlanacak nesneyi bir bütün olarak gösterir.

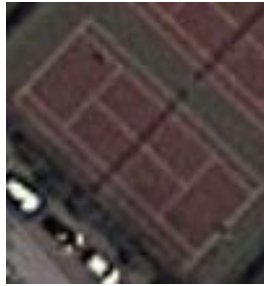
Bütünsel yöntemler, iki boyutlu ilinti ve bütünsel öznitelik vektörleri olmak üzere iki kısımdan oluşmaktadır.

3.1.1 İki boyutlu ilinti

İki boyutlu ilinti yönteminde, öncelikle bulunacak nesneyi temsil eden şablon görüntü elde edilir. Ardından tüm görüntü üzerinde bulunması istenen nesne şablonu piksel piksel gezdirilerek maksimum benzerlik elde edilmeye çalışılır. Bulunması istenen nesne ile görüntü arasındaki benzerlik denklem 3.1’de kullanılarak hesaplanmaktadır.

$$\rho(a,b) = \frac{\sum_{x=0}^W \sum_{y=0}^H (I_s(x+a, y+b) - \bar{I}_s) \cdot (I_T(x, y) - \bar{I}_T)}{\sqrt{\sum_{x=0}^W \sum_{y=0}^H (I_s(x+a, y+b) - \bar{I}_s)^2 \cdot \sum_{x=0}^W \sum_{y=0}^H (I_T(x, y) - \bar{I}_T)^2}} \quad (3.1)$$

$\rho(a,b)$, $[a,b]$ yer değiştirmesindeki şablon görüntü ile tüm görüntü arasındaki normalize edilmiş çapraz ilinti katsayısıdır. $I_s(x,y)$ ve $I_T(x,y)$, $[x,y]$ pozisyonundaki görüntü ve şablonun yoğunluğunu göstermektedir. W ve H , şablonun genişlik ve yüksekliğini göstermektedir. Payda, normalizasyon vazifesi gören terim olduğu için ρ -1 ile 1 arasında değişmektedir. Yüksek pozitif değerler, görüntü ve şablonun çok benzer olduğunu gösterirken; sıfır değeri görüntü ve şablonun içeriklerinin ilintisiz olduğunu gösterir. Ayrıca, negatif değerler ters içeriğin göstergesidir. Şekil 3.1 ile şablon, Şekil 3.2 ile üzerinde şablonun aranacağı görüntü gösterilmiştir.



Şekil 3.1 Şablon görüntüsü



Şekil 3.2 Şablonun aranacağı görüntü

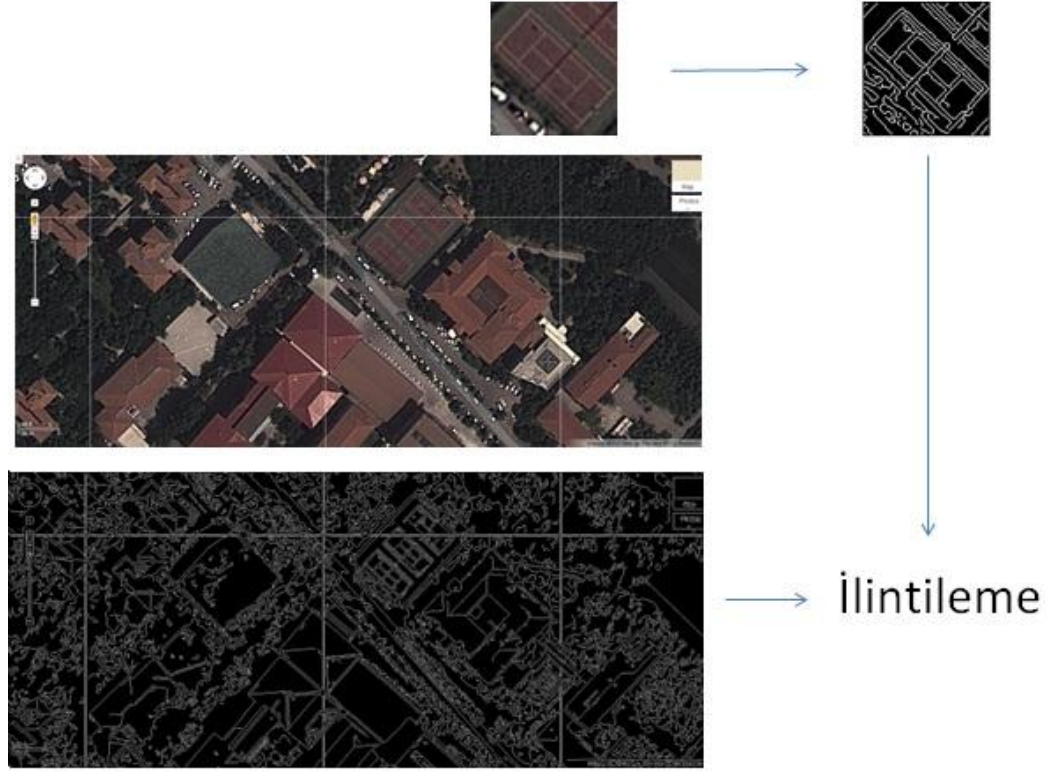
İki boyutlu ilintinin avantajları, kolay uygulanabilir ve görünümünden bağımsız olmasıdır. Ayrıca kontrasttaki doğrusal değişimlerden ve doğrusal ışıklılık değişimlerinden bağımsız olarak çalışabilmektedir. Dezavantajları ise yüksek oranda tanıma için nesnenin çeşitli açılardaki rotasyonları ve çeşitli oranlardaki ölçeklemelerinin olduğu şablonlar görüntü veri tabanına eklenmelidir. Şablon sayısındaki artış, hesap yükünü arttırmaktadır. Bu yöntemler, arka plan gürültüsüne ve kapatmaya (occlusion) dayanıksızdır. Ayrıca görüntülerde genellikle doğrusal olmayan ışıklılık değişimleri olduğu için bu yöntem ile çok başarılı sonuçlar elde edilememektedir.

İki boyutlu ilinti başarımını arttırmak için şablon görüntülerinin çeşitli açılardaki rotasyonları ve çeşitli oranlardaki ölçeklemelerinin kullanılması gerekmektedir. Bu işlemi hesaplama yükünde artışa sebep olmaktadır. Dayanıklılığı arttırmak ve hesaplama süresini kısaltmak için çeşitli yöntemler bulunmaktadır. Bu ilintileme yöntemleri arasında önişleme, alt örnekleme veya görüntü piramitleme ve faz ilintisi yer almaktadır.

3.1.1.1 Ön işleme

Önişleme kullanılarak iki boyutlu ilintileme işleminde, öncelikle şablon ve görüntünün kenarları tespit edilir. Ardından kenarları bulunan şablon ve görüntü arasındaki benzerlik hesaplanır. Yüksek ilinti katsayısı değerleri elde etmek için ince kenarlı

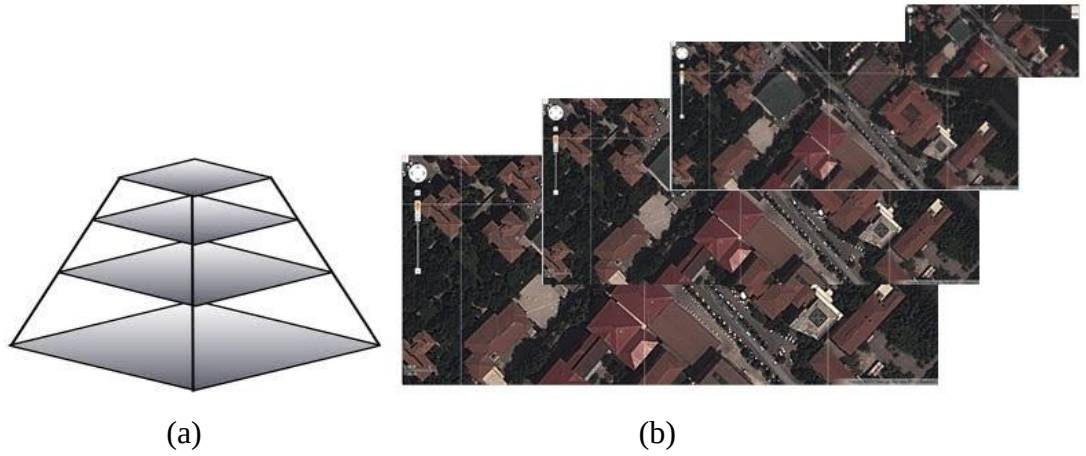
bölgelerin örtüşmesi gerekir. Bu yöntem, daha başarılı sonuçlar alınmasına ve doğrusal olmayan ışıklılık değişimlerine dayanıklılığa olanak vermektedir. Ayrıca kısmi kapatmaya (occlusion) daha dayanıklıdır. Bu yöntem kullanılarak gerçekleştirilen bir örnek, Şekil 3.3'te gösterilmiştir.



Şekil 3.3 Önişleme kullanarak iki boyutlu ilintileme

3.1.1.2 Alt örnekleme veya görüntü piramitleme

Bu yöntem ile şablonun aranacağı görüntü, öncelikle Şekil 3.4'te gösterildiği gibi alt seviyelere örneklenir. Bu işleme, görüntü piramitleme de denilmektedir.



Şekil 3.4 Görüntü alt örnekleme

Şekil 3.4’te gösterilen görüntülerde, en yüksek çözünürlüğe sahip görüntü sıfırıncı seviyeyi oluşturmakta ve görüntü çözünürlüğü azaldıkça seviye artmaktadır. Görüntü alt seviyelere örneklendikten sonra şablon, en üst seviyedeki görüntü ile ilintilendirilir. En yüksek ilinti katsayısının elde edildiği pozisyon etrafında daha aşağı seviyedeki görüntü ile şablon ilintilendirilir. Alt seviyelerdeki ayrıntılı görüntülerde üst seviyede elde edilen pozisyon bilgisi etrafında ilintileme gerçekleştirilerek pozisyon bilgisi iyileştirilmeye çalışılır. Böylece gereksiz yere tüm görüntü üzerinde şablon gezdirilmemiş ve sadece olabilirliği yüksek pozisyonlarda ilintileme yapılmış olmaktadır.

Görüntü alt örnekleme, hızlı işlem kabiliyeti sağlamasının yanında geleneksel yöntemle karşılaştırılabilir doğruluk sunmaktadır. Temel avantajı, görüntüdeki önemli kısımların üst seviyelerde bulunabilmesidir. Böylelikle düşük seviyelerde işlem yapmaya gerek duyulmayabilir. Sağladığı hızlanma yardımıyla şablonun rotasyonlu ve ölçeklendirilmiş versiyonları da kullanılabilir.

3.1.1.3 Faz ilintisi

Faz ilintisi, genellikle görüntü tescili için kullanılmakla birlikte nesne tanımda da kullanılmaktadır. Bu yöntemde, görüntü öncelikle uzamsal (uzaysal) bölgeden frekans bölgesine dönüştürülmelidir. Bu işlem, hızlı Fourier dönüşümü ile yapılır. Görüntünün

ve şablonun faz bilgisi denklem 3.2, 3.3 ve 3.4 kullanılarak elde edilir. Daha sonra faz bilgileri ilintilendirilir.

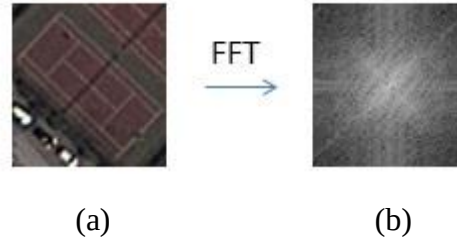
$$F_I = A_I(\omega_1, \omega_2) \cdot e^{\theta_I(\omega_1, \omega_2)} \quad (3.2)$$

$$R(\omega_1, \omega_2) = A_S(\omega_1, \omega_2) \cdot A_T(\omega_1, \omega_2) \cdot e^{\theta(\omega_1, \omega_2)} \quad (3.3)$$

$$\theta(\omega_1, \omega_2) = \theta_S(\omega_1, \omega_2) - \theta_T(\omega_1, \omega_2) \quad (3.4)$$

Uzaysal bölgedeki $I(x,y)$ görüntüsü, frekans bölgesinde yukarıdaki denklemde F_I ile ifade edilir. A_I ifadesi genlik bileşenini göstermekle birlikte üstel ifade faz bileşenini göstermektedir. $R(\omega_1, \omega_2)$, iki görüntünün çapraz spektrumunu göstermektedir. $\theta(\omega_1, \omega_2)$ ise iki spektrumun faz farkını göstermektedir.

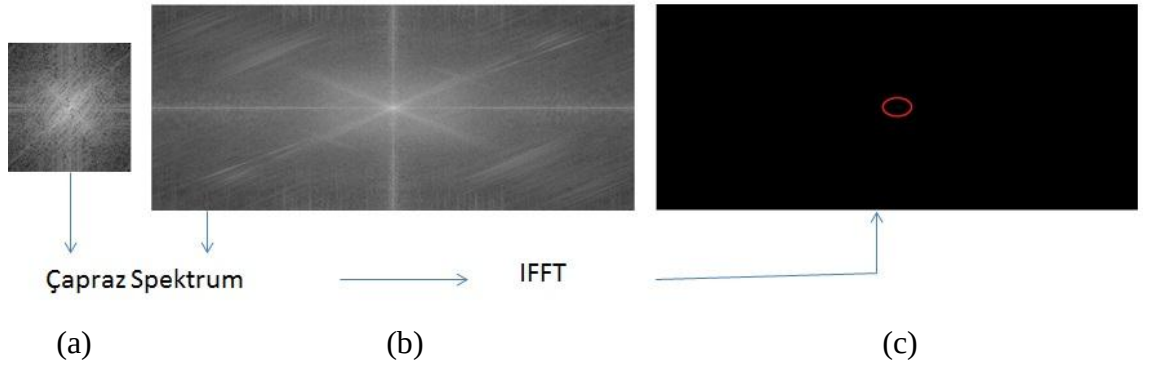
Faz ilintisi kullanılarak yapılan örnek Şekil 3.5'te gösterilmiştir.



Şekil 3.5 Şablonun uzaysal bölgedeki ve frekans bölgesindeki görünümü



Şekil 3.6 Görüntünün uzaysal bölgedeki ve frekans bölgesindeki görünümü



Şekil 3.7 Faz ilintisi sonucu tespit edilen nesne

3.1.2 Bütünsel öznitelik vektörleri

Bütünsel öznitelik vektörleri kullanılarak nesnenin alan, çevre, dairesellik ($\text{çevre}^2 / \text{alan}$), moment, ortalama gri değeri gibi öznitelikler çıkartılır. Tüm görüntü üzerinden benzer özniteliklerin varlığı tespit edilmeye çalışılır.

Öznitelik çıkartılması sırasında iyi öznitelik seçimi önemlidir. Farklı türdeki nesneler arasındaki değişimin yüksek olduğu ve aynı türdeki nesneler arasındaki değişimin düşük olduğu öznitelikler, iyi özniteliklerdir. Öznitelik vektörünün boyutu arttırılarak yüksek nesne tanıma oranları elde edilmekle birlikte bu vektörün boyutunun fazla miktarda arttırılması karmaşıklığı neden olmaktadır.

Bu yöntemin avantajı, hızlı olmasıdır. Ayrıca nesne gösterimi değişimlere karşı daha dayanıklı olarak yapılabilmektedir. Bu yöntemin dezavantajı ise bölütleme işlemine ihtiyaç duymasındır. Ayrıca kapatmaya (occlusion) karşı dayanıklı değildir.

Çizelge 3.1 ile optik karakter tanıma uygulaması için kullanılan moment özniteliği bilgileri verilmiştir. Bazı öznitelikler, bazı karakterler için ayırt edici değerler göstermektedir. Mesela, I karakterinin anizometrisi diğer karakterlere kıyasla daha fazladır. μ_{11} momenti incelendiği zaman P ve G karakterlerinin değerleri, simetri

özelliğinin eksikliğinden ötürü diğerlerinden farklıdır. Diğer taraftan bazı momentler çok fazla bilgi içermezler. μ_{02} momentinin değerleri tüm karakterler için benzerdir.

Çizelge 3.1 Optik karakter tanıma işlemi için moment özniteliği

Karakter	A	I	G	M	Y	P
Anizometri	1.09	3.21	1.05	1.24	1.27	1.40
μ_{11}	1.46	0	-20.5	-0.148	-0.176	113
μ_{20}	368	47.9	481	716	287	331
μ_{02}	444	494	507	465	460	410
μ_{22}	164.784	23.670	168.234	351.650	134.691	113.767

3.2 Dönüşüm - Arama Tabanlı Yöntemler

Dönüşüm – arama tabanlı yöntemlerde nesnenin gösterimi, nesneye ait sonlu noktalar kümesi ve noktaların pozisyonu kullanarak yapılır. Öncelikle görüntüden nokta kümesi çıkartılır. Model görüntüsü nokta görünüsü ile sahne görüntüsü nokta kümesi arasındaki dönüşüm parametreleri hesaplanır. Son olarak benzerliği maksimize eden veya uzaklığı minimize eden dönüşüm parametreleri hesaplanır. Bu işlem, dönüşüm uzayı üzerinde arama yapılarak gerçekleştirilir.

Kapatma (occlusion) ve arka plan gürültüsü bazı nokta kümelerinde kayıplara sebep olmakla birlikte dönüşüm – arama tabanlı yöntemlerin kısmi kapatma ve arka plan gürültüsüne karşı dayanıklılığı bulunmaktadır.

3.3 Öznitelik Tabanlı Yöntemler

Öznitelik tabanlı yöntemler, öznitelik tespiti ve öznitelik eşleştirme olmak üzere iki bölümden oluşmaktadır. Öznitelik tespiti ve öznitelik eşleştirme, birçok bilgisayar görüşü uygulamasında ihtiyaç duyulan bileşenlerdir. Şekil 3.8’de gösterilen iki görüntü çiftini göz önüne alalım. İlk görüntü çifti için iki görüntüyü bir mozaik içerisinde saydam bir şekilde birleştirmek için iki görüntüyü hizalamak isteyebiliriz. İkinci görüntü çifti için bir 3 boyutlu (3B) model oluşturmak sebebiyle veya bir ara görüntü oluşturmak sebebiyle benzerlik kümesi oluşturmak isteyebiliriz. Her iki durumda da hangi özniteliklerin tespit edileceği ve bir hizalama veya uyumluluk kümesi elde etmek için hangi eşleştirme yöntemi kullanılacağı cevap verilmesi gereken sorulardır.



(a)



(b)



(c)

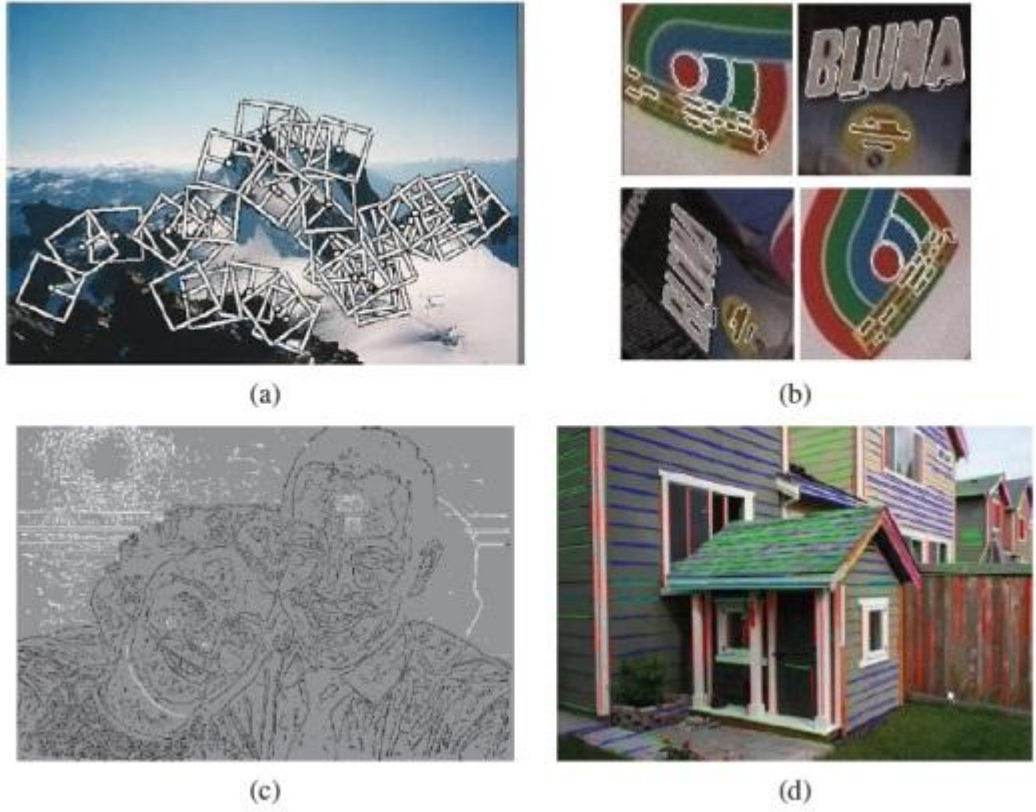


(d)

Şekil 3.8 Eşleştirilecek iki görüntü çifti

Öncelikli göze çarpan öznitelikler, görüntülerdeki dağ zirveleri, bina köşeleri, kapı girişleri veya karın ilginç olarak şekillenmiş parçaları gibi spesifik lokasyonlardır. Bu tür öznitelikler, anahtar öznitelikler veya ilgi noktaları olarak adlandırılırlar. Bu noktaların etrafındaki piksel parçalarının görünümüyle tanımlanırlar. Önemli özniteliklerden bir diğeri de kenarlardır. Şekil 3.8'deki dağın gökyüzüne karşı olan profili örnek verilebilir. Bu tür öznitelikler, oryantasyonları ve yerel görünümleri (kenar profili gibi) kullanılarak eşleştirilebilirler. Ayrıca bu öznitelikler, nesne sınırlarının ve görüntü dizisindeki kapatma durumlarının iyi bir göstergesi olabilirler. Kenarlar, daha uzun eğrilere ve düz çizgi parçalarına gruplanabilirler. Bu sayede direkt olarak eşleştirilebilirler veya gözden kaybolan noktaların bulunması için analiz edilebilirler.

Şekil 3.9'da görüntülerin analizinde, tanımlanmasında ve eşleştirilmesinde kullanılabilecek çeşitli öznitelik bulucular ve tanımlayıcılar gösterilmektedir. Şekil 3.9.a görüntüsünde nokta benzeri ilgi operatörü (Brown vd. 2005), Şekil 3.9.b görüntüsünde bölge benzeri ilgi operatörü (Matas vd. 2004), Şekil 3.9.c görüntüsünde kenarlar (Elder ve Goldberg 2001) ve Şekil 3.9.d görüntüsünde düz çizgiler (Sinha vd. 2008) öznitelik olarak seçilmiştir.



Şekil 3.9 Görüntülerin analiz, tanımlama ve eşleştirmelerinde kullanılabilecek çeşitli öznitelikler

Bu bölümde, çeşitli öznitelikler bulmak için kullanılan pratik yaklaşımlardan bahsedilecektir. Ayrıca farklı görüntüler arasında öznitelik benzerliklerinin nasıl saptanacağı ele alınacaktır. Nokta öznitelikleri, iyi tespit edilme ve bunlardan bazı algoritmalar geliştirilmesi açısından uygulama alanı bulmaktadır. Kenarlar ve çizgiler, hem anahtar noktalara hem de bölge-tabanlı tanımlayıcılara bilgi sağlarlar. Aynı zamanda nesne sınırları ve insan yapımı olan nesneleri tanımlamak için çok uygundur.

3.3.1 Öznitelik algılayıcılar

Nokta öznitelikleri, farklı görüntülerde bu noktalara karşılık gelen konumların bir seyrek kümesini bulmak için kullanılabilirler. Görüntülerden panoramik bir görüntü elde edileceği veya video stabilizasyonu gerçekleştirileceği uygulamalarda farklı görüntülerin hizalanmasında kullanılabilirler. Aynı zamanda nesne ve kategori tanıma

gerçekleřtirmek için de kapsamlı olarak kullanılırlar. Anahtar noktalar kullanmanın en önemli avantajı, gürültü olduđu durumlarda, ölçek ve rotasyon deęişikliklerinde bile eşleřtirmeye olanak sağlamasıdır.

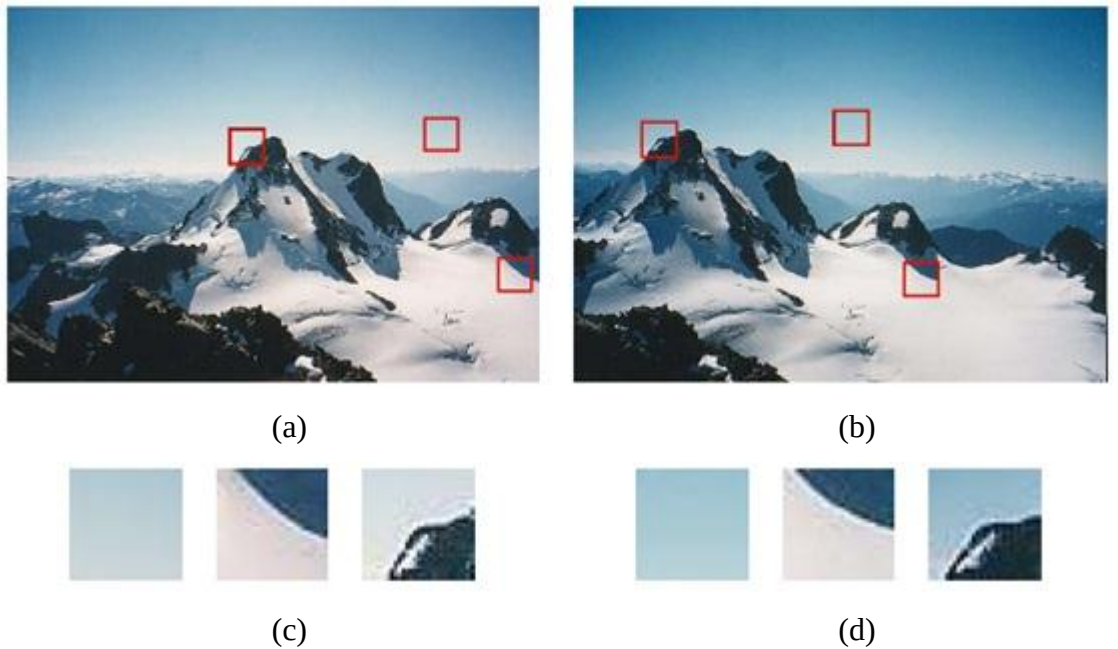
Öznitelik tabanlı benzerlik teknięi, stereo eşleřtirmenin (Hannah 1974, 1988) ilk günlerinden itibaren kullanılmaktadır (Moravec 1983). Son zamanlarda panoramik görüntü uygulamaları (Zoghلامي vd. 1997, Brown vd. 2007) ve tamamen otomatikleřtirilmiř 3B modelleme uygulamaları (Beardsley vd. 1996, Schaffalitzky ve Zisserman 2002, Brown ve Lowe 2003, Snavely vd. 2006) sayesinde popülerlik kazanmıřtır.

Öznitelik noktalarını ve benzerliklerini bulmak için kullanılan iki temel yaklařım vardır. Birinci yaklařım, ilinti ve en küçük kareler gibi yerel arama teknikleri kullanılarak hatasız veya az hatalı olarak takip edilebilecek özniteliklerin bir görüntüde bulunmasıdır. İkinci yaklařım ise tüm görüntülerdeki özniteliklerin birbirinden bağımsız olarak tespit edilmesi ve yerel görünümlerine göre eşleřtirilmesi esasına dayanır. İlk yaklařımın görüntülerin farklı fakat birbirine yakın bakıř açılarında veya video gibi diziler halinde elde edildięi durumlarda kullanılması daha uygundur. İkinci yaklařım ise birbiri ile ilintili birden fazla görüntüden panoramik bir görüntü elde edilmesi (Brown ve Lowe 2007) gibi büyük miktarda hareket veya görünüm deęiřiklięi beklenen durumlar ve nesne tanıma uygulamalarında (Fergus vd. 2007) kullanılması daha uygundur.

Bu bölümde, nokta tespit etme ve eşleřtirme iřlemi üç ayrı ařamada ele alınacaktır. Öznitelik tespit etme ařamasında, her bir görüntü için dięer görüntülerde en iyi eşleřtirilecek konumlar bulunur. Öznitelik tanımlama ařamasında, tespit edilen noktaların etrafındaki her bir bölge, dięer tanımlayıcılar ile eşleřtirilebilecek daha özlü ve dengeli (deęiřimsiz) tanımlayıcılara dönüřtürülür. Öznitelik eşleřtirme ařaması, dięer görüntüler üzerinde muhtemel eşleřebilecek adayların etkili olarak arařtırıldıęı ařamadır.

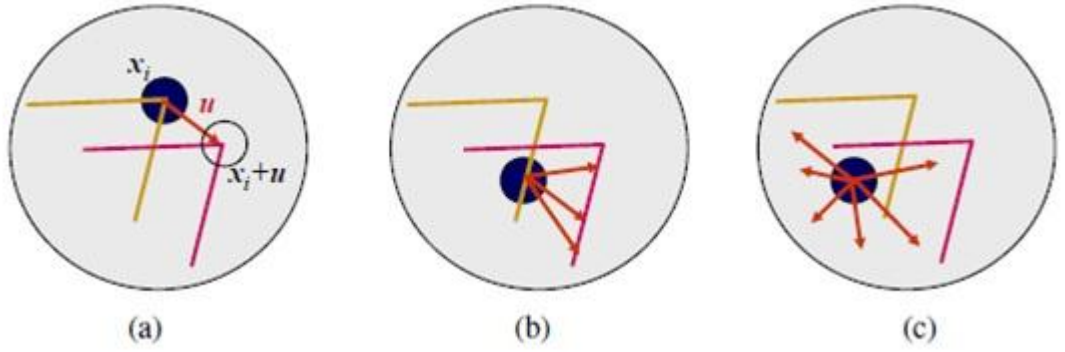
Tüm aşamaların içinde bulunduğu bir örnek, ölçekleme değişimsiz öznitelik dönüşümü (SIFT) çalışmasında mevcuttur (Lowe 2004). Hem öznitelik tespit etme (Schmid vd. 2000, Mikolajczyk vd. 2005, Tuytelaars ve Mikolajczyk 2007) hem de öznitelik tanımlayıcıları (Mikolajczyk ve Schmid 2005, Shi ve Tomasi 1994, Triggs 2004) üzerine inceleme ve değerlendirme kapsamında alternatif çalışmalar yapılmıştır. Ayrıca öznitelik tespit tekniklerinin gözden geçirilmesi kapsamında da çalışmalar yapılmıştır (Shi ve Tomasi 1994, Triggs 2004).

Başka bir görüntüde benzerliklerin güvenilir bir şekilde aranacağı konumlar, bir görüntüde nasıl tespit edilir? Takip edilecek iyi öznitelikler üzerine çeşitli çalışmalar yapılmıştır (Shi ve Tomasi 1994, Triggs 2004). Şekil 3.10'da birbirleri ile ilintili iki görüntü ve bunlardan çıkartılan görüntü parçaları verilmiştir. Tahmin edilebileceği gibi içinde hiçbir doku bulunmayan görüntü parçalarının yerini tayin etmek imkansızdır. Tek yöndeki düz çizgi bölümlerinin açıklık probleminde muzdarip olmasına (Horn ve Schunck 1981, Lucas ve Kanade 1981, Anandan 1989) rağmen fazla kontrast (gradyan) değişimlerine sahip görüntü parçalarının yerini tayin etmek daha kolaydır.



Şekil 3.10 Görüntü çiftleri ve görüntülerden çıkartılan parçalar

Şekil 3.11’de farklı görüntü parçaları için açıklık problemi verilmiştir. Şekil 3.11.a kararlı (köşe benzeri) akış; şekil 3.11.b klasik açıklık problemi; şekil 3.11.c dokusuz bölgeyi ifade etmektedir. Sarı renk ile gösterilen I_0 ve kırmızı renkle gösterilen I_1 görüntüleri üst üste bindirilmiştir. Kırmızı u vektörü, görüntü parçalarının merkezi ile koyu çember ile gösterilen $w(x_i)$ ağırlık fonksiyonu (görüntü parçası penceresi) arasındaki yer değiştirmeyi gösterir.



Şekil 3.11 Farklı görüntü parçaları için açıklık problemi

Görüntü parçaları sadece normalin yönü boyunca kenar yönü ile hizalanabilir (Şekil 3.11.b). En az iki farklı oryantasyonda gradyana sahip görüntü parçalarının yerini tayin etmek en kolaydır (Şekil 3.11.a).

Bu sezgisel ifadeler, iki görüntü parçasını karşılaştırmak için en basit olası eşleştirme kriterine bakarak formülize edilebilir. Ağırlıklandırılmış toplam kare farkı:

$$\sum_{wSSD} (u) = \sum_i w(x_i) [I_1(x_i + u) - I_0(x_i)]^2 \quad (3.5)$$

I_0 ve I_1 karşılaştırılacak iki görüntü, $u = (u,v)$ yer değiştirme vektörü, $w(x)$ uzaysal olarak değişen ağırlık veya pencere fonksiyonu ve i ise görüntü parçasındaki tüm pikselleri ifade etmektedir.

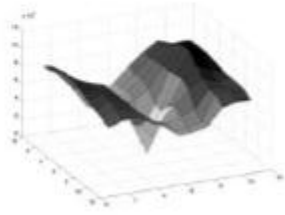
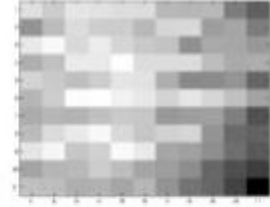
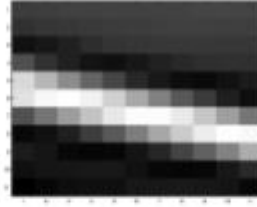
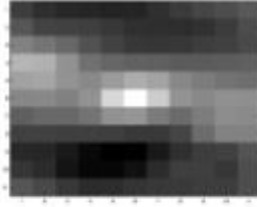
Öznitelik tespit etme işlemini gerçekleştirirken diğer hangi görüntü konumlarına karşı özniteliklerin karşılaştırılacağı bilinmiyor. Bu yüzden, biz sadece bu metriklerin Δu pozisyonundaki küçük değişimler karşısında ne kadar kararlı olduğunu oto-korelasyon fonksiyonu veya düzlemi olarak da bilinen görüntü parçasının kendisi ile karşılaştırarak hesaplayabiliriz.

$$\sum_{AC}(\Delta u) = \sum_i w(x_i)[I_0(x_i + \Delta u) - I_0(x_i)]^2 \quad (3.6)$$

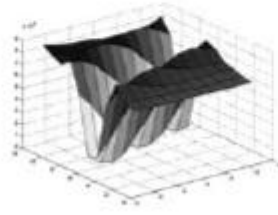
Şekil 3.12 ile hem gri seviye görüntü ve hem de düzlem çizimi olarak gösterilen üç oto-korelasyon düzlemi verilmiştir. Oto-korelasyon yüzeylerinin hesaplandığı üç nokta, orijinal görüntü üzerinde kırmızı çarpı ile işaretlenmiştir (Şekil 3.12.a). Şekil 3.12.b ile iyi, eşsiz ve minimum olan çiçekten alınan görüntü parçasına ait görüntü ve yüzey gösterilmiştir. Bir dokuya sahip çiçek için oto-korelasyon yüzeyi, yerinin çok iyi tayin edildiğini gösteren kuvvetli bir minimumdur. Şekil 3.12.c ile bir boyutlu açıklık problemi olan çatı kenarından alınan görüntü ve yüzey gösterilmiştir. Çatı kenarına karşılık gelen ilinti yüzeyi, bir yön boyunca kuvvetli bir belirsizliğe sahiptir. Şekil 3.12.d ile iyi tepe noktası olmayan buluttan alınan bir görüntü ve yüzeyi verilmiştir. Buluta karşılık gelen ilinti yüzeyi, kararlı bir minimuma sahip değildir. Her bir çizim, Δu 'nın bir değeri içindir.



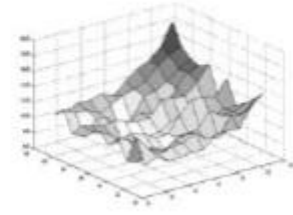
(a)



(b)



(c)



(d)

Şekil 3.12 Gri seviye görüntü ve düzlem çizimi olarak gösterilen üç oto-korelasyon düzlemi

Oto-korelasyon yüzeyine,

$$I_0(x_i + \Delta u) \approx I_0(x_i) + \nabla I_0(x_i) \cdot \Delta u \quad (3.7)$$

görüntü fonksiyonuna bir Taylor serisi açılımı kullanarak aşağıdaki gibi yaklaşabiliriz (Lucas ve Kanade 1981, Shi ve Tomasi 1994).

$$\begin{aligned}
\sum_{AC}(\Delta u) &= \sum_i w(x_i)[I_0(x_i + \Delta u) - I_0(x_i)]^2 \\
\sum_{AC}(\Delta u) &\approx \sum_i w(x_i)[I_0(x_i) + \nabla I_0(x_i) \cdot \Delta u - I_0(x_i)]^2 \\
\sum_{AC}(\Delta u) &= \sum_i w(x_i)[\nabla I_0(x_i) \cdot \Delta u]^2 \\
\sum_{AC}(\Delta u) &= \Delta u^T A \Delta u
\end{aligned} \tag{3.8}$$

$$\nabla I_0(x_i) = \left(\frac{\partial I_0}{\partial x}, \frac{\partial I_0}{\partial y} \right)(x_i) \tag{3.9}$$

3.9 denklemi ile x_i noktasındaki görüntü gradyanı belirtilmiştir. Bu gradyan, çeşitli teknikler kullanılarak hesaplanabilir (Schmid vd. 2000). Klasik Harris bulucusu (Harris ve Stephens 1988), $[-2 \ -1 \ 0 \ 1 \ 2]$ filtresini kullanır. Ama daha modern çeşitleri (Schmid vd. 2000, Triggs 2004) görüntüyü bir Gauss'un (genel olarak $\sigma = 1$) yatay ve düşey türevleriyle katlayarak hesaplar.

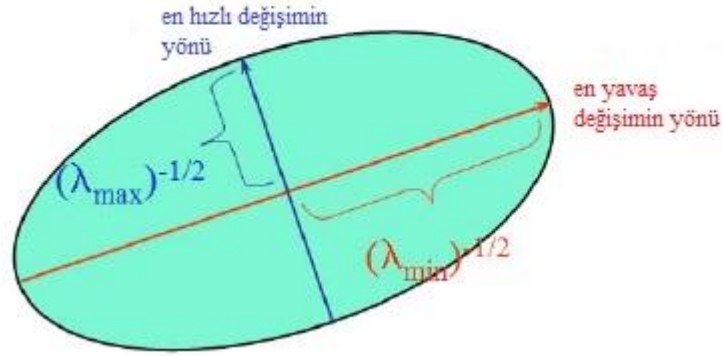
Oto-korelasyon matrisi A aşağıdaki gibi hesaplanabilir:

$$A = w * \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \tag{3.10}$$

Ağırlıklı toplamaları, ağırlıklandırma fonksiyonu w 'ya sahip ayırık katlama ile yer değiştirilir. Gradyanlar, ağırlıklandırma fonksiyonu w ile katlanarak oto-korelasyon fonksiyonunun yerel şeklinin her bir piksel için kestirimi hesaplanır (3.10).

A matrisinin tersi, eşleştirilen görüntü parçasının konumunda belirsizlik üzerinde daha daraltılmış bir alan sunar. Bu yüzden, hangi görüntü parçalarının kararlı olarak eşleştirildiğinin yararlı bir göstergesidir. Bu belirsizliği görselleştirmenin en kolay yöntemi, A oto-korelasyon matrisinin sonucunda iki öz değer ve iki öz vektör elde edilen, öz değer analizini gerçekleştirmektir (Şekil 3.13). Daha büyük belirsizlikler daha

küçük öz değerlere bağlı olduğundan takip edilecek iyi özniteliklerin yerini tayin etmek için daha küçük öz değerdeki maksimumu bulmak mantıklıdır (Shi ve Tomasi 1994).



Şekil 3.13 Oto-korelasyon matrisinin bir öz değer analizine karşılık gelen belirsizlik elipsi

Anandan vd. oto-korelasyon matrisinin belirsiz yapısını analiz ederken optik akış ölçümleriyle de ilişkilendirmişlerdir. Förstner (1986), Harris ve Stephens (1988), seyrek öznitelik eşleştirme amacıyla noktaların yerini tayin etmek için oto-korelasyon matrisinden türetilen rotasyondan bağımsız skaler ölçülerdeki yerel maksimumları kullanmayı ilk önerenlerdir. Schmid vd. (2000) ve Triggs (2004) öznitelik tespit algoritmalarının tarihsel gözden geçirmesi üzerine daha fazla detaylar vermişlerdir. Bu tekniklerin her ikisi de önceden kullanılan kare pencereler yerine Gauss ağırlıklandırma penceresi kullanmayı önermişlerdir. Çünkü kare pencereler, düzlem içindeki görüntü rotasyonlarına duyarsızdırlar.

En düşük λ_0 öz değeri, anahtar nokta bulmak için kullanılan tek nicelik değildir (Shi ve Tomasi 1994). Harris ve Stephens (1988) tarafından önerilen daha basit nicelik $\alpha = 0,06$ olacak şekilde aşağıdaki gibi hesaplanır.

$$\det(A) - \alpha \cdot \text{trace}(A)^2 = \lambda_0 \lambda_1 - \alpha (\lambda_0 + \lambda_1)^2 \quad (3.11)$$

Öz değer analizinden farklı olarak bu nicelik karekök kullanımını gerektirmez. Fakat hala rotasyona değişimsizdir. Triggs (2004), $\alpha = 0,05$ olacak şekilde aşağıdaki formülü önermiştir.

$$\lambda_0 - \alpha \cdot \lambda_1 \quad (3.12)$$

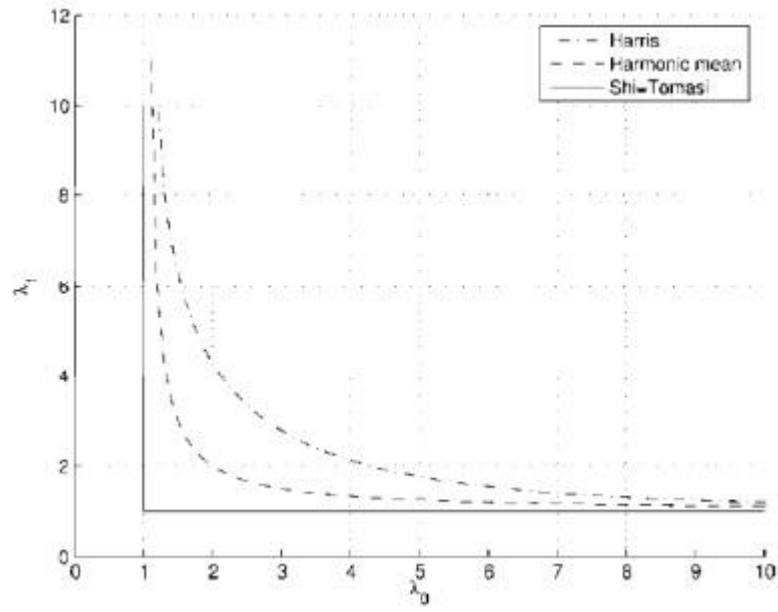
Bu denklem, örtüşme hatalarının küçük öz değerlerin bazen şişmesine sebep olduğu 1B kenarlara olan yanıtı da düşürme avantajı sağlar. Noktaların ölçekleme ve rotasyondan bağımsız olarak yerlerinin tayin edilebilmesi için temel 2 x 2 Hessian'ın parametrik hareketlere nasıl genişletileceğini göstermiştir. Diğer taraftan, Brown vd. (2005) $\lambda_0 \approx \lambda_1$ olduğu bölgelerde daha yumuşak geçişlere sahip aşağıda belirtilen harmonik ortalamayı kullanır.

$$\frac{\det A}{\text{tr } A} = \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1} \quad (3.13)$$

Farklı ilgi noktaları operatörlerinin eş çevreleri, Şekil 3.14'te gösterilmiştir. Grafikteki operatörler, son ilgi değeri belirlemek için iki öz değerın nasıl kademeli olarak geçtiğini göstermektedir. Her bir algılayıcı,

$$A = w * \nabla I \nabla I^T \quad (3.14)$$

denkleminin λ_0 ve λ_1 öz değerlerinin büyük olduğu noktaları araştırır.



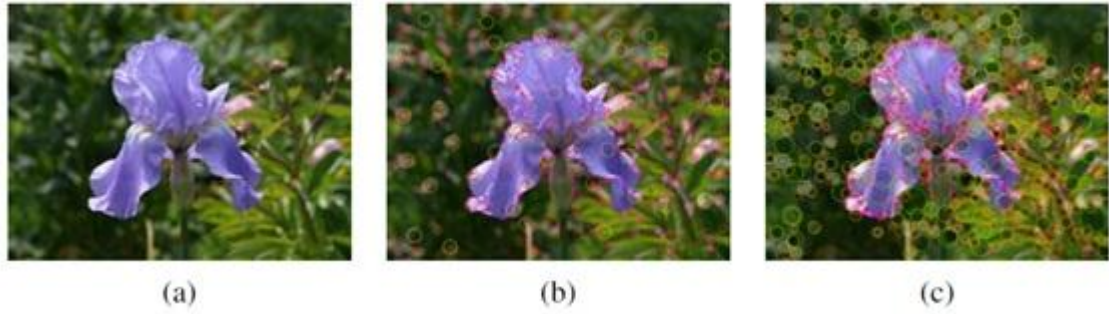
Şekil 3.14 Popüler anahtar nokta tespit etme fonksiyonların eş çevreleri

Temel oto-korelasyon tabanlı anahtar nokta bulucu algoritmasının aşamaları, Çizelge 3.2’de özetlenmiştir.

Çizelge 3.2 Temel öznitelik tespit etme algoritması ana hatları

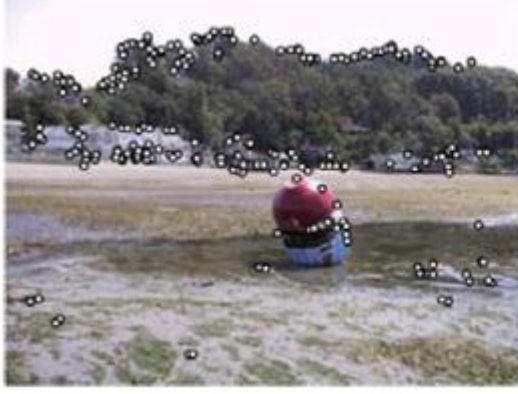
1. Orijinal görüntüyü Gauss türevleri ile katlayarak I_x ve I_y görüntülerinin yatay ve düşey türevlerini hesapla.
2. Bu gradyanların çarpımına karşılık gelen üç görüntüyü hesapla. (A matrisi simetriktir. Bu yüzden sadece üç girdiye ihtiyaç duyulur.)
3. Bu görüntülerin her birini daha büyük bir Gauss ile katla.
4. Yukarıda anlatılan formüllerden birini kullanarak skalar ilgi ölçüsünü hesapla.
5. Bir eşik değeri aşan yerel maksimumları bul ve bunları tespit edilen öznitelik nokta konumları olarak bildir.

Şekil 3.15'te örnek bir görüntüye uygulanan Harris bulucusu ve sonraki bölümde bahsedilecek Gauss fark (Difference of Gaussian - DoG) bulucusu sonuçları gösterilmiştir. Şekil 3.15.a örnek görüntüyü, şekil 3.15.b Harris sonucunu, şekil 3.15.c ise DoG sonucunu göstermektedir. Çember boyutları ve renkler, tespit edilen her bir ilgi noktasındaki ölçeği gösterir. İki algılayıcının tamamlayıcı konumlarda cevap verme eğiliminde olduğu görülmektedir.



Şekil 3.15 İlgi operatörlerinin yanıtları

Birçok öznitelik bulucu, sade bir şekilde ilgi fonksiyonundaki yerel maksimumları araştırırken bu durum görüntüdeki öznitelik noktalarının düzgün olmayan bir şekilde dağılımına sebep olabilmektedir. Daha yüksek karışıklığın bulunduğu bölgelerde noktaların daha yoğun olması buna örnek verilebilir. Brown vd. (2005), bu problemi hafifletmek için sadece hem yerel maksimum olan hem de yanıt değeri bir r yarıçapı içindeki komşularından %10 daha fazla olan noktaları öznitelik olarak tespit etmektedir. Şekil 3.16 ile en iyi n sayıda öznitelik seçimi ve ANMS kullanımının niceliksel karşılaştırılması gösterilmiştir. Brown vd. (2005) verimli bir yöntem tasarlamışlardır. Bu yöntemde bastırma yarıçapı ve tüm yerel maksimumlar kullanılarak öncelikle öznitelikleri güçlülüklerine göre sıralamışlardır. Ardından bastırma yarıçapını azaltarak ikinci bir liste oluşturmuşlardır. Üstteki iki görüntü, en güçlü 250 ve 500 ilgi noktasını gösterirken alttaki iki görüntü uyarlamalı maksimum olmayan noktaların bastırılması ve r yarıçapındaki noktalar bastırılarak seçilen ilgi noktalarını gösterir. Bu yüzden, alttaki öznitelikler görüntü üzerinde daha düzenli bir dağılım gösterir.



(a) En güçlü 250



(b) En güçlü 500



(c) ANMS 250, $r = 24$



(d) ANMS 500, $r = 16$

Şekil 3.16 Uyarlamalı Maksimum Olmayan Noktaların Bastırımı (ANMS)

3.3.1.1 Öznitelik özellikleri

3.3.1.1.1 Tekrarlanabilirlik miktarı

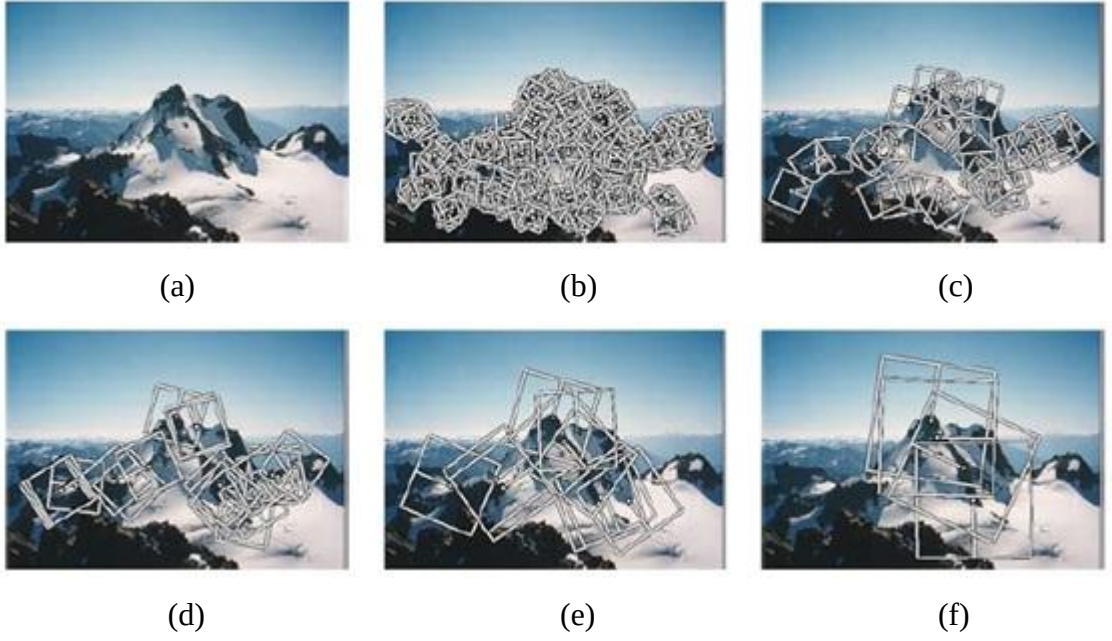
Bilgisayar görüşü için çok sayıda öznitelik bulucu geliştirilmiştir. Bunlar arasından hangisinin kullanılacağına dair çeşitli çalışmalar yapılmıştır. Schmid vd. (2000) öznitelik bulucuların tekrarlanabilirliklerinin miktarını öneren ilk çalışmayı yapanlardır. Bir görüntüde bulunan anahtar noktalardan hangilerinin, dönüştürülmüş görüntüdeki karşılık gelen konumundan ϵ sayıda ($\epsilon = 2$ gibi) piksel etrafında sıklıkla ortaya çıktığını tanımladılar. Çalışmalarında orijinal görüntüye rotasyon, ölçekleme, ışıklılık, bakış açısı değişiklikleri ve gürültü ekleyerek görüntüyü dönüştürmüşlerdir. Ayrıca her bir tespit

edilen öznitelik noktasının içerdiği bilgi miktarını ölçmüşlerdir. Bunun için rotasyondan bağımsız yerel gri seviye tanımlayıcılar kümesinin entropisini kullanmışlardır. Araştırdıkları teknikler sonucunda $\sigma_d = 1$ (türev Gauss'unun ölçeği) ve $\sigma_i = 2$ (integrasyon Gauss'unun ölçeği) olan Harris operatörünün geliştirilmiş (Gauss türevi) versiyonunun en iyi sonucu verdiğini bulmuşlardır.

3.3.1.1.2 Ölçekten bağımsızlık

Birçok durumda, en kararlı ölçekte öznitelikleri tespit etmek mümkün olmayabilir. Örneğin, yüksek frekans bilgi içermeyen görüntülerde (bulut gibi) eşleştirme yapılmak istendiğinde iyi ölçek öznitelikleri mevcut olmayabilir.

Farklı ölçeklerde öznitelik çıkartımı bu probleme bir çözümdür. Bir piramit üzerinde farklı çözünürlüklerde aynı operasyonu gerçekleştirerek ve aynı seviyeden öznitelik eşleştirme yapılarak bu işlem gerçekleştirilir. Bu tür bir yaklaşım, büyük ölçek değişimlerine maruz kalmamış eşleştirilmek istenen görüntüler için uygundur. Uçaklardaki kameralardan alınan ve ard arda gelen görüntüler eşleştirilmek istendiğinde veya sabit odak uzaklığına sahip kameralardan alınan görüntülerden panoramik bir görüntü elde edilmek istendiğinde bu tür bir yaklaşım kullanılabilir. Brown vd. (2005) çoklu ölçek yönelimli görüntü parçası (multi-scale oriented patches - MOPS) türünden bir algılayıcı kullandığı yaklaşım Şekil 3.17'de verilmiştir. Bu yaklaşımda, beş piramit seviyesinde beş farklı ölçek için çoklu ölçek yönelimli görüntü parçaları (MOPS) elde edilmiştir. Kutular, öznitelik yönelimini ve tanımlayıcı vektörlerinden örneklenen bölgeyi göstermektedir.



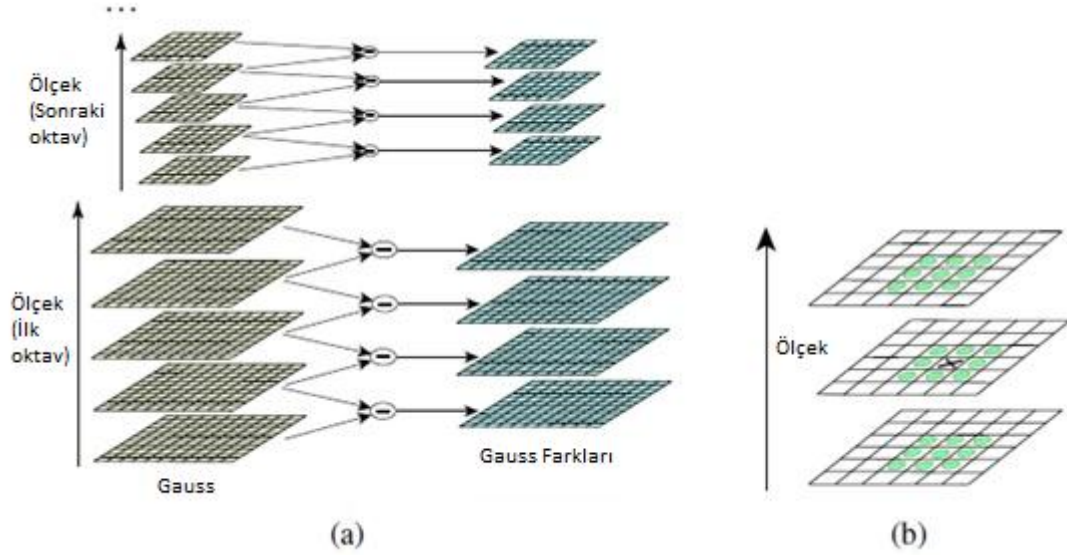
Şekil 3.17 Çoklu ölçek yönelimli görüntü parçaları

Nesne tanıma uygulamalarında genellikle görüntüdeki nesnenin ölçeği bilinmez. Birçok farklı ölçekte öznitelik çıkartmak ve bunların hepsini eşleştirmek yerine hem konumda hem de ölçekte kararlı olan özniteliklerin çıkartılması daha etkilidir (Lowe 2004, Mikolajczyk ve Schmid 2004).

3.3.1.1.2.1 Ölçekleme Değişmezlikli Öznitelik Dönüşümü (SIFT)

Lindeberg (1993, 199b) ölçek seçimi konusunda gerçekleştirilen ilk çalışmalarda ilgi noktası olarak Gauss'un Laplace işleci (Laplacian of Gaussian - LoG) fonksiyonlarındaki uç değer kullanımını önermiştir. Bu çalışma doğrultusunda Lowe (2004), Gauss filtrelerin farklarının (Difference of Gaussian - DoG) alt oktavlarının kümesini hesaplamayı (Şekil 3.18.a, Lowe 2004) ve bu yapı içinde 3B (ölçek + uzam) maksimumların araştırılmasını (Şekil 3.18.b, Lowe 2004) önermiştir. Ardından bir karesel oturtma kullanarak piksel altı uzam ve ölçek konumları hesaplanmıştır (Brown ve Lowe 2002). Şekil 3.18.a'da Gauss fark görüntüleri oluşturmak için oktav altı Gauss piramidin komşu seviyelerinin birbirinden çıkartıldığı gösterilmektedir. Şekil 3.18.b'de bir pikselin, 26 komşusu ile karşılaştırılarak uç noktaların (maksimum ve minimum) tespit edildiği gösterilmiştir. Deneysel çalışmalar sonucunda oktav altı seviyesi üç

olarak belirlenir. Bu da Triggs (2004) tarafından kullanıldığı gibi çeyrek oktav piramide karşılık gelir.



Şekil 3.18 Gauss farklarının (DoG) alt oktav piramidi kullanılarak ölçek – uzam öznelilik tespiti (Lowe 2004)

Gösterge fonksiyonunun (bu durum için Gauss Farkları) yerel eğrisindeki kuvvetli asimetri bulunan pikseller, Harris operatöründeki gibi, reddedilir. Bu işlem öncelikle fark görüntüsünün (D) yerel Hessian değeri (H) hesaplanarak gerçekleştirilir.

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (3.15)$$

Aşağıdaki denkleme göre anahtar noktalardan uygun olmayanlar elenir.

$$\frac{Tr(H)^2}{Det(H)} > 10 \quad (3.16)$$

Lowe'un ölçekten bağımsız öznelilik dönüşümü (SIFT) pratikte başarıyla oto-korelasyon tabanlı algılayıcıların kullandığı gibi aynı maksimum uzamsal kararlılık

teorik esasından farklıdır. Aslında SIFT'in bulduğu konumlar maksimum uzamsal kararlılık türünden tekniklerden elde edilenler konumlar ile genellikle tamamlayıcı niteliktedir. Bu yüzden diğer yaklaşımlarla birlikte kullanılabilir. Mikolajczyk ve Schmid (2004), Harris köşe algılayıcıya ölçek seçim mekanizmasını eklemek için her bir bulunan Harris noktasında (çoklu ölçek piramidinde) Gauss'un Laplace işlecini test etmiştir. Sadece Laplace işlecinin hem daha genel hem de daha ayrıntılı seviyelerdeki değerlerden daha büyük veya daha küçük olan uç değerleri seçmiştir. İsteğe bağlı yinelemeli bir ayrıştırma hem ölçek hem de konum için önerilmiş ve değerlendirilmiştir. Ölçekten bağımsız bölge algılayıcılar için ek örnekler Mikolajczyk vd. (2005), Tuytelaars ve Mikolajczyk (2007) tarafından verilmiştir.

3.3.1.1.3 Rotasyon değişmezlik ve yönelim kestirimi

Birçok görüntü eşleştirme ve nesne tanıma algoritması, ölçek değişimlerinin üstesinden gelmenin yanında en azından görüntü içindeki rotasyonlarla da başa çıkmalıdır. Bu problem ile başa çıkmanın bir yöntemi, dönüden bağımsız tanımlayıcıların tasarlanmasıdır (Schmid ve Mohr 1997). Ama bu tür tanımlayıcılar, zayıf ayırım kabiliyetine sahiptir. Farklı görünümdeki görüntü parçalarını aynı tanımlayıcı ile gösterebilmektedir.

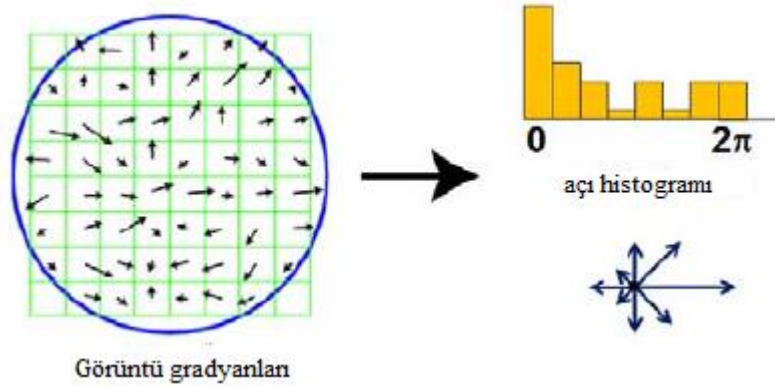
Tespit edilen her bir anahtar noktadaki bir baskın yönelim kestirimi daha iyi bir yöntemdir. Bir anahtar noktanın ölçeği ve yerel yönelimi kestirildikten sonra tespit edilen nokta etrafında ölçekli ve yönelimli bir görüntü parçası çıkartılır ve bir öznitelik tanımlayıcı oluşturmak için kullanılır (Şekil 3.17 ve Şekil 3.23).

En basit olası yönelim kestirimi, anahtar nokta etrafındaki bir bölge içinde gradyanların ortalamasını hesaplamaktır. Eğer bir Gauss ağırlıklandırma fonksiyonu kullanılırsa (Brown vd. 2005) ortalama gradyan birinci dereceden yönlendirilebilir bir filtreye eşittir. Bu işlem, Gauss filtresinin yatay ve düşey türevleriyle görüntü katlaması yapılarak hesaplanabilir (Freeman ve Adelson 1991). Bu kestirimi daha güvenilir yapmak için tespit penceresi (Brown vd. 2005) yerine daha geniş bir kurgulama

penceresi kullanımı genellikle tercih edilebilir. Şekil 3.17’deki kare kutuların yönelimi bu teknik kullanılarak hesaplanmıştır.

3.3.1.1.3.1 Ölçekleme Değişmezlikli Öznitelik Dönüşümü (SIFT)

Bir bölge içindeki ortalama gradyan küçük olabilir ve güvenilir olmayan bir yönelim göstergesi olabilir. Daha güvenilir bir teknik ise anahtar nokta etrafında yönelimin histogramına bakmaktır. Lowe (2004), hem gradyan büyüklükleri hem de merkeze olan Gauss uzaklıkları ile ağırlıklandırılan kenar yönelimlerinin 36 seleye bölünmüş histogramını hesaplar. En büyük değerin %80’lik kısmında kalan tüm tepeleri bulur. Ardından üç seleye parabolik oturtma kullanarak daha doğru bir yönelim kestirimi hesaplar (Şekil 3.19, Lowe 2004). Şekil 3.19’da tüm gradyan yönelimlerinin (ağırlıklandırılarak veya küçük gradyanlar ile eşiklenerek) histogramını oluşturarak ve ardından bu dağılımdaki anlamlı tepeleri bularak bir baskın yönelim kestirimi hesaplanır.



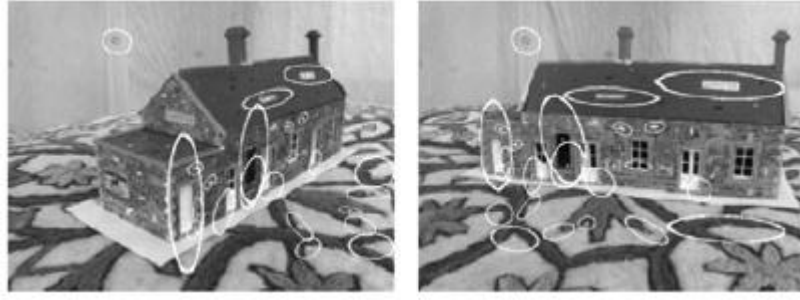
Şekil 3.19 Baskın yönelim kestirimi (Lowe 2004)

3.3.1.1.4 İlgin dönüşümünden bağımsızlık:

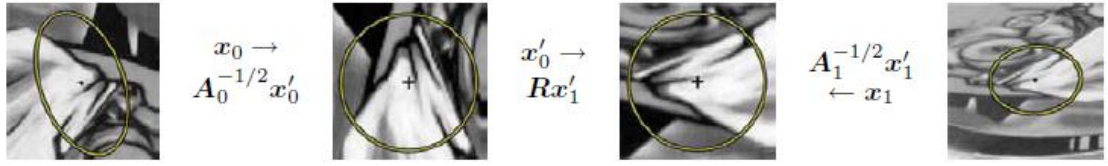
Ölçekleme ve rotasyon değişmezlik birçok uygulama için karşılanması gereken isterlerken stereo eşleştirme (Pritchett ve Zisserman 1998, Schaffalitzky ve Zisserman 2002) veya konum tanımlama (Chum vd. 2007) gibi bazı uygulamalar da tam ilgin dönüşüme karşı değişmezliğe ihtiyaç duyarlar.

İlgin bağımsız algılayıcılar, sadece ölçek ve yönelim değişimlerinden sonra tutarlı konumlar hesaplamakla kalmaz, perspektif gibi ilgin bozulmaları karşısında da tutarlı cevaplar vermektedir (Şekil 3.20). Şekil 3.20’de farklı bakış açılarından alınan iki görüntüyü eşleştirmek için kullanılan ilgin bölge algılayıcı kullanımı gösterilmiştir (Mikolajczyk ve Schmid 2004). Aslında yeteri kadar küçük görüntü parçaları için sürekli görüntü bükme, ilgin deformasyon kullanarak iyi bir şekilde yakınsanabilir.

İlgin bağımsızlığı açıklamak için oto-korelasyona veya Hessian matrise (öz değer analizi kullanarak) bir elips oturtma önerilmiştir. Ardından temel eksenler ve bunların oranı kullanılarak ilgin koordinat çerçevesi olarak oturtulmuştur (Lindeberg ve Garding 1997, Baumberg 2000, Mikolajczyk ve Schmid 2004, Mikolajczyk vd. 2005, Tuytelaars ve Mikolajczyk 2007). Şekil 3.21’de yerel görüntü parçalarını rotasyon benzeri çerçevelere dönüştürmek için moment matrisinin karekökünün nasıl kullanılabileceği gösterilmiştir. Mikolajczyk vd. (2005) çalışmasında belirtildiği gibi ikinci moment matrisi kullanarak yapılan ilgin normalizasyonu Şekil 3.21 ile gösterilmiştir. Görüntü koordinatları, $A_0^{-1/2}$ ve $A_1^{1/2}$ matrisleri kullanılarak dönüştürüldükten sonra saf bir R dönüşü ile ilişkilendirilir.



Şekil 3.20 İlgin bölge algılayıcı kullanımı



Şekil 3.21 İlgin normalizasyonu

3.3.1.1.4.1 Maksimum kararlı uç bölge

Bir diğer önemli ilgin bağımsız bölge algılayıcı, Matas vd. (2004) tarafından geliştirilen maksimum kararlı uç bölge (Maximally Stable Extremal Region – MSER) algılayıcısıdır. MSER’leri bulmak için görüntü tüm olası gri seviyelerde eşiklenerek ikili bölgeler hesaplanır. Dolayısıyla bu teknik sadece gri seviyeli görüntülere uygulanabilir. Bu işlem, öncelikle tüm piksellerin gri değerlerine göre verimli bir şekilde sıralandığı ve sonrasında eşik değeri değiştikçe her bir bağlı bileşene artırımlı olarak piksel ilave edilerek gerçekleştirilir (Nister ve Stewenius 2008). Eşik değeri değiştikçe her bir bileşenin (bölgenin) alanı izlenir. Bir eşik değere göre alan değişim hızı minimum olan bölgeler maksimum kararlı olarak tanımlanır. Bu yüzden, bu bölgeler hem ilgin geometri dönüşümlerine hem de fotometrik dönüşümlerine karşı duyarsızdır. Şekil 3.22’de çeşitli görüntülerden MSER kullanılarak çıkartılmış ve eşleştirilmiş bölgeler gösterilmektedir (Matas vd. 2004). Eğer istenirse bir ilgin koordinat çerçevesi, her bir tespit edilen bölgeye moment matrisi kullanarak oturtulabilir.



Şekil 3.22 MSER kullanılarak çıkartılan ve eşleştirilen bölgeler

Öznitelik nokta algılayıcıları konusu, her yıl önemli bilgisayar görüşü konferanslarında görülen çalışmalar ışığında hala güncelleğini korumaktadır (Xiao ve Shah 2003, Koethe 2003, Carneiro ve Jepson 2005, Kenney vd. 2005, Bay vd. 2006, Platel vd. 2006, Rosten ve Drummond 2006). Mikolajczyk vd. (2005), popüler bir takım ilgin bölge algılayıcı araştırmışlardır. Bunun sonucunda, ölçekleme, rotasyon, gürültü ve bulanıklaşma gibi ortak görüntü dönüşümlerine karşı olan değişmezliklerin deneysel karşılaştırmasını yapmışlardır.

Çakıştırılmış görüntüler bulmak için kullanılan öznitelikler sadece anahtar noktalar değildir. Zoghlami vd. (1997), görüntü çiftleri arasındaki eş grafik (homographies) kestirimi için nokta benzeri öznitelikler kadar iyi çizgi bölütleri de kullanmışlardır. Diğer taraftan, Bartolli vd. (2004), 3B yapı ve hareket kestirimi için kenarlar boyunca yerel karşıtlıklı çizgi bölütlerini kullanmışlardır. Tuytelaars ve Van Gool (2004), stereo eşleştirme uygulamasında benzerlikleri tespit etmek için ilgin değişimsiz bölgeler kullanırken; Kadir vd. (2004), görüntü entropisinin ve ölçekleme ile değişim hızının yerel olarak maksimum olduğu dikkat çeken bölgeleri tespit etmişlerdir. Corso ve Hager (2005), türdeş bölgeleri tespit etmek amacıyla 2B yönelimli Gauss çekirdekler oturtmak için benzer bir teknik kullanmışlardır. Eğriler, çizgiler ve bölgeler bulmak ve eşleştirmek için kullanılan bazı teknikler de sonraki bölümlerde verilmiştir.

3.3.1.2 Öznitelik tanımlayıcılar

Öznitelikleri (anahtar noktaları) bulduktan sonra onları eşleştirmeliyiz. Hangi özniteliklerin farklı görüntülerdeki karşılık gelen konumlardan geldiğini belirlemeliyiz. Video dizileri (Shi ve Tomasi 1994) veya doğrultulmuş stereo çiftleri (*rectified stereo patches*) (Zhang vd. 1995, Loop ve Zhang 1999, Scharstein ve Szeliski 2002) gibi bazı

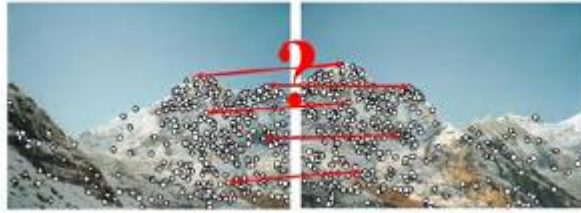
durumlarda her bir öznitelik noktası etrafındaki yerel hareket genellikle öteleme hareketidir. Bu durumlarda karesel farkların toplamı veya normalize edilmiş çapraz ilinti gibi basit hata metrikleri, her bir öznitelik noktası etrafındaki küçük görüntü parçalarındaki ışıklılığı karşılaştırmak için direk olarak kullanılabilir. (Mikolajczyk ve Schmid (2005), karşılaştırmalı çalışmalarında çapraz ilintiyi kullanmışlardır.) Çünkü öznitelik noktaları direk olarak bulunamayabilir. Artımlı hareket ayrıştırması uygulayarak daha doğru eşleştirme puanları hesaplanabilir. Ancak bu işlem zaman alabilir. Hatta bazen performansı düşmesine bile sebep olabilir (Brown vd. 2005).

Ancak birçok durumda özniteliklerin yerel görünümü yönelim, ölçekleme ve bazen de ilgin deformasyonuna maruz kalarak değişecektir. Yerel bir ölçek, yönelim veya ilgin çerçeve kestirimi çıkartımı ve sonra bunun öznitelik tanımlayıcıları biçimlendirmeden önce görüntü parçalarını yeniden örneklemek için kullanımı genellikle tercih edilir (Şekil 3.23). Şekil 3.23'te yanlılık ve kazanç ile normalize edilmiş ışıklılık değerlerinin 8 x 8 örnekleme kullanılarak MOPS tanımlayıcılar biçimlendirilmiştir (Brown vd. 2005). Düşük frekansta örnekleme, ilgi noktası konum hatasına karşı özniteliklere bazı dayanıklılıklar sağlar. Tespit ölçekleme seviyesinden daha yüksek piramit seviyesinde örneklenerek bu işlem gerçekleştirilmektedir.



Şekil 3.23 MOPS tanımlayıcılar

Bu deęişiklikler dengelendikten sonra bile görüntü parçalarının yerel deęişimi hala görüntüden görüntüye deęişiklik gösterecektir. Farklı görüntü parçaları arasındaki ayrımsamayı korurken görüntü tanımlayıcıların belirtilen deęişikliklere karşı daha fazla deęişmez olmasını nasıl sağlayabiliriz (Şekil 3.24). Mikolajczyk ve Schmid (2005), bakış açısı deęişikliklerine karşı deęişmeyen bazı yerel görüntü tanımlayıcıları gözden geçirmişlerdir. Deneysel olarak performanslarını da karşılaştırmışlardır.



Şekil 3.24 Öznitelik eşleştirme

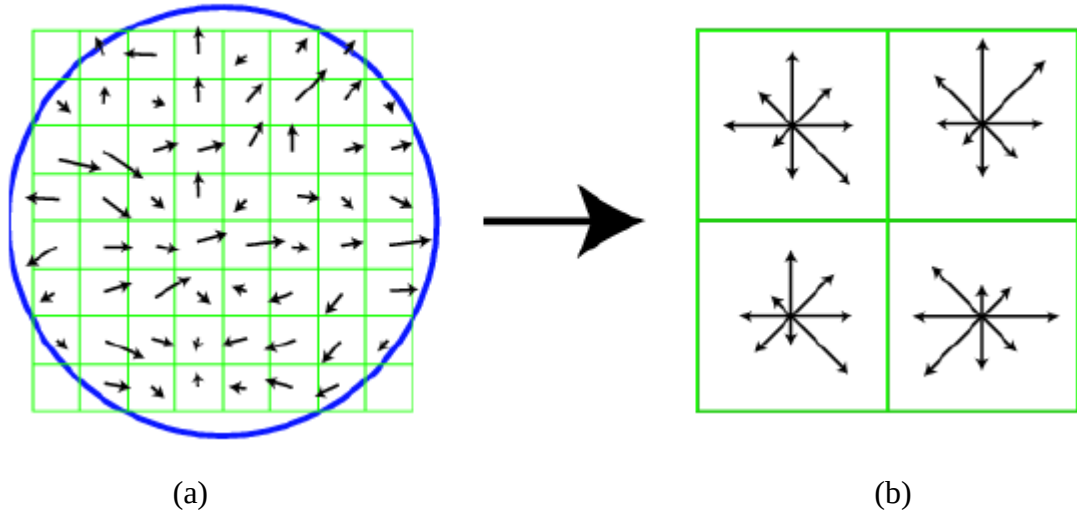
Sonraki bölümde, bazı tanımlayıcılar daha detaylı olarak anlatılmıştır.

3.3.1.2.1 Yanlılık ve kazanç normalizasyonu

Basit ışıklılığı normalize edilmiş görüntü parçaları, görüntü piramitleme gibi uygulamalar için gerçekleştirmesi basittir ve iyi sonuçlar vermektedir (Brown vd. 2005). Çok ölçekli yönelimli görüntü parçaları (MOPS), öznitelik nokta algılayıcıdaki konum, yönelim ve ölçekleme gibi az miktardaki hataları gidermek için tespit ölçeğine göre beş piksel aralıkta örneklenir. Örtüşmeyi engellemek için görüntü piramidinin daha yukarı seviyesi kullanılır. Görüntü parçaları, ilgin foto metrik deęişimleri (doğrusal ışıklılık deęişimleri, yanlılık veya kazanç) gidermek için ortalamaları sıfır ve deęişimleri (*variance*) bir olacak şekilde yeniden ölçeklenir.

3.3.1.2.2 SIFT

SIFT öz nitelikleri, tespit edilen anahtar nokta etrafında 16×16 boyutlarındaki çerçevenin her bir pikselinde gradyan hesaplayarak biçimlendirilir. Bu işlem için tespit edilen anahtar noktadaki Gauss piramidinin uygun seviyesi kullanılır. Gradyan büyüklükleri, merkezden uzak olan noktalar küçük hizalama hatalarından fazla etkilendiği için, merkezden uzak olan gradyanların etkisini azaltmak için Gauss azaltma fonksiyonu kullanarak düşük ağırlıklandırılır. Gaussian fonksiyonları, Şekil 3.25'te mavi çember ile gösterilmiştir. Gradyan yönelimleri ve büyüklükleri her bir pikselde hesaplanır. Ardından Gauss azaltma fonksiyonu ile ağırlıklandırılır. Bu işlemler sonucunda Şekil 3.25.a'daki (Lowe 2004) gibi bir sonuç elde edilir. Ardından Şekil 3.25.b'de (Lowe 2004) gösterildiği gibi doğrusal (trilinear) interpolasyon kullanarak her bir alt bölgedeki ağırlıklandırılmış bir gradyan yönelim histogramı hesaplanır. Şekil 3.25, 8×8 görüntü parçasını ve 2×2 tanımlayıcı dizisini gösterirken, Lowe'un kendi uygulamasında 16×16 görüntü parçaları ve 4×4 sekiz seleli histogram dizisi kullanmıştır.



Şekil 3.25 SIFT tanımlayıcı gösterimi

a. Görüntü Gradyanı, b. Anahtar Nokta Tanımlayıcı

Bir gradyan yönelim histogramı, sekiz yönelim histogram selesinin birine ağırlıklandırılmış gradyan değeri ekleyerek biçimlendirilir. Baskın yönelim hatalı kestirimini ve konum etkilerini azaltmak için ağırlıklandırılmış 256 gradyan büyüklüğünün her biri interpolasyon kullanarak $2 \times 2 \times 2$ histogram selesine eklenir.

Ortaya çıkan 128 pozitif değer SIFT tanımlayıcı vektörün ham bir versiyonunu oluşturmaktadır. 128B (128-Boyutlu) vektör, karışıklık ve kazanç etkilerini düşürmek için birim uzunluğa normalize edilir. Toplanır değişiklikler, gradyan ile zaten ortadan kaldırılmıştır. Değerler, tanımlayıcıları foto metrik değişimlere karşı daha dayanıklı yapmak için 0,2 değerine kırpılır. Ortaya çıkan vektör, birim uzunluğa bir kez daha normalize edilir.

3.3.1.2.3 PCA – SIFT

Ke ve Sukthankar (2004), tanımlayıcıları hesaplamak için SIFT'ten esinlenerek daha basit bir yöntem önermişlerdir. 39×39 boyutlarında görüntü parçası üzerinde x ve y gradyanları hesaplarlar. Ardından temel bileşen analizi (PCA), kullanarak ortaya çıkan 3042 boyutlu vektörü 36'ya düşürür.

3.3.1.2.4 SURF

SURF, SIFT'in bir diğer popüler versiyonudur. Bay vd. (2006), anahtar noktanın yönelimlerini tanımlamak için SIFT içerisinde kullanılan türevlere ve integrallere yakınsamak için kutu filtresi kullanırlar. Gradyan benzeri yerel operatörlerin histogramları ise öznitelik tanımlayıcı olarak tanımlanır. SIFT tanımlayıcı olarak 128 boyutlu bir vektör kullanmasına rağmen SURF tanımlayıcı olarak 64 boyutlu bir vektör kullanır. Bu yüzden hesapsal maliyet açısından SIFT yöntemine kıyasla daha ucuzdur. Bununla birlikte SIFT yöntemine göre başarımlar bazı uygulamalarda daha düşük olabildiği gibi bazı uygulamalarda da SIFT'e yakın hatta daha yüksek başarı elde edilebilir.

3.3.1.2.5 OpponentSIFT - OpponentSURF

SIFT ve SURF tanımlayıcı, sadece yoğunluk kanalındaki bilgileri kullanırlar. Bilgi teorisinden görüntüdeki nadir renk geçişlerinin çok ayırt edici olabileceği bilinmektedir. Bu yüzden, Sande vd. (2008), renk bilgisinin öznitelik tanımlayıcı için kullanımını önermişlerdir. Bu anlamda renkli görüntü (*red, green, blue - RGB*) aşağıdaki eşitlik kullanılarak karşı renk uzayına (*opponent color space*) dönüştürülür.

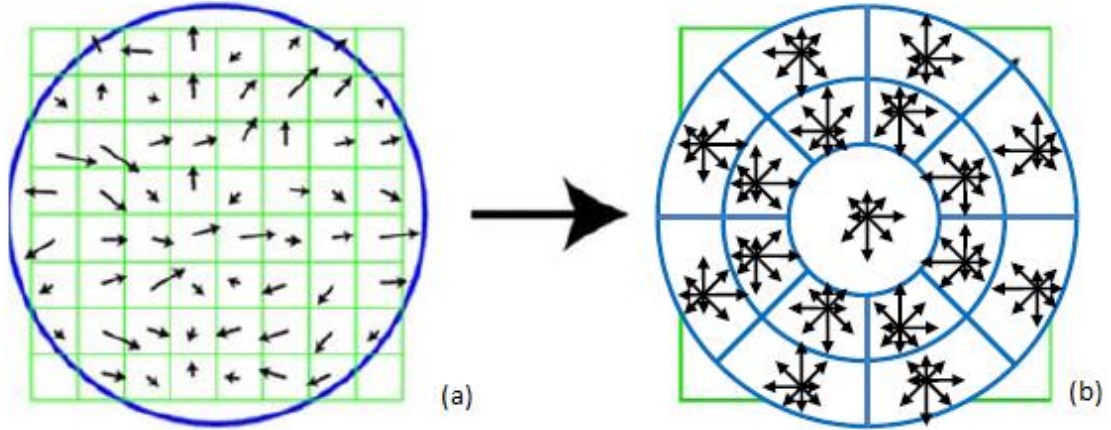
$$\begin{pmatrix} O_1 \\ O_2 \\ O_3 \end{pmatrix} = \begin{pmatrix} \frac{R-G}{\sqrt{2}} \\ \frac{R+G-2B}{\sqrt{6}} \\ \frac{R+G+B}{\sqrt{3}} \end{pmatrix} \quad (3.17)$$

Üçüncü kanal O_3 , yoğunluk bilgisini vermektedir. O_1 ve O_2 kanalları ise kırmızı – yeşil ve yeşil – mavi renk bilgilerini içermektedir. Ancak, renk bilgisi içeren kanallar yoğunluk bilgisini de içermektedirler. Bu yüzden ışıklılık değişimlerine karşı değişimsiz değildir. Her bir kanal, SIFT veya SURF tanımlayıcı kullanılarak tanımlanır. Bu yöntemler, OpponentSIFT ve OpponentSURF olarak adlandırılmaktadır.

3.3.1.2.6 Gradyan konum – yönelim histogramı

Bu tanımlayıcı, Mikolajczyk ve Schmid (2005) tarafından SIFT yönteminin değişik bir biçimi olarak geliştirilmişlerdir. Lowe (2004) tarafından kullanılan dört kuadrant yerine bir log - polar seleleme yapısı kullanmışlardır. Önerilen yöntem, Şekil 3.26’da gösterilmiştir. Uzamsal seleler, merkez bölgesi hariç sekiz açısız sele ile toplamda 17 uzamsal sele ve 16 yönelim selesi için 6, 11 ve 15 yarıçapları kullanılarak hesaplanır. 272 boyutlu histogram, geniş bir veritabanı üzerinde eğitilmiş temel bileşen analizi kullanılarak 128 boyutlu tanımlayıcıya izdüşümü alınmıştır. Mikolajczyk ve Schmid

(2005), değerlendirmeleri sonucunda GLOH yönteminin, SIFT yöntemini ufak bir farkla geçerek genelde en iyi performansı verdiğini bulmuşlardır.



Şekil 3.26 Gradyan konum – yönelim histogram tanımlayıcısı

a. Görüntü gradyanı, b. Anahtar nokta tanımlayıcı

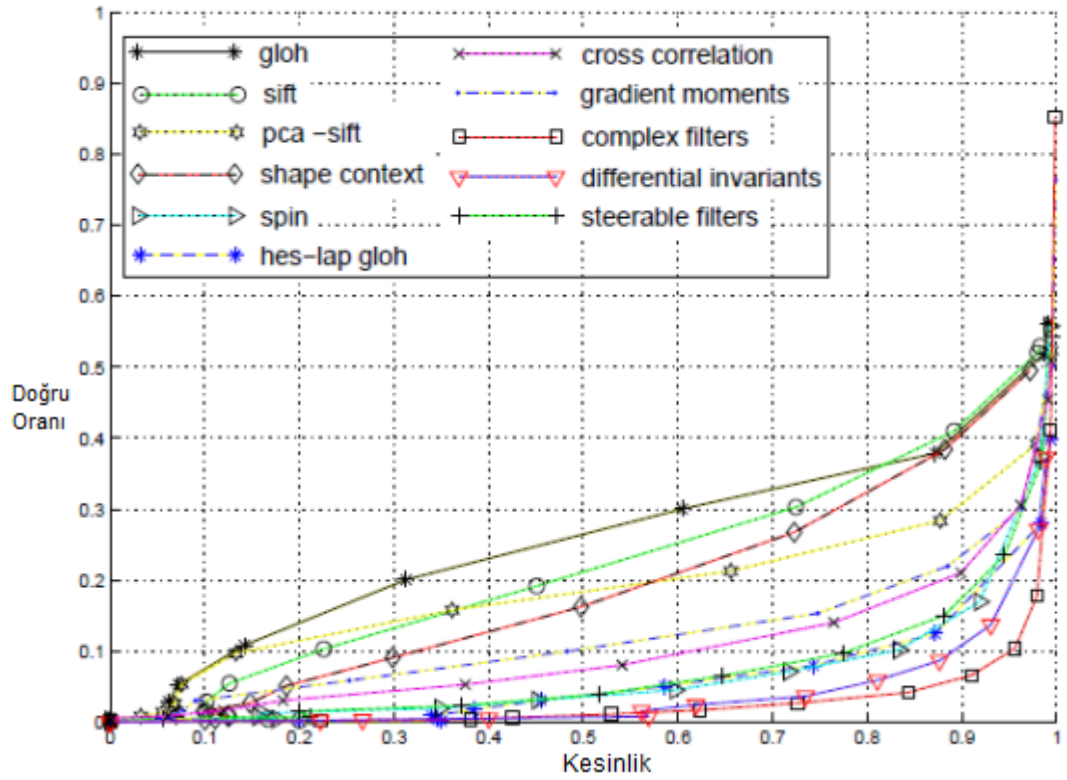
3.3.1.2.7 Yönlendirilebilir filtreler:

Yönlendirilebilir filtreler, tüm olası yönelimlerdeki simetrik ve simetrik olmayan kenar ve köşe benzeri özniteliklerin hızlı hesabına olanak sağlayan Gauss filtrelerinin türevlerinin kombinasyonlarıdır (Freeman ve Adelson 1991). Makul ölçülerde geniş Gauss'lar kullandıkları için yer tayin etme ve yönelim hatalarına da oldukça duyarsızdır.

3.3.1.2.8 Yerel tanımlayıcıların performansı:

Mikolajczyk ve Schmid (2005) tarafından karşılaştırılan yerel tanımlayıcılar arasında en iyi performansı GLOH vermiştir. Ardından az farkla SIFT gelmektedir. Öznitelik tanımlayıcısı karşılaştırma sonuçları Şekil 3.27'de verilmiştir.

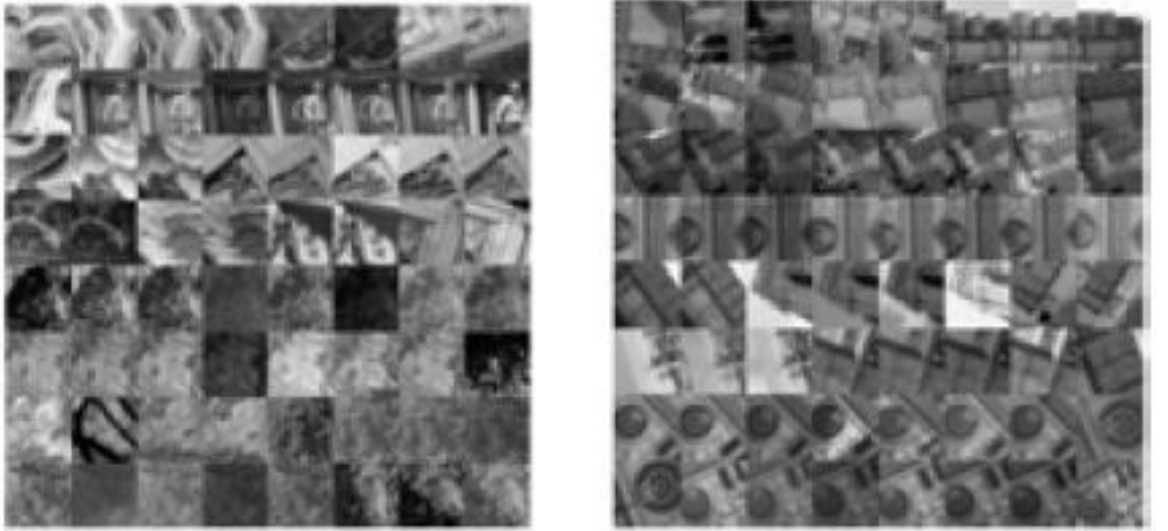
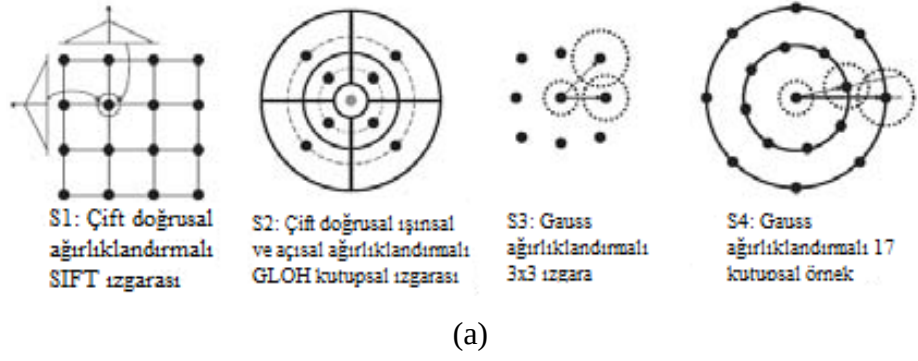
Öznitelik tanımlama alanı, yerel renk bilgisine bakan daha yeni tekniklerle (Van de Weijer ve Schmid 2006, Abdel-Hakim ve Farag 2006) gelişimini sürdürmektedir. Winder ve Brown (2007), hem SIFT hem de GLOH yöntemlerini içeren (Şekil 3.28.a) bir öznitelik tanımlayıcısı hesaplamak için çok aşamalı bir yapı geliştirmişlerdir. Önceden elle ayarlanan tanımlayıcılardan daha iyi performans veren yeni tanımlayıcılar için en iyi parametrelerin öğrenmesine izin verir.



Şekil 3.27 Öznitelik tanımlayıcı performans değerlendirme

Hua vd. (2007), en iyi ayırtma gücüne sahip daha yüksek boyutsal öznitelik tanımlayıcıların daha düşük boyutsal izdüşümlerini öğrenerek bu çalışmayı genişletmişlerdir. Bu çalışmalardan her ikisi de gerçek dünya görüntü parçalarının (Şekil 3.28.b) veri tabanını kullanmaktadırlar.

Bu teknikler, tüm nesne sınıflarına karşı tekrar kullanabilmek için en iyileştirilen öznitelik tanımlayıcılar oluştururken spesifik sınıflara karşı ayımsamayı maksimum yapacak özel bir sınıfa veya örneğe bağlı öznitelik tanımlayıcılar geliştirmek de mümkündür.



(b)

Şekil 3.28 SIFT ve GLOH yöntemleri için uzamsal toplama blokları

3.3.1.3 Öznitelik eşleştirme

İki veya daha çok görüntüden öznitelikler ve özniteliklerin tanımlayıcıları çıkartıldıktan sonra yapılması gereken bu görüntüler arasında bazı ön öznitelik eşleştirmeler tespit etmektir. Bu bölümde, eşleştirme stratejisi ve bu stratejilerin hata oranları verilmiştir.

3.3.1.3.1 Eşleştirme stratejisi ve hata oranı

Bir sonraki işlemde kullanmak için hangi öznitelik eşleştirmenin makul olacağına karar vermek eşleştirmenin gerçekleştirileceği içeriğe bağlıdır. Büyük ölçüde örtüşen iki görüntü verilmiş olsun. Bazı özniteliklerin, olabilecek örtme ve görünümündeki fazla miktardaki değişimler sebebiyle bazı özniteliklerde eşleşme olmayabilmesine rağmen birçok özniteliğin diğer görüntülerde muhtemelen eşleşecektir.

Şekil 3.29'daki gibi gürültülü bir ortamda bilinen nesnelerden kaç tane bulunduğunu tespit etmek istiyorsak özniteliklerin birçoğu eşleşmeyebilir. Şekil 3.29'un sol tarafında bulunmak istenen görüntülerden iki tanesi gösterilmiştir. Bu görüntüler, ortada bulunan gürültülü sahne ile SIFT özniteliği kullanılarak eşleştirilmiştir. Tespit edilen nesneler, sağdaki görüntüde çerçeve içine alınarak gösterilmektedir. Her bir tanımlanmış veri tabanı görüntüsünün ilgin bükülmesine uğraması sonucu değişen nesneler, sağdaki görüntüde daha geniş bir çerçeve olarak gösterilir.

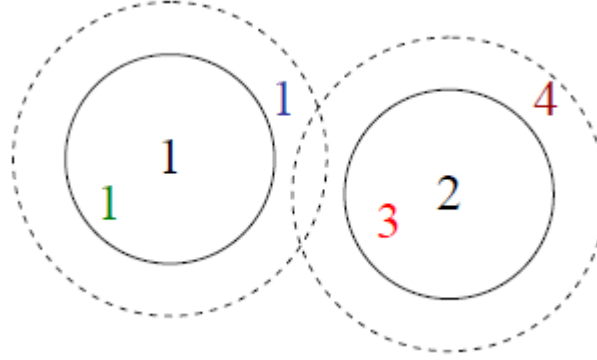


Şekil 3.29 Gürültülü ortamda nesne tanıma

Başlangıç olarak öznitelik uzayındaki Öklid (vektör büyüklüğü) uzaklıklarının, potansiyel eşleştirmelerin sıralaması için direk olarak kullanılabildiği öznitelik tanımlayıcıların tasarlandığını varsayalım. Bir tanımlayıcıdaki belirli parametreler (eksenler), diğerlerinden daha güvenilir olursa genellikle bu eksenleri yeniden ölçeklemek tercih edilir. Bu işlem, diğer bilinen iyi eşleştirmeler ile karşılaştırıldığı zaman ne kadar değiştiğine karar vererek yapılır (Hua vd. 2007). Daha genel bir işlem

ise beyazlatıcı (*whitening*) olarak adlandırılan ve öznitelik vektörlerinin ölçeklendirilmiş yeni bir tabana dönüştürülmesidir.

Bir Öklid uzaklığı metriği verilmiş olsun. En basit eşleştirme stratejisi, bir eşik değer (maksimum uzaklık) belirlemek ve bu eşik değeri içinde kalan diğer görüntülerden tüm eşleştirmeleri elde etmektir. Eşik değeri, çok büyük bir değer belirlemek birçok istenmeyen nesnelerin tespiti (doğru olmayan eşleştirme, yanlış kabul, *false positive*) ile sonuçlanır. Eşik değeri çok düşük bir değer belirlemek ise birçok yanlış ret hatasına (ikinci tür hata, *false negative*) sebep olur (Şekil 3.30).



Şekil 3.30 Yanlış kabuller ve redler

Şekil 3.30’da siyah renkli 1 ve 2, diğer görüntülerdeki özniteliklerin veri tabanı karşısında eşleştirilen özniteliklerdir. Kesintisiz çember ile gösterilen eşik değeri karşısında elde edilen yeşil 1 iyi eşleştirmeyi (*true positive*) gösterirken mavi 1 yanlış ret hatasını (*false negative*) ve kırmızı 3 ise doğru olmayan eşleştirmeyi (yanlış kabul, *false positive*) göstermektedir. Eğer daha yüksek bir eşik değeri belirlenirse (kesikli çember ile gösterilen) mavi 1 doğru eşleştirme olur. Ama bu sefer de kahverengi 4 yanlış kabul olur.

Özel bir eşik değeri üzerindeki eşleştirme algoritmasının performansı, aşağıdaki tanım kullanılarak doğru ve yanlış eşleştirmelerin ve eşleştirme başarısızlıklarının sayısını sayarak ölçülür (Fawcett 2006).

- Doğru Kabul (*True Positive*) : Hedef varken eşleştirme var

- Yanlış Ret (*False Negative*) : Hedef varken eşleştirme yok
- Yanlış Kabul (*False Positive*) : Hedef yokken eşleştirme var
- Doğru Ret (*True Negative*) : Hedef yokken eşleştirme yok

Çizelge 3.3, yukarıda bahsedilen sayıları içeren örnek bir hata matrisini gösterir. Öznitelik eşleştirme algoritması kullanarak kestirilen doğru ve yanlış eşleştirme sayıları ifade edilmiştir. Bu çizelgede, TP doğru tespitlerin sayısını, FP yanlış kabullerin sayısını, FN yanlış retlerin ve TN ise nesne yokluğunun doğru olmasının sayısı göstermektedir. P ve N, doğru (pozitif) ve yanlış (negatif) koşulların toplamını gösterirken P' ve N', tahmin edilen doğru ve yanlış koşulların toplamını gösterir. TPR, doğru pozitif oranını; FPR, yanlış pozitif oranını; PPV, pozitif tahmin edilen değeri; AAC ise doğruluğu gösterir.

Çizelge 3.3 Doğru ve yanlış eşleştirme sayıları

	Doğru Eşleştirme		Doğru Eşleştirmeme	
Eşleştirme Tahmini	TP = 18	FP = 4	P' = 22	PPV = 0.82
Eşleştirmeme Tahmini	FN = 2	TN = 76	N' = 78	
	P = 20	N = 80	Toplam = 100	ACC = 0.94

Aşağıdaki ifadeleri tanımlayarak bu sayıları birim orana çevirebiliriz (Fawcett 2006):

- Doğru Pozitif Oranı (TPR – *True Positive Rate*),

$$TPR = \frac{TP}{TP + FN} = \frac{TP}{P} \quad (3.18)$$

- Yanlış Pozitif Oranı (FPR – *False Positive Rate*),

$$FPR = \frac{FP}{FP + TN} = \frac{FP}{N} \quad (3.19)$$

- Pozitif Tahmin Değeri (PPV – *Positive Predicted Value*),

$$PPV = \frac{TP}{TP + FP} = \frac{TP}{P'} \quad (3.20)$$

- Doğruluk (ACC - *Accuracy*),

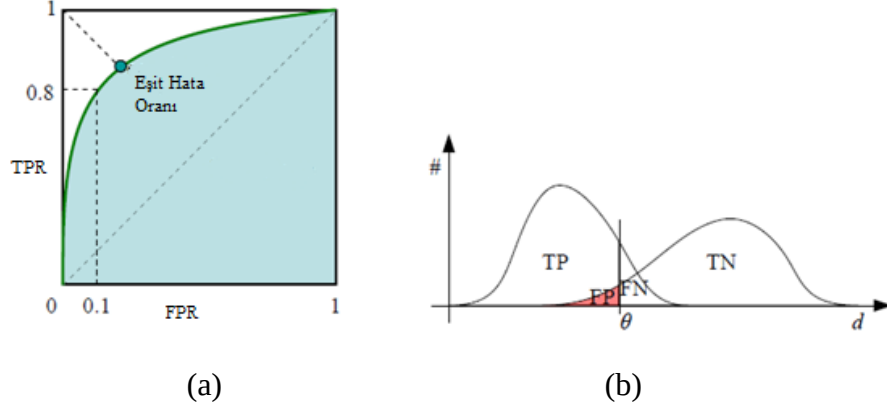
$$ACC = \frac{TP + TN}{P + N} \quad (3.21)$$

Bilginin geri kazanımı literatüründe (Baeza-Yates ve Riveiro-Neto 1999, Manning vd. 2008), pozitif tahmin değeri (PPV) yerine duyarlık (*precision*) terimi kullanılır. Bu terim, geri kazanılan bilgilerden kaç tanesinin yararlı (ilgili) olduğunu belirtir. Doğru pozitif oranı (TPR) yerine ise geri getirme (*recall*) terimi kullanılır. Bu terim, yararlı (ilgili) bilginin bulunma oranını ifade eder.

Özel bir eşleştirme stratejisi, özel bir eşik seviyesi veya parametre ayarları ile doğru pozitif oranı (TPR) ve yanlış pozitif oranı (FPR) ile değerlendirilebilir. İdealde doğru pozitif oranı 1 değerine yakın, yanlış pozitif oranı ise 0 değerine yakın olur. Eşleştirme eşik değerini değiştirdiğimiz sürece, topluca ROC (*receiver operating characteristic*) eğrisi denilen bir noktalar kümesi elde ederiz. Şekil 3.31’de ROC eğrisi gösterilmiştir. ROC eğrisi, öznitelik çıkartımı ve eşleştirme algoritmalarının özel bir kombinasyonu için yanlış pozitif oran karşısında doğru pozitif oran değişimi çizilmiştir. ROC eğrisi altında kalan alan (*Area Under Curve* - AUC), algoritma performansının tek sayıl (scaler) ölçümü olarak sık sık kullanılır. ROC eğrisi, sol üst köşeye yaklaştıkça (eğri altında kalan alan (AUC) arttıkça) performans artar. Bazen, eşit hata oranı (*equal error rate*) alternatif olarak kullanılmaktadır. Şekil 3.31.b’de öznitelikler arası d uzaklığının bir fonksiyonu olarak pozitiflerin ve negatiflerin dağılımı verilmiştir. θ eşik değeri arttıkça doğru pozitiflerin ve yanlış pozitiflerin sayısı artar. Ardından bu eğriler kullanılarak ROC eğrisi çizilir. ROC eğrisi, ortalama doğruluğu (*mean average precision*) hesaplamak için de kullanılabilir.

Sabit bir eşik değeri kullanılmadaki problem, kullanılacak değeri belirleme zorluğudur. Öznitelik uzayının farklı parçalarına hareket ettikçe eşik değerlerin faydalı olan aralığı büyük miktarda değişebilir (Mikolajczyk ve Schmid 2005). Bu gibi durumlardaki daha

iyi bir strateji, öznitelik uzayındaki en yakın komşuyu basitçe eşleştirmektir. Bazı öznitelikler eşleşmediği (arka plan gürültüsünden veya kapatma sebebiyle) için yanlış pozitiflerin sayısını azaltmak için hala bir eşik değeri kullanılabilir.



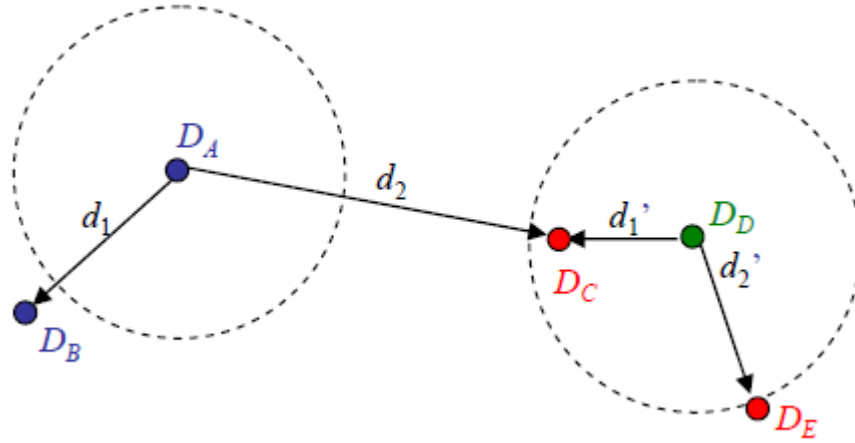
Şekil 3.31 ROC eğrisi ve ilgili oranlar

İdeal olarak bu eşik değeri, öznitelik uzayının farklı bölgelerine kendini adapte edecektir. Eğer yeteri kadar eğitim verisi varsa (Hua vd. 2007), farklı öznitelikler için farklı eşik seviyeleri öğrenmek bazen mümkündür. Ancak genellikle bu kadar veri sağlanamaz. Bu durumda, faydalı bir sezgisel yöntem, en yakın komşu uzaklığını, tercihen hedef eşleşmenin olmadığı bilindiği bir görüntüden (veri tabanındaki farklı bir görüntü gibi) ikinci en yakın komşu ile karşılaştırmak olabilir (Brown ve Lowe 2002, Lowe 2004). En yakın komşu uzaklık oranı, aşağıdaki gibi tanımlanabilir (Mikolajczyk ve Schmid 2005):

$$NNDR = \frac{d_1}{d_2} = \frac{\|D_A - D_B\|}{\|D_A - D_C\|} \quad (3.22)$$

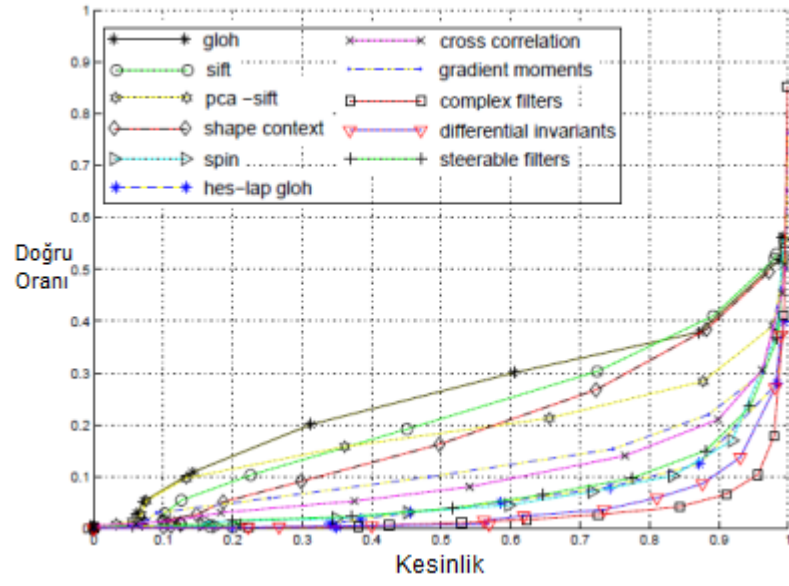
Yukarıdaki denklemde d_1 ve d_2 en yakın ve ikinci en yakın komşu uzaklığı, D_A hedef tanımlayıcı, D_B ve D_C ise D_A 'ya en yakın iki komşudur. Şekil 3.32'de üç farklı eşleştirme kriteri gösterilmiştir. D_A tanımlayıcısı, kesikli çizgi ile gösterilen sabit uzaklık eşik

seviyesinde D_B ile eşleşmekte başarısız olur. D_D ise yanlışlıkla D_C ve D_E ile eşleşir. Eğer en yakın komşu seçilirse D_A doğru olarak D_B ile eşleşir. Ama D_D yanlışlıkla D_C ile eşleşir. En yakın komşu uzaklık oranı (NNDR) eşleştirmesi kullanarak küçük NNDR d_1/d_2 D_A 'yı doğru olarak D_B ile eşleştirir ve büyük NNDR d_1'/d_2' D_D için eşleştirmeleri doğru olarak reddeder.

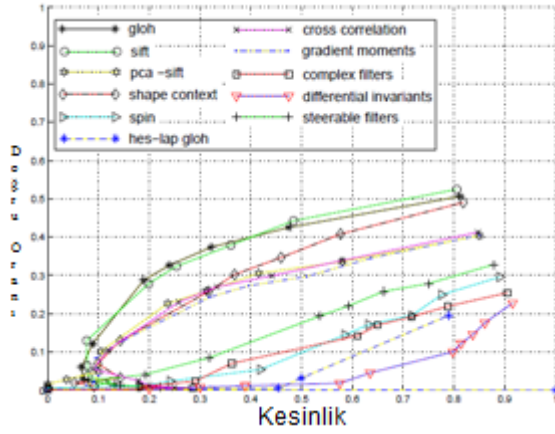


Şekil 3.32 Sabit eşik değeri, en yakın komşu ve en yakın komşu uzaklık oranı eşleştirmesi

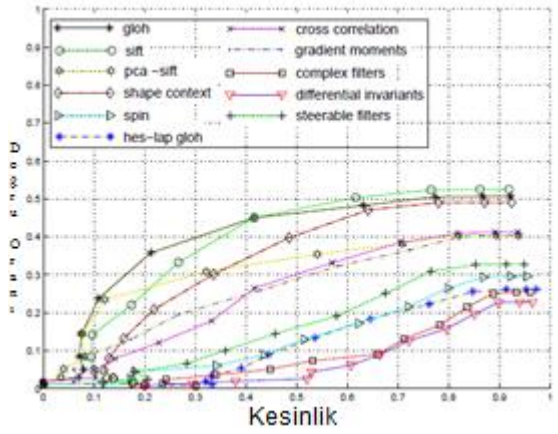
Öznitelik tanımlayıcılar için üç farklı eşleştirme stratejisi kullanmanın etkisi, Şekil 3.33'te karşılaştırmalı olarak verilmiştir (Mikolajczyk ve Schmid 2005). Şekil 3.33.a ile sabit eşik değeri, Şekil 3.33.b ile en yakın komşu ve Şekil 3.33.c ile en yakın komşu uzaklık oranı (NNDR) kullanılarak elde edilen sonuçlar verilmiştir. Algoritmaların sırasının pek değişmediğini ama farklı eşleştirme stratejileri arasında genel başarımın önemli ölçüde arttığı görülmektedir. Görülebileceği gibi en yakın komşu ve en yakın komşu uzaklık oranı (NNDR) stratejileri, daha iyi ROC eğrileri vermiştir.



(a)



(b)



(c)

Şekil 3.33 Öznitelik tanımlayıcıların eşleştirme stratejisine göre performans değerlendirmesi

3.4 Kullanılan Yöntem

Tez kapsamında rotasyon, ölçekleme ve öteleme değişmezlikli, arka plan ve kapatma bozucu etkilerine karşı dayanıklı ayrıca düşük yanlış kabul oranına sahip bir nesne tanıma yöntemi kullanılmıştır. Bu yöntemde öncelikle eğitim verileri üzerinde öznitelik

çıkartımı gerçekleştirilir. Öznitelik çıkartımı işlemi sonucunda elde edilen öznitelik vektörlerinin içerdiği bilgi çeşitliliği fazladır. Bu durum rotasyon ve ölçekleme gibi bir dönüşüme maruz kalan görüntülerden elde edilen özniteliklerin yanlışlıkla yakın başka bir özniteliğe eşleşmesine sebep olmaktadır. Bu yüzden öznitelik vektörlerinin daha az sayıdaki (az çeşitlilikteki) kümelerle gruplanması yanlışlığı (bias) azaltacaktır. Bu gruplama işlemi bir eğitimsiz öğrenme (unsupervised learning) tekniği olan k-ortalama++ (k-means++) kümeleme algoritması ile gerçekleştirilecektir. Bu algoritma, standart k-ortalama algoritmasına göre hem iterasyon süresini kısaltmakta hem de daha iyi gruplama sağlayabilmektedir. K-ortalama++ algoritması, kümelerin merkezlerinin ilk değerlerine atanmasını sağlayan bir başlangıç prosedürüne sahiptir. İlk değerlemenin ardından standart k-ortalama algoritması ile gruplama işlemi gerçekleştirilecektir. Bilgisayar görüşü literatüründe özniteliklerin bu şekilde gruplanması ile elde edilen her bir gruba veya özniteliğe görsel kelime; tüm gruplara ise görsel kelime sözlüğü (Bag of Visual Words) denilmektedir.

Görsel kelime sözlüğü kullanılarak öznitelik tanımlayıcılar hesaplanır. Bunun için öncelikle görüntü ve görüntüdeki öznitelik noktalar kümesi kullanılarak öznitelik tanımlayıcılar hesaplanır. Her bir öznitelik tanımlayıcı için sözlükteki en yakın kelime bulunur. Görüntüde karşılaşılan görsel kelimelerin normalleştirilmiş histogramı olarak görüntü tanımlayıcılar hesaplanır. Histogramdaki n-inci sele, görüntüdeki sözlüğün n-inci kelime sıklığıdır. Görsel kelime sözlüğü sayesinde öznitelik çeşitliliği azaltıldıktan sonra her bir özniteliğin ait olduğu sınıfı gösteren etiketleme işlemi eğitmenli (Supervised) olarak gerçekleştirilir.

Eğitim işleminin son aşaması, öznitelikler ve eğitmen tarafından belirlenen etiketler kullanılarak iki farklı sınıflandırıcı eğitilmesidir (Supervised Learning). Sınıflandırıcı olarak Destek Vektör Makinaları (Support Vector Machines - SVM) ve Bayes sınıflandırıcı seçilmiştir. Destek Vektör Makinalarının seçilmesinin sebebi, az sayıda veri ile yüksek sınıflandırma başarısının elde edilmesine olanak sağlamasıdır. Destek Vektör Makinalarının sınıflandırma başarısını karşılaştırmak için basit ve etkin bir yöntem olması sebebiyle Bayes sınıflandırıcı seçilmiştir.

Eğitim işleminin ardından test verilerinin sınıf tahmini eğitilen sınıflandırıcı kullanılarak gerçekleştirilmektedir. Bu işlem için öncelikle test verilerinden görsel kelime sözlüğündeki kelimelere karşılık gelen öznitelikler çıkartılmaktadır. Bu öznitelikler kullanılarak sınıf tahmini gerçekleştirilmektedir. Sınıf tahmininin dolayısıyla sınıflandırıcının başarısını ölçmek için kesin referans (ground truth) sınıflar ile tahmin edilen sınıflar arasında karşılaştırma yapılmaktadır. Kesin referans sınıflar, test verilerinin eğitmen tarafından etiketlenmesiyle elde edilmektedir. Deneysel testler sonucunda nesne tanıma performansına etki eden parametreler detaylı olarak Bölüm 4’te belirtilmektedir. Kullanılan yöntemi oluşturan bölümler aşağıda detaylandırılmıştır.

3.4.1 Öznitelik çıkartımı

Görüntü gösterimi rotasyon, ölçekleme, öteleme ve ilgin değişmezlikli farklı öznitelikler kullanılarak gerçekleştirilmiştir. Öznitelik bulucu olarak kullanılan yöntemler aşağıda verilmiştir:

- SIFT,
- SURF,
- PyramidSIFT,
- PyramidSURF,
- GridSIFT,
- GridSURF
- MSER.

GridSIFT/GridSURF yönteminde kaynak görüntü ızgaralara bölünür ve her bir hücredeki noktalar SIFT/SURF kullanılarak tespit edilir. PyramidSIFT/PyramidSURF yönteminde ise Gauss piramidinin birçok seviyesi üzerinde SIFT veya SURF noktaları tespit edilir. Öznitelik tanımlayıcı olarak da aşağıda belirtilen yöntemler kullanılmıştır:

- SIFT,
- SURF,
- OpponentSIFT,
- OpponentSURF.

3.4.2 Görsel kelime sözlüğü

Kelime sözlüğü (*bag-of-word* - BOW) modeli, doğal dil işleme (*natural language processing*) ve bilgi kazanım (*information retrieval*) alanlarında gösterimi basitleştirmek için kullanılır. Bu modelde, bir metin (bir cümle veya doküman) dil bilgisi ve kelime sırasından bağımsız olarak kelimelerin sırasız koleksiyonu olarak gösterilir. BOW modeli, sınıflandırıcı eğitimi için her bir kelimenin meydana gelme sıklığının öznitelik olarak kullanıldığı doküman sınıflandırma yöntemlerinde çoğunlukla kullanılır.

BOW modeli, son zamanlarda bilgisayar görüşü alanında görüntü sınıflandırma gerçekleştirmek için kullanılmaktadır. Görüntü öznitelikleri, modeldeki *kelimeler* olarak davranır. Bu yüzden bilgisayar görüşü alanında kullanılan BOW yerini görsel kelime sözlüğü terimine (Bag-of-Visual-Word - BOVW) bırakır. BOVW, yerel görüntü özniteliklerinin meydana gelme sayılarının seyrek vektörüdür.

Bir görüntüyü BOVW modeli ile göstermek için görüntüye bir dokümanmış gibi davranılabilir. BOW modeline benzer şekilde görüntüdeki kelimeler de tanımlanmalıdır. Bunu başarmak için genellikle üç aşamalı bir yapı kullanılır (Fei-Fei ve Perona 2005).

- Öznitelik tespiti
- Öznitelik tanımlama
- Kod tablosu oluşturma

BOW modelin diğer bir tanımı da bağımsız öznitelik tabanlı histogram gösterimidir (Fei-Fei vd. 2007).

Öznitelik Gösterimi

Öznitelikler tespit edildikten sonra her bir görüntü birkaç yerel görüntü parçası ile özetlenir. Öznitelik gösterim metotları, görüntü parçalarının rakamsal vektörler olarak nasıl ifade edilebilebileceğinin üstesinden gelirler. Bu vektörler öznitelik tanımlayıcılar olarak ifade edilirler. İyi bir tanımlayıcı bir seviyeye kadar yoğunluk (*intensity*), rotasyon, ölçekleme ve ilgin (*affine*) değişimlerinin üstesinden gelebilmelidir. Ölçekleme değişmezlikli öznitelik dönüşümü (Scale-invariant feature transform – SIFT)

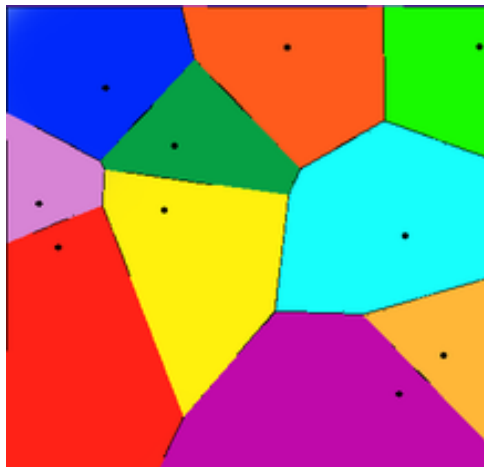
en ünlü tanımlayıcılardan bir tanesidir. SIFT, her bir görüntü parçasını 128-boyutlu vektöre çevirir. Bu aşamadan sonra her bir görüntü farklı vektörlerin sırasının önemli olmadığı aynı (128) boyutlu vektörlerin koleksiyonudur.

Kod tablosu oluşturma

BOW modeli için son aşama vektörlerle gösterilen görüntü parçalarını kod tablosunu (codebook) oluşturan kod kelimelerine (*codewords*) çevirmektir. Bir kod kelimesi benzer birkaç görüntü parçasının bir temsili olarak düşünülebilir. Basit bir metot, tüm vektörler üzerinde k-ortalama kümeleme (*k-means clustering*) gerçekleştirmektir (Leung ve Malik 2001). K-ortalama kümeleme, Bölüm 3.4.2.1’de anlatılmaktadır. Bu işlem neticesinde kod kelimeleri öğrenilen kümelerin merkezleri olarak tanımlanırlar. Kümelerin sayısı kod tablosunun (codebook) boyutudur. Bu nedenle, bir görüntüdeki her bir görüntü parçası, kümeleme işlemi sayesinde belirli bir kod kelimesine eşlenir. Görüntü, kod kelimelerinin histogramı ile temsil edilir.

3.4.2.1 K-ortalama kümelemesi

K-ortalama kümeleme, kümeleme analizinde (clustering analysis) kullanılan bir yöntemidir. Amaç, her bir gözlemin en yakın ortalama kümeye ait olduğu n sayıda gözlemin k sayıda kümeye bölünmesidir. Bu işlem, veri uzayının Voronoi hücrelerine bölünmesi ile sonuçlanır (Şekil 3.34). Voronoi diyagramı, uzayı belirli bir sayıda bölgeye ayırmaktadır (Şekil 3.34).



Şekil 3.34 Voronoi hücreleri

Her bir gözlemin d-boyutlu olduğu vektör olduğu bir gözlem uzayı $(x_1, x_2, \dots, x_n)$ verilmiş olsun. K-ortalama kümeleme, küme-içi değişintiyi minimum yapmak için n gözlemi k kümeye ayırmaktır.

$$\arg \min_S \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_i\|^2 \quad (3.23)$$

Denkleminde μ_i , S_i kümesindeki noktaların ortalamasıdır.

K-ortalama kümeleme üzerine en yaygın algoritma bir yinelemeli düzeltme tekniğidir. Her yerde karşılaşıldığı için sıklıkla k-ortalama algoritması olarak bilinmesine rağmen bilgisayar bilimi topluluğunda *Lloyd'un algoritması* olarak adlandırılır. Bu algoritma için öncelikle k ortalamaların başlangıç kümesi $m_1, \dots, m_k$ olarak verildiğini varsayalım. Algoritma, iki aşama arasında yinelemeli olarak gerçekleşir (MacKay 2003):

1- Atama aşaması:

Her bir gözlem, ortalaması gözleme en yakın olan kümeye atanır. Örneğin, gözlemler ortalamalarla şekillendirilen Voronoi diyagramına (Şekil 3.34) bölünür.

$$S_i^{(t)} = \{x_p : \|x_p - m_i^{(t)}\| \leq \|x_p - m_j^{(t)}\| \forall 1 \leq j \leq k\} \quad (3.24)$$

Her bir x_p sadece bir $S^{(t)}$ atanır.

2- Güncelleme aşaması:

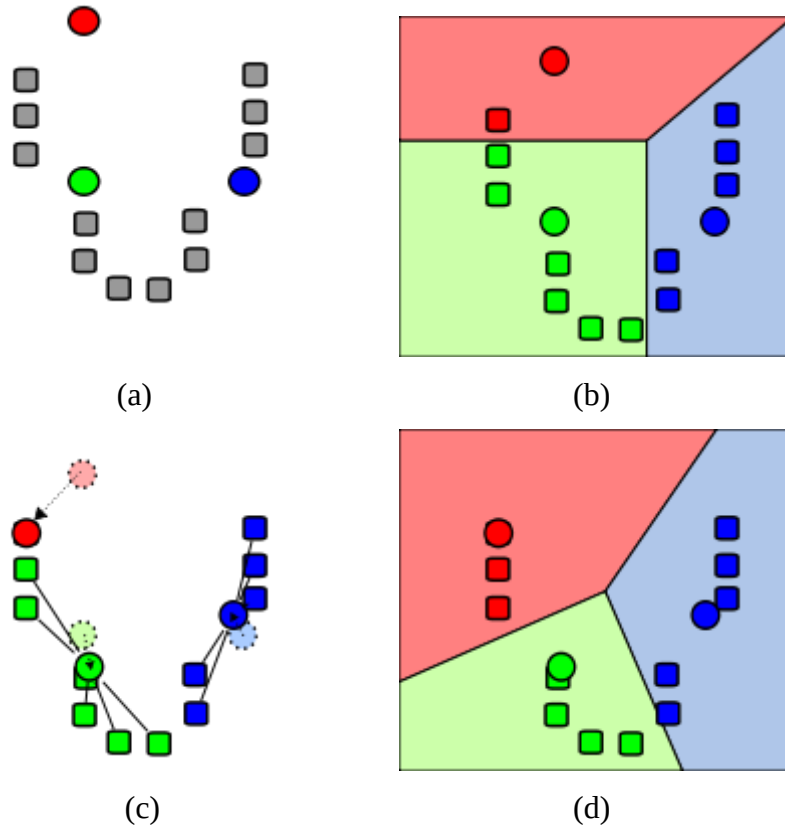
Yeni ortalamalar, yeni kümelerdeki gözlemlerin ağırlık merkezi olarak hesaplanır.

$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j \quad (3.25)$$

Atamalar daha fazla deđiřmediđi zaman algoritma sonlanır.

İklendirme Yöntemi

İklendirme yöntemi olarak genellikle Forgy ve rastgele bölme (*random partition*) kullanılır (Hamerly ve Elkan 2002). Forgy yöntemi, veri kümesinden rastgele k gözlem seçer ve bunları ilk ortalamalar olarak kullanır. Rastgele bölme yöntemi, her bir gözlemi rastgele bir kümeye atar ve ardından güncelleme aşamasını gerçekleştirir. Bu şekilde kümenin rastgele atanan noktalarının ağırlık merkezi ilk ortalamalar olarak hesaplanır. Forgy yöntemi, ilk ortalamaları yaymak eğilimdeyken rastgele bölme yöntemi tüm gözlemleri veri kümesinin merkezine yakın olarak yerleştirir.



Şekil 3.35 k-ortalama gösterimi

Şekil 3.35.a'da gösterilen renkli veriler, tüm veriler arasından rastgele olarak oluşturulan k sayıdaki ($k=3$) ilk ortalamaları göstermektedir. Şekil 3.35.b'de k sayıda küme, kendilerine en yakın veri kullanılarak oluşturulmuştur. Burada gösterilen bölümler, ortalamalarla oluşturulan Voronoi diyagramlarını göstermektedir. Her bir kümenin ağırlık merkezi, Şekil 3.35.c'de gösterildiği gibi yeni ortalama olarak güncellenir. Son iki aşama yakınsama sağlanıncaya kadar tekrar edilir.

3.4.2.2 K-ortalama++ kümeleme

K-ortalama algoritmasının dezavantajı, optimum kümelemeye kıyasla kötü yaklaşımlar bulabilmesidir. Bu dezavantajı ortadan kaldırmak için k-ortalama++ algoritması kullanılacaktır. Bu algoritmanın avantajı, standard k-ortalama en iyileme tekrarlamasına başlamadan önce kümelerin merkezlerini ilklendirmektir. Bu algoritma aşağıdaki gibidir:

- 1- Veri noktaları arasından bir merkez rastgele düzenli olarak seçilir.
- 2- Her bir veri için veri ile en yakın merkez arasındaki uzaklık hesaplanır.
- 3- Yeni bir merkez olarak yeni veri noktası rastgele seçilir.
- 4- Aşama 2 ve 3 k sayıda merkez seçilene kadar tekrar edilir.
- 5- İlk merkezler seçildikten sonra standard k-ortalama kümeleme kullanılır.

Bu ilklendirme yöntemi, k-ortalamanın son hata miktarında kayde değer bir azalma sunmaktadır. Algoritmadaki ilk seçim işlemi fazladan zaman almasına rağmen ilklendirmeden sonra yakınsamayı çok hızlı gerçekleştirir. Sonuç olarak k-ortalamaya kıyasla hesaplama süresi azalır.

3.4.3 Sınıflandırma

Sınıflandırma yöntemi olarak Bayes ve destek vektör makinaları (Support Vector Machines – SVM) kullanılmıştır.

3.4.3.1 Bayes

Bayes sınıflandırıcı yanlış sınıflandırma olasılığını minimize eder (Devroye vd. 1996). Öznitelik uzayındaki öznitelik vektörlerinin dağılımının istatistiksel özellikleri kullanılarak Bayes kuralı göre sınıflandırma gerçekleştirilir. Bir x öznitelik vektörü, bir $p(x)$ olasılık yoğunluk fonksiyonu (*probability density function* - *PDF*) ile modellenenir. Bu modelleme, verilen bir x gözleminin k sınıfından oluşma olasılığını hesaplayarak sınıflandırma problemini çözmeye çalışır. Eğer $p(k|x)$ koşullu olasılığı tüm sınıflar için $k \in [1, 2, \dots, K]$ bilinirse x gözlemi $p(k|x)$ (*posterior probability*) değerini maksimum yapan sınıfa atanabilir.

Birçok uygulamada $p(k|x)$ genellikle bilinmeyen bir değerdir. Ancak, her bir örneğin ait olduğu sınıfa etiketlendiği bir eğitim veri kümesinde bir k sınıfına ait tüm eğitim verilerinin dağılımı hesaplanarak eğitim aşamasında $p(x|k)$ koşullu olasılıklarını hesaplamak mümkündür. Bayes kuralına göre $p(k|x)$ aşağıdaki formül kullanarak hesaplanabilir:

$$p(k | x) = \frac{p(x | k) p(k)}{p(x)} \quad (3.26)$$

Spesifik k sınıfının meydana gelme olasılığı $p(k)$, eğitim aşamasında k sınıfına ait eğitim verilerinin sayısının tüm veri kümesinin sayısına oranı hesaplanarak tahmin edilebilir. Aşağıdaki formül kullanılarak $p(x)$ hesaplanabilir:

$$\sum_{k=1}^K p(x | k) \cdot p(k) \quad (3.27)$$

Böylece tanıma sırasında verilerin hangi sınıfa ait olduğunu hesaplamak için gerekli olan tüm değişkenler elde edilmiş olur.

3.4.3.2 Destek Vektör Makinaları

Destek vektör makinaları, verilerin ilişkilendirilmiş öğrenme algoritmaları ile analiz edildiği ve örüntü tanımanın gerçekleştirildiği bir eğitmenli eğitimidir (Cortes ve Vapnik 1995). Sınıflandırma ve bağlanım (regression) analizi için kullanılırlar. Temel SVM sınıflandırıcı, bir grup girdi verisini alır ve verilen her bir grup için iki olası sınıftan hangisinin çıktığı biçimlendirdiğini tahmin eder. Bu işlem, SVM’i olasılıksal olmayan ikili doğrusal sınıflandırıcı yapmaktadır. SVM kullanılırken öncelikle her biri iki sınıftan bir tanesine ait olarak etiketlenmiş olan eğitim veri kümesi elde edilir. SVM öğrenme algoritması bu verilerden bir model oluşturur. Test verileri oluşturulan bu modele göre değerlendirilir ve iki sınıftan hangisine ait olduğu tahmin edilir. Doğrusal sınıflandırmaya ek olarak SVM çekirdek hilesi (*kernel trick*) olarak adlandırılan yöntemi kullanarak doğrusal olmayan sınıflandırmayı da verimli olarak gerçekleştirmektedir.

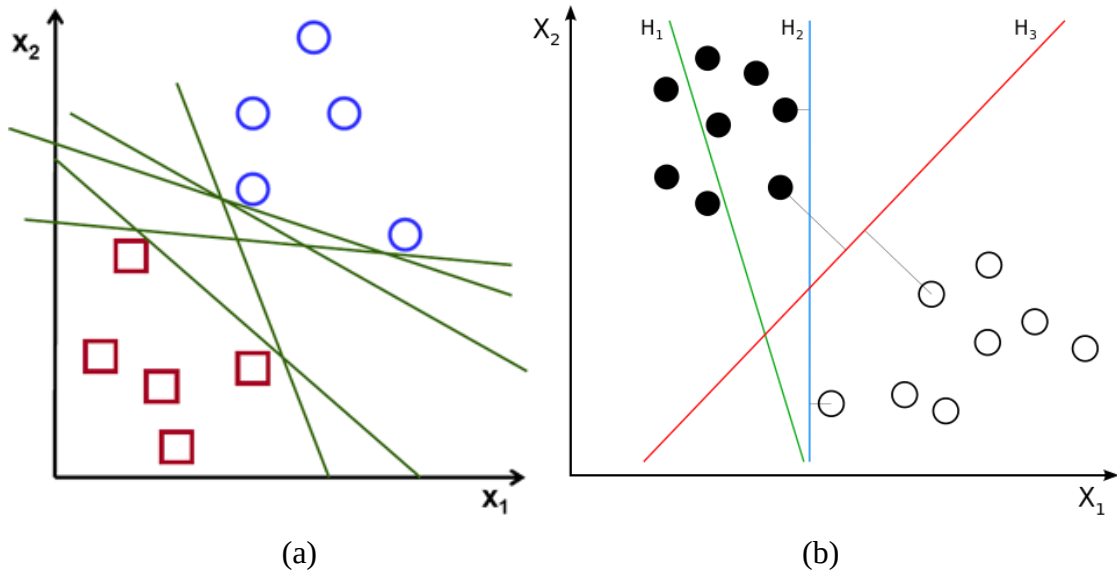
Bir destek vektör makinası, sınıflandırmada ve regresyonda kullanılacak bir veya bir grup çok boyutlu (high-dimensional) hiper-düzlem (hyperplane) oluşturur. İyi bir ayırım, marjın (margin) olarak adlandırılan ve her bir sınıfın en yakın eğitim verisi noktasına olan uzaklığını maksimum yapan bir hiper-düzlem ile başarılı bir şekilde gerçekleştirilebilir. Genellikle marjın arttırılarak sınıflandırıcı genelleme hataları (generalization error) azaltılmaktadır.

Orijinal problem sonlu boyutsal uzayda belirtiliyorken günlük hayatta genellikle problemin bulunduğu uzayda doğrusal olarak ayrılamayan verilerle karşılaşılır. Bu sebeple, ayırımın daha zor olduğu orijinal sonlu boyutlu uzay, ayırımın daha kolay yapılabildiği daha yüksek boyutlu bir uzaya (*map*) eşleştirilir. Orijinal uzaydaki değişkenlerin iç çarpımının (*dot product*), hesap yükünü makul tutarak, kolaylıkla hesaplamasını sağlamak için SVM tarafından probleme uygun bir çekirdek fonksiyon $K(x,y)$ tanımlanarak eşleştirmeler yapılır. Daha yüksek boyutlu uzaydaki hiper-düzlemler, noktalar kümesinin aynı uzaydaki bir vektör ile iç çarpımının sabit olması olarak tanımlanır. Hiper-düzlemi tanımlayan vektörler, veri tabanında bulunan öz nitelik vektörlerinin α_i parametrelerinin doğrusal kombinasyonu olarak seçilebilir.

Hiperdüzleme eşleştirilen öznitelik uzayındaki x noktaları aşağıdaki formül ile tanımlanır:

$$\sum_i \alpha_i K(x_i, x) = \text{sabit} \quad (3.28)$$

$K(x,y)$ ifadesi y değeri, x değerinden fazla miktarda büyüdükçe küçülmektedir. Toplamadaki her bir terim, x test noktasının x_i temel veri noktasına göre yakınlığının derecesidir. Bu şekilde yukarıdaki kernellerin toplamı, her bir test noktasının ayırım sağlanacak olan diğer kümelerin veri noktalarına olan göreceli yakınlığını hesaplamak için kullanılır (Cortes ve Vapnik 1995). Herhangi bir hiper-düzlem ile eşleştirilen ve orijinal uzayda dışbükey olmayan kümeler arası çok karmaşık ayırma olanak sağlayan x noktalarının kümesi bu şekilde kolaylıkla ayrıştırılabilir.



Şekil 3.36 Sınıflar arası ayırım sağlayan doğrular

Şekil 3.36 ile her biri iki sınıftan birine ait olan veri noktaları verilmiş olsun. Burada amaç, bir veri noktasının hangi sınıfa dahil olduğuna karar vermektir. Destek vektör makinaları açısından bir veri noktası p -boyutlu bir vektör olarak görülür. Bu verilerin

ayrımını sağlayacak (p-1)-boyutlu bir hiper-düzlemin varlığı tespit edilmeye çalışılır. Bu işlem, doğrusal sınıflandırıcı (linear classifier) olarak adlandırılır. Şekil 3.36 ile gösterildiği gibi verileri sınıflandırmak için birden fazla doğru (veya düzlem) bulunabilir. Şekil 3.36.b’de H1 sınıflar arası ayırım sağlamazken H2 küçük bir sınır (margin) ile ayırım sağlar. Ancak gürültüye çok duyarlıdır. H3 ise sınıflar arası maksimum ayrımı sağlamaktadır. Ayrıca gürültüye de dayanıklıdır. Sınıflar arasındaki en büyük ayrımı, sınırı (margin) sağlayan doğru veya düzlem en iyisi olarak seçilir. Diğer bir ifadeyle, ayırım sağlayan düzlemin her bir taraftaki veri noktasına olan uzaklığını maksimum yapan düzlem seçilir (Cortes ve Vapnik 1995). Bu düzlem, Şekil 3.36.b’de H3 ile gösterilmektedir. Eğer böyle bir düzlem (hiper-düzlem) varsa bu düzlem maksimum-sınır hiper-düzlemi (maximum-margin hyperplane) olarak adlandırılır. Bu düzlemi tanımlayan sınıflandırıcı da maksimum-sınır sınıflandırıcı (maximum-margin classifier) olarak bilinir.

D eğitim verileri, n noktalarının kümesi olarak verilmiş olsun:

$$D = \{(x_i, y_i) \mid x_i \in \mathbb{R}^p, y_i \in \{-1, 1\}\}_{i=1}^n \quad (3.29)$$

Yukarıda y_i , x_i noktasının ait olduğu sınıfı -1 ve 1 türünden göstermektedir. SVM, $y_i = 1$ olan noktaları $y_i = -1$ olan noktalardan ayırt eden maksimum-sınır düzlemini bulmaya çalışır. Bir düzlem, aşağıdaki denklemi sağlayan x noktalarının kümesi olarak yazılabilir:

$$w \cdot x - b = 0 \quad (3.30)$$

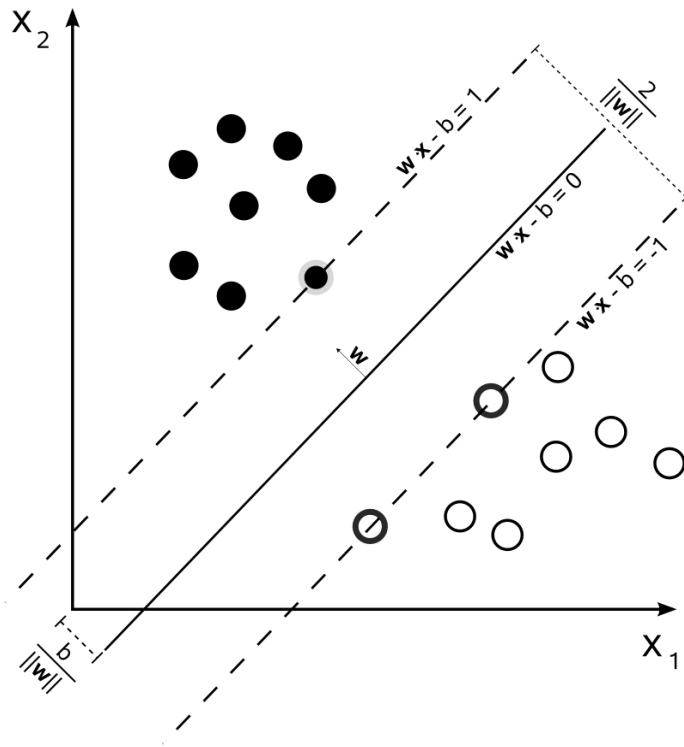
Burada $\cdot$ iç çarpımı (*dot product*) ve w ise düzleme dik olan vektörü (normal vector) göstermektedir. Şekil 3.37 ile maksimum-sınır hiperdüzlemi ve iki sınıfın verileri ile eğitilen bir SVM için sınırlar (margins) verilmiştir. Sınırlar üzerindeki ve yuvarlak ile işaretlemiş verilen destek vektör (*suppor vector*) olarak adlandırılır (Cortes ve Vapnik 1995). Şekil 3.37’de $b/\|w\|$ hiper-düzlemin w dik vektörü boyunca orijine olan kaymasını göstermektedir. Eğitim verisi doğrusal olarak ayrıştırılabilir ise veriyi ayıran

ve arasında hiçbir veri noktası bulunmayan iki hiper-düzlem seçilebilir. Ardından seçilen bu iki hiper-düzlem arasındaki uzaklık maksimum yapılmaya çalışılır. Düzlemlerle sınırlandırılan bölge sınır (*margin*) olarak isimlendirilir. Bu düzlemler, şekil 3.37’de gösterilen ve aşağıda verilen eşitliklerle ifade edilir (Cortes ve Vapnik 1995):

$$w \cdot x - b = 1$$

$$w \cdot x - b = -1$$

(3.31)



Şekil 3.37 Maksimum-sınır hiperdüzlemi

Bu iki düzlem arasındaki uzaklık (M) aşağıda eşitlikler kullanılarak hesaplanabilir:

$$\frac{|w \cdot x - b|}{\|w\|} = \frac{1}{\|w\|}$$

$$M = \frac{2}{\|w\|} \quad (3.32)$$

M değerini maksimum yapmak için $\|w\|$ değeri minimum yapılır. Ayrıca verileri sınır (*margin*) içine düşmesini engellemek için aşağıdaki sınır eşitlikler kullanılır:

$$w \cdot x_i - b \geq 1$$

$$w \cdot x_i - b \leq -1 \quad (3.33)$$

Bu eşitlikler kullanılarak ifade aşağıdaki gibi genelleştirilebilir:

$$y_i (w \cdot x_i - b) \geq 1 \quad (3.34)$$

Eğitim verilerinin ait olduğu sınıfların her biri y_i ile gösterilir. Artık problem yukarıdaki denkleme göre $\|w\|$ değerini minimum yapmaktır. Bu optimizasyon problemi w değerinin normuna, $\|w\|$, bağlı olduğundan ve bu da karekök gerektirdiğinden çözmek zordur. Neyseki, $\|w\|$ ifadesi yerine $\frac{1}{2}\|w\|^2$ kullanmak, orijinal ifadenin minimumu ile değiştirilen ifadenin minimum değeri aynı w ve b değerlerine sahip olduğu için, mümkündür (Cortes ve Vapnik 1995). Minimum yapılacak değer artık aşağıdaki ifadedir:

$$\frac{1}{2} \|w\|^2 \quad (3.35)$$

Bu ifade, Lagrange çarpımlar (α) kullanarak aşağıdaki biçime dönüştürülür.

$$\min_{w,b} \max_{a \geq 0} \left\{ \frac{1}{2} \|w\|^2 - \sum_{i=1}^n a_i [y_i (w \cdot x_i - b) - 1] \right\} \quad (3.36)$$

Çözüm, eğitim vektörlerinin doğrusal kombinasyonları olarak ifade edilebilir.

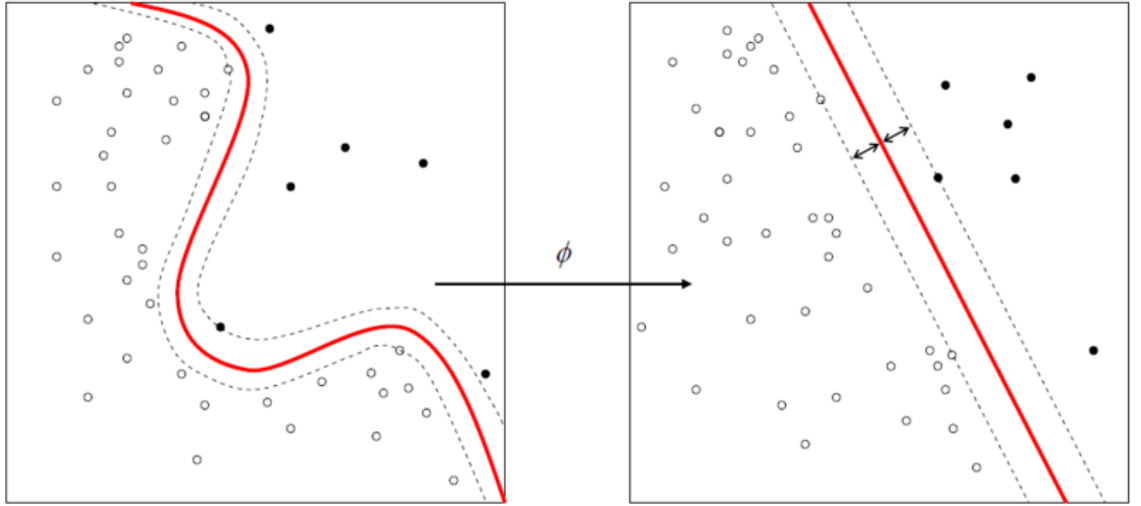
$$w = \sum_{i=1}^n a_i y_i x_i \quad (3.37)$$

Sadece bir kaç tane a_i sıfırdan büyük olacaktır. Karşılık gelen x_i noktaları sınır (*margin*) üzerinde bulunan ve aşağıdaki eşitliği destekleyen destek vektörlerdir.

$$y_i (w \cdot x_i - b) = 1 \quad (3.38)$$

3.4.3.2.1 Doğrusal-olmayan sınıflandırma

Boser vd. (1992), maksimum-sınır hiperdüzlemi için çekirdek hilesi (kernel trick) kullanarak doğrusal-olmayan sınıflandırıcı oluşturmayı önermişlerdir. Her iç çarpımın doğrusal-olmayan bir kernel fonksiyonu ile yer değiştirmesi haricinde orijinal algorithmada değişiklik yapmadan problem çözülebilir. Bu işlem, dönüştürülmüş bir öznitelik uzayında maksimum-sınır hiper-düzlemine oturtmaya olanak sağlar. Orijinal veri uzayında problem doğrusal-olmayan olsa bile dönüştürüldükten sonra yüksek-boyutlu (*high-dimensional*) öznitelik uzayında doğrusal ayrıştırılabilir bir hiper-düzlem olabilir (Şekil 3.38).



Şekil 3.38 Kernel makinası

Kullanılabilecek bazı kerneller Çizelge 3.4’te verilmiştir.

Çizelge 3.4 Kernel fonksiyonları

Polinomsal (türdeş)	$k(x_i, x_j) = (x_i \cdot x_j)^d$
Polinomsal (türdeş olmayan)	$k(x_i, x_j) = (x_i \cdot x_j + 1)^d$
Gauss radyal taban fonksiyonu (<i>Gaussian radial basis function</i>)	$k(x_i, x_j) = \exp(-\gamma \ x_i - x_j\ ^2)$ $\gamma = \frac{1}{2\sigma^2}$

Kernel fonksiyonu aşağıdaki eşitlik ile ifade edilir.

$$k(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j) \quad (3.39)$$

Doğrusal-olmayan ayrıştırma problemlerinde radyal taban fonsiyonu genellikle başarılı sonuçlar vermektedir.

4. DENEYSEL SONUÇLAR

Kullanılan yöntemin başarımını değerlendirmek için öncelikle üzerinde öznitelik çıkartımı gerçekleştirilecek olan bir eğitim veri kümesi oluşturulmuştur. Bu veri kümesi, çeşitli yönelim ve ölçeklemelere sahip nesnelerin bulunduğu görüntülerle birlikte arka plan gürültüsü ve kapatma da içermektedir. Veri kümeleri, hedef ve hedef olmayan olmak üzere iki sınıftan oluşmaktadır. Uygulama alanı olarak uydu görüntüleri; tanımlanacak nesne olarak uçak seçilmiştir. Farklı sayılarda eğitim ve değerlendirme verileri kullanılarak veri sayısının sınıflandırma performansına etkisi incelenmiştir. Eğitim veri kümesini oluşturan örnek hedef sınıfı görüntüleri şekil 4.1’de verilmiştir.



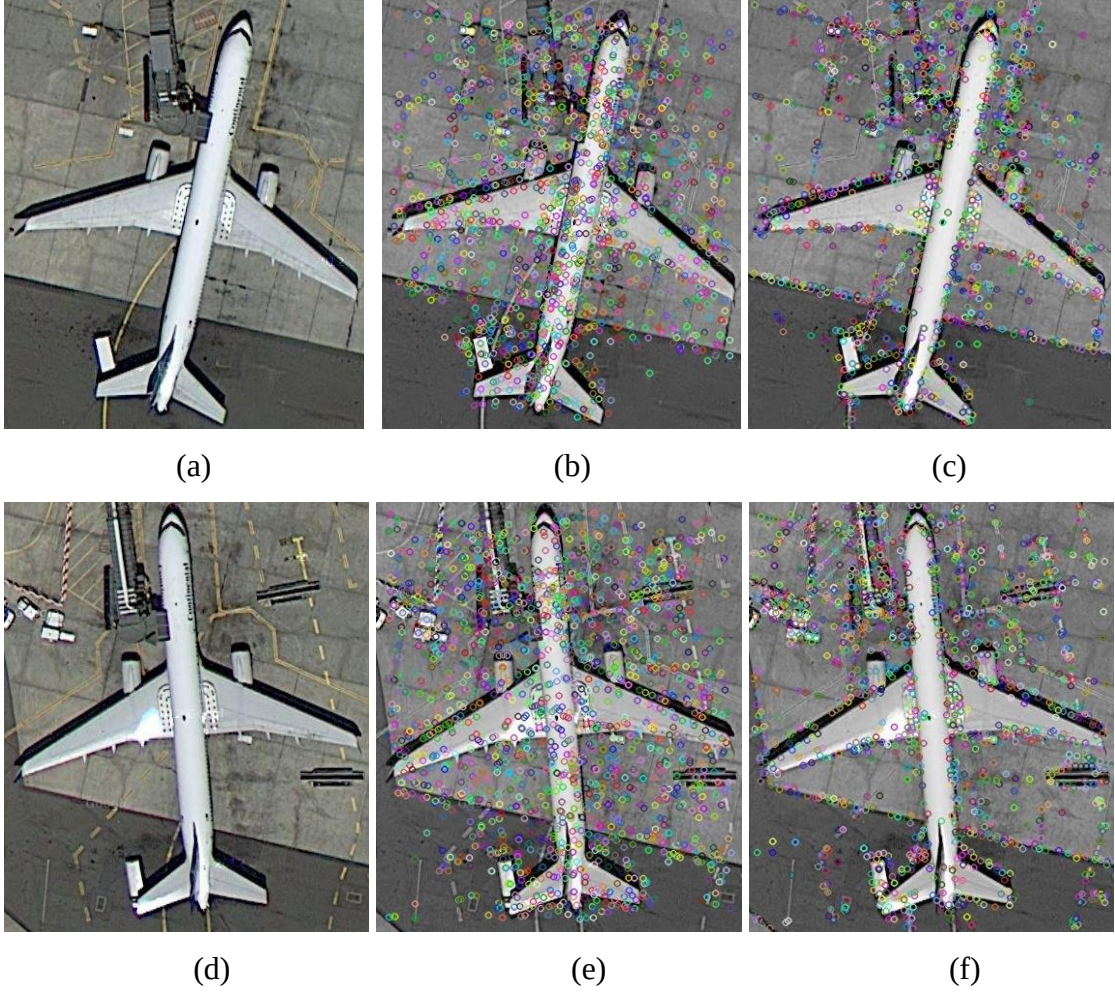
Şekil 4.1 Eğitim veri kümesi

Değerlendirme veri kümesini oluşturan hedef-olmayan sınıfa ait örnek görüntüler ise Şekil 4.2’de verilmiştir.



Şekil 4.2 Değerlendirme veri kümesi

Öznitelik çıkartımı olarak Bölüm 3.4.1’de belirtilen yöntemler kullanılmıştır. Şekil 4.3’te görüntülerde tespit edilen SIFT ve SURF öznitelikleri görüntüler üzerinde konumlandırılmıştır. Şekil 4.3.b’de ve Şekil 4.3.e’de görüntülerdeki SURF öznitelikleri gösterilmiştir. Şekil 4.3.c’de ve Şekil 4.3.f’de görüntülerdeki SIFT öznitelikleri gösterilmiştir.



Şekil 4.3 Görüntülerde öznitelik gösterimi

Öznitelik çıkartımı yapıldıktan sonra öznitelikler, BOVW (Bag of Visual Word) modeline uygun olarak modellenmiştir. Modelleme işlemi, k-ortalama (k-means++) kullanılarak gerçekleştirilmiştir. Bu modelde öznitelikler, görsel kelimelere karşılık gelmektedir. 20, 50, 100 ve 250 gibi farklı sayıda görsel kelimeler kullanılarak modelleme işlemi gerçekleştirilmiştir.

Bayes ve destek vektör makinaları (SVMs) sınıflandırıcı olarak kullanılmıştır. Destek vektör makinası ile doğrusal ve doğrusal-olmayan (non-linear) ayırım gerçekleştirilmiştir. Doğrusal-olmayan ayırım gerçekleştirilirken çekirdek (kernel) fonksiyon olarak radyan taban fonksiyonu (Radial Basis Function - RBF) kullanılmıştır. Sınıflandırma tahmininin başarımını ölçmek için kesin referans (*ground truth*) veriler ile karşılaştırma yapılmıştır. Kesin referans verileri, test verileri etiketlenerek elde

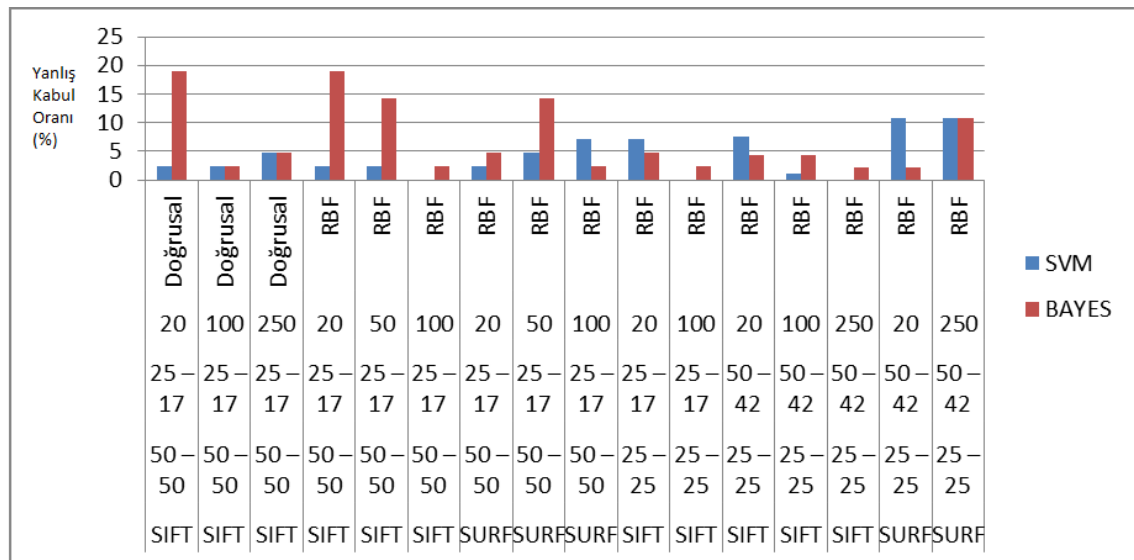
edilmiştir. Çizelge 4.1 ve Çizelge 4.3’de farklı özniteliklerin, görsel kelime sayılarının ve sınıflandırma çeşitlerinin sınıflandırma performansına etkisi belirtilmiştir.

Çizelge 4.1’de aynı ölçekleme seviyesindeki (10 metre çözünürlükte) eğitim ve test verileri kullanılarak gerçekleştirilen deneylerin sonuçları verilmiştir.

Çizelge 4.1 Aynı ölçekleme seviyesindeki deneysel sonuçlar

Öznitelik Çıkartımı	#Eğitim Verileri (Hedef - Hedef olmayan)	#Test Verileri (Hedef olmayan - Hedef)	#Görsel Kelime	SVM çekirdek çeşidi	SVM Yanlış Kabul Oranı (%)	Bayes Yanlış Kabul Oranı (%)
SIFT	50 – 50	25 – 17	20	Doğrusal	2.38	19.05
SIFT	50 – 50	25 – 17	100	Doğrusal	2.38	2.38
SIFT	50 – 50	25 – 17	250	Doğrusal	4.76	4.76
SIFT	50 – 50	25 – 17	20	RBF	2.38	19.05
SIFT	50 – 50	25 – 17	50	RBF	2.38	14.29
SIFT	50 – 50	25 – 17	100	RBF	0	2.38
SURF	50 – 50	25 – 17	20	RBF	2.38	4.76
SURF	50 – 50	25 – 17	50	RBF	4.76	14.29
SURF	50 – 50	25 – 17	100	RBF	7.14	2.38
SIFT	25 – 25	25 – 17	20	RBF	7.14	4.76
SIFT	25 – 25	25 – 17	100	RBF	0	2.38
SIFT	25 – 25	50 – 42	20	RBF	7.60	4.35
SIFT	25 – 25	50 – 42	100	RBF	1.09	4.35
SIFT	25 – 25	50 – 42	250	RBF	0	2.17
SURF	25 – 25	50 – 42	20	RBF	10.87	2.17
SURF	25 – 25	50 – 42	250	RBF	10.87	10.87

Çizelge 4.1’e ait grafik gösterimi aşağıda verilmiştir.



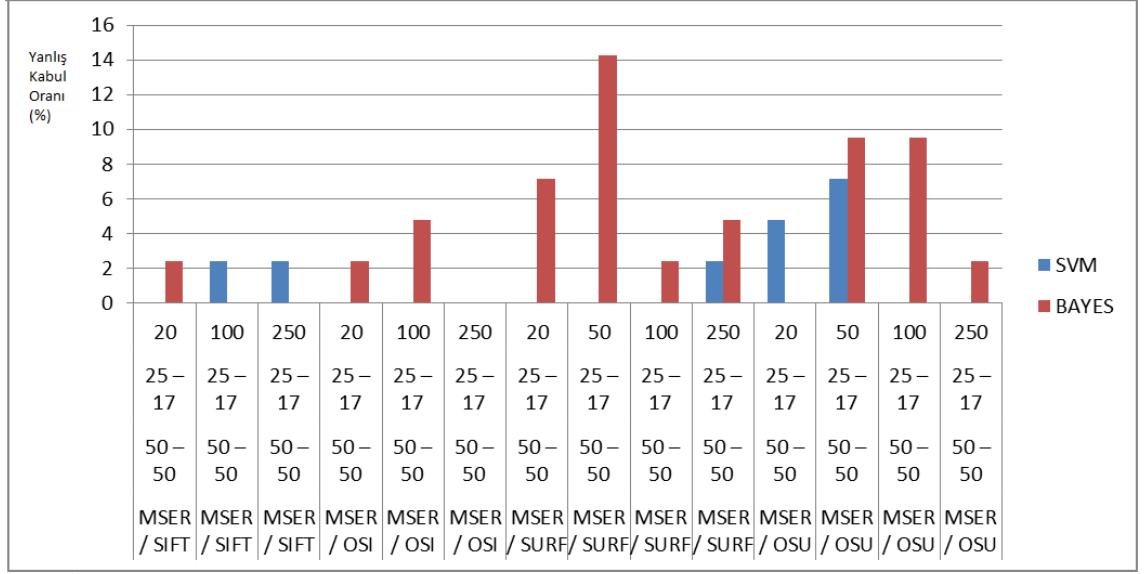
Şekil 4.4 Aynı ölçeklemedeki nesne tanıma hata oranları

İlgin-değişimsiz öznitelik bulucu olan MSER çeşitli öznitelik tanımlayıcılar ile birlikte kullanılarak nesne tanımlama performansı test edilmiştir. Testler aynı ölçekleme seviyesinde gerçekleştirilmiştir. Test sonuçları Çizelge 4.2’de verilmiştir.

Çizelge 4.2 İlgin-değişimsiz öznitelik bulucuyla aynı ölçekleme seviyesindeki deneysel sonuçlar

Öznitelik Bulucu / Tanımlayıcı	#Eğitim Verileri (Hedef - Hedef olmayan)	#Test Verileri (Hedef olmayan - Hedef)	#Görsel Kelime	SVM Yanlış Kabul Oranı (%)	Bayes Yanlış Kabul Oranı (%)
MSER / SIFT	50 – 50	25 – 17	20	0	2.38
MSER / SIFT	50 – 50	25 – 17	100	2.38	0
MSER / SIFT	50 – 50	25 – 17	250	2.38	0
MSER / OpponentSIFT	50 – 50	25 – 17	20	0	2.38
MSER / OpponentSIFT	50 – 50	25 – 17	100	0	4.76
MSER / OpponentSIFT	50 – 50	25 – 17	250	0	0
MSER / SURF	50 – 50	25 – 17	20	0	7.14
MSER / SURF	50 – 50	25 – 17	50	0	14.29
MSER / SURF	50 – 50	25 – 17	100	0	2.38
MSER / SURF	50 – 50	25 – 17	250	2.38	4.76
MSER / OpponentSURF	50 – 50	25 – 17	20	4.76	4.76,
MSER / OpponentSURF	50 – 50	25 – 17	50	7.14	9.52
MSER / OpponentSURF	50 – 50	25 – 17	100	0	9.52
MSER / OpponentSURF	50 – 50	25 – 17	250	0	2.38

Çizelge 4.2 ile aynı ölçekleme seviyesinde MSER kullanılarak elde edilen nesne tanıma performansının SIFT ve SURF öznitelik bulucularından başarılı sonuçlar verdiği gösterilmektedir. Ayrıca SVM sınıflandırma yöntemi ile elde edilen iyileşme Bayes yöntemine göre daha iyidir. Çizelge 4.2’ye ait grafiksel gösterim aşağıda verilmiştir. Grafikte OpponentSURF yönteminin kısaltması OSU, OpponentSIFT yönteminin kısaltması ise OSI olarak gösterilmiştir.



Şekil 4.5 İlgin-değişimsiz öznitelikle aynı ölçeklemede elde edilen nesne tanıma hata oranları

Çizelge 4.3'te farklı ölçekleme seviyesindeki eğitim verileri ve test verileri kullanılarak gerçekleştirilen sınıflandırma sonuçları verilmiştir. Deneysel çalışmalar sırasında 10 metre çözünürlüklü eğitim verileri ve 20 metre çözünürlüklü değerlendirme verileri kullanılmıştır.

Çizelge 4.3 Farklı ölçekleme seviyesindeki deneysel sonuçlar

Öznitelik Bulucu	Öznitelik Tanımlayıcı	#Eğitim Verileri (Hedef - Hedef olmayan)	#Test Verileri (Hedef olmayan - Hedef)	#Görsel Kelime	SVM çekirdek çeşidi	SVM Yanlış Kabul Oranı (%)	Bayes Yanlış Kabul Oranı (%)
SIFT	SIFT	67 – 58	17 – 53	20	RBF	4.29	45.71
SIFT	SIFT	67 – 58	17 – 53	50	RBF	4.29	42.86
SIFT	SIFT	67 – 58	17 – 53	100	RBF	4.29	7.14
SURF	SURF	67 – 58	17 – 53	20	RBF	18.57	52.86
SURF	SURF	67 – 58	17 – 53	50	RBF	18.57	55.71
SURF	SURF	67 – 58	17 – 53	100	RBF	10.00	5.71
SIFT	OpponentSIFT	67 – 58	17 – 53	20	RBF	10	28.57
SIFT	OpponentSIFT	67 – 58	17 – 53	50	RBF	7.14	44.29
SIFT	OpponentSIFT	67 – 58	17 – 53	100	RBF	2.86	7.14
Pyramid SIFT	SIFT	67 – 58	17 – 53	20	RBF	8.57	47.14
Pyramid SIFT	SIFT	67 – 58	17 – 53	50	RBF	4.29	47.14
Pyramid SIFT	SIFT	67 – 58	17 – 53	100	RBF	1.43	4.29
GridSIFT	SIFT	67 – 58	17 – 53	20	RBF	11.43	58.57
GridSIFT	SIFT	67 – 58	17 – 53	50	RBF	7.14	61.43
GridSIFT	SIFT	67 – 58	17 – 53	100	RBF	8.57	4.29
Pyramid SIFT	OpponentSIFT	67 – 58	17 – 53	20	RBF	5.71	41.43
Pyramid SIFT	OpponentSIFT	67 – 58	17 – 53	50	RBF	4.29	8.57

Çizelge 4.3 Farklı ölçekleme seviyesindeki deneysel sonuçlar (devam)

Pyramid SIFT	OpponentSIFT	67 – 58	17 – 53	100	RBF	0	7.14
Pyramid SURF	Opponent SURF	67 – 58	17 – 53	20	RBF	11.43	11.43
Pyramid SURF	Opponent SURF	67 – 58	17 – 53	100	RBF	10.00	12.86
SURF	Opponent SURF	67 – 58	17 – 53	20	RBF	1.29	61.43
SURF	Opponent SURF	67 – 58	17 – 53	100	RBF	8.57	4.29
Pyramid SURF	SURF	67 – 58	17 – 53	20	RBF	18.57	52.86
Pyramid SURF	SURF	67 – 58	17 – 53	100	RBF	7.14	7.14
Grid SURF	SURF	67 – 58	17 – 53	20	RBF	12.90	45.16
Grid SURF	SURF	67 – 58	17 – 53	100	RBF	16.13	3.23
MSEr	SIFT	67 – 58	17 – 53	20	RBF	59.99	24.29
MSEr	SIFT	67 – 58	17 – 53	100	RBF	55.71	58.57
MSEr	OpponentSIFT	67 – 58	17 – 53	20	RBF	58.57	40
MSEr	OpponentSIFT	67 – 58	17 – 53	100	RBF	54.29	64.29
MSEr	Opponent SURF	67 – 58	17 – 53	20	RBF	40.00	44.29
MSEr	Opponent SURF	67 – 58	17 – 53	100	RBF	24.29	31.43
MSEr	SURF	67 – 58	17 – 53	20	RBF	44.29	61.43
MSEr	SURF	67 – 58	17 – 53	100	RBF	20.00	37.14

Çizelge 4.3 incelendiği zaman en başarılı sonucun %0 yanlış kabul oranı (False Alarm Rate - FAR) ile SVM sınıflandırma kullanılarak elde edildiği görülmektedir. Öznitelik bulucu ve tanımlayıcı olarak kullanılan PyramidSIFT ve OpponentSIFT ikilisi diğer tanımlayıcılardan üstün nesne tanıma başarısı göstermiştir. Görsel kelime sayısı seçiminde ise 100 değeri ile iyi sonuçlar alınmıştır. Diğer taraftan Bayes sınıflandırıcı için en başarılı sonuç, %3.23 yanlış kabul oranı ile GridSURF ve SURF öznitelik bulucu ve tanımlayıcı ikilisi ile elde edilmiştir. Bayes sınıflandırıcıda da görsel kelime sayısı olarak 100 seçimi ile öznitelik dağılımının daha kararlı yapıldığı görülmektedir. Genel sonuçlara bakıldığı zaman SVM yöntemi, sınıflandırma başarısı açısından Bayes yöntemine nazaran üstün sonuçlar vermiştir.

5. SONUÇLAR

Bütünsel yöntemler, önbilgi veya varsayımlar gerektirmediği için avantajlıdır. Bu yüzden birçok alanda kullanılmak için elverişlidir. Nesnenin spesifik özelliklerine uygun olarak öznitelik vektörleri kullanılabilir. Bu özelliklerinin yanında kullanımı da basit ve anlaşılırdır. Ancak bu basitlik verimsizliği de beraberinde getirmektedir. Nesnelerin tanımlanma olasılığını arttırmak için ölçeklendirilmiş ve dönülü şablonların kullanımı gerekmektedir.

Bütünsel yöntemler, karmaşık arka plan senaryoları için başarılı sonuçlar verememektedir. Açık olarak ayırt edilebilir nesneleri ayırt etmek için kullanılırlar. Ayrıca bu yöntemler, nesnenin arka planından doğru olarak bölütlendiğini varsayarlar.

Dönüşüm – arama tabanlı yöntemler, çeşitliliği fazla olan nesneler için uygulanabilirler. Bu yöntemler, uzaklık veya benzerlik ölçüsü tabanlıdır. Kapatmanın (occlusion) üstesinden gelebilirler. Aynı zamanda arka plandaki parazit karışımı için de dayanıklıdır.

Dönüşüm – arama tabanlı yöntemde, nesnelerin gösterimi sonlu sayıdaki nokta kümesi ile sınırlanmıştır. Modele ait her bir noktanın ölçüye katkısı bulunmaktadır. Nesne gösteriminin nokta tabanlı olması, modele ait bazı noktaların olmamasının tolere edilmesini sağlamaktadır. Bu yüzden, dönüşüm – arama tabanlı yöntemlerde bölütleme ihtiyaç duyulan bir işlem değildir.

Nesne gösteriminin sonlu sayıdaki nokta kümesi ile sınırlandırılması, kenar tespiti ile de gerçekleştirilir. Bu işlem, ışıklılık değişimlerine karşı kısmi dayanıklılık sağlar. Kontrast seviyesindeki küçük farklar, eğitim aşamasında bazı kenarların tespit edilmesini ancak tanıma sürecinde gradyan eşik seviyesinin altında kalması sebebiyle tespit edilememeye neden olur. Bu, kısmi kapatma (occlusion) ile aynı etkiye sebep olmaktadır.

Tez kapsamında, uydu görüntüleri üzerinde düşük yanlış kabul oranında nesne tanıma gerçekleştirmek için öznitelik tabanlı bir yöntem kullanılmıştır. Yapılan deneysel çalışmalar doğrultusunda sistemin performansını etkileyen parametreler belirlenmiştir. Sözlüğü oluşturan görsel kelime sayısının, öznitelik çıkartım yönteminin ve sınıflandırma yönteminin nesne tanıma performansını etkilediği deneysel çalışmalarla tespit edilmiştir. Bu anlamda, SIFT tabanlı yöntemler, SURF tabanlı yöntemler karşısında daha başarılı sonuçlar vermektedir. Ancak SURF tabanlı yöntemlerin SIFT tabanlı yöntemlerle benzer hatta daha iyi sonuçlar verdiği durumlarda bulunmaktadır. SURF öznitelik tanımlama yönteminin 64 boyutlu vektörler kullanması SIFT yöntemine kıyasla hızlı sonuçlar elde edilmesini sağlamıştır. Ayrıca anahtar nokta tespiti için alternatif olarak kullanılan GridSIFT, GridSURF, PyramidSIFT, PyramidSURF yöntemleri ile anahtar nokta tanımlayıcı için alternatif olarak kullanılan OpponentSIFT ve OpponentSURF yöntemleri ile daha başarılı sonuçlar alınabilmektedir.

İlgin-değişimsiz MSER yöntemi ile yapılan çalışmalar, eğitim ve değerlendirme verileri aynı ölçekleme seviyesinde iken daha başarılı sonuçlar elde edildiğini; farklı ölçekleme seviyelerinde ise MSER yönteminin başarımının düştüğünü göstermektedir. Elde edilen sonuçlar neticesinde farklı ölçekleme seviyelerinde MSER kullanımı önerilmemektedir.

SIFT ve SURF yöntemleri, rotasyon ve ölçekleme değişmezlikli olması yanında 45° - 60° arasında ilgin-değişmezlik de sağlamaktadır. Uydu görüntülerinde genellikle 45° dereceden düşük bakış açısında elde edilen görüntüler kullanıldığından öznitelik çıkartım tekniği olarak SIFT ve SURF tercih edilmiştir. Uydu görüntülerinde nesne tanıma için bakış açısı ihtiyacı göz önünde bulundurularak 85° 'ye kadar ilgin-değişmezlik sağlayan hesapsal karmaşıklığı yüksek olan ASIFT yöntemi tercih edilmemiştir.

Sözlükte bulunan görsel kelime sayısı ile doğru orantılı olarak nesne tanıma performansı değişmektedir. Sözlükteki kelime sayısının az olması, görüntü parçalarının yetersiz gösterimine neden olmakta ve yetersiz örtüşme (underfitting - yüksek hata) problemine neden olmaktadır. Diğer taraftan, fazla sayıda kelime kullanımı ise aşırı örtüşme (overfitting - yüksek değişim (variance)) problemine ve nicemleme hatalarına

neden olmaktadır. Bu yüzden optimum sayıda görsel kelime seçimi çok önemlidir. Ayrıca görsel kelime sayısının aşırı arttırımı durumunda nesne tanıma performansı da incelenmiştir. Görsel kelime sayısının 1000 seçilerek yapılan çalışmalar nesne tanıma performansını düşük miktarda iyileştirebilmektedir. Ancak bazı durumlarda hiçbir iyileşme sağlanamadığı gibi daha kötü tanıma oranları elde edilmiştir. Ayrıca bu işlem hesap karmaşıklığının çok fazla artmasına sebep olmuştur. Özellikle sınıflandırma eğitimi aşamasında eğitim süresi nesne tanıma yöntemi kullanılamayacak kadar artmıştır. Bu yüzden bu değerle yapılan testlere devam edilmemiştir.

Sınıflandırma yöntemi olarak kullanılan destek vektör makinaları (SVM), genellikle Bayes sınıflandırma yönteminden başarılı sonuçlar vermektedir. Ayrıca sınıflandırıcılar kısıtlı eğitim verisi kullanılarak eğitildiği zaman SVM yönteminin daha başarılı ve göze çarpan sonuçlar verdiği görülmüştür. Buna rağmen, Bayes sınıflandırma yönteminin nadir de olsa daha iyi sonuçlar verdiği durumlar bulunmaktadır. Bayes yöntemi ile daha başarılı sonuçlar elde etmek için nesne tanıma algoritmasını oluşturan tüm parametreler detaylı olarak incelenmeli ve en iyi parametreler seçilmelidir. Bu sonuç, karmaşık sınıflandırma yöntemleri kullanmak yerine daha iyi öznelilik çıkartım yöntemleri ve uygun sayıda görsel kelime kullanımının daha iyi sonuçlar verebileceğini göstermektedir.

KAYNAKLAR

- Abdel-Hakim, A. E. and Farag, A. A. 2006. CSIFT: A SIFT descriptor with color invariant characteristics, In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2006), pp. 1978–1983, New York City, NY.
- Alexandrov, A., Hristova, T., Ivanova, K., Koeva, M., Madzharova, T. and Petrova, V., 2004. Application of QUICKBIRD Satellite Imagery For Updating Cadastral Information, XXth ISPRS Congress, İstanbul-Turkey.
- Anandan, P. 1989. A computational framework and an algorithm for the measurement of visual motion, *International Journal of Computer Vision*, 2(3), 283–310.
- Arbel'aez, P., Maire, M., Fowlkes, C., and Malik, J., 2010, Contour Detection and Hierarchical Image Segmentation, Technical Report UCB/EECS-2010-17, EECS Department, University of California, Berkeley. Submitted to PAMI.
- Baeza-Yates, R. and Ribeiro-Neto, B., 1999, *Modern Information Retrieval*. Addison Wesley.
- Baumberg, A., 2000, Reliable feature matching across widely separated views, In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2000), pp. 774–781, Hilton Head Island.
- Bay, H., Tuytelaars, T. and Van Gool, L., 2006, SURF: Speeded up robust features, In Ninth European Conference on Computer Vision (ECCV 2006), pp. 404–417.
- Beardsley, P., Torr, P., and Zisserman, A., 1996, 3D model acquisition from extended image sequences, In Fourth European Conference on Computer Vision (ECCV'96), pp. 683–695, Cambridge, England
- Boser, B. E., Guyon, M. and Vapnik, N. 1992. A training algorithm for optimal margin classifiers. In Haussler, David (editor); 5th Annual ACM Workshop on COLT, pages 144–152, Pittsburgh, PA, ACM Press
- Bowyer, K. W., Kranenburg, C., and Dougherty, S. 2001. Edge detector evaluation using empirical ROC curves, *Computer Vision and Image Understanding*, 84(1):77–103.
- Brown, M. and Lowe, D. 2002. Invariant features from interest point groups, In British Machine Vision Conference, pp. 656–665, Cardiff, Wales.

- Brown, M. and Lowe, D. 2003. Unsupervised 3D object recognition and reconstruction in unordered datasets, In International Conference on 3D Imaging and Modelling, pp. 1218–1225, Nice, France.
- Brown, M. and Lowe, D. 2007. Automatic panoramic image stitching using invariant Features, International Journal of Computer Vision, 74(1), 59–73.
- Brown, M., Szeliski, R., and Winder, S. 2005. Multi-image matching using multi-scale oriented patches, In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2005), pp. 510–517, San Diego, CA.
- Carneiro, G. and Jepson, A. 2005. The distinctiveness, detectability, and robustness of local image features. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2005), pp. 296–301, San Diego, CA.
- Chum, O., Philbin, J., Sivic, J., Isard, M., and Zisserman, A. 2007. Total recall: Automatic query expansion with a generative feature model for object retrieval, In Eleventh International Conference on Computer Vision (ICCV 2007), Rio de Janeiro, Brazil.
- Congalton R.G., Green, K. 2009. Assessing the Accuracy of Remotely Sensed Data, 2.Baski, CRC Press, New York.
- Corso, J. and Hager, G. 2005. Coherent regions for concise and stable image description, In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2005), pp. 184–190, San Diego, CA.
- Cortes, C., Vapnik, V. 1995. Support Vector Network, Machine Learning, 20(3): 273–297.
- Devroye, L., Györfi, L., Lugosi, G. 1996. A probabilistic theory of pattern recognition, Springer. ISBN 0-3879-4618-7.
- Douglas, D. H. and Peucker, T. K. 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, The Canadian Cartographer, 10(2):112–122.
- Eisenbeiss, H., Baltsavias, E., Pateraki, M., Zhang, L. 2004. Potential of IKONOS and QUICKBIRD Imagery for Accurate 3D Point Positioning, Orthoimage and DSM Generation, XXth ISPRS Congress, İstanbul-Turkey
- Elder, J. H. and Goldberg, R. M. 2001. Image editing in the contour domain, IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(3), 291–296.

- Fawcett, T. 2006. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874.
- Fei-Fei, L., ve Perona P. 2005. A Bayesian Hierarchical Model for Learning Natural Scene Categories. *Proc. of IEEE Computer Vision and Pattern Recognition*. pp. 524–531.
- Fei-Fei, L., Fergus, R., ve Torralba, A. 2007. Recognizing and Learning Object Categories, *CVPR 2007 short course*.
- Ferencz, A., Learned-Miller, E. G., and Malik, J. 2008. Learning to locate informative features for visual identification, *International Journal of Computer Vision*, 77(1-3):3–24.
- Fergus, R., Perona, P., and Zisserman, A. 2007. Weakly supervised scale-invariant learning of models for visual recognition, *International Journal of Computer Vision*, 71(3), 273–303.
- Förstner, W. 1986. A feature-based correspondence algorithm for image matching. *Intl. Arch. Photogrammetry & Remote Sensing*, 26(3):150–166.
- Freeman, W. T. and Adelson, E. H. 1991. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(9), 891–906.
- Hamerly, G., ve Elkan, C. 2002. Alternatives to the k-means algorithm that find better clusterings". *Proceedings of the eleventh international conference on Information and knowledge management (CIKM)*.
- Hannah, M. J. 1974. *Computer Matching of Areas in Stereo Images*, Ph.D. thesis, Stanford University.
- Hannah, M. J. 1988. Test results from SRI's stereo system. In *Image Understanding Workshop*, pp. 740–744, Cambridge, Massachusetts.
- Harris, C. and Stephens, M. J. 1988. A combined corner and edge detector. In *Alvey Vision Conference*, pp. 147–152.
- Horn, B. K. P. and Schunck, B. G. 1981. Determining optical flow. *Artificial Intelligence*, 17, 185–203.
- Hua, G., Brown, M., and Winder, S. 2007. Discriminant embedding for local image descriptors, In *Eleventh International Conference on Computer Vision (ICCV 2007)*, Rio de Janeiro, Brazil.

- Huertas, A. ve Nevatia, R. 1988. Detecting Buildings in Aerial Images, *Computer Vision, Graphics, and Image Processing*, 41, s.131-152.
- İşcan, L., Aksu, O., Önder, M., Atak, V.O., Lenk, O., Gürdal, M.A. 2004. Accuracy Assessment of High Resolution Satellite Images, XXth ISPRS Congress, İstanbul-Turkey.
- Kadir, T., Zisserman, A., and Brady, M. 2004. An affine invariant salient region detector, In *Eighth European Conference on Computer Vision (ECCV 2004)*, pp. 228–241, Prague.
- Ke, Y. and Sukthankar, R. 2004. PCA-SIFT: a more distinctive representation for local image descriptors, In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2004)*, pp. 506–513, Washington, DC.
- Kenney, C., Zuliani, M., and Manjunath, B. 2005. An axiomatic approach to corner detection, In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2005)*, pp. 191–197, San Diego, CA.
- Kim, T., Lee, T.Y. ve Kim, K.O. 2006. Semiautomatic Building Line Extraction from Ikonos Images Through Monoscopic Line Analysis, *Photogrammetric Engineering and Remote Sensing*, 72(5), s.541-549.
- Kim, Z. ve Nevatia, R. 2004. Automatic Description of Complex Buildings from Multiple Images, *Computer Vision and Image Understanding*, 96, s.60-95.
- Koethe, U. 2003. Integrated edge and junction detection with the boundary tensor, In *Ninth International Conference on Computer Vision (ICCV 2003)*, pp. 424–431, Nice, France.
- Leung, T., ve Malik, J. 2001. Representing and recognizing the visual appearance of materials using three-dimensional textons, *International Journal of Computer Vision* 43 (1): 29–44. doi:10.1023/A:1011126920638.
- Liang, L., Liu, C., Xu, Y.-Q., Guo, B., and Shum, H.-Y. 2001. Real-time texture synthesis by patch-based sampling, *ACM Transactions on Graphics*, 20(3), 127–150.
- Lin, C., Huertas, A. ve Nevatia, R. 1994. Detection of Buildings Using Perceptual Grouping and Shadows, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Seattle, A.B.D., s. 62-69.
- Lindeberg, T. 1993. Detecting salient blob-like image structures and their scales with

- a scale-space primal sketch: a method for focus-of-attention. *International Journal of Computer Vision*, 11(3), 283–318.
- Lindeberg, T. 1998. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2), 79–116.
- Lindeberg, T. and Garding, J. 1997. Shape-adapted smoothing in estimation of 3-D shape cues from affine deformations of local 2-D brightness structure. *Image and Vision Computing*, 15(6):415–434.
- Loop, C. and Zhang, Z. 1999. Computing rectifying homographies for stereo vision, In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'99)*, pp. 125–131, Fort Collins.
- Lowe, D. G. 2004. Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, 60(2), 91–110.
- Lucas, B. D. and Kanade, T. 1981. An iterative image registration technique with an application in stereo vision, In *Seventh International Joint Conference on Artificial Intelligence (IJCAI-81)*, pp. 674–679, Vancouver.
- Manning, C. D., Raghavan, P., and Schütze, H. 2008. *Introduction to Information Retrieval*. Cambridge University Press.
- Matas, J., Chum, O., Urban, M., and Pajdla, T. 2004. Robust wide baseline stereo from maximally stable extremal regions, *Image and Vision Computing*, 22(10), 761–767.
- MacKay, D. 2003. *An Example Inference Task: Clustering, Information Theory, Inference and Learning Algorithms*. Cambridge University Press, pp. 284–292. ISBN 0-521-64298-1. MR 2012999.
- Moravec, H. 1983. The Stanford cart and the CMU Rover, *Proceedings of the IEEE*, 71(7), 872–884.
- Mikolajczyk, K. and Schmid, C. 2004. Scale & affine invariant interest point detectors. *International Journal of Computer Vision*, 60(1), 63–86.
- Mikolajczyk, K. and Schmid, C. 2005. A performance evaluation of local descriptors, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10):1615–1630.
- Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F.,

- Kadir, T., and Van Gool, L. J. 2005. A comparison of affine region detectors. *International Journal of Computer Vision*, 65(1-2), 43–72.
- Nister, D. and Stewenius, H. 2006. Scalable recognition with a vocabulary tree, In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2006)*, pp. 2161–2168, New York City, NY.
- Nister, D. and Stewenius, H. 2008. Linear time maximally stable extremal regions, In *Tenth European Conference on Computer Vision (ECCV 2008)*, pp. 183–196, Marseilles.
- Özbalmumcu, M. 1999. Coğrafi Bilgi Sistemi Oluşturulmasında Kullanılan Veri Kaynakları, Veri Toplama Sistemleri ve Konumsal Veri Toplama Yöntemlerinin Araştırılması, Doktora Tezi, Yıldız Teknik Üniversitesi, İstanbul.
- Özbalmumcu, M., Erdoğan, M. 2001. Uzaktan Algılama Amaçlı Uydu Görüntüleme Sistemleri, *Harita Dergisi*, Sayı: 125, Ankara.
- Platel, B., Balmachnova, E., Florack, L., and ter Haar Romeny, B. 2006. Top-points as interest points for image matching. In *Ninth European Conference on Computer Vision (ECCV 2006)*, pp. 418–429.
- Rosten, E. and Drummond, T. 2006. Machine learning for high-speed corner detection, In *Ninth European Conference on Computer Vision (ECCV 2006)*, pp. 430–443.
- San, D. ve Turker, M. 2004. Automatic Building Detection and Delineation from High Resolution Space Images Using Model-Based Approach, *Proceedings of the XX ISPRS Kongresi, Komisyon I*, İstanbul.
- Sande, E.A.K., Gevers, T., ve Snoek, C. 2008. Color descriptors for object category recognition, *CGIV 2008*.
- Schaffalitzky, F. and Zisserman, A. 2002. Multi-view matching for unordered image sets, or “How do I organize my holiday snaps?”, In *Seventh European Conference on Computer Vision (ECCV 2002)*, pp. 414–431, Copenhagen
- Scharstein, D. and Szeliski, R. 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, *International Journal of Computer Vision*, 47(1), 7– 42.
- Schmid, C. and Mohr, R. 1997. Local grayvalue invariants for image retrieval, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(5), 530–534.
- Schmid, C., Mohr, R., and Bauckhage, C. 2000. Evaluation of interest point detectors,

- International Journal of Computer Vision, 37(2), 151–172.
- Shi, J. and Tomasi, C. 1994. Good features to track, In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'94), pp. 593–600, Seattle.
- Sinha, S. N., Steedly, D., Szeliski, R., Agrawala, M., and Pollefeys, M. 2008. Interactive 3D architectural modeling from unordered photo collections, ACM Transactions on Graphics (Proc. SIGGRAPH Asia 2008), 27(5).
- Snaveley, N., Seitz, S. M., and Szeliski, R. 2006. Photo tourism: Exploring photo collections in 3D. ACM Transactions on Graphics (Proc. SIGGRAPH 2006), 25(3), 835–846.
- Şirinyıldız, T. 2004. Change Detection Analysis by Using IKONOS Imagery, XXth ISPRS Congress, İstanbul-Turkey.
- Triggs, B. 2004, Detecting keypoints with stable position, orientation, and scale under illumination changes, In Eighth European Conference on Computer Vision (ECCV2004), pp. 100–113, Prague.
- Tuytelaars, T. and Mikolajczyk, K. 2007. Local invariant feature detectors, Foundations and Trends in Computer Graphics and Computer Vision, 3(1).
- Tuytelaars, T. and Van Gool, L. 2004. Matching widely separated views based on affine invariant regions, International Journal of Computer Vision, 59(1):61–85.
- Van de Weijer, J. and Schmid, C. 2006. Coloring local feature extraction, In Ninth European Conference on Computer Vision (ECCV 2006), pp. 334–348.
- Vapnik, V.N. 1995. The Nature of Statistical Learning Theory, Springer-Verlag, New York.
- Vapnik, V.N. 2000. The Nature of Statistical Learning Theory, 2. Baskı, Springer-Verlag, New York.
- Winder, S. and Brown, M. 2007. Learning local image descriptors, In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007), Minneapolis, MN.
- Xiao, J. and Shah, M. 2003. Two-frame wide baseline matching, In Ninth International Conference on Computer Vision (ICCV 2003), pp. 603–609, Nice, France.

- Zhang, Z., Deriche, R., Faugeras, O., and Luong, Q. 1995. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry, *Artificial Intelligence*, 78:87–119.
- Zoghalmi, I., Faugeras, O., and Deriche, R. 1997. Using geometric corners to build a 2D mosaic from a set of images, In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'97)*, pp. 420–425, San Juan, Puerto Rico.

ÖZGEÇMİŞ

Adı Soyadı : Yusuf SOYMAN

Doğum Yeri : MALATYA

Doğum Tarihi : 25.08.1984

Medeni Hali : Bekar

Yabancı Dili : İngilizce

Eğitim Durumu

Lise : Erzincan Fen Lisesi (1998 – 2000)

Ankara Kurtuluş Lisesi (2000 – 2002)

Lisans : Çankaya Üniversitesi, Bilgisayar Mühendisliği (2002-2008)

Çankaya Üniversitesi, Elektronik - Haberleşme Mühendisliği,
(2004 – 2008)

Yüksek Lisans: Ankara Üniversitesi, Fen Bilimleri Enstitüsü,

Elektrik – Elektronik Mühendisliği, (Şubat 2011 – Eylül 2013)

Çalıştığı Kurumlar ve Yıl

AYDIN Yazılım (06/2005 – 09/2005)

AYDIN Yazılım (06/2005 – 08/2005)

ROKETSAN A.Ş. (02/2008 – ...)

Yayınlar

SOYMAN, Yusuf. *Robust Automatic Target Recognition in FLIR imagery.*

Automatic Target Recognition XXII, Proceedings of SPIE, 2012.

SOYMAN, Yusuf., ILGIN, Hakkı Alparslan. Rotation, Scaling and Translation

Invariant Object Recognition. 13th International Conference on Knowledge in

Telecommunication Technologies and Optics, 2013.