

BAYES TABANLI SAHNE SINIFLANDIRMASI

BAYESIAN SCENE CLASSIFICATION

AJDA KARABULUT

Hacettepe Üniversitesi

Lisansüstü Eğitim – Öğretim ve Sınav Yönetmeliğinin

BİLGİSAYAR Mühendisliği Anabilim Dalı İçin Öngördüğü

YÜKSEK LİSANS TEZİ

olarak hazırlanmıştır.

2006

Fen Bilimleri Enstitüsü Müdürlüğü'ne,

Bu çalışma jürimiz tarafından **BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI'nda YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

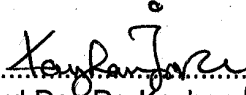
Başkan(Danışman)


.....
Prof.Dr. Ali SAATÇI

Üye(Danışman)


.....
Yard.Doç.Dr. Selim AKSOY

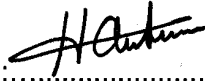
Üye


.....
Yard.Doç.Dr. Kayhan İMRE

Üye


.....
Dr. Mustafa EGE

Üye


.....
Yard.Doç.Dr. Harun ARTUNER

ONAY

Bu tez, /..... / tarihinde Enstitü Yönetim Kurulunca kabul edilmiştir.

Prof. Dr. Ahmet R. ÖZDURAL
FEN BİLİMLERİ ENSTITÜSÜ MÜDÜRÜ

BAYES TABANLI SAHNE SINIFLANDIRMASI

Ajda Karabulut

ÖZ

Günümüzde teknolojinin hızlı bir şekilde ilerlemesi, sayısal ortamda saklanan resim ve video arşivlerini ve bu arşivleri kullanan araştırma alanlarını hızlı bir şekilde arttırmıştır. Bu arşivlerden en iyi şekilde yararlanılabilmesi için otomatik olarak değerlendirme ve çözümleme yapılması zorunlu bir koşul olarak ortaya çıkmıştır. Bu konuda, bu güne kadar yapılan çalışmalarda, resimlerin bütününden elde edilen öznitelikler kullanılarak çeşitli istatistiksel yöntemler uygulanmıştır. Ancak, bu yöntemlerin başarısının sınırlı kalması, resimlerdeki alt alanların incelenmesi ve resimlerdeki uzamsal bilginin dikkate alınması gerektiğini göstermiştir. Bu tez çalışmasında, uzamsal bilgi, resmin alt alanlara bölünmesiyle modellenmiştir. Alt alanlardan çıkarılan öznitelik bilgileriyle öbeklendirme yapılmış ve resimler, farklı öbeklere atanan alt alanların toplamı olarak üç farklı biçimde modellenmiştir. Önerilen modeller, kar, bitki örtüsü, su, gökyüzü, dağ ve şehir sahnelerinin sınıflandırılması için TRECVID2005 verisi ile test edilmiş ve uzamsal bilgiyi kullanmayan modellere göre sınıflandırma başarımının arttığı gözlenmiştir.

Anahtar Kelimeler: resim ve video arşivleri, uzamsal bilgi, öbeklendirme, sınıflandırma.

Danışman: Yard.Doç.Dr. Selim AKSOY, Bilkent Üniversitesi, Bilgisayar Mühendisliği Bölümü

Danışman: Prof.Dr. Ali SAATÇI, Hacettepe Üniversitesi, Bilgisayar Mühendisliği Bölümü

BAYESIAN SCENE CLASSIFICATION

Ajda Karabulut

ABSTRACT

The rapid development of technology has increased the digital image and video archives and enhanced the research areas that use these archives. To benefit from these archives most efficiently, automatically evaluating and analysing them has arisen as an obligation. At all the researches about this topic which are done up to now, many statistical methods were applied that use the features extracted from the whole image. But, the limited success of these methods has shown that, it is required to examine the subregions of the images and to consider the spatial information. In this thesis, spatial information has been modelled by the division of image into subregions. Clustering has been done with the features that are extracted from the subregions and the images have been modelled as the sum of distinctly clustered subregions in three different ways. The proposed models have been tested with TRECVID2005 dataset for the classification of snow, vegetation, water, sky, mountain and city scenes and it has been observed that the classification performance increased according to the models that do not use spatial information.

Keywords: image and video archives, spatial information, clustering, classification.

Advisor: Asst.Prof.Dr. Selim AKSOY, Bilkent University, Department of Computer Engineering

Advisor: Prof.Dr. Ali SAATÇI, Hacettepe University, Department of Computer Engineering

TEŐEKKÜR

Yazar, bu alıőmanın gerekleőmesinde katkılarından dolayı, aőađıda adı geen kiőilere itenlikle teőekkür eder.

Sayın Yard.Do.Dr. Selim AKSOY (tez danıőmanı), alıőmanın sonuca ulaőtırılmasında ve karőtılaőtılan gülüklerin aőtılmasında yön gösterici olmuőtur.

Sayın Prof.Dr. Ali SAATİ (tez danıőmanı), tez metninin son halini almasına katkıda bulunmuőtur.

Sayın Tayfun DALKILI, tez süresi boyunca manevi desteęini esirgememiőtir.

İÇİNDEKİLER DİZİNİ

Sayfa

ÖZ	i
ABSTRACT	ii
TEŞEKKÜR	iii
İÇİNDEKİLER DİZİNİ	iv
ŞEKİLLER DİZİNİ	vi
ÇİZELGELER DİZİNİ	vii
SİMGELER VE KISALTMALAR DİZİNİ	viii
1. GİRİŞ	1
2. SAHNE SINIFLANDIRMASI	3
2.1. Uygulama Alanları	3
2.2. TRECVID Yarışması	4
2.3. Sahne Sınıflandırmasına Genel Bakış (Tarihçe)	5
2.4. Sahne Sınıflarının Uzamsal İlişkiler ile Modellenmesi	10
2.5. Organizasyon	11
3. GÖRSEL ÖZİNİTELİKLER	12
3.1. Renk Öz niteliği	12
3.1.1. Renk uzayları	12
3.1.1.1. RGB renk uzayı	12
3.1.1.2. HSV renk uzayı	14
3.2. Doku Öz niteliği	16
3.2.1. Gabor süzgeçler	17
3.3. Kenar Öz niteliği	18
3.3.1. Canny kenar detektörü	18
4. İSTATİSTİKSEL ÖRÜNTÜ TANIMA YÖNTEMLERİ	20
4.1. Örüntü Tanıma	20
4.2. Öbeklendirme	23
4.3. Bayes Karar Kuralı	25
5. BAYES TABANLI SAHNE SINIFLANDIRMASI	30
5.1. Resimlerin Modellenmesi	30
5.2. Tek Alt Alanlı Konumdan Bağımsız Model	32
5.3. Tek Alt Alanlı Konuma Bağımlı Model	35
5.4. Çift Alt Alanlı Görelî Konumlu Model	35
5.5. Önemli Parametrelerin Seçilmesi	36
6. DENEYLER	39
6.1. Veri Kümesi	39
6.2. Öz niteliklerin Çıkarılması	40
6.3. Alt Alanların Öbeklendirilmesi	42
6.4. Tek Alt Alanlı Konumdan Bağımsız Model	47
6.5. Tek Alt Alanlı Konuma Bağımlı Model	48
6.6. Çift Alt Alanlı Görelî Konumlu Model	52
6.7. Önemli Parametrelerin Seçilmesi	56
6.8. Sonuçların Geleneksel Yöntemlerle Karşılaştırılması	58

7. SONUÇLAR	59
KAYNAKLAR.....	62

ŞEKİLLER DİZİNİ

Sayfa

Şekil 3.1. RGB renk küpü	14
Şekil 3.2. HSV renk uzayının koni biçiminde gösterimi.....	15
Şekil 3.3. Çeşitli doku örnekleri	17
Şekil 3.4. Farklı ölçek ve yönelimlerdeki süzgeçler	17
Şekil 3.5. Canny kenar detektöründen farklı parametreler ile alınan sonuçlar.....	19
Şekil 4.1. Örüntü tanıma dizgeleri	22
Şekil 4.2. Örüntü tanıma dizgelerinin tasarımı.....	23
Şekil 4.3. Farklı şekil ve büyüklüklerdeki öbekler	24
Şekil 4.4. Farklı önsel olasılıklar için eşik değerleri	27
Şekil 4.5. Bayes hata oranı	28
Şekil 5.1. Resmin alt alanlara bölünmesi.....	30
Şekil 5.2. Alt alanların öbeklere ayrılması	31
Şekil 5.3. Resimlerin alt alanların toplamı olarak modellenmesi.....	32
Şekil 5.4. Alt alanların gösterimi	32
Şekil 5.5. Alt alan çiftlerinin gösterimi	36
Şekil 6.1. Sahne sınıflarına ait resim örnekleri	39
Şekil 6.2. Bir resmin 4x4'lük alt alanlara ayrılması	41
Şekil 6.3. Bir resmin 7x5'lik alt alanlara ayrılması.....	41
Şekil 6.4. Kenar yönelim histogramları kullanılarak elde edilen öbek örnekleri	43
Şekil 6.5. HSV histogramı ve k-ortalama algoritması ile elde edilen öbek örnekleri ..	44
Şekil 6.6. Doku özniteliği kullanılarak öbeklerin alt öbeklere ayrılması.....	45
Şekil 6.7. Bazı öbek örnekleri	46
Şekil 6.8. Öbeklerin ayrıntılı gösterimi	46
Şekil 6.9. Öbeklerin sınıflardaki dağılımını gösteren olasılık tablosu.....	47
Şekil 6.10. Tek alt alanlı konumdan bağımsız model kullanıldığında doğru ve yanlış olarak sınıflandırılan resim örnekleri.....	49
Şekil 6.11. Bir sınıf için her konumdaki öbek dağılımını gösteren olasılık tablosu.....	50
Şekil 6.12. Tek alt alanlı konuma bağımlı model kullanıldığında doğru ve yanlış olarak sınıflandırılan resim örnekleri	51
Şekil 6.13. Alt alan çiftlerinin sınıflardaki dağılımını gösteren olasılık tablosu	52
Şekil 6.14. Çift alt alanlı görel konumlu model kullanıldığında doğru ve yanlış olarak sınıflandırılan resim örnekleri	54
Şekil 6.15. Tek alt alanlı konuma bağımlı model çerçevesinde seçilen öbeklerden bazıları.....	57

ÇİZELGELER DİZİNİ

	<u>Sayfa</u>
Çizelge 6.1. TRECVID 2005 yarışması veri kümesi[16]	40
Çizelge 6.2. Tek alt alanlı konumdan bağımsız model için hata matrisi	48
Çizelge 6.3. Tek alt alanlı konuma bağımlı model için hata matrisi	50
Çizelge 6.4. Çift alt alanlı görelı konumlu model için hata matrisi	53
Çizelge 6.5. Tek alt alanlı konumdan bağımsız model ile seçme yaparak gerçekleştirilen sınıflandırmaya ilişkin hata matrisi	56
Çizelge 6.6. Tek alt alanlı konuma bağımlı model ile seçme yapılarak gerçekleştirilen sınıflandırmaya ilişkin hata matrisi	57
Çizelge 6.7. QDC ile gerçekleştirilen sınıflandırma için hata matrisi	58

SİMGELER VE KISALTMALAR DİZİNİ

NIST	Amerikan Standartlar Enstitüsü
RGB	Red Green Blue
HSV	Hue Saturation Value
CMY	Cyan Magenta Yellow
bpp	bit per pixel
SIFT	Scale Invariant Feature Transform
BOV	bag-of-visual-words
DCT	Discrete Cosine Transform
GB	Gigabayt
QDC	Quadratic Discriminant Classifier

1. GİRİŞ

Son yıllarda, teknolojik gelişmelerin, resimlerin sayısallaştırılması, saklanması ve iletiminde ucuz ve etkili yollar ortaya çıkarmasıyla çok sayıda arşiv değişik alanlardaki kişiler tarafından kullanılır olmuştur. Birçok ticari kuruluş geniş resim ve video arşivlerine sahiptir. Sayısal fotoğraf makinelerinin kullanılmaya başlanmasıyla da pek çok insan kendi kişisel fotoğraflarını bilgisayarında saklama olanağına kavuşmuştur. Pek çok müze, tabloların ve sanat eserlerinin resimlerini sayısallaştırarak sanal müzeler oluşturmaktadır. Böylelikle hem erişimi kolay, hem de asıl eserleri koruyan güvenli bir uygulama meydana getirmektedirler. Seyahat acentaları da tatil yerlerinin resimlerini kullanarak kendilerine sayısal arşivler oluşturmaktadırlar. Bir kullanıcı, seyahatini planlarken bu arşivler üzerinde sorgulama yapabilmektedir. Ayrıca, hepimizin sıkça kullandığı İnternet üzerindeki resimler, uydu görüntüsü verileri, biomedikal veriler, haber fotoğrafları, haber ve belgesel videoları, film arşivleri, güvenlik videoları özümsemeyecek miktarda veri içermektedir. Verinin büyüklüğü hakkında bir fikir vermesi açısından, dünyanın yörüngesindeki tek bir uydudan her gün yüzlerce GB veri gönderildiğini söyleyebiliriz[1]. Yukarıda bahsedilen veri tabanları geleceğe ait bir rüya değil, aksine günümüzde kullanılmaya başlanan sistemlerdir. Veri tabanları, farklı kişiler tarafından farklı şekillerde kullanılsa da açık olan şey bu veri tabanlarının ancak verimli bir şekilde arama ve tarama yapılabilirdiği, otomatik olarak değerlendirilebildiği sürece faydalı olacaktır.

Eski sistemlerde, resimler veri tabanına kaydedilirken, resmin içinde neler olduğuna dair birtakım anahtar sözcükler de resimle birlikte kaydediliyordu. Kullanıcı, sistemde sorgulama yapmak istediğinde, aradığı resimle ilgili anahtar sözcükler girmesi isteniyor ve bu sözcüklerle eşleşen resimler veri tabanından getiriliyordu. Bu yöntem, resimler hakkında tek tek elle metinsel bilgi girilmesini gerektirdiğinden oldukça zaman alıcı ve insan gücü gerektiren etkisiz bir yöntemdi. Ayrıca resimle ilgili yalnızca birkaç anahtar sözcük eşleştirildiğinden, resmin tam anlamıyla tasvir edilmesi mümkün olmuyordu.

Listeleme ve erişim için metinsel bilginin yetersizliği, resimlerin içeriğine bağlı sistemlerin (content-based retrieval) geliştirilmesine neden olmuştur. Bu tür sistemlerde genel yaklaşım, resmin içeriğine ait renk, doku, şekil gibi bilgilerin kullanılmasıdır. Resimle birlikte, resme ait renk, doku, şekil özelliklerine ait çeşitli istatistiksel (ortalama, standart sapma, histogram) bilgiler de kaydedilir. Kullanıcıdan, aradığı resimdeki renk, doku gibi özellikleri bir paletten seçmesi istenir ve bu özellikleri taşıyan en yakın resimler veri tabanından seçilerek görüntülenir. Ya da kullanıcıdan aradığı resme benzeyen örnek bir resim girmesi istenir ve bu resme en çok benzeyen resimler veri tabanından getirilir.

Ancak, kullanıcı isteklerini inceleyen araştırmalar göstermiştir ki, aranan şey daha çok resimdeki anlamsal bilgiyle ilgilidir. Resimlerden elde edilen düşük seviyeli öznitelikler ile kullanıcıların bu resimlerden çıkardığı yüksek seviyeli anlamsal yorumlar arasında çok büyük bir uçurum vardır. Bu nedenle, önemi gittikçe artan resim ve video arşivleri üzerinde etkin bir erişim ve değerlendirme yapabilmenin ana koşulu anlamsal bilgilerin çıkarılması ve ulaşılabilir hale getirilmesidir.

On binlerce resim içeren arşivlerde etkili bir dizin yapısı geliştirmek içerik-tabanlı resim erişimi için çok önemli bir konu haline gelmiştir[2]. Bu sorun günümüzdeki veri tabanlarında henüz çözülebilmemiş değildir. Bundan daha vahim olan konu ise düşük seviyeli öznitelikler kullanarak resimleri nasıl dizinleyeceğimiz ya da anlamlı sınıflara ayırabileceğimiz sorundur. Başarılı bir uygulama, resim erişimi sırasında ilgisiz resimlerin elenmesini sağlayarak içerik-tabanlı aramalarda başarıyı oldukça iyi oranlarda etkileyecektir[3].

2. SAHNE SINIFLANDIRMASI

Son yıllarda, görüntü işleme alanında resimlerin anlamsal olarak sınıflara ayrılması konusu sahne sınıflandırması (scene classification) adı altında incelenmektedir.

Sahne sınıflandırması, görüntü işlemenin çok önemli problemlerinden biridir. Herhangi bir resim verildiğinde, ne tür bir anlamsal sahne içerdiğini tarif etmek isteriz. Ancak, bu konu üzerinde günümüze kadar çok az çalışma yapılmıştır. Bunun nedeni, problemin doğasının oldukça zor olmasının yanısıra, sahnelerin adlandırılmasının bir belirsizlik taşımasıdır. Örneğin, aynı resimler iki farklı kişiye gösterildiğinde, birinci kişi bunları yedi farklı sahne altında gruptandırırken, ikinci kişi onyediy farklı grup oluşturabilmektedir.

Sahne sınıflandırması oldukça zor bir problemdir ve nesne tanıma (object recognition) ya da resim erişimi (image retrieval) gibi konulardan oldukça farklıdır. Örneğin, nesne tanımda, resimlerde, önceden belli olan çeşitli geometrik şekillerin birbirleriyle olması gereken ilişkileri incelenirken, sahne sınıflandırması için böyle bir imkan bulunmaz. Çünkü, bir sahne çok farklı nesnelerin (araba, ev, ağaç, gökyüzü, toprak...) tahmin edilemez bir şekilde resmin üzerine yerleşiminden oluşur. Bu nedenle belirli türden bir sahne sınıfı kendi içinde çok değişik resimler içerebilir. Örneğin, bir orman sahnesi için %70'i ağaç, %30'u da gökyüzüdür gibi bir kural koyulamaz, çünkü çok farklı oranlardaki ağaç, gökyüzü ve değişik nesnelere bir orman sahnesi oluşturabilir.

2.1. Uygulama Alanları

Sahne sınıflandırması probleminin çeşitli uygulama alanları mevcuttur. Sahne sınıflandırması, görüntü işleme kapsamında gerçekleştirilen sınıflandırma, bölütleme gibi birçok görev için yardımcı olabilecek bir yöntemdir.

Sahne sınıflandırması ile resimlerden çıkarılan anlamsal bilgiler, nesne tanıma gibi diğer süreçlerde yol gösterici olabilir[4]. Örneğin, resim içinde araba nesnelerini ararken, öncelikle yol sahnelerini incelemek mantıklı olacaktır, ya da bir tost makinesi bulmak için aramaya mutfak sahnelerinden başlamak etkili ve kısa sürede sonlanan

sonuçlara götürecektir. Bu örneklerden de anlaşılacağı üzere, nesnelere ile bulunabilecekleri sahneler arasında ilişki kurmak mümkündür. Bu nedenle resimdeki sahne hakkında bilgi edinmek, o resimde ne tür nesnelere olabileceği hakkında da fikir sahibi olunmasını sağlar.

Sahne sınıflandırması, resim erişimi konusunda da oldukça faydalı olabilir[5]. Örneğin, bir plaj sahnesinin arandığı düşünüldüğünde, böyle bir resmi öncelikle açık hava resimleri içinde aramak işlemi oldukça kolaylaştırır. Resim ve video arşivlerini daha küçük sayıda alt sınıflara ayırmak ve etkili bir dizin yapısı geliştirmek, kısa zamanda erişim, arama ve tarama yapılmasını sağlayacaktır. Böylelikle akıllı veri tabanları oluşturulacaktır. Kişisel ya da kurumsal birçok resim arşivinin etkin kullanımı için sahne sınıflandırması uygulanmalıdır.

Bir diğer uygulama alanı ise oldukça ilginçtir. Öncelikle sahne hakkında bir bilgi edinmek, daha akıllı görüntü işleme tekniklerinin geliştirilmesini sağlar. Örneğin, herhangi bir film çekildikten sonra negatifleri elde edilirken, renk ayarları her kare için aynı olmak üzere otomatik olarak yapılır. Ancak eğer resmin içeriğine bağlı anlamsal bir bilgi edinilirse, resimdeki sahneye göre farklı ayarlamalar yapılabilir. Örneğin bina içindeki sahnelerle, açık hava sahnelerinin ayarlamaları farklı yapılabilir.

2.2. TRECVID Yarışması

Video erişim sistemlerinin başarısının ölçülebilmesi için son yıllarda NIST (Amerikan Standartlar Enstitüsü) tarafından TRECVID adı altında bir yarışma düzenlenmekte, bu yarışma sayesinde dünya çapında bu konuda yapılan çalışmaların başarısı karşılaştırılabilmektedir[16]. Bu yarışma kapsamında, yarışmacılara çeşitli türlerde sorgular yöneltilmektedir. Yarışmanın bir bölümü de resimlerdeki üst düzeyli anlamsal bilgilerin çıkarılmasına yönelik sorguların bulunduğu “üst düzey öznelik çıkarımı (high-level feature extraction)” bölümüdür. Yarışmacılardan, videolardaki ofis, bina, çöl, dağ, yol, eğlence, spor aktiviteleri gibi çok çeşitli sahnelerin bulunması istenmektedir. Tez kapsamında ele alınan sahne sınıfları bu yarışmadaki sorgulardan alınmıştır.

2.3. Sahne Sınıflandırmasına Genel Bakış (Tarihçe)

Martin Szummer ve Rosalind Picard 1998'de yayınladıkları çalışmalarında, alt düzey resim özniteliklerini kullanarak üst düzey sahne bilgilerini elde etmeye çalıştılar[5]. İlgilendikleri konu ise, daha çok bina içi/açık hava (indoor/outdoor) sahnelerinin sınıflandırılmasıydı. Ancak bunu yaparken resmin tümünden elde edilen öznitelikler yerine, resmin alt alanlarından ayrı ayrı elde edilen öznitelikleri sınıflandırdılar ve sahne hakkında karar verebilmek için bu sınıflandırmaları birleştirdiler.

Szummer ve Picard'ın üzerinde çalıştıkları bina içi/açık hava sahnelerinin sınıflandırılması, diğer sahne çeşitleriyle kıyaslandığında belirsizlik içermeyen bir sınıflandırmaydı. Ancak, yine de bu ayrımın tam yapılamadığı resimler mevcuttu. Örneğin, bir pencereden bakılarak çekilmiş, pencere çerçevelerinin görüldüğü resimler, yağlı boya tablo resimleri ya da çok yakından çekilmiş insan yüzü resimleri sınıflandırılması belirsiz olan resimlerdi.

Szummer ve Picard'ın çalıştıkları resim kümesi Kodak müşterileri tarafından çekilen 1343 resimden oluşuyordu. Bu resimler farklı zamanlarda, farklı kişiler tarafından çekilmiş aile ve tatil resimleriydi. Bu nedenle, resim kümesi çok farklı türden resmi birarada bulunduruyordu. Bu resimler, ekipten olmayan iki ayrı kişi tarafından incelendi ve elle %52'si bina içi ve %48'i de açık hava olarak etiketlendi. Resimlerden yalnızca 19 tanesi belirsiz olarak işaretlendi ve çalışma kapsamına alınmadı.

Szummer ve Picard, çalışmalarında renk, doku, ve frekans olmak üzere üç tür öznitelik kullandılar. Resimleri 4x4'lük alt alanlara ayırdılar ve her bir alt alan için bu öznitelikleri hesapladılar. Renk özneliği ile ilgili olarak, RGB renk uzayının bir tür dönüşümü olan Ohta renk uzayında hesaplanan histogramları kullandılar. Doku özneliği ile ilgili olarak, resimlerin siyah-beyaz gösterimlerinden hesaplanan 15-boyutlu bir öznitelik vektörü, frekans özneliği ile ilgili olarak ise düzenli olarak tekrar eden doku örüntülerini temsil eden katsayılar hesapladılar.

Resimlerin alt alanları için bu öznitelikler ayrı ayrı hesaplandıktan sonra, bunların bir şekilde birleştirilmesi gerekiyordu. Bu aşamada, çok-safhalı sınıflandırma yapmaya karar verdiler. Önce, alt alanların her birini bina içi/açık hava olarak sınıflandırdılar ve daha sonra resmin tümünü bu sınıflandırmalara dayanarak etiketlediler. Bir çoğunluk sınıflandırıcısı (majority classifier) uygulayarak, resmin bütününde en çok hangi sınıfa ait alt alan varsa, resmin kendisini de o sınıfa dahil ettiler.

Aditya Vailaya, Anil Jain ve H. Jiang Zhang, 1998'de, açık hava sahnelerinin sınıflandırılmasının daha ileri bir aşaması olan şehir/kırsal alan sahnelerinin sınıflandırılmasını gerçekleştirdikleri bir çalışma yayınladılar[3]. Bu sınıflandırmayı gerçekleştirirken, ayırt edicilik özelliği yüksek olan öznitelikleri belirlemeye çalıştılar ve belirledikleri bu öznitelikleri kullanarak etkili bir sınıflandırma yöntemi geliştirdiler.

Çalışmalarında kullandıkları resim veri tabanı, farklı kaynaklardan elde edilmiş şehir ve kırsal alan sahnelerini içeren toplam 2716 resimden oluşuyordu. Resimlerin büyüklükleri birbirinden farklıydı ve tüm resimler, tek bir kişi tarafından, elle, 1128'i şehir ve 1588 tanesi de kırsal alan olarak etiketlenmişti.

Şehir ve kırsal alan sınıflarını birbirinden en iyi ayıran özniteliği belirlemek için, çeşitli özniteliklerin sınıf-içi ve sınıflar-arası uzaklık dağılımlarını incelediler. Sınıf-içi uzaklık dağılımını belirlerken, aynı sınıftaki tüm resim çiftlerinin birbirlerine olan uzaklığını; sınıflar-arası uzaklık dağılımını belirlerken de şehir sınıfındaki her bir resim ile kırsal alan sınıfındaki her bir resim arasındaki uzaklık değerlerini hesapladılar. Sınıf-içi ve sınıflar-arası uzaklık dağılımlarının kesişme oranlarını inceleyerek, en ayırt edici özniteliği belirlediler. İnceledikleri öznitelikler, renk histogramları, renk eşvrelilik histogramları, resmin DCT (Discrete Cosine Transform) katsayıları, kenar yönelim histogramları ve kenar yönelim eşvrelilik histogramlarını içeriyordu. Yaptıkları inceleme sonunda, kenar yönelimlerine ait özniteliklerin, diğer özniteliklere oranla daha ayırt edici olduğuna karar verdiler. Bir şehir sahnesindeki insan yapımı nesnelere keskin yatay ve dikey kenarlara sahipken, kırsal alan sahneleri çeşitli yönlerdeki

rastgele kenar dağılımları içerir. Kenar yönelimlerine dayanan bir öznitelik, bu iki sınıf arasında önemli bir ayırt ediciliğe sahip olabilir.

Bir resmi, şehir ya da kırsal alan olarak sınıflandırırken, k-en yakın komşu sınıflandırıcısı kullandılar. En iyi sonuçlar ise, kenar yönelim eşvrelilik histogramı ve 5-en yakın komşu sınıflandırıcısı kullanıldığında elde edildi.

Bir sonraki aşamada ise, kırsal alan resimlerini, orman, dağ ve günbatımı&gündoğumu olarak sınıflandırmaya çalıştılar. Bu problemde ise, en ayırt edici öznitelik, renge dayalı öznitelikler olarak tespit edildi. Doğal nesnelere, belli renk dağılımlarına sahip olma eğilimindedirler. Bu nedenle, doğal sahneler içeren resimleri sınıflandırırken renk özneliği önemli ipuçları sunabilir. Vailaya ve arkadaşları da, bu bulguya dayanarak renkten yararlandılar. Üç-sınıflı bir sınıflandırıcı kullanmak yerine, öncelikle, bir resmi, günbatımı&gündoğumu ya da orman/dağ sınıflarından birine atadılar. Daha sonra da, orman/dağ sınıfı içinde yer alan resimleri, orman ya da dağ sahnesi olarak sınıflandırdılar.

Li Fei-Fei ve Pietro Perona 2005'de yayınladıkları çalışmalarında, doğal sahne sınıflarının öğrenilmesi ve tanınması için yeni bir yaklaşım geliştirdiler[6]. Geçmiş çalışmalarda yapıldığı üzere eğitim verisine el ile açıklayıcı bilgilerin eklenmesi gerekmiyordu. Bunun yerine, bir sahne, öğretmensiz öğrenme (unsupervised learning) yöntemiyle elde edilen ve kod sözcüğü (codeword) denilen resimdeki çeşitli alt alanların toplamından oluşan bir gösterimle temsil ediliyordu. Her alan, bir temayla ifade ediliyordu. Önceki çalışmalarda, bu temalar, uzmanların elle girdiği açıklamalar yardımıyla öğreniliyordu; ancak, Fei-Fei ve Perona'nın çalışmasında, temaların sahnelerdeki dağılımı ve kod sözcüklerinin temalardaki dağılımı öğretmensiz olarak belirlendi. Bu yöntemle, 13 ayrı karmaşık sahne sınıfının etiketlenmesini sağlayan bir uygulama geliştirdiler.

Uyguladıkları yöntem şu şekilde çalışıyordu; bir resim birçok alt alanın toplamı olarak modelleniyor, ve her alan, kod sözcüklerinden oluşan büyük bir sözlükte yer alan bir

kod sözcüğü ile temsil ediliyordu. Öğrenme aşamasındaki temel amaç ise bu kod sözcüklerinin her bir sınıftaki dağılımını gösteren bir model elde edebilmektir. Tanıma aşamasında ise, öncelikle etiketlenmek istenen resimdeki tüm kod sözcükleri teşhis ediliyor ve resimdeki bu kod sözcüğü dağılımına en iyi uyan sınıf modeli bulunuyordu.

Uygulanan bu yöntem, Blei ve arkadaşlarının[7], belge modellerinin öğrenilmesi ve gösterilmesinde kullandıkları yaklaşımın görüntü işleme problemine uyarlanması olarak ele alınabilir. Bu yaklaşımda, alt alanlar öncelikle farklı ara temalardan oluşan öbeklere ayrılıyor ve daha sonra da bu ara gösterimler kullanılarak sınıflara ayırma işlemi gerçekleştiriliyor. Alt alanların ve ara temaların öğrenilmesi herhangi bir insan müdahalesi olmadan otomatik olarak gerçekleştiriliyor.

Fei-Fei ve Perona'nın kullandıkları veri kümesi anayol (260 resim), şehir (308 resim), yüksek binalar (356 resim), caddeler (292 resim), banliyö yerleşimleri (241 resim), orman (328 resim), sahil (360 resim), dağ (374 resim), kırsal bölge (410 resim), yatak odası (174 resim), mutfak (151 resim), oturma odası (289 resim), ve ofis (216 resim) olmak üzere toplam 13 sınıftan oluşuyordu. Her bir resmin büyüklüğü ise ortalama 250 x 300 piksel kadardı. Hem öğrenme, hem de tanıma aşamasında yalnızca siyah-beyaz resimler kullandılar. Her bir sınıftaki resimleri rastgele olarak iki bölüme ayırdılar. Her sınıf için, 100 resim öğrenme aşamasında, kalanı ise sınamada kullanıldı. Eğitim kümesindeki resimlerden elde edilen alt alanları kullanarak kod sözcüklerinden oluşan bir sözlük oluşturdular ve her sınıf için kod sözcüğü dağılımını belirten bir model öğrendiler. Kod sözcüğü oluşturmak için çıkarılan öznitelik ise resimlerdeki kenar yönelimlerini gösteren histogram bilgileriydi. Bir sınama resmi verildiğinde, karar, en yüksek olabilirlik olasılığını veren sınıf modeline göre yapılıyordu. Sınıflandırmada toplam 40 ara tema gösterimi kullandılar. Uygulanan yöntemin ortalama başarı oranı %64 olarak gerçekleşti. En yüksek hata oranı veren sınıflar genelde ev içinde olan yatak odası, oturma odası, mutfak ve ofis sınıflarıydı. Bu dört sınıfın hepsinde de keskin yatay ve dikey kenarlar bulunduğu için, bu sınıflardaki tema ve kod sözcüğü dağılımlarının birbiriyle benzerlik gösterdiği ve bunun karışıklığa yol açtığı gözlemlendi.

Yine, 2005 yılında yayınlanan bir başka çalışma ise Quelhas, Monay ve ekibinin benzer bir yaklaşım kullanarak gerçekleştirdikleri sahne sınıflandırmasıydı[8]. Bu çalışmada, Quelhas ve ekibi görsel sahneleri modellemek için yerel değişimsiz (invariant) öznitelikler kullanıyorlardı. Bu özniteliklerle hesaplanan yerel tanımlayıcılarla oluşturulan 'görsel terim torbası(bag-of-visual-words(BOV))' gösteriminin ikili ya da çok sınıflı sahne sınıflandırmaları için başarılı sonuçlar ürettiğini gösterdiler.

Verilen bir resim için BOV oluşturma süreci birçok basamak içermekteydi. Kısaca anlatmak gerekirse; resimlerdeki ilgilenilen bölgeler otomatik olarak belirlenmekte, ve bu bölgelerde yerel tanımlayıcılar (local descriptors) hesaplanmaktaydı. Bu hesaplamalardan elde edilen sonuçlarla görsel terimler belirlenmekte ve bu terimlerin resimlerdeki yerel sıklıkları sayılarak (BOV) yapıları oluşturulmaktaydı.

İlgi noktası algılayıcısının (interest point detector) temel amacı, resimdeki geometrik ve fotometrik dönüşümlere karşı değişimsiz karakteristik noktaları otomatik olarak tespit etmektir. Quelhas ve ekibinin, bu çalışmalarında kullandıkları algılayıcı ise ölçek, aydınlatma, döndürme, çevirme gibi değişikliklere karşı değişimsiz olan Gaussianlar farkı nokta algılayıcısıydı (difference of Gaussians point detector)[18].

Belirlenen her ilgi noktasının etrafındaki bölgede yerel tanımlayıcılar hesapladılar. Bu hesaplamayı yaparken kullandıkları ise SIFT (Scale Invariant Feature Transform) olarak adlandırılan, yönelim farklarına karşı değişimsiz ve resimlerin siyah-beyaz gösterimleri üzerinde uygulanan bir tanımlayıcıydı[19]. Bu tanımlayıcının iyi bir başarı oranı gösterdiği daha önce yapılan çalışmalarda kanıtlanmıştı. SIFT tarafından hesaplanan öznitelikler, ilgi bölgesinin çeşitli kısımlarında hesaplanan kenar yönelimlerine ait yerel histogramlardı.

Tüm görsel terimlerin oluşturduğu kümeye sözlük adını verdiler. Sözlük oluşturulması öbekleme metodlarıyla gerçekleştirildi. Başka bir deyişle, eğitim resimlerinden hesaplanan yerel tanımlayıcılar k-ortalama (k-means) algoritması ile öbeklere ayrıldı

ve bu öbeklerin orta noktaları da görsel terim olarak tanımlandı. Daha sonra her bir resim için, hangi terimin ne sıklıkta yer aldığı sayılarak BOV denilen gösterimler oluşturuldu.

Çalışma kapsamında üç deney grubu düzenlenmişti: bina içi/açık hava, şehir/kırsal alan, ve bina içi/şehir/kırsal alan sınıflandırmaları üzerinde çalıştılar. Bina içi/açık hava sınıflandırması için renk özniteliklerini, şehir/kırsal alan sınıflandırması için ise kenar özniteliklerini kullandılar. Resimler renk özniteliklerine göre bina içi/açık hava olarak sınıflandırıldıktan sonra, doğru olarak sınıflandırılan açık hava resimleri de kenar histogramları kullanılarak şehir ya da kırsal alan olarak sınıflandırıldı.

2.4. Sahne Sınıflarının Uzamsal İlişkiler ile Modellenmesi

Yapılan çeşitli deneyler, insanların bir resimdeki nesnelere tespit edebilmeleri için, resme çok kısa süre bakmalarının yeterli olduğunu göstermiştir. İnsanlar, resimlerdeki nesnelere tespit etmede oldukça hızlıdır, peki aynı hızla resimdeki anlamsal sahneleri anlamakta da gösterebilirler mi? Bu doğrultuda yapılan çalışmalar, insanların resimdeki sahneyi adlandırmalarının da, nesnelere tespit etmeleri kadar hızlı olduğunu göstermiştir. Bunu yaparken de resimdeki uzamsal yerleşimin algılamada birincil payı olduğu saptanmıştır[9].

Literatürdeki sahne sınıflandırması çalışmalarında, resimlerin çeşitli alt alanlarından çıkarılan öznitelikler kullanılmış ancak bu alt alanların birbirleri ile olan uzamsal ilişkileri dikkate alınmamıştır. Bu tez kapsamında, alt alanların birbirleri ile olan komşuluk ilişkileri modellenmiş, ve bu ilişkilerin sahne yapılarının anlaşılmasında önemli ipuçları sunduğu anlaşılmıştır.

Resimler eşit büyüklükteki alt alanların toplamı olarak ifade edilmiş, daha sonra bu alt alanlardan elde edilen öznitelik vektörleri kullanılarak, alanlar doğal gruplara ayrılmıştır. Bir resimdeki alt alan çiftlerinin uzamsal ilişkileri incelenmiş ve hangi sınıfta hangi tür komşuluk ilişkilerinin bulunduğu öğrenilmeye çalışılmıştır. Öğrenilen komşuluk ilişkileri kullanılarak, verilen bir resmin hangi sahne sınıfına ait olduğu doğru

olarak bulunmaya çalışılmıştır. Bunu gerçekleştirmek için önerilen modeller, çeşitli deneyler kapsamında sınanmıştır. Sonuç olarak, bu yaklaşımın, sınıflandırma başarımını arttırdığı gözlenmiştir.

2.5. Organizasyon

Tez raporu kapsamında, 3. bölümde, genel olarak görüntü işleme konusunda kullanılan görsel özniteliklerden ve bu öznitelikleri çıkarmak için kullanılan bazı yöntemlerden bahsedilmiştir. 4. bölümde, genel olarak, öbeklendirme, Bayes karar kuramı gibi temel örüntü tanıma tekniklerine değinilmiştir, 5. bölümde ise bu tekniklerin, tez kapsamında, resimlerin modellenmesi için nasıl uygulandığı anlatılmıştır. 6. bölümde sahne sınıflandırması ile ilgili olarak gerçekleştirilen deneyler ele alınmış ve deney sonuçları üzerinde durulmuştur. 7. bölümde ise tez çalışması ile ilgili genel sonuç ve öneriler yer almaktadır.

3. GÖRSEL ÖZİNİTELİKLER

Resimlerin insanlar tarafından algılanması çok karmaşık bir süreçtir. Ancak, günümüzde şu bilgi kabul edilmiştir ki, retinadaki sinir hücreleri, resimleri hızla dış hatlar, doku ve renk gibi alt bileşenlere ayırırlar. Bu parçalanmış bilgi daha sonra beyinde tutarlı bir resim oluşturacak şekilde yeniden birleştirilir ve beyindeki bu resimden içeriğe ilişkin bilgi çıkarılır. Görüntü işleme tekniklerinde de benzer olarak resimlerden, renk, doku, dış hatlar (kenarlar) gibi alt düzey bilgiler çıkarılarak kullanılır. Bilgisayarla, resimlerden bu tip özneliklerin çıkarılması kolaylıkla gerçekleştirilebilirken, bu bilgileri kullanarak anlamsal sonuçlara ulaşmak işin zor olan kısmıdır. Çeşitli yöntemlerle resimlerden elde edilen renk, doku, şekil, kenar gibi özneliklere genel olarak görsel öznelikler denilir. İnsan görme sisteminde olduğu gibi, bu alt düzey görsel öznelikler kullanılarak üst düzey yorumlar elde edilmeye çalışılır. Resimlerin görsel öznelikleri çıkartılarak çok-boyutlu öznelik vektörleri ile ifade edilir. Renk, doku ve kenar bilgileri en çok kullanılan görsel özneliklerdir.

3.1. Renk Özneliği

Renk özneliği, görüntü işleme alanında en çok kullanılan görsel özneliktir. Rengin üç-boyutlu yapısı, bir resmin tek-boyutlu gri değerlerine oranla çok üstün bir ayırt etme gücü sağlamaktadır. Uygun bir renk tanımlaması seçmeden önce, kullanılacak renk uzayı belirlenmelidir.

3.1.1. Renk uzayları

Renkli görüntüyü ifade etmek için, saklama ortamları ve uygulamaların gereksinimlerine uygun olacak şekilde çeşitli renk uzayları kullanılmaktadır. Kullanılacak uzayın seçimi, yapılacak işleme göre değişmektedir. Bir rengi, bir uzayda temsil etmek için temel renk bileşenleri, parlaklık, renk tonu, doygunluk gibi öğeler kullanılır. RGB,HSV,CMY,YUV,YIQ en çok kullanılan renk uzaylarına örnek olarak verilebilir[20].

3.1.1.1. RGB renk uzayı

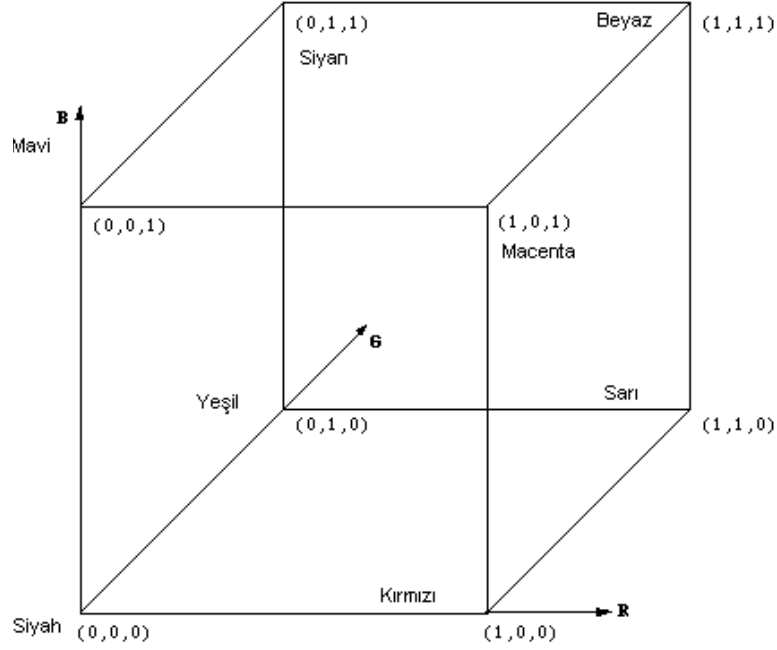
RGB renk uzayı, kırmızı, mavi ve yeşil ışığın çeşitli oranlarda birleşerek diğer renkleri meydana getirdiği toplamalı bir renk modelidir. RGB kısaltması üç temel öğe olan

kırmızı (red), yeşil (green) ve mavi (blue) rengin baş harflerinden oluşur. RGB renk uzayı serisindeki diğer tüm renkler, belli miktarlarda kırmızı, mavi ve yeşil ışığın birleşmesinden oluşur. RGB'nin toplamalı bir renk uzayı oluşunun nedeni, iki farklı frekansa sahip ışık biraraya geldiğinde bu olayı gözlemleyen kişi için bu renklerin toplamından oluşan yeni bir renk olarak algılanmasıdır. Kırmızı, mavi ve yeşil renklerin seçilme sebebi ise, her birinin insan gözündeki renge duyarlı algılayıcı türlerinden birine karşılık gelmesidir. Kırmızı, mavi ve yeşil renklerin uygun karışımlarıyla insanın algıladığı renklerin hepsini oluşturmak mümkündür. Örneğin; saf kırmızı ve yeşilin birleşiminden sarı; kırmızı ve mavinin birleşiminden macenta; mavi ve yeşilin birleşiminden siyan; ve üçünün birleşiminden de beyaz renk elde edilir. RGB renk uzayının genel kullanım alanlarından biri, bilgisayar grafiğinin ekran üstünde görüntülenmesidir. Ekran üstündeki her bir piksel, bilgisayar belleğinde kırmızı, mavi ve yeşil için ayrı değerler halinde tutulur. Her piksel, kırmızı, mavi ve yeşil bileşenlerin her biri için 8 bit olmak üzere toplam 24 bitlik bilgi ile temsil edilir. Bu sistemle yaklaşık 16.7 milyon farklı renk elde edilebilmektedir ve bu insanın ayırt edebileceği renk sayısından çok daha fazladır. Yazılı olarak, RGB değerleri, sırasıyla kırmızı, mavi ve yeşili temsil eden üç tamsayı (24 bpp, 0-255 arasında) ile ifade edilir. Örneğin;

- (0, 0, 0) siyah;
- (255, 255, 255) beyaz;
- (255, 0, 0) kırmızı;
- (0, 255, 0) yeşil;
- (0, 0, 255) mavi;
- (255, 255, 0) sarı;
- (0, 255, 255) siyan;
- (255, 0, 255) macentadır.

RGB renk uzayı genellikle bir küp ile gösterilir. Her renk (kırmızı, mavi, yeşil), üç boyutlu uzayda birbirine dik olan koordinat eksenlerinden birine atanır. Renk küpünün her ekseninde, renkler, rengin hiç bulunmadığı durumdan, tam doygun renge doğru uzanır. Renk küpü katıdır, yani küpün üzerindeki herhangi bir nokta r,g,b üçlüsüyle ifade edilebilir. Siyahtan (0, 0, 0) beyaza (1, 1, 1) doğru olan köşegen tüm

gri renkleri temsil eder, bu köşegen üzerindeki her noktada kırmızı, yeşil ve mavi bileşenleri eşit miktardadır. Kırmızı, mavi, yeşil renklerin 0-255 arasında değişmesi, üç boyutlu renk küpünün doğrusal olarak ölçeklendirilmesidir. Örnek bir renk küpü aşağıdaki şekilde görülmektedir (Şekil 3.1), bazı anahtar renk ve koordinatları küpün üstünde gösterilmiştir.

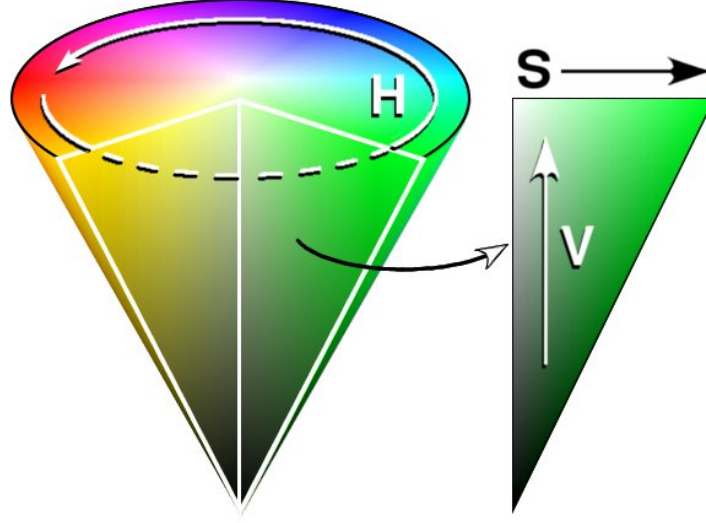


Şekil 3.1. RGB renk küpü

3.1.1.2. HSV renk uzayı

HSV (hue saturation value) renk modeli, üç bileşenden oluşan bir renk uzayı tanımlar. H(renk tonu) bileşeni, kırmızı, mavi ya da yeşil gibi temel rengin dalga boyunu belirler ve değeri 0 - 360 arasında değişir. S (doygunluk), rengin saflık ölçüsüdür ve değeri 0 – 1 arasında değişir. Bir rengin doygunluğu ne kadar düşükse, içinde o kadar grilik bulunur ve renk de o kadar solgun görünür. Doygunluk değeri 1 olan renkler tam-doygundur, 0 olan renkler ise gridir. V (değer) bileşeni ise rengin parlaklığını verir ve değeri de 0 – 1 arasında değişir. HSV modeli, 1978 yılında Alvy Ray Smith tarafından oluşturulmuştur. HSV renk uzayı, insanın algılamasına uygun olarak açısız bir koordinat sistemine sahiptir. RGB renk uzayının doğrusal olmayan bir çevrimidir. HSV modelinin gösterimi için genellikle koni kullanılır (Şekil 3.2.). Bu gösterimde, H bileşeni, renklerden oluşan bir çemberin koni biçiminde üç boyutlu

biçimlendirilmesiyle gösterilir. Doygunluk, bu koninin bir dairesel kesitinin merkezinden uzaklık, değ er ise koninin uç noktasına olan uzaklıktır.



Şekil 3.2. HSV renk uzayının koni biçiminde gösterimi

HSV modeli, Şekil 3.2'de gösterilen koniye benzer biçimde silindirik olarak da ifade edilebilir. Renk tonu, silindirin dış çevresi boyunca değişirken, doygunluk konideki benzer biçimde silindirin bir dairesel kesitinin merkezine olan uzaklıkla değişir. Değ er ise yine aşağıdan yukarıya doğru değişmektedir. Bu tür bir gösterim, matematiksel olarak HSV modelinin en doğru gösterimi kabul edilse de, insanın renk algılamasının sınırlı oluşu, konik gösterimi birçok durumda daha kullanışlı hale getirir.

RGB renk uzayında tanımlı bir renk verildiğinde ($R, G, B \in [0, 1]$), birtakım formüllerle bu rengin HSV modelindeki gösterimini bulmak mümkündür. (R, G, B) değ erlerinden en büyüğüne MAX, en küçüğüne de MIN denirse, formüller aşağıdaki gibi olacaktır (Eş. 3.1).

$$H = \begin{cases} 60 \times \frac{G - B}{MAX - MIN} + 0 & \text{eğer } MAX = R \text{ ve } G \geq B \\ 60 \times \frac{G - B}{MAX - MIN} + 360 & \text{eğer } MAX = R \text{ ve } G < B \\ 60 \times \frac{B - R}{MAX - MIN} + 120 & \text{eğer } MAX = G \\ 60 \times \frac{R - G}{MAX - MIN} + 240 & \text{eğer } MAX = B \end{cases}$$

$$S = \frac{MAX - MIN}{MAX} \quad \text{ve} \quad V = MAX \quad (3.1)$$

Elde edilen sonuçlar HSV formundadır ve H'nin değeri 0-360 derece aralığında değişen ve renk çemberi etrafında renk tonunun yerleşimini gösteren açı değeridir. S ve V değerleri ise 0 en düşük değer ve 1 en yüksek değer olmak üzere 0-1 aralığında değişir. H'nin 0-360 aralığı dışındaki değerleri 360'a göre mod alınarak bu aralık içine sokulur. Örneğin -30 değeri 330'a, 480 değeri ise 120'ye karşılık gelir.

3.2. Doku Özniteliği

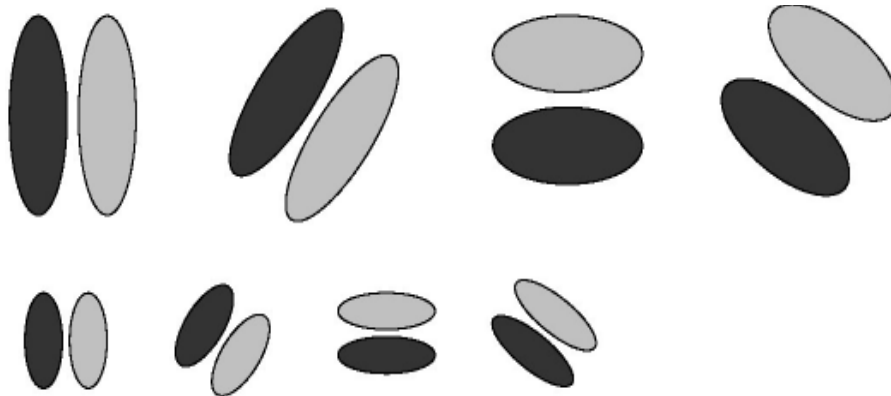
Resimlerin bir başka önemli özneliği dokudur. Doku kolayca farkedilebilen, fakat zor tarif edilen bir özneliktir (Şekil 3.3). Dokular, anlamlı şekillerde tekrar eden oldukça biçimli alt öğelerden oluşurlar. Dokuları ifade etmek için, dokuyu oluşturan alt öğeleri bulmak ve bu öğeler arasındaki istatistiksel ilişkileri ifade etmek gerekir. Alt öğeleri bulmak için çeşitli filtreler uygulanır ve bu filtrelerden alınan tepkiler yorumlanır. Uygulanan filtreler farklı yönelim ve ölçeklerdeki çubuk ve noktalardan oluşabilir. Filtrelerden alınan sonuçları yorumlamak için ortalama ve standart sapma gibi istatistiksel yöntemler uygulanır[20].



Şekil 3.3. Çeşitli doku örnekleri

3.2.1. Gabor süzgeçler

Doku özniteliğinin çıkarılmasında en çok kullanılan yöntemlerden biri Gabor süzgeçlerdir. Gabor süzgeçlerin değişik yönelim ve ölçeklerdeki çekirdekleri (Şekil 3.4) kullanılarak çeşitli özelliklere karşı değişimsiz öznitelik vektörleri elde edilebilmektedir[24]. Uygulanan ölçek ve yönelim parametrelerine bağlı olarak, herbir ölçek ve yönelim değerine ilişkin olarak, resimdeki her pikselin etrafında bulunan alana ait doku örüntülerine ilişkin istatistiksel sonuçlar üretirler. Bu istatistiksel değerler doku özniteliğini ifade eden öznitelik vektörleri olarak kullanıldığında başarılı sonuçlar verirler.



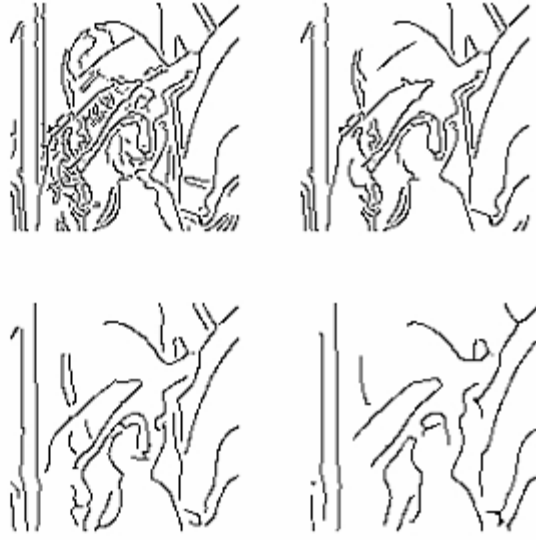
Şekil 3.4. Farklı ölçek ve yönelimlerdeki süzgeçler

3.3. Kenar Özniteliği

Bir resimdeki kenarları bulmak, veri miktarını büyük ölçüde azaltıp resim hakkında önemli ipuçları sunarken, faydasız bilginin de elenmesini sağlar. Resimde parlaklığın keskin bir şekilde değiştiği noktalar kenarları oluşturur. Kenar bulma işlemi oldukça basit hesaplamalar ile gerçekleştirilebilir. Birçok kenar bulma yöntemi iki aşamalı olarak gerçekleşir. Birinci aşamada, parlaklık değerlerinde eğimde sürekliliğin olmadığı, kenar olması muhtemel pikseller bulunur. İşaretlenen bu aday noktalar genelde kenar noktaları olarak isimlendirilir. Herhangi bir işlevde, keskin değer değişimlerinin olduğu noktalar, işlevin birinci türevi alındığında yerel maksimum noktaları olarak bulunur. Kenar bulma işleminde de piksellerin parlaklık değerine ilişkin işlevin birinci türevi alınarak yerel maksimum noktaları bulunur ve kenar noktaları olarak işaretlenir. İkinci aşamada ise, doğrular, eğriler şeklinde kenarlar oluşturabilmek için bulunan bu kenar noktaları birleştirilir.

3.3.1. Canny kenar detektörü

Canny kenar detektörü en iyi kenar bulma yöntemi olarak gösterilmektedir. Canny, kenar bulma yöntemini geliştirirken bazı prensiplerden yararlanmışır[21]. Bunlardan birincisi, resimdeki tüm kenarların bulunmasıdır. Aynı zamanda, kenar oluşturmayan yerler de yanlış bir şekilde kenar olarak işaretlenmemelidir. İkinci prensibe göre ise, kenarların yerleri doğru olarak tespit edilmelidir. Üçüncü ve de son prensibe göre de bir kenar sadece bir kere belirtilmelidir. Canny'nin bu prensipler ışığında gerçekleştirdiği kenar detektörü, oldukça başarılı sonuçlar vermektedir. Canny kenar detektörü bazı parametreleri kullanarak kenar noktaları hakkında karar verir. Bu parametreler, resimlerin gürültüden arındırılması için uygulanan filtreyle ve kenar noktalarını belirlerken kullanılan eşik değerleri ile ilgilidir. Farklı parametre değerleri uygulanarak, farklı ayrıntılarda kenar değerleri elde edilebilir(Şekil 3.5).



Şekil 3.5. Canny kenar detektöründen farklı parametreler ile alınan sonuçlar

4. İSTATİSTİKSEL ÖRÜNTÜ TANIMA YÖNTEMLERİ

4.1. Örüntü Tanıma

İnsanların etraflarında olan biten olayları kavrama ve buna göre hareket edebilmelerini sağlayan çok gelişmiş yetenekleri vardır. Örneğin, beş yaşına geldiğinde, birçok çocuk, harfleri ve rakamları tanıyabilmektedir. Küçük ya da büyük harfler, elyazısı, makine yazısı, karmaşık bir arkaplan, buruşuk bir kağıt ya da harflerin bir kısmının görünmemesi gibi özellikler bu yeteneği etkilememektedir. Makinalara da, insanlarda olan bu yetenekler kazandırılmaya çalışılmaktadır.

Örüntü tanıma, makinaların, çevreyi gözlemlene, ilgilenilen örüntüleri ayırt etmeyi öğrenme, örüntü sınıfları hakkında anlamlı kararlar verebilmelerinin nasıl sağlanacağı konusunda yapılan çalışmalardır[10]. Bu alanda yapılan yaklaşık 50 yıllık araştırmalara rağmen, genel amaçlı bir örüntü tanıma sisteminin tasarlanması hala ulaşılmaması zor bir amaç olarak varlığını sürdürmektedir.

Birçok durumda, en iyi örüntü tanıma başarımı gösterenler insanlardır, ancak, insanların bunu nasıl başardıkları hakkında kesin bir bilgi bulunmamaktadır. İnsanların karar verme süreçlerindeki en kritik noktalardan birinin örüntü tanıma olduğu bilinmektedir: Konuyla ilgili ne kadar çok örüntü kullanılırsa, verilen kararlar da o kadar iyi olacaktır[11]. Bilgisayarlarla da örüntü tanıma teknikleri uygulanabildiğinden, başarılı bilgisayar programları, bankalarda kredi başvurularının puanlanması, doktorlar için hastalıkların teşhis edilmesi ya da pilotların güvenli bir şekilde iniş yapabilmelerini sağlamak için örüntü tanıma tekniklerinden yararlanarak birçok alanda kullanılabilirler. Bu nedenle örüntü tanıma hakkında çok daha fazla dikkat harcanması ve öneminin daha iyi anlaşılması gereklidir.

Örüntülerin, otomatik olarak tanınması, betimlenmesi, sınıflandırılması ve gruplandırılması, birçok mühendislik alanında ve biyoloji, psikoloji, tıp, pazarlama, görüntü işleme, yapay zeka ve uzaktan algılama gibi birçok bilimsel disiplin için çok önemli bir problemdir. Örüntü, kargaşanın tersi olarak, bir isim verilebilen, üstü kapalı

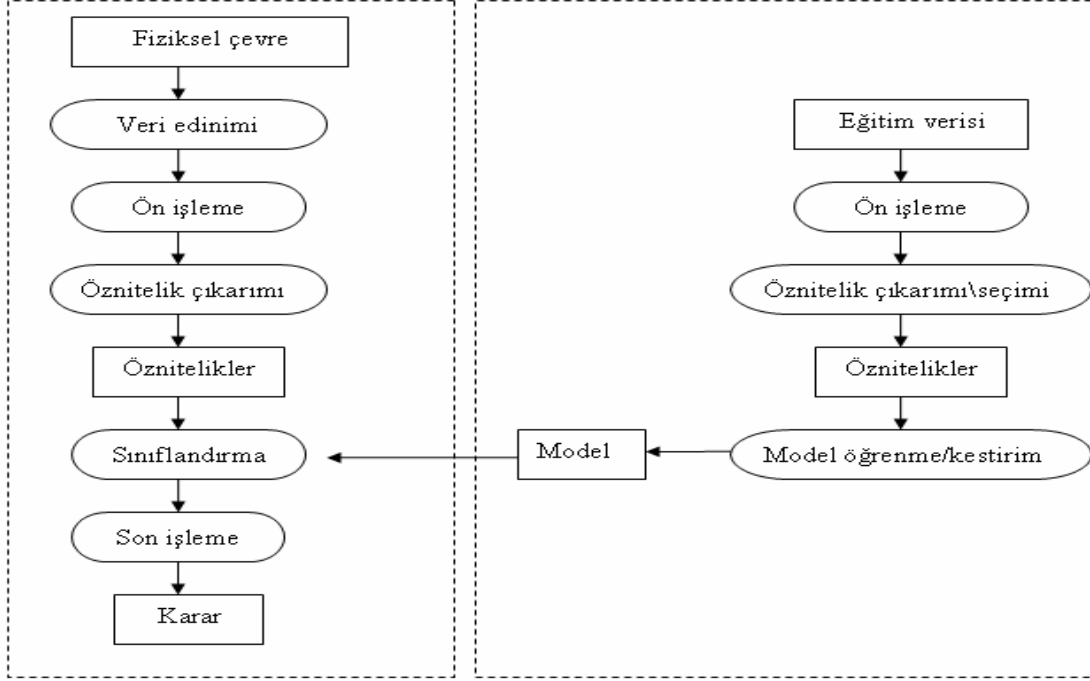
halde tanımlanmış bir varlıktır[12]. Parmakizi, elyazısıyla yazılmış bir sözcük, bir insan yüzü ya da bir konuşma sinyali örüntüye örnek olarak verilebilir. Verilen bir örüntü için, bu örüntünün tanınması / sınıflandırılması iki yöntemden biri kullanılarak gerçekleştirilebilir: 1) öğretmenle öğrenme yönteminde, eğitim kümesindeki örüntülerin sınıfları dizgeye verilir, 2) öğretmensiz öğrenme yönteminde ise örüntüler, dizge tarafından belirlenen sınıflara atanır.

Örüntü tanıma alanındaki çalışmalar, karmaşık ve zor uygulamaların ortaya çıkmasıyla yeniden canlanmıştır. Bu uygulamalar arasında, veri madenciliği, belge sınıflandırması, finansal tahmin, çoğulortam veri tabanlarının örgütlenmesi ve erişimi, yüz ve parmakizi gibi fiziksel özellikler kullanılarak şahısların tanınması anlamına gelen biometrik gibi uygulamalar sayılabilir.

Hızla büyüyen ve erişimi kolay hale gelen bilgisayar gücü, çok büyük veri kümelerinin hızlı bir şekilde işlenmesini sağlarken, veri çözümü ve sınıflandırması için çok çeşitli ve ayrıntılı yöntemlerin kullanılmasını da mümkün hale getirmiştir. Aynı zamanda, büyük veri tabanlarının kullanılmaya başlanması ve uyulması zorunlu başarımların gereksinimlerinin (hız, doğruluk, maliyet gibi) ortaya çıkması, otomatik örüntü tanıma sistemlerine olan ihtiyacı büyük oranlarda arttırmıştır. Ortaya çıkan uygulamaların çoğunluğuna bakıldığında, sınıflandırma için tek bir yaklaşımın kullanılmasının en iyi yöntem olmadığı görülmektedir ve bu nedenle birçok yöntem ve yaklaşımın aynı anda kullanılması zorunluluğu ortaya çıkmaktadır. Sonuç olarak, birçok modeli ve sınıflandırıcıyı birleştirmek, günümüzde oldukça sık kullanılan bir yaklaşımdır.

Örüntü tanıma dizgeleri çeşitli aşamalardan oluşur (Şekil 4.1). Veri edinimi, fiziksel değişkenlerin ölçülmesidir. Bu aşamada dikkat edilecek hususlar arasında band genişliği, çözünürlük, duyarlılık, bozulma, gecikme süresi gibi etkenler bulunur. Ön- işleme aşaması verinin gürültüden arındırılması, ilgilenilen örüntülerin arka plandan ayrılması gibi işlemleri içerir. Öznitelik çıkarma, özneliklere dayalı yeni gösterim biçimlerinin bulunmasına dayanır. Model öğrenme ve kestirim aşamasında ise

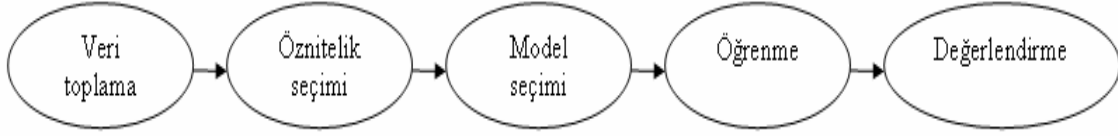
özniteliklerle örüntü grupları ve sınıflar arasında bağ kurulmaya çalışılır. Sınıflandırma aşamasında öznitelikler ve öğrenilen modellere dayanarak bir sınıfa bir örüntü eşlenir. Son-işleme aşamasında da verilen kararların güvenilirliği hesaplanır, başarımın artırılmasına çalışılır.



Şekil 4.1. Örüntü tanıma dizgeleri

Örüntü tanıma dizgelerinin tasarımı da belirli aşamalardan oluşur (Şekil 4.2). Veri toplama aşaması, eğitim ve sınama için kullanılacak verinin toplanmasıdır. Ancak, ne kadar verinin yeterli olacağını kestirmek bir sorun olarak karşımıza çıkar. Öznitelik seçim aşamasında ise bazı noktalara dikkat edilmesi gerekir. İlgilenilen alana bağımlılığı ve ön bilgi edinimi, hesaplama maliyeti ve olurluğu, ayırt edicilik özelliği (aynı örüntüler için benzer değerler, farklı örüntüler için farklı değerleri olmalı), çevirim, ölçek ve döndürme gibi değişimlerden bağımsızlığı, bozulma, ortamdaki değişimlere karşı güvenilirliği dikkate alınmalıdır. Model seçim aşamasında da dikkat edilecek hususlar vardır. İlgilenilen alana bağımlılığı ve ön bilgi edinimi, tasarım ölçütlerinin tanımları, parametrik ya da parametrik olmayan model seçimi, elde bulunmayan özniteliklerin ele alınması, hesaplama karmaşıklığı, model türleri (istatistiksel, sözdizimsel ya da yapısal) dikkat edilmesi gereken noktalardır. Eğitim

aşamasında kullanılacak iki yöntem mevcuttur: Öğretmenle öğrenme metodunda, bir öğretmen eğitim kümesindeki her bir örüntü için bir sınıf etiketi sağlar; öğretmensiz öğrenmede ise sistem girdi örüntülerine göre kendisi doğal gruplandırmalar oluşturur.

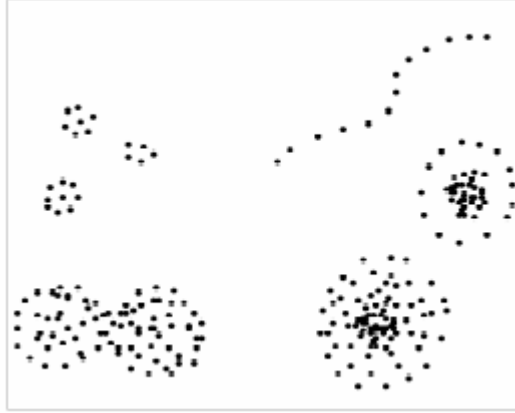


Şekil 4.2. Örüntü tanıma dizgelerinin tasarımı

Örüntü tanıma için bilinen en iyi dört yaklaşım: şablon eşleştirme (template matching), istatistiksel sınıflandırma (statistical classification), sözdizimsel ya da yapısal eşleştirme (syntactic or structural matching), ve sinir ağlarıdır (neural networks)[10].

4.2. Öbeklendirme

Öğretmensiz öğrenme de örüntü tanımadaki kullanılan yöntemlerdendir. Öğretmensiz öğrenme yönteminin kullanılmasının çeşitli sebepleri olabilir. Öncelikle çok sayıda örnek örüntü toplamak ve bunları etiketlemek maliyetli bir iştir. Verinin ait olduğu doğru sınıfın belirlenmesi de her zaman mümkün olmayabilir. Çok sayıda etiketlenmemiş veri ile eğitim yapılabilir ve bunun sonucunda oluşan gruplar kişi tarafından etiketlenebilir. Böylelikle verinin doğasının ve yapısının daha iyi kavranması sağlanabilir. Öbeklendirme de etiketlenmemiş örnekleri kullanan bir öğretmensiz öğrenme yöntemidir. Öbeklendirme oldukça zor bir problemdir, çünkü veri çok çeşitli şekillerde ve büyüklükte öbekler içerebilir (Şekil 4.3). Veri içindeki öbeklerin sayısı, çoğunlukla verinin görüldüğü çözünürlüğe bağlıdır. Bir öbek, biraraya toplanmış ya da gruplaşmış, birbirine benzeyen belli sayıda nesneden oluşur. Öbeğin ne olduğunu tanımlamak için çeşitli tanımlar ileri sürülmüştür. Bu tanımlardan birine göre aynı öbekteki örüntüler, birbirine, farklı öbeklerdeki örüntülerden daha çok benzerler. Bir diğer tanıma göre ise, öbekler, birbirlerinden, az yoğunlukta noktalar içeren kısmen boş bölgelerle ayrılan çok yoğunlukta nokta bölgeleridir. Bu işlevsel tanımlar bile kesin bir öbek tanımı yapılmasını sağlayamamaktadır. En önemli konulardan birisi de öbekleri belirleyebilmek için hem veriye hem de içeriğe bağımlı uygun bir benzerlik ölçütü seçebilmektir.



Şekil 4.3. Farklı şekil ve büyüklüklerdeki öbekler[10]

Öbeklendirme çok önemli ve kullanışlı bir yöntemdir. Hızlı, güvenilir ve tutarlı bir öbeklendirme algoritması, çok büyük miktarlarda verinin düzenlenmesini sağlayarak, veri madenciliği[13], imge bölütleme[14], sinyal işleme[15] gibi uygulamalarda kullanım alanı bulabilir. Bu nedenle, literatürde birçok öbeklendirme algoritması ileri sürülmüş ve günümüzde de sürülmeye devam etmektedir. Ancak öbeklendirme algoritmalarının büyük çoğunluğu iki yaygın tekniğe dayanmaktadır: döngülü hata kareleri bölümlenmesi (iterative squared-error partitioning) ve sıradüzensel öbeklendirme (hierarchical clustering).

Hata kareleri bölümlenmesi, sınıf-içi saçılımı en aza indirmeye ya da sınıflar-arası saçılımı en yükseğe çıkarmaya çalışır. En iyi çözümü garantilemek için, d-boyutlu n tane örüntünün tüm k (verilen bir k değeri için) öbeklik olasılıklarının incelenmesi gerekir. Ancak, bu yaklaşımın hesaplama maliyetinin çok büyük olması nedeniyle, maliyeti azaltıcı yöntemler uygulanır, bu durumda da en iyi çözüm garanti edilemez.

Hata kareleri bölümlenmesi için genel bir algoritma şu şekildedir:

1. Başlangıçta k öbekten oluşan bir ilk bölümlenme yapılır ve 2'den 5'e kadar olan basamaklar, öbeklerdeki elemanlar sağlaştıkana kadar uygulanır.
2. Her bir örüntüyü en yakınında hangi öbek merkezi bulunuyorsa o öbeğe atayarak yeni bir bölümlenme elde edilir.
3. Yeni öbek merkezleri hesaplanır.

4. Ölçüt olarak kullanılan işlevin en iyi değeri (örneğin bir yerel minimum noktası bulununca ya da önceden belirlenen sayıda tekrar tamamlanınca) bulunana kadar 2 ve 3 numaralı basamaklar tekrar edilir.
5. Öbek sayısı, bazı öbekler ayrılarak, bazı öbekler birleştirilerek, küçük ya da aykırı öbekler ayıklanarak ayarlanmaya çalışılır.

Son basamak dışında kalan diğer basamaklar k-ortalama algoritması olarak bilinmektedir. k-ortalama algoritması çeşitli koşullar bulunduğunda, hızlı hesaplanan ve iyi sonuçlar veren bir algoritmadır. Ancak k' yı seçmek ve başlangıçta ilk bölümlenmeyi yapmak konusundaki belirsizlikler bu algoritmanın dezavantajlarıdır. k , genellikle sezgisel olarak ya da veriye ait ilk bilgilere dayanılarak seçilir. İlk bölümlenme ise veri içinde düzenli olarak dağılmış rastgele k tane nokta seçilerek ya da sadece rastgele k tane nokta seçilerek yapılır.

Bazı uygulamalarda, belirli bir düzeyden bakıldığında bazı örüntü gruplarının benzer karakteristik özellikler taşıdığı görülür. Sıradüzensel öbeklendirme, sıradüzensel gösterimlere dayanarak, bu çok-düzeyleli gruplaşmaları tespit etmeye çalışmaktadır. Bir ağaç yapısı, sıradüzensel öbeklendirme algoritmalarının sonuçlarını göstermek için kullanılabilir.

4.3. Bayes Karar Kuralı

Bayes karar kuralı, çeşitli kararlar arasındaki farkları, bu kararlara eşlik eden olasılıkları ve maliyetleri kullanarak ölçen istatistiksel bir yaklaşımdır[22].

Bir gözlemcinin, yürüyen bir bant üzerinde ilerleyen meyveleri sınıflandırmak gibi bir problemi olduğu varsayalım. Bant üzerinde, elma ve portakal olmak üzere iki çeşit meyve bulunsun. Gözlemlenen(state of nature) meyve türünü ifade etmek üzere rastgele bir ω değişkeni tanımlanırsa; elma için $\omega=\omega_1$ ve portakal için de $\omega=\omega_2$ olur. Bu durumda $P(\omega_1)$ bir sonraki meyvenin elma olması için ve $P(\omega_2)$ de bir sonraki meyvenin portakal olması için önsel olasılıklar (prior probability) olacaklardır. Önsel olasılıklar, meyveyi gerçekten görmeden önce o meyvenin görülebilme olasılığı hakkında ne kadar bilgi olduğunu yansıtan değerlerdir. Önsel olasılıkları belirlerken

kesin kurallar koymak zordur, sezgisel davranılır. Örneğin meyvelerin görülebilme olasılıkları eşitse $P(\omega_1) = P(\omega_2)$ olarak atanır. Ya da bulunulan bölge ya da yılın hangi zamanında bulunduğu gibi değişkenlere bakılarak farklı değerler atanır. Başka meyve çeşidi olmadığı için $P(\omega_1) + P(\omega_2) = 1$ olacaktır. Birinin, meyvelerin %80'inin elma, geri kalanının da portakal olduğunu söylediği varsayılırsa, sadece önsel olasılıkları kullanarak karar verme şu şekilde yapılabilir.

$$\text{Karar} \begin{cases} \omega_1 & \text{eğer } P(\omega_1) > P(\omega_2) \\ \omega_2 & \text{diğer durumlar} \end{cases}$$

Bu karara ilişkin hata olasılığı ise $P(\text{hata}) = \min \{ P(\omega_1), P(\omega_2) \}$ olacaktır.

Bu aşamada, renk ölçümleri (x) kullanılarak karar geliştirilebilir. x rastgele sürekli değişken olsun, bu durumda $p(x | \omega_j)$ sınıf-koşullu olasılık yoğunluğu olarak tanımlanır ($j = 1,2$ için gözlemlenen değer ω_j olarak verildiğinde x 'in olma olasılığı). $p(x | \omega_1)$ ve $p(x | \omega_2)$ ise elma ve portakal topluluklarının renk dağılımını ve bu topluluklar arasındaki renk farkını tanımlar. $j = 1,2$ için $P(\omega_j)$ ve $p(x | \omega_j)$ bilindiği ve bir meyvenin renginin x olarak ölçüldüğü varsayıldığında, sonsal olasılık (posterior probability) olarak tanımlanan $P(\omega_j | x)$ (öznitelik değerinin x olarak ölçüldüğü verildiğinde gözlemlenen değer ω_j olma olasılığı) değeri hesaplanabilir. Önsel olasılığı, sonsal olasılığa çevirmek için Bayes formülü kullanılabilir (Eş. 4.1).

$$P(\omega_j | x) = \frac{p(x | \omega_j)P(\omega_j)}{p(x)} \quad (4.1)$$

Bu formülde $p(x)$ (Eş. 4.2);

$$p(x) = \sum_{j=1}^2 p(x | \omega_j)P(\omega_j) \quad (4.2)$$

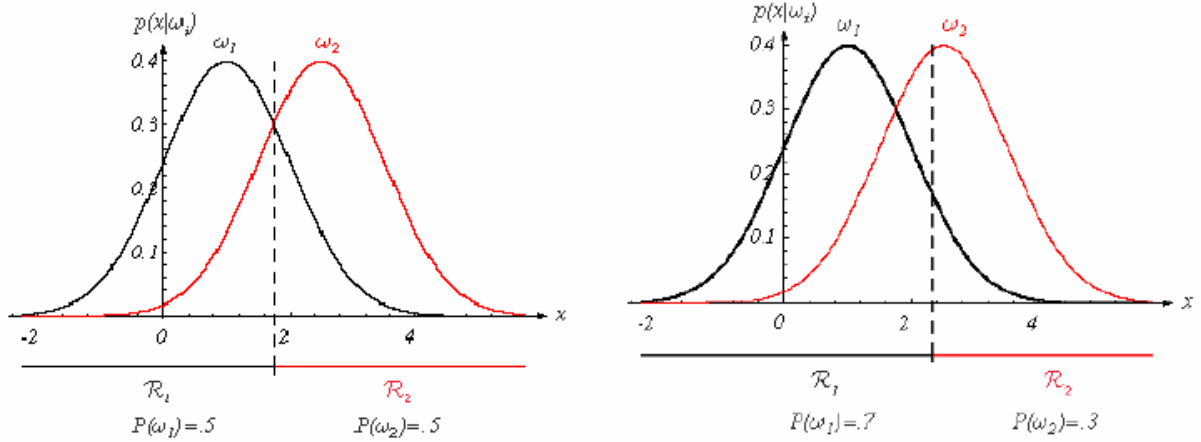
şeklinde hesaplanır. $p(x | \omega_j)$ olabilirlik, $p(x)$ ise kanıt olarak adlandırılır. x 'in değerini gözlemledikten sonra şu şekilde karar verilebilir:

$$\text{Karar} \begin{cases} \omega_1 & \text{eğer } P(\omega_1 | x) > P(\omega_2 | x) \\ \omega_2 & \text{diğer durumlarda} \end{cases}$$

Formülü yeniden düzenlersek:

$$\text{Karar} \begin{cases} \omega_1 & \text{eğer } \frac{p(x | \omega_1)}{p(x | \omega_2)} > \frac{P(\omega_2)}{P(\omega_1)} \\ \omega_2 & \text{diğer durumlarda} \end{cases}$$

Örnelemek gerekirse farklı önsel olasılıklar için en iyi eşik değeri aşağıdaki gibi olacaktır (Şekil 4.4).



Şekil 4.4. Farklı önsel olasılıklar için eşik değerleri[22]

Verilen karardaki hata olasılığı ise şu şekildedir:

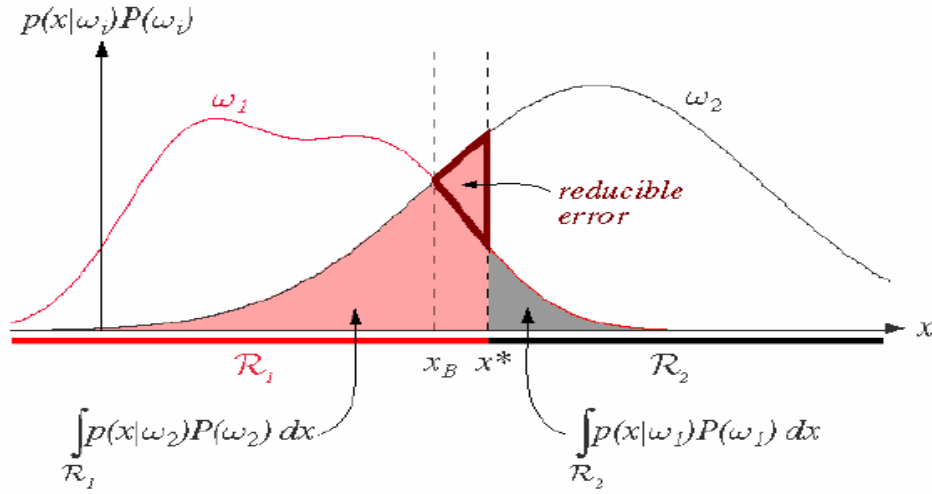
$$P(\text{hata} | x) = \begin{cases} P(\omega_1 | x) & \text{eğer karar } \omega_2 \text{ ise} \\ P(\omega_2 | x) & \text{eğer karar } \omega_1 \text{ ise} \end{cases}$$

Ortalama hata olasılığı ise (Eş. 4.3):

$$P(\text{hata}) = \int_{-\infty}^{\infty} p(\text{hata}, x) dx = \int_{-\infty}^{\infty} P(\text{hata} | x) p(x) dx \quad (4.3)$$

şeklinde hesaplanır. Bayes karar kuralı bu hatayı en aza indirir (Eş. 4.4).

$$P(\text{hata} | x) = \min\{ P(\omega_1 | x), P(\omega_2 | x) \} \quad (4.4)$$



Şekil 4.5. Bayes hata oranı[22]

Yukarıdaki şekilde (Şekil 4.5) eşit önsel olasılıklar ve en iyi olmayan x^* karar noktası için hata olasılığı bileşenleri görülmektedir. En iyi karar noktası olan x_B ise toplam taralı alanı azaltır ve Bayes hata oranını verir.

Bayes karar kuralını birden fazla öznitelik için genelleştirmek gerektiğinde sayıl x yerine öznitelik vektörü x konur. İki den fazla gözlenen durum olması halinde ise gösterimdeki bir değişiklik yeterlidir. Sadece karar verme eylemi değil, reddetme eylemi de eklenebilir. Verilen karardaki farklı riskleri hesaplamak için de her eylemin maliyeti belirlenmelidir.

$\{ \omega_1, \dots, \omega_c \}$ olarak tanımlanmış c adet gözlenen durum yani sınıfın olduğu ve x 'in öznitelik vektörü olarak adlandırılan d -bileşenli rastgele değişken olduğu varsayılırsa, tüm hatalar aynı maliyetli olduklarında en küçük-hata karar kuralı şöyle tanımlanır:

$$\text{Karar } \omega_i \text{ eğer } P(\omega_i | x) > P(\omega_j | x) \quad \forall j \neq i$$

Sonuçta elde edilen hata oranı, Bayes hata oranıdır ve elde edilebilecek en iyi sonuçtur. Olasılıkların gerçek değerlerinin bilindiği varsayıldığında, Bayes karar kuralı en iyi karar kuralı olarak ortaya çıkar. Bilinmeyen $p(x | \omega_j)$, $j=1, \dots, c$ değerine yönelik olarak gerçeğe uygun bir kestirim yapmak için parametrik ya da parametrik olmayan modeller kullanılabilir. Parametrik modeller, yoğunluk işlevlerinin biçiminin bilindiği varsayımına dayanır. Bunlar; yoğunluk modelleri (örn. Gauss), karışım modelleri, (örn. mixture of Gauss), Saklı Markov modelleri ve Bayes güven ağlarıdır. Parametrik olmayan modellerin ise biçime ilişkin bir varsayımı yoktur. Parametrik olmayan modeller histogram-tabanlı kestirim ve en yakın komşu kestirimi gibi modellerdir.

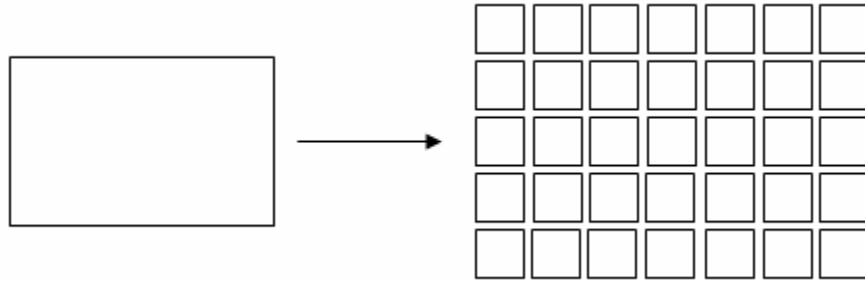
Bayes karar kuralı, önsel olasılıklar $P(\omega_i)$ ve sınıf-koşullu yoğunluklar $p(x | \omega_i)$ bilindiği takdirde, en iyi sınıflandırıcının nasıl tasarlanması gerektiğini gösterir. Ancak, genelde olasılıksal yapı hakkında tam bir bilgi bulunmaz. Buna rağmen, sınıflandırılmak istenen örüntülerin temsillerini içeren tasarım örnekleri ya da eğitici veriler bulunabilir.

5. BAYES TABANLI SAHNE SINIFLANDIRMASI

5.1. Resimlerin Modellenmesi

Resmin bütününden çıkarılan öznitelik değerleri, resmin geneli hakkında alt düzey bilgiler edinilmesini sağlar, ancak, resmin çeşitli alanlarında yer alan ayrıntılar hakkında bilgi taşımaz. Halbuki, resmin içeriği hakkında anlamsal bilgiler elde edebilmek için resmin çeşitli alanlarından ayrı ayrı çıkarılan öznitelik değerleri ele alınıp yorumlanmalı ve hatta bu alanların birbirleriyle olan ilişkileri modellenmelidir.

Tez kapsamında, resimlerdeki ayrıntıların daha iyi ifade edilebilmesi için, resimlerin belli sayıda alt alanlara bölünmesi gerçekleştirilmiştir. Alt alanlara bölme işlemi için, her resme, a sütun sayısı ve b satır sayısı olmak üzere, axb boyutlu ızgaraya (grid) bölme işlemi uygulanmıştır (Şekil 5.1). Sonuçta meydana gelen m tane ızgara hücresinden herbiri resmin bir alt alanını oluşturmaktadır. Resimler modellenirken, resimlerin ızgaraya bölünmesi sonucu elde edilen eşit büyüklükteki bu alt alanlardan çıkarılan öznitelik vektörleri kullanılmıştır.

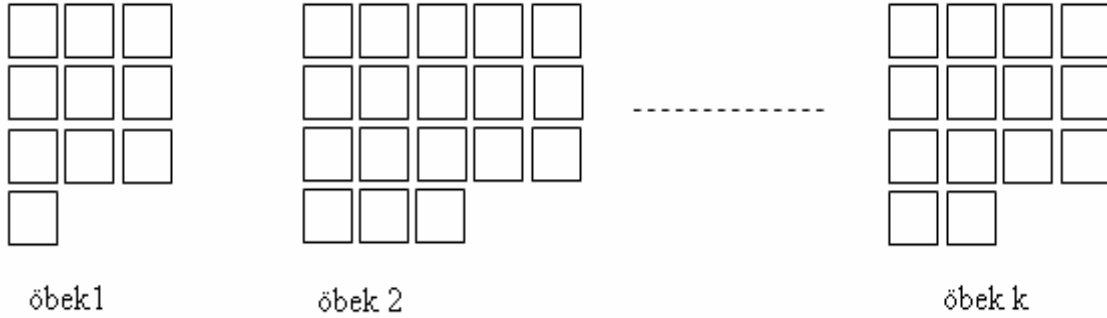


Şekil 5.1. Resmin alt alanlara bölünmesi

Bir resimde yer alan her bir alt alan için, ayrı ayrı çeşitli öznitelik değerleri hesaplanmıştır. Renk özneliği ile ilgili olarak, RGB ve HSV renk uzaylarında, ortalama ve standart sapma gibi çeşitli istatistiksel değerler ile histogramlar hesaplanmıştır. Doku özneliği için, Gabor süzgeçler kullanılarak çeşitli ölçek ve yönelimlerdeki doku örüntülerini temsil eden istatistiksel değerler elde edilmiştir. Kenar özneliği ile ilgili olarak ise, Canny kenar detektörü ile elde edilen kenar

yönelimlerine ait histogramlar hesaplanmıştır. Bu öznitelik değerleri, tüm resimlerde yer alan tüm alt alanlar için ayrı ayrı hesaplanmıştır.

Alt alanlardan elde edilen bu öznitelik vektörleri ile, daha önce bölüm 4.2.'de bahsedilen öbeklendirme yöntemleri kullanılarak, öbeklendirme işlemi gerçekleştirilmiştir. Böylelikle, benzer öznitelik değerlerine sahip olan alt alanların biraraya gelmesi sağlanarak, doğal gruplandırmalar oluşturulmuştur (Şekil 5.2). Benzer alt alanların biraraya toplanması ile oluşan bu öbekler, su, gökyüzü, ağaç, bina gibi çeşitli anlamsal öğeleri temsil etmektedirler. Örneğin, d numaralı öbek, çeşitli resimlerde gökyüzünü oluşturan alt alanların biraraya toplanması ile oluşmuş olabilir. Diğer öbekler de benzer şekilde, aynı anlamsal alanları oluşturan alt alanların gruplanması ile oluşur. Buna karşılık, herhangi bir anlam ifade etmeyen alt alanlar da, benzer öznitelik değerlerinden dolayı belli bir öbeğin içinde yer alabilirler. Ortaya çıkan öbeklerin sayısına k denirse; her alt alan bu k öbekten birine atanmıştır ve bundan sonraki bölümlerde, alt alanlar atandıkları öbek numarasını temsil eden etiketlerle modellenmişlerdir.



Şekil 5.2. Alt alanların öbeklere ayrılması

Alt alanlar öbeklere ayrıldıktan sonra, bir resimde yer alan alt alanların hangi öbeklere atandığı saptanmıştır. Resimdeki her alt alan atandığı öbek numarasıyla eşleştirilmiş, böylelikle her resim, hangi öbeklerden eleman içerdiğine dair bilginin modellenmesini sağlayan, etiketlenmiş alt alanların toplamı olarak ifade edilmiştir(Şekil 5.3).

34	16	81	3	6	25	41
34	34	2	3	25	16	16
12	21	2	65	65	4	19
12	12	38	5	46	10	11
12	38	38	5	22	35	40

resim 1

21	21	45	7	7	7	45
10	45	45	10	10	7	38
52	52	4	52	53	10	9
15	33	26	25	17	9	9
15	15	12	41	56	56	63

resim N

Şekil 5.3. Resimlerin alt alanların toplamı olarak modellenmesi

Gerçekleştirilen tez çalışmasında amaç, alt alanlarla modellenen her bir resmi $\{\omega_1, \dots, \omega_c\}$ olarak gösterilen c tane sahne sınıfından birine atamaktır. Bu amacı gerçekleştirmek için tez kapsamında önerilen modele göre, alanların resimdeki dağılımı üç şekilde modellenmektedir.

5.2. Tek Alt Alanlı Konumdan Bağımsız Model

Bu modelde, her bir alt alan, resimde bulunduğu konumdan bağımsız olarak modellenir. Örneğin, 5 numaralı öbeğe atanmış bir alt alanın resimde hangi konumda bulunduğu, yani solda üst köşede, ortada ya da sağ alt köşede bulunması, bu model için farketmez. Sonuçta, her resim içerdiği alt alanların konumdan bağımsız listesi ile modellenir. m tane alt alana sahip bir resimde, bu alt alanlar, $\{x_1, \dots, x_m\}$ olarak gösterilir(Şekil 5.4).

x_1	x_2	x_3
...
...
...
...	x_m

Şekil 5.4. Alt alanların gösterimi

Alt alanlarının öbek ataması belli olan bir resmin j . sınıfa ait olma olasılığı;

$$p(\omega_j | x_1, \dots, x_m) \quad (5.1)$$

olarak bulunur (Eş. 5.1). Bayes karar kuramına göre bu olasılık şu şekilde hesaplanır (Eş. 5.2):

$$p(\omega_j | x_1, \dots, x_m) = \frac{p(\omega_j)p(x_1, \dots, x_m | \omega_j)}{p(x_1, \dots, x_m)} \quad (5.2)$$

$p(\omega_j)$, j . sınıf için önsel olasılığı, $p(x_1, \dots, x_m | \omega_j)$ ise sınıf-koşullu olasılığı gösterir. Bu değerler bilindiği takdirde, bir resmin hangi sınıfa ait olduğunu hesaplamak için sonsal olasılıklara bakılır (Eş. 5.3).

$$j^* = \arg \max_{j=1, \dots, c} p(\omega_j | x_1, \dots, x_m) \quad (5.3)$$

$p(x_1, \dots, x_m | \omega_j)$ olasılığı, k^m tane olası durum içerir. Bu kadar fazla sayıda olasılık içeren bir değer sınırlı sayıda örneklerden öğrenilmesi mümkün değildir. Böyle durumlarda, modeli basitleştirmek için en çok kullanılan yöntem, x 'lerin birbirinden bağımsız olduğunu varsayan Naive Bayes varsayıdır. Bu durumda, $p(x_1, \dots, x_m | \omega_j)$ sınıf-koşullu olasılığı, birbirinden bağımsız olduğu varsayılan $p(x_i | \omega_j), i \in \{1, \dots, m\}$, olasılıklarının birbirleriyle çarpımı sonucu hesaplanır (Eş. 5.4).

$$p(x_1, \dots, x_m | \omega_j) = \prod_{i=1}^m p(x_i | \omega_j) \quad (5.4)$$

Yukarıdaki eşitlikte, bilinmeyen her bir $p_{iu} = p(x_i = u | \omega_j), u \in \{1, \dots, k\}$, olasılığı çok-terimli dağılımdan (multinomial distribution) yararlanılarak bulunabilir. (Gösterimi

basitleştirmek için j kullanılmamıştır. Normalde, p_{iu} değerleri, her bir sınıf için, $j \in \{1, \dots, c\}$, ayrı ayrı öğrenilir). Çok-terimli dağılımı belirleyen çeşitli özellikler vardır:

- deney, n defa tekrarlanan denemelerden oluşur,
- her deneme, belli sayıda, olması muhtemel sonuçlardan biri ile sonuçlanır,
- herhangi bir denemede belli bir sonucun ortaya çıkma olasılığı sabittir,
- denemeler birbirinden bağımsızdır.

Tez kapsamında gerçekleştirilen, resimlerin modellenmesi, çok-terimli dağılım ile birebir örtüşmektedir. Model çerçevesinde, bir sınıfa ait n tane birbirinden bağımsız resim bulunur ve bu resimlerde $x_i = u$ alt alanının bulunma olasılığının iki tane muhtemel sonucu vardır (bu alt alan resimde bulunur ya da bulunmaz). Konumdan bağımsız olunca, model i 'den bağımsız olur ve bu durumda her bir sınıf için k parametrelili çok-terimli dağılım modeli öğrenilir. Öğrenilen bu modelin parametreleri p_u , $u \in \{1, \dots, k\}$, olmak üzere k tanedir.

- p_u , u . sonucun ortaya çıkma olasılığı, $\sum_{u=1}^k p_u = 1$
- n_u , u . sonuçla sonuçlanan deneme sayısı, $\sum_{u=1}^k n_u = n$
- p_u için en büyük olabilirlik kestirimi (maximum likelihood estimate)

$$p_u = \frac{n_u}{n} \quad (5.5)$$

olacaktır (Eş. 5.5). Yukarıdaki eşitlikte, n_u , bir sınıfa ait örnek resimlerde u . öbeğe atanan alt alan sayısı, n ise o sınıfa ait tüm örnek resimlerden elde edilen alt alan sayısını ifade etmektedir. Bu parametreler, örneklerden öğrenildikten sonra, x_1, \dots, x_m alt alanlarına sahip olduğu bilinen bir resmin hangi sınıfa ait olduğuna karar vermek için, her bir sınıf ve her bir x_i parametresi için hesaplanan p_u değeri alınır ve sınıflandırma en yüksek sonsal olasılığı veren sınıf lehine yapılır (Eş. 5.6).

$$j^* = \arg \max_{j=1, \dots, c} p(\omega_j) \prod_{i=1}^m p(x_i | \omega_j) \quad (5.6)$$

5.3. Tek Alt Alanlı Konuma Bağımlı Model

Model konuma bağımlı olunca, bir önceki bölümde anlatılan modelden farklı olarak, her i . konum için k parametrelı ayrı bir çok-terimli dağılım modeli öğrenilir. Bu durumda, öğrenilen modelin parametre sayısı, toplam $m \cdot k$ tane olacaktır.

$p_{iu} = p(x_i = u | \omega_j), i \in \{1, \dots, m\}, u \in \{1, \dots, k\}$, olasılıkları için en büyük olabilirlik kestirimi (Eş. 5.7).

$$p_{iu} = \frac{n_{iu}}{n_i} \quad (5.7)$$

şeklinde hesaplanır. Yukarıdaki eşitlikte, n_{iu} , resimde i . konumda bulunup, u . öbeğe atanan alt alan sayısı, n_i ise, $n_i = \sum_{u=1}^k n_{iu}$ olmak üzere, i . konumda bulunan tüm alt alanların sayısını ifade etmektedir (bir önceki bölümde (bölüm 5.2) olduğu gibi burada da parametreler her bir sınıf için ayrı ayrı öğrenilir). Sınıflandırma yapmak için ise, konumdan bağımsız modelde olduğu üzere, x_1, \dots, x_m alt alanlarına sahip bir resim verildiğinde, her bir sınıf ve her bir x_i için hesaplanan p_{iu} değeri alınıp, sonsal olasılıklar hesaplanır ve en büyük sonsal olasılığı veren sınıf seçilir.

5.4. Çift Alt Alanlı Görelı Konumlu Model

Bu modelde, resimler alt alan çiftleri ile modellenir. Alt alan çiftlerinin görelı konumları bu alanların alt-üst ilişkisi ile modellenir (Şekil 5.5). Bu ilişkinin seçilmesinin sebebi ise alt alanların karşılık gelebileceği anlamsal öğelerin (örneğin, gökyüzü, ağaç, bina,...) değişik sahne sınıfları için, gerçek dünyada da bu türlü bir ilişki içinde olmalarıdır.

x_1	x_2	x_3
...
...	x_r
...	x_s
...	x_m

Şekil 5.5. Alt alan çiftlerinin gösterimi

Bu modelde, $p(x_1, \dots, x_m | \omega_j) = \prod_{(r,s) \in \Omega} p(x_{rs} | \omega_j)$ sınıf-koşullu olasılığını öğrenmek için, Ω

alt alan çiftlerini temsil eden bir küme olarak tanımlanırsa; $\Omega = \{(r, s) | s = r + a, 1 \leq r \leq m - a, 1 + a \leq s \leq m\}$, a , alt alanlara ayrılan resim için sütun sayısı olmak üzere, toplam k^2 parametrelili çok-terimli dağılım modeli öğrenilir.

Bilinmeyen $p_{uv} = p(x_r = u, x_s = v | \omega_j) = p(x_{rs} = (u, v) | \omega_j), u \in \{1, \dots, k\}, v \in \{1, \dots, k\}$ olasılıkları için en büyük olabilirlik kestirimi (Eş. 5.8);

$$p_{uv} = \frac{n_{uv}}{n} \quad (5.8)$$

olur. Bu eşitlikte, n_{uv} , birincisi u , ikincisi v olan öbek atamalarına sahip olan çiftlerinin sayısını, n ise $n = \sum_{u,v} n_{uv}$ olmak üzere tüm alan çiftlerinin sayısını ifade etmektedir.

Öbek atamaları belli olan alt alan çiftlerini içeren bir resmin hangi sahne sınıfına ait olduğunu bulmak için de benzer şekilde, her bir sınıf ve her bir x_{rs} değeri için bulunan olasılıklar kullanılarak hesaplanan sonsal olasılıklara bakılır. En büyük sonsal olasılığı veren sınıf, resmin ait olduğu sınıf olarak seçilir.

5.5. Önemli Parametrelerin Seçilmesi

Sonuçta, tek alt alanlı konumdan bağımsız model k parametrelili, tek alt alanlı konuma bağımlı model $k \cdot m$ parametrelili, çift alt alanlı görelili konumlu model ise k^2

parametrelidir. k deęerinin büyüklüęü arttıkça, parametre sayısı da buna baęlı olarak çok büyük oranlarda artmaktadır. Modelin öğrenilmesinde kullanılan sınırlı sayıda örnek olması nedeniyle çoęu parametrenin deęeri sıfır olarak bulunabilir (5 numaralı eşitlikteki n_u deęeri, 7 numaralı eşitlikteki n_{iu} deęeri ve 9 numaralı eşitlikteki n_{uv} deęeri). Bu deęerlerin sıfır olması, çarpımların sıfır olmasına ve bunun sonucunda da sonsal olasılıkların sıfır çıkmasına neden olur. Bunu önlemek için çok sayıdaki parametre arasından önemli olan alt alanlara karşılık gelenlerin seçilmesi gerekir.

Önemli bir alt alan, belirli bir sınıfa ait örneklerde tutarlı bir şekilde sıklıkla bulunan, buna karşılık dięer sınıflarda nadiren bulunan alt alandır. Bu tür alt alanlar bir sınıfı dięerlerinden ayırt etmek için kullanılabilir. Bu alt alanların belirlenmesi için ise hipotez sınaması (hypothesis testing) gibi istatistiksel yöntemlerden yararlanılabilir.

İki veri kümesi verildiğinde, bu iki kümenin aynı dağılımdan mı yoksa farklı dağılımlardan mı geldięi incelenebilir. Yani, iki veri kümesinin aynı dağılımdan geldiğini öne süren bir sıfır hipotezi (null hypothesis) oluşturulup, bu hipotez çürütülmeye çalışılabilir [17]. Bu hipotezi çürütmek, hipotezin karşıt hipotezi olan iki veri kümesinin farklı dağılımlardan geldięi hipotezini kanıtlar. Ancak, sıfır hipotezi çürütülemezse, bu durum iki veri kümesinin aynı dağılımdan geldiğini kanıtlamaz. Bu durumda, yalnızca, iki veri kümesinin tutarlı olduęu söylenebilir. Çünkü, iki veri kümesinin aynı dağılımdan geldiğini ispatlamak mümkün deęildir. Bunun nedeni ise, uygulamada, bunu ispatlayabilecek miktarlarda veri kullanabilmenin olanaksız olmasıdır.

İki tane gözeli (binned) dağılımın birbirlerinden farklı olup olmadığını sınamak için genel olarak uygulanan test , Ki-Kare testidir (Chi-Square test). Sıfır hipotezi, iki dağılımın aynı olduęunu, karşıt hipotez ise iki dağılımın farklı olduęunu öne sürer. Bu hipotezleri sınamak için uygulanacak Ki-Kare istatistięi, R_i , birinci dağılımın i . gözesindeki eleman sayısı ve S_i de ikinci dağılımın i . gözesindeki eleman sayısı ve N_G de göze sayısı olmak üzere;

$$\chi^2 = \sum_{i=1}^{N_G} \frac{(R_i - S_i)^2}{R_i + S_i} \quad (5.9)$$

şeklinde hesaplanır (Eş. 5.9). İki dağılım için, toplam eleman sayılarının eşit olduğu durumda, serbestlik derecesi, göze sayısının bir eksiğine eşit olur ($N_G - 1$). Serbestlik derecesine göre Ki-Kare tablosundan kritik değer bulunur. Eğer hesaplanan Ki-Kare istatistiğinin değeri, kritik değerden büyük çıkarsa, sıfır hipotezi reddedilir. Yani, iki dağılımın farklı olduğu ispatlanmış olur.

Tez kapsamındaki modeller için, herhangi bir alt alanın, önemli olup olmadığına karar verebilmek için Ki-Kare testinden yararlanılmıştır. Tek alt alanlı konumdan bağımsız modelde, öbeklerin sınıflardaki dağılımı, tek alt alanlı konuma bağımlı modelde, resimdeki konuma göre sınıflardaki öbek dağılımı ve çift alt alanlı görelî konumlu modelde, alan çiftlerinin sınıflardaki dağılımı Ki-Kare testi ile test edilmiştir. Bu dağılımlar, her bir sınıf için, o sınıfa ait örneklerde her bir öbeğe ait alt alanların görülme sayısının histogramı kullanılarak kestirilmiştir. Örneğin, birinci modelde, herhangi bir öbeğin dağılımı, sınıflar arasında önemli farklılıklar içeriyorsa, bu öbek önemli olarak belirlenmiştir. Bu farklılıklar bulunurken, Ki-Kare testi uygulanarak, öbeğin bir sınıftaki dağılımı, diğer sınıflardaki dağılımlarıyla karşılaştırılmıştır. Diğer modellerde de, benzer şekilde, alt alanların ve alan çiftlerinin dağılımları, sınıflar arasında karşılaştırılarak, farklı dağılıma sahip olan alt alanlar belirlenmiştir.

6. DENEYLER

6.1. Veri Kümesi

Tez çalışması kapsamında kullanılan veri kümesi, kar, bitki örtüsü, su, gökyüzü, dağ ve şehir olmak üzere toplam altı tane sahne sınıfından oluşmaktadır (Şekil 6.1). Her sahne sınıfı için 100 resim olmak üzere, toplam 600 resim kullanılmıştır. Resimler TRECVID 2005 yarışması veri kümesi içinden seçilmiştir. Bu resimler, TRECVID 2005 veri kümesi içinde yer alan videolardan elde edilen anahtar resimlerdir (keyframe) (Çizelge 6.1). Her sahne sınıfı için rastgele 100 resim seçilmiştir. Tüm resimlerin boyutu birbirine eşit ve 352x240 pikseldir. Resimler *jpeg* formatındadır. Tüm resimler renklidir ve her bir piksel 24 bit ile ifade edilmiştir. Resimler üzerinde herhangi bir ön-işleme uygulanmamıştır. Her resim, altı sahne sınıfından birine atanmıştır, çalışma kapsamında bir resmin hiçbir sahne sınıfına ait olmaması durumu ele alınmamıştır.



Şekil 6.1. Sahne sınıflarına ait resim örnekleri

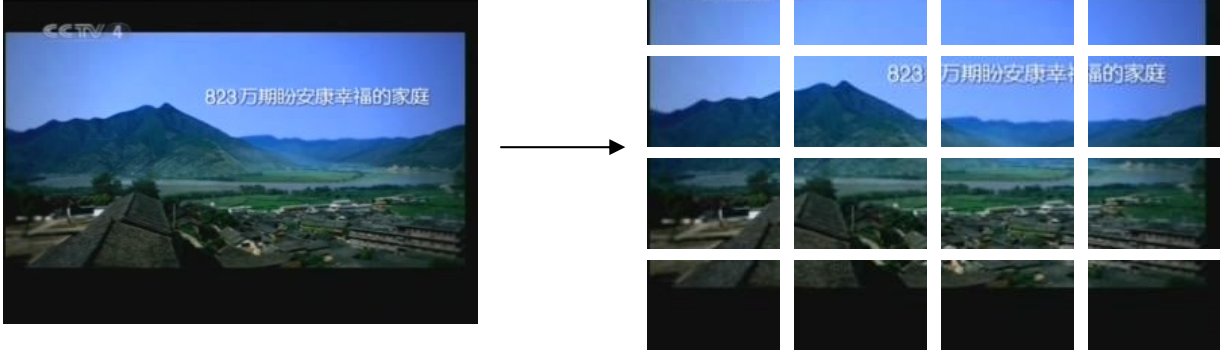
Çizelge 6.1. TRECVID 2005 yarışması veri kümesi[16]

TRECVID-use	Lang	Source	Program	Hours	
tv5	tv6	Eng	NBC	NIGHTLYNEWS	9.0
tv5	tv6	Eng	CNN	LIVEFROM	14.5
	tv6	Eng	CNN	COOPER	8.3
	tv6	Eng	MSN	NEWSLIVE	14.5
-----					46.3
tv5	tv6	Chi	CCTV4	DAILY_NEWS	9.2
	tv6	Chi	PHOENIX	GOODMORNCN	7.5
	tv6	Chi	NTDTV	ECONFRNT	7.8
	tv6	Chi	NTDTV	FOCUSINT	5.2
-----					29.7
tv5	tv6	Ara	LBC	LBCNAHAR	35.3
tv5	tv6	Ara	LBC	LBCNEWS	39.5
	tv6	Ara	ALH	HURRA_NEWS	7.8
-----					82.6
-----					158.6 hours
					(136,026,556,416 bytes)

6.2. Özniteliklerin Çıkarılması

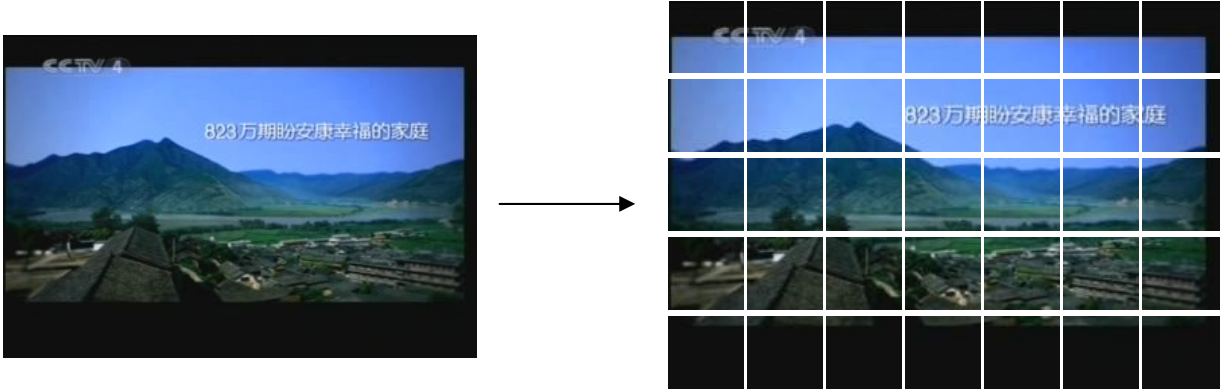
Daha önceki bölümlerde de anlatıldığı gibi, resimlerdeki ayrıntıların daha iyi modellenebilmesi için tüm resimler belli sayıda alt alanlara ayrılmıştır. Tez kapsamında, önceleri, resimler 4x4'lük ızgaraya bölme işlemi uygulanarak toplam 16 adet alt alana ayrılmıştır (Şekil 6.2). Bu bölme işlemi sonunda oluşan her bir alt alan birbirine eşit ve 88x60 piksel boyutlarındadır. Her alt alan için, kenar yönelimlerini ifade eden 20 gözelik histogramlar hesaplanmıştır. Canny kenar detektörü ile bir alt alandaki her bir piksel için 0-360 derece arasındaki kenar yönelimini ifade eden değerler elde edilmiş ve bu değerler 20 gözelik histogram değerlerine dönüştürülmüştür. Böylelikle, her alt alan için kenar yönelimlerine ait, 20-boyutlu öznitelik vektörleri oluşturulmuştur.

Resimlerin 4x4'lük alt alanlara ayrılması, oluşan alt alanların boyutlarının büyük olmasına, bu nedenle de, tek bir alt alan içinde birçok farklı ayrıntı bulunmasına



Şekil 6.2. Bir resmin 4x4'lük alt alanlara ayrılması

sebebi olmuştur. Benzer öznelik değerlerine sahip olan daha küçük boyutlu alanlara daha iyi odaklanabilmek için, resimlere 7x5'lik ızgaraya bölme işlemi uygulanmıştır ve böylelikle, her resim için toplam 35 alt alan elde edilmiştir (Şekil 6.3). Ortaya çıkan bu yeni alt alanlar, resmin farklı bölgelerinde bulunan değişik ayrıntıları daha iyi ifade edebilme gücüne sahip olduğundan, resimlerin daha etkili bir şekilde modellenmesi sağlanmıştır. Ayrıca, haber videolarında yer alan resimlerde bulunabilen alt ve üst siyah bantlar da, bu yeni alt alan sayısı ile daha etkin bir şekilde resimlerden ayrılabilmiştir. Oluşan yeni alt alanların boyutları ise 50x48 pikseldir.



Şekil 6.3. Bir resmin 7x5'lik alt alanlara ayrılması

Resimlerden elde edilen her alt alan için, HSV renk uzayında 54 gözellik histogramları hesaplanmıştır. Bir alt alanda yer alan her piksel için elde edilen RGB değeri,

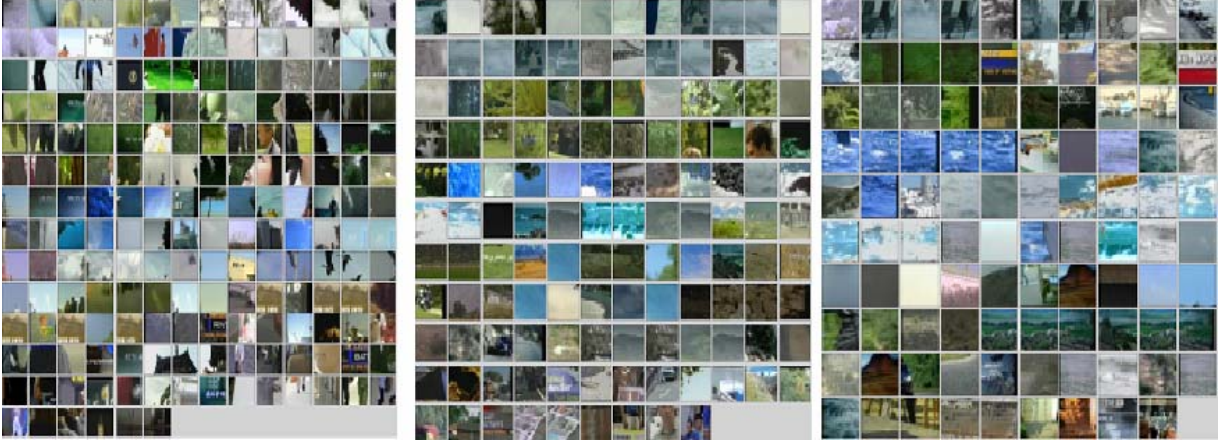
dönüştürüm formülleri uygulanarak, öncelikle HSV değerine dönüştürülmüş, daha sonra da H,S ve V değerleri için sırasıyla, 6x3x3'lük histogram dönüştürüm değerleri uygulanarak, HSV renk özneliğine ait 54-boyutlu öznelik vektörleri elde edilmiştir.

Alt alanlar için hesaplanan bir diğer öznelik değeri ise, doku özneliğine ait histogramlardır. Gabor süzgeçler kullanılarak, bir alt alandaki her piksel için, dört farklı ölçek ve üç farklı yönelimde olmak üzere toplam oniki farklı süzgeç uygulanmıştır ve tüm pikseller için elde edilen bu değerlerin ortalaması ve standart sapması hesaplanarak, her bir alt alan için, doku özneliğine ait istatistiksel değerleri içeren 24-boyutlu öznelik vektörleri elde edilmiştir.

6.3. Alt Alanların Öbeklendirilmesi

Tez kapsamında ele alınan sahne sınıflandırması sorununu çözmeye yönelik olarak gerçekleştirilen bir sonraki aşamada, resimlerdeki alt alanlar için hesaplanan öznelik vektörleri kullanılarak öbeklendirme işlemi gerçekleştirilmiştir.

Resimlerin 4x4'lük alt alanlara ayrılması ile hesaplanan kenar yönelim histogramları kullanılarak, temel öbeklendirme yöntemlerinden biri olan k-ortalama algoritması ile öbeklendirme yapılmıştır. Bu algoritmaya parametre olarak verilecek öbek sayısı 100 olarak belirlenmiştir. Öbeklendirme algoritmasının uygulanması ile, tüm resimlerden elde edilen alt alanlar, 100 öbeğe ayrılarak, doğal bir gruplandırma elde edilmiştir. Ancak, ortaya çıkan öbekler incelendiğinde, öbeklerin kendi içinde benzer alt alanlardan oluşmadıkları tespit edilmiştir. Bir öbekte yer alan alt alanlar, beklendiği üzere, ortak bir anlamsal öğeyi temsil etmek yerine, birbirleriyle ilgisiz ve rastgele bir dağılım göstermişlerdir (Şekil 6.4). Buradan hareketle, kenar yönelimlerine ait özneliklerin, tez kapsamında ele alınan sahne sınıflarının ayrılması için yeterli ayırt ediciliğe sahip olmadığı görülmüş ve renge dayalı öznelik değerlerinin kullanılmasına karar verilmiştir.



Şekil 6.4. Kenar yönelim histogramları kullanılarak elde edilen öbek örnekleri

HSV renk uzayında hesaplanan renk histogramlarıyla gerçekleştirilen öbeklendirme işleminde, k-ortalama algoritması uygulanarak, alt alanlar yine 100 öbeğe ayrılmıştır (Şekil 6.5). Oluşan öbekler, genel olarak incelendiğinde, benzer alt alanların biraraya toplandıkları görülmüştür. Ancak, öbekler ayrıntılı bir şekilde gözden geçirildiğinde, farklı anlamsal öğeleri temsil eden alt alanların (su, kar, gökyüzü gibi) benzer renk özelliklerinden dolayı aynı öbek içinde yer alarak, birbirlerine karıştıkları tespit edilmiştir.

Öbeklerin daha homojen hale getirilebilmesi için, renk özniteliğine ek olarak, doku özniteliğinden de yararlanılabileceği düşünülmüştür. Böylece, renk histogramları kullanılarak elde edilen öbekler, kendi içlerinde, doku özniteliğine ait istatistiksel değerler kullanılarak gerçekleştirilen alt öbeklendirme işlemi ile, belli sayıda (iki, üç, dört...) alt öbeğe ayrılmıştır. Bu yöntemle, aynı öbekte yer alan farklı anlamsal öğelerin birbirlerinden ayrılarak, farklı alt öbeklere dağılması amaçlanmıştır. Ancak, elde edilen alt öbekler incelenerek, bu öğelerin ayrılmasında, doku özniteliğinin başarılı sonuçlar üretmediği anlaşılmıştır (Şekil 6.6). Bu aşamada, daha tutarlı öbekler elde edebilmek için, k-ortalama algoritması yerine, daha etkin bir öbeklendirme yöntemi kullanılmasına karar verilmiştir. Tez kapsamında gerçekleştirilen sahne sınıflandırması yöntemi için öbeklendirme kilit bir rol oynamaktadır ve bu nedenle öbeklerin olabildiğince tutarlı olması, sınıflandırma başarımını da oldukça etkileyecektir. Başarılı bir sınıflandırma gerçekleştirebilmek

için, benzer alt alanların birarada öbekenmesi ve farklı alt alanların birbirinden ayrılması kaçınılmaz olarak gerçekleştirilmelidir.



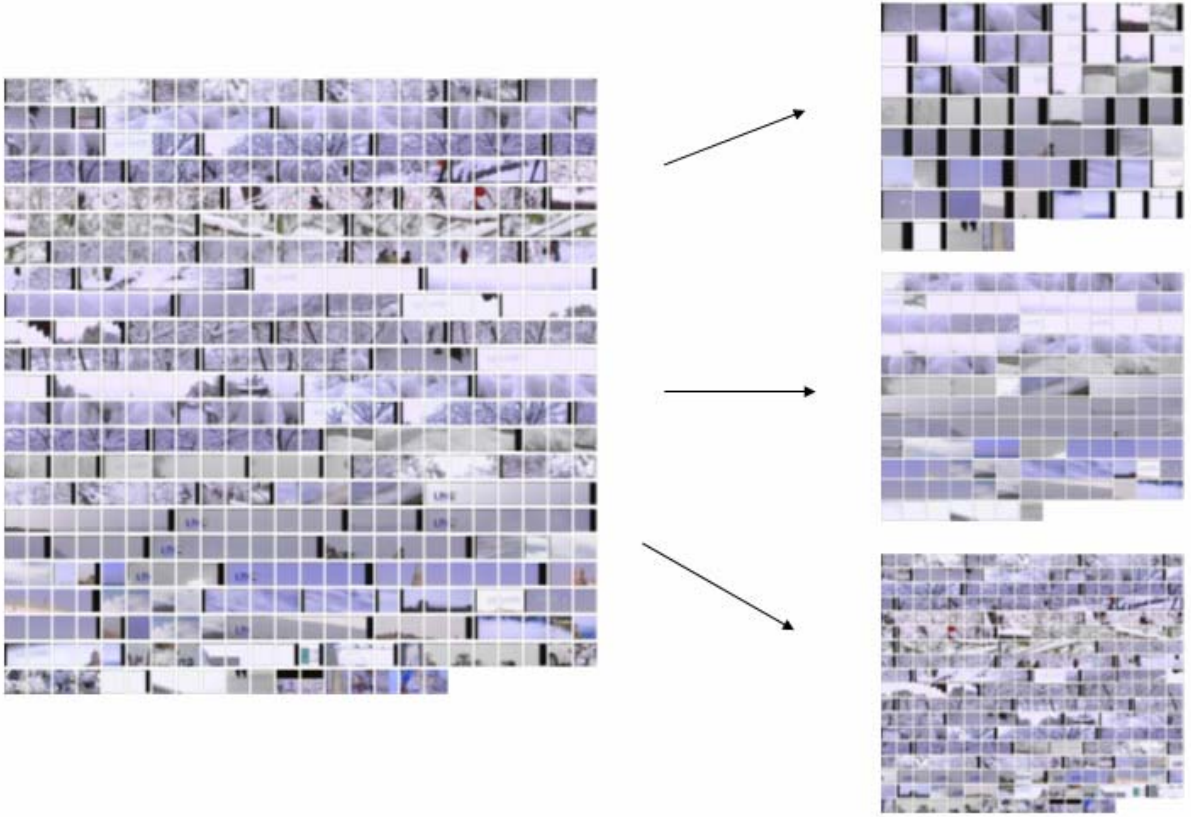
Şekil 6.5. HSV histogramı ve k-ortalama algoritması ile elde edilen öbek örnekleri

k-ortalama algoritmasının çeşitli sorunları bulunmaktadır. Bu algoritma, verinin yapısına bakmaksızın, parametre olarak verilen öbek sayısını esas alarak, mutlaka algoritmaya verilen sayıda öbek oluşturmaktadır. Halbuki, verinin doğası gereği, parametre olarak verilen öbek sayısından daha az ya da çok sayıda öbek meydana gelebilir. Eğer veriye ait çeşitli parametreler dikkate alınarak öbek sayısı dengelenirse, daha doğal bir öbektendirme gerçekleştirilebilir. Bu nedenle, öbektendirme aşamasında, çeşitli parametreler eklenerek, öbek sayısının, verinin yapısına uygun olarak ortaya ortaya çıkması sağlanmaya çalışılmıştır. Belirlenen parametreler, öbeklerdeki alt alan sayısı, öbeklerin standart sapması ve öbeklerin birbirine olan uzaklığı ile ilgili değerlerdir. Bu parametreler kullanılarak;

- alt alan sayısı belli bir eşik değerinden daha az olan öbekler elenmiştir. Az sayıda alt alan içeren öbeklerin önemsiz öbekler oldukları varsayılmıştır.

- standart sapması belli bir eşik değerinden daha büyük olan öbekler ikiye bölünmüştür. Standart sapması büyük olan öbeklerin birbirinden farklı alt alanlar içerdikleri düşünülmüştür.

- birbirine olan uzaklığı, belli bir eşik değerinden daha küçük olan öbekler birleştirilmiştir. Birbirine çok yakın olan öbeklerin benzer alt alanlar içerdikleri düşünülmüştür.

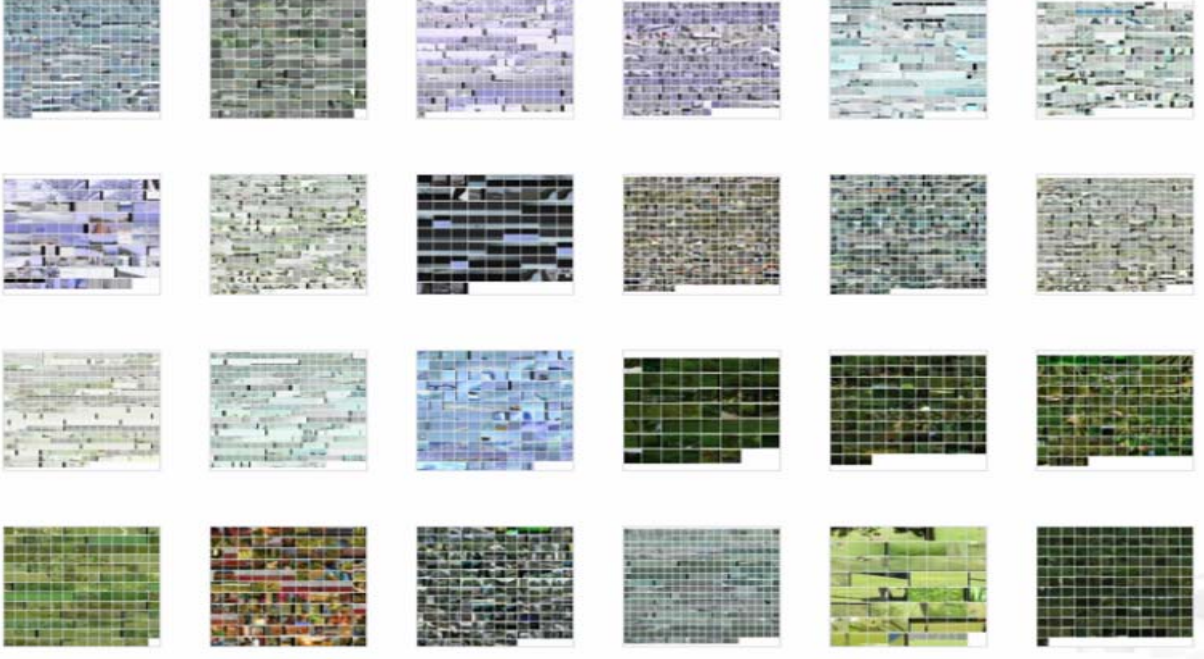


Şekil 6.6. Doku özneliği kullanılarak öbeklerin alt öbeklere ayrılması

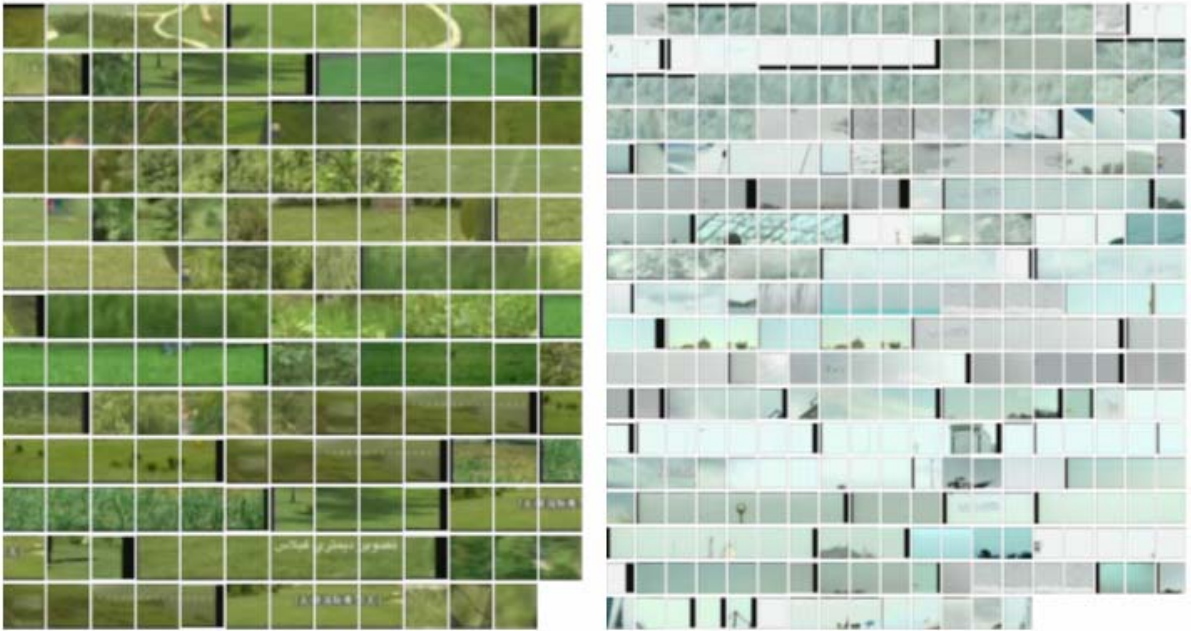
Bu denetimler, öbeklerdeki alt alanlar sağlamlaşana kadar devam etmiştir. Sonuçta ortaya çıkan öbekler, verinin doğasına uygun olarak oluşmuş ve daha tutarlı yapıdaki öbeklerdir. Bu yöntem, literatürde ISODATA öbeklendirme algoritması olarak da adlandırılmaktadır[23].

Yeni öbeklendirme yöntemi kullanılarak, alt alanlara ait HSV renk histogramları ile öbeklendirme işlemi gerçekleştirilmiştir. Kullanılan farklı parametreler ile

gerçekleştirilen öbeklendirme çalışmalarında, en iyi öbek sayısı 115 olarak bulunmuştur (Şekil 6.7, Şekil 6.8).



Şekil 6.7. Bazı öbek örnekleri



Şekil 6.8. Öbeklerin ayrıntılı gösterimi

6.4. Tek Alt Alanlı Konumdan Bağımsız Model

Bu modelin sınanması için gerçekleştirilen deney çerçevesinde, veri kümesindeki tüm resimler için, resimlerdeki alt alanların öbek dağılımlarını konumdan bağımsız olarak modelleyen k (öbek sayısı) boyutlu histogramlar hesaplanmıştır. Bu histogramlar kullanılarak, her sınıf için ayrı ayrı, k parametrelili, çok-terimli dağılım modeli öğrenilmiştir. Modelin parametreleri, en büyük olabirlik kestiriminden yararlanılarak hesaplanmış ve tüm sınıflar için, her öbeğin o sınıfta bulunma olasılığını gösteren olasılık tablosu oluşturulmuştur(Şekil 6.9).

	1	2	3	4	k ($k=115$)
kar					
bitki					
su					
gökyüzü					
dağ					
şehir					

Şekil 6.9. Öbeklerin sınıflardaki dağılımını gösteren olasılık tablosu

Verilen herhangi bir resmin hangi sınıfa ait olduğuna karar vermek için ise, Bayes kuralında, her sınıf için ayrı ayrı olmak üzere, tablodan elde edilen olasılık değerleri formülde yerine konularak sonsal olasılıklar elde edilmiş ve en yüksek sonsal olasılığı veren sınıf, resmin ait olduğu sınıf olarak seçilmiştir.

Modelin başarımını sınamak için, bağımsız geçerlilik sınaması (cross validation) yöntemi kullanılmıştır. Bu yöntemde göre, tüm resimler, eşit büyüklükte ve aynı sınıf dağılımına sahip on ayrı kümeye ayrılmıştır. Daha sonra, her defasında farklı bir küme olmak üzere, kümelerden bir tanesi sınama için ayrılmış ve kalan dokuz küme, parametrelerin öğrenilmesi için kullanılmıştır. Böylelikle, her defasında farklı bir küme için sınama yapılarak, birbirinden bağımsız olan bu sınamaların ortalaması alınmış ve modelin başarım oranı olarak kabul edilmiştir. Çizelge 6.2'de, modelin başarımına

ilişkin hata matrisi görülmektedir. Buna göre, modelin başarımı %52,3 olarak gerçekleşmiştir.

Çizelge 6.2. Tek alt alanlı konumdan bağımsız model için hata matrisi

		Atanan						başarım(%)
		kar	bitki	su	gökyüzü	dağ	şehir	
Doğru	kar	51	3	9	17	8	12	51
	bitki	8	74	0	5	4	9	74
	su	22	1	47	12	9	9	47
	gökyüzü	16	7	17	38	17	5	38
	dağ	19	3	13	7	37	21	37
	şehir	8	9	3	4	9	67	67

Şekil 6.10'da ise, model tarafından doğru ve yanlış olarak sınıflandırılan resim örnekleri görülmektedir.

6.5. Tek Alt Alanlı Konuma Bağımlı Model

Bu modelin, yukarıda anlatılan modelden farkı, belirli bir öbeğe atanan bir alt alanın, resimde bulunduğu konumla ilgili olarak modellenmesidir. Bu nedenle, öbeklerin sınıflara göre olasılıklarını hesaplarken, önceki modelden farklı olarak, resimdeki her bir konum için, o konumda herhangi bir öbeğin bulunması olasılığı ayrı ayrı hesaplanmıştır. Yani, her konum için k parametrelili çok-terimli dağılım modeli öğrenilmiştir. Bu modelle öğrenilen parametreler kullanılarak, bir sınıf için, her konumdaki öbek dağılımını gösteren olasılık tablosu hesaplanmıştır. Bu tablo, tüm sınıflar için ayrı ayrı oluşturulmuştur (Şekil 6.11).

	1	2	3	4	k (k=115)
1					
2					
3					
4					
5					
.					
.					
.					
35					

Şekil 6.11. Bir sınıf için her konumdaki öbek dağılımını gösteren olasılık tablosu

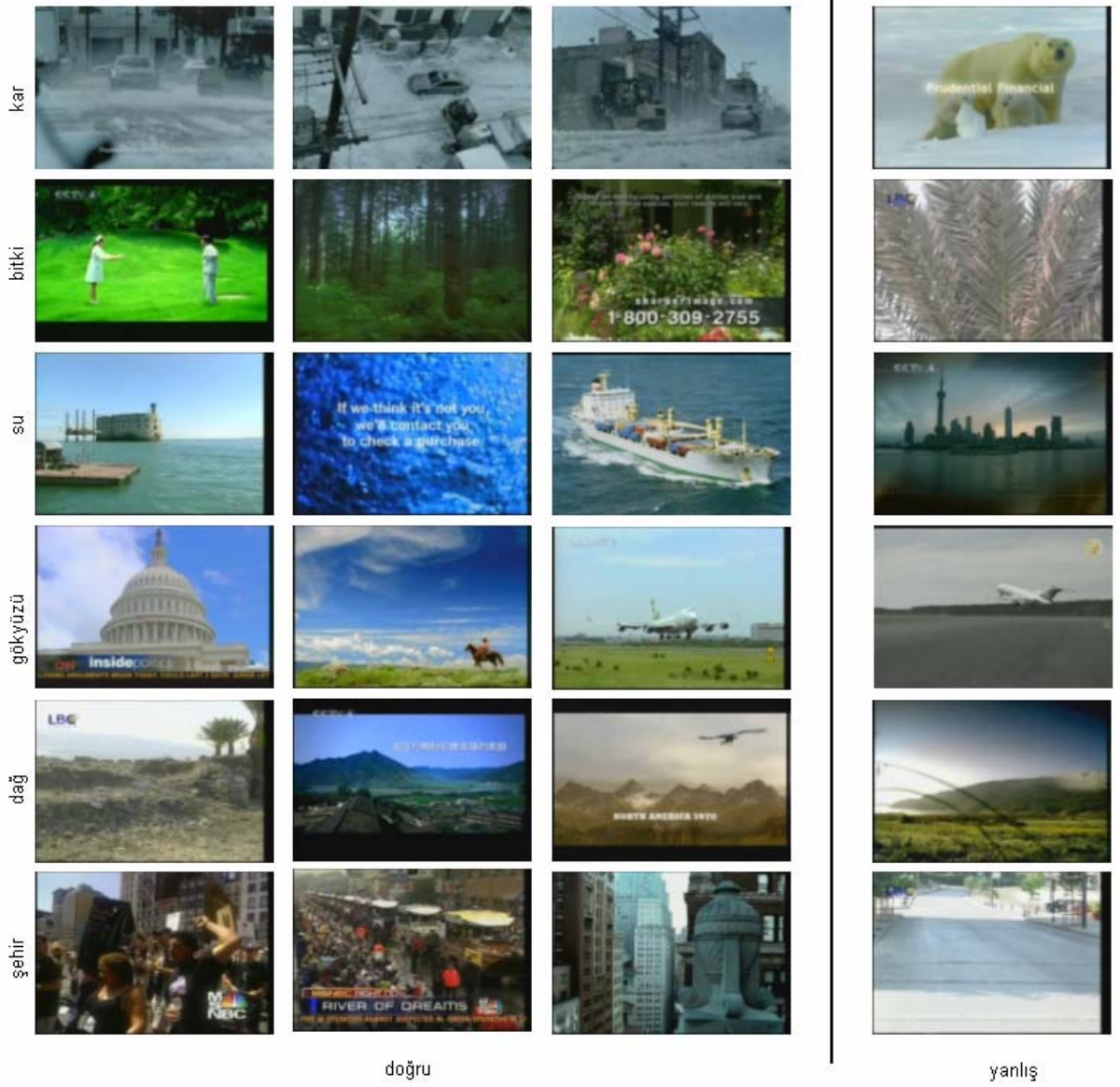
Bir resmin hangi sınıfa ait olduğuna karar vermek için ise, yine Bayes kuralı uygulanmıştır. Tablodan elde edilen olasılık değerleri kullanılarak, her sınıf için sonsal olasılıklar hesaplanmış ve en yüksek olasılığı veren sınıf seçilmiştir.

Modelin başarımını sınamak için de, benzer şekilde bağımsız geçerlilik sınaması kullanılmıştır. Sınamaya ilişkin hata matrisi Çizelge 6.3'de görülmektedir. Modelin başarım oranı ise % 55.6 olarak gerçekleşmiştir.

Çizelge 6.3. Tek alt alanlı konuma bağımlı model için hata matrisi

		Atanan						başarım(%)
		kar	bitki	su	gökyüzü	dağ	şehir	
Doğru	kar	56	4	11	13	5	11	56
	bitki	3	72	1	6	12	6	72
	su	17	4	52	12	7	8	52
	gökyüzü	8	7	15	51	16	3	51
	dağ	7	7	14	21	36	15	36
	şehir	6	7	3	9	8	67	67

Şekil 6.12'de ise doğru ve yanlış olarak sınıflandırılan resimlere örnekler verilmiştir.



Şekil 6.12. Tek alt alanlı konuma bağımlı model kullanıldığında doğru ve yanlış olarak sınıflandırılan resim örnekleri

6.6. Çift Alt Alanlı Göreli Konumlu Model

Tez çalışmasında gerçekleştirilen deneyler kapsamında kullanılan tüm resimler için, alt alan çiftlerinin alt-üst komşuluğunu modelleyen ve olası tüm komşuluk ilişkilerini içeren k^2 boyutlu histogramlar oluşturulmuştur. Daha sonra, oluşturulan bu histogramlar kullanılarak, her sınıftaki tüm alt alan çiftleri için olmak üzere, toplam k^2 parametrelili çok-terimli dağılım modeli öğrenilmiştir. Böylelikle, herhangi bir alt alan çiftinin, herhangi bir sınıfta bulunma olasılığını içeren bir olasılık tablosu elde edilmiştir (Şekil 6.13).

	1	2	3	4	$k*k$ ($k=115$)
kar					
bitki					
su					
gökyüzü					
dağ					
şehir					

Şekil 6.13. Alt alan çiftlerinin sınıflardaki dağılımını gösteren olasılık tablosu

Verilen herhangi bir resmin hangi sınıfa ait olduğuna karar vermek için ise, önceki modellerde olduğu üzere, tablodan elde edilen olasılık değerleri, Bayes formülünde yerine konularak her sınıf için sonsal olasılık değerleri hesaplanmıştır. Bu işlemin sonucunda da, resim, en büyük sonsal olasılığa sahip olan sınıfa atanmıştır.

Modelin başarımını sınamak için, aynı şekilde, bağımsız geçerlilik sınaması (cross validation) yöntemi kullanılmıştır. Alt alanların ayrıldığı öbek sayısı, k , 115 olarak belirlendikten sonra yapılan sınamaya ilişkin hata matrisi Çizelge 6.4'de görülmektedir. Buna göre, modelin başarım oranı % 62,6'dır.

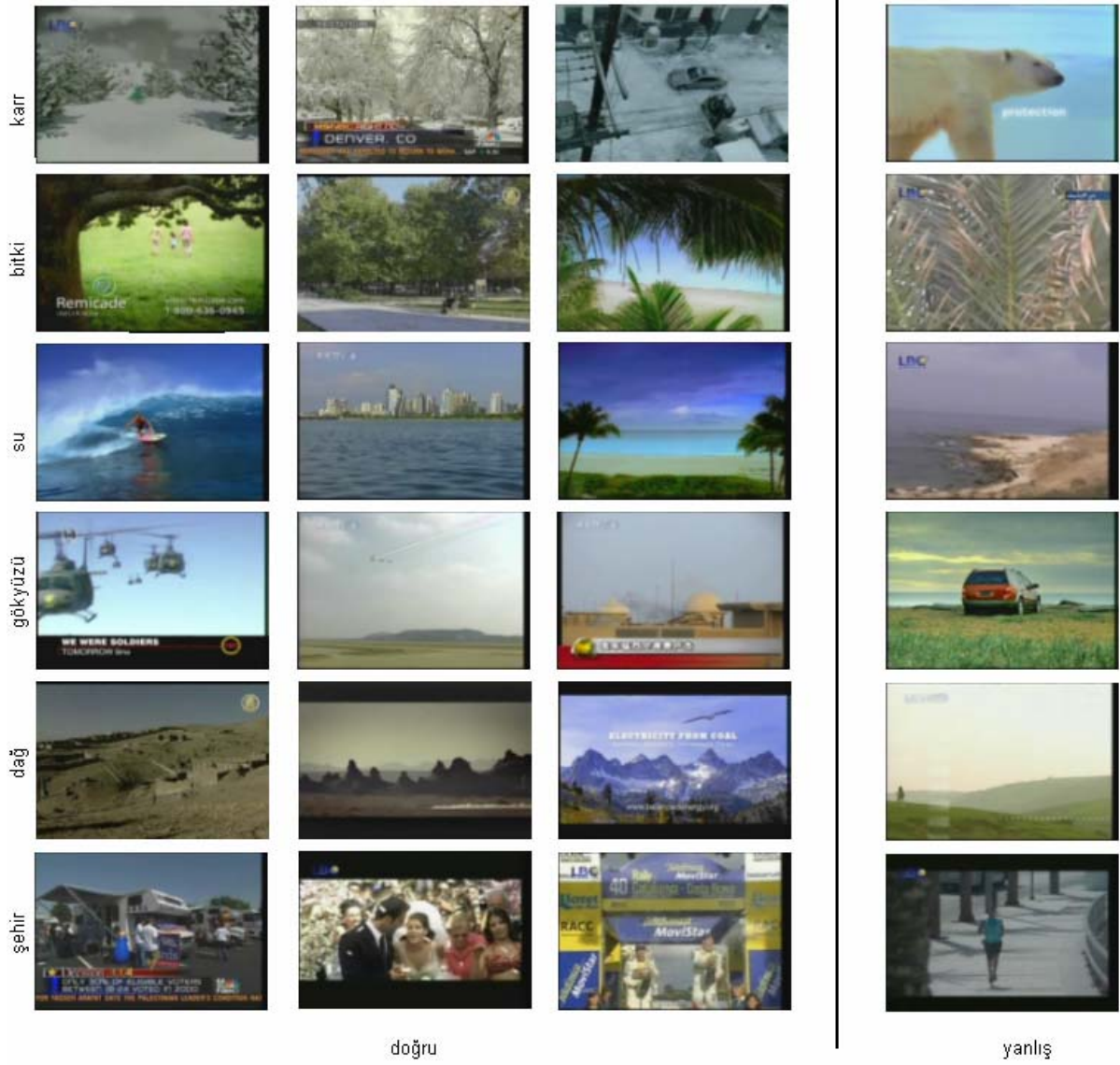
Çizelge 6.4. Çift alt alanlı görel konumlu model için hata matrisi

		Atanan						başarım(%)
		kar	bitki	su	gökyüzü	dağ	şehir	
Doğru	kar	62	4	8	13	4	9	62
	bitki	6	77	3	5	4	5	77
	su	13	3	61	10	6	7	61
	gökyüzü	10	6	14	55	12	3	55
	dağ	11	5	10	11	52	11	52
	şehir	10	4	5	6	6	69	69

Şekil 6.14'de model tarafından yanlış ve doğru olarak sınıflandırılan resimler için örnekler görülmektedir.

Sonuçlardan da görüldüğü üzere, tüm modellerde en yüksek başarımları bitki sınıfı için gerçekleşmiştir. Bu durum, bitki sınıfındaki resimlerde hakim olan yeşil rengin, diğer sınıflardan ayırt edici bir etkisi olmasıyla açıklanabilir. Bu sınıfın, her üç modelde de başarılı bir şekilde sınıflandırılmış olması, sınıflandırılması kolay bir sınıf olduğu yorumunun yapılmasına olanak vermektedir.

Tüm modellerde, sınıflandırma başarımları yüksek olan bir diğer sınıf da şehir sınıfıdır. Diğer beş sınıfın doğal sahneler içermesi ve şehir sınıfındaki resimlerde insan yapımı nesnelerin ve insanların çoğunlukla yer alması, bu sınıftaki renk dağılımının ve anlamsal öğelerin diğer sınıflardan farklı olmasına yol açmaktadır. Bu da, şehir sınıfını, sınıflandırması kolay bir sınıf haline getirmektedir. Diğer sınıflar için böyle ayırt edici renklerden ya da öğelerden bahsedilemediği için, özellikle kar, su, gökyüzü ve dağ sınıfları, benzer renk dağılımları (mavi, beyaz) ve benzer anlamsal öğelere sahip olmalarından dolayı birbirlerine karıştıkları gözlenmektedir. Bu sınıflar, sınıflandırılmaları oldukça zor olan, ayırt edici karakteristik özellikleri belirgin olmayan sınıflardır. Bu nedenle de, başarımları bitki ve şehir sınıflarına göre düşük olmaktadır.



Şekil 6.14. Çift alt alanlı görelî konumlu model kullanıldığında doğru ve yanlış olarak sınıflandırılan resim örnekleri

Sınıflandırma başarımının çok yüksek olmamasının çeşitli sebepleri olabilir. Öncelikle, problemin doğasının zorluğu nedeniyle, sadece renk özneliğinin kullanılması, karmaşık sahne sınıflarının ayırımında çok yüksek başarımlar için yeterli olmamaktadır. Özellikle, bazı sınıfların renk dağılımlarının birbirine yakın olması bu yetersizliği daha çok belirginleştirmektedir.

Modellerin başarımlarını düşüren bir diğer sebep de öbeklendirme sonuçlarındaki hatalardır. Herhangi bir öbek, genel olarak incelendiğinde, benzer renklerdeki alt alanların biraraya gelmesiyle oluştuğu görülmektedir. Ancak, daha ayrıntılı incelendiğinde, birçok farklı sınıftan gelen alt alanın aynı öbek içinde toplanarak birbirlerine karıştıkları farkedilmektedir. Öbek sayısının artırılması, benzer alt alanların farklı öbeklere dağılmasına sebep olurken, öbek sayısının azaltılması da, farklı alt alanların aynı öbek içinde bulunmalarına neden olmaktadır. Bu durumda, öbeklendirme işlemi için hatasız bir sonuç elde edilemeyeceği için, bu süreçteki hatalar sınıflandırma aşamasına da etki etmektedir. Problemin en önemli noktalarından biri olan öbeklendirme algoritması ne kadar başarılı olursa, sonuçlar da o oranda başarılı olacaktır.

Başarımı etkileyen faktörlerden bir diğeri ise resmin bölüneceği alt alan sayısıdır. Eğer bu sayı az tutulursa, alt alanların boyutu büyük olmaktadır ve birbirinden farklı birçok ayrıntı aynı alt içinde yer alabilmektedir. Bu durumda, alt alanların etkili bir biçimde öbeklendirilmesi de zorlaşmaktadır. Buna karşılık, alt alan sayısının çok olması da, resimdeki ayrıntıların çok fazla bölünmesine ve anlamsal bütünlüğün parçalanmasına neden olmaktadır. Genel olarak, en iyi alt alan sayısından bahsetmek zordur. Bu durumdan kaynaklanan hatalar, sınıflandırmaya da yansımaktadır.

Tez çalışmasında kullanılan sahne sınıfları, içiçe geçmiş yani birbirini içeren sınıflardır. Örneğin, sadece gökyüzü sınıfındaki resimlerde değil, diğer tüm sınıflardaki resimlerde de genelde gökyüzü bulunmaktadır. Ya da tüm sınıflarda bitki örtüsü ya da su içeren resimler bulunabilmektedir. Çalışma kapsamında bir resim tek

bir sınıfa atanabildiğinden, bu durum hatalara yol açmaktadır. Örneğin, içinde bitki örtüsü ve gökyüzü de bulunan su sınıfına ait bir resim, bitki ya da gökyüzü sınıflarına atanırsa, bu sonuç hata olarak değerlendirilmektedir.

Çalışmada kullanılan veri kümesinin sınırlı sayıda olması, renk histogramları oluşturulurken renk bilgisinin belli oranlarda kaybedilmesi gibi sebepler de modellerin başarımını düşürmektedir.

6.7. Önemli Parametrelerin Seçilmesi

Geliştirilen modeller kapsamında, önemli yani ayırt edici alt alanların belirlenmesi için Ki-Kare testi kullanılmıştır. Tek alt alanlı konumdan bağımsız model için, bir öbeğin, bir sınıftaki dağılımı, tüm sınıflar için ikili olarak test edilmiştir. Yani bu dağılımın, söz konusu iki sınıf için farklı olup olmadığı belirlenmeye çalışılmıştır. Eğer dağılım farklı çıkarsa, öbek, ayırt edici olarak işaretlenmiştir. Gerçekleştirilen çeşitli deneyler sonucunda, dağılımı en az üç sınıf arasında farklılık gösteren öbekler, belli bir eşik değerine göre seçilmiştir. Buna göre seçilen öbek sayısı 72'dir. Sadece seçilen öbekleri kullanarak yapılan sınıflandırmaya ilişkin hata matrisi Çizelge 6.5'de verilmiştir. Sınıflandırmanın ortalama başarımı %54'dür. Bu değer, bu modelle seçme yapılmadan gerçekleştirilen sınıflandırmaya göre daha yüksek bir değerdir. Buradan da, yöntemin başarılı olduğu görülmektedir.

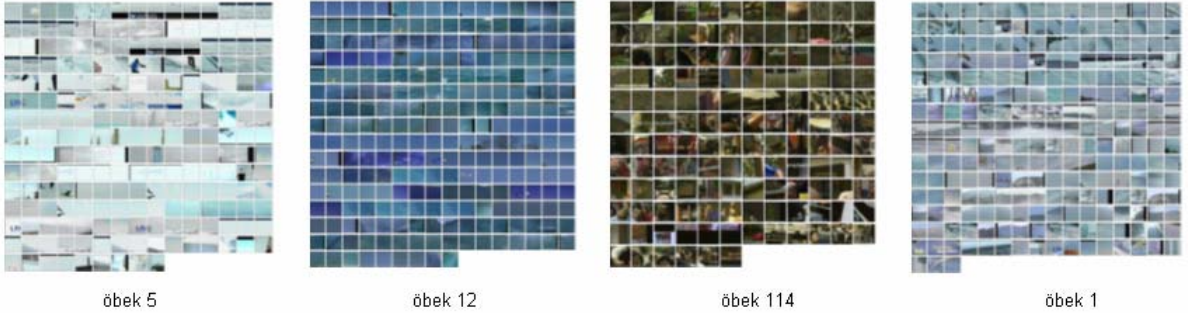
Çizelge 6.5. Tek alt alanlı konumdan bağımsız model ile seçme yaparak gerçekleştirilen sınıflandırmaya ilişkin hata matrisi

		Atanan						başarım(%)
		kar	bitki	su	gökyüzü	dağ	şehir	
Doğru	kar	52	4	9	14	5	16	52
	bitki	7	76	0	5	5	7	76
	su	20	3	53	11	3	10	53
	gökyüzü	15	10	17	44	7	7	44
	dağ	20	5	13	10	30	22	30
	şehir	6	13	3	5	4	69	69

Tek alt alanlı konuma bağımlı model için de benzer şekilde ayırt edici öbekler belirlenmiştir. En az üç sınıf için farklı dağılıma sahip olan öbekler belli bir eşik değerine göre seçilmiş ve sadece seçilen öbekler kullanılarak sınıflandırma yapılmıştır. Seçilen öbek sayısı 100 olarak bulunmuştur. Buna ilişkin hata matrisi ise Çizelge 6.6'da görülmektedir. Sınıflandırmanın başarımlı oranı %62,6'dır. Bu oran, seçme yapılmadan önceki oranla karşılaştırıldığında, büyük bir artış olduğu görülmektedir. Şekil 6.15'de ise seçilen öbekler için bazı örnekler verilmiştir.

Çizelge 6.6. Tek alt alanlı konuma bağımlı model ile seçme yapılarak gerçekleştirilen sınıflandırmaya ilişkin hata matrisi

		Atanan						başarımlı(%)
		kar	bitki	su	gökyüzü	dağ	şehir	
Doğru	kar	69	2	9	9	4	7	69
	bitki	4	76	1	5	10	4	76
	su	18	2	61	11	2	6	61
	gökyüzü	7	4	18	57	13	1	57
	dağ	10	4	15	18	45	8	45
	şehir	5	4	4	6	13	68	68



Şekil 6.15. Tek alt alanlı konuma bağımlı model çerçevesinde seçilen öbeklerden bazıları

Çift alt alanlı görelı konumlu model için de seçme işlemi gerçekleştirilerek sınıflandırma yapılmıştır. Ancak bu modelde kayda değer bir iyileşme gözlenememiştir. Bunun çeşitli sebepleri olabilir. Örneğın, eşik değerini geçen öbek sayısı az olduđu için, bu sayı sınıflandırma için yetersiz kalmıştır denebilir.

Sonuç olarak, önemli parametrelerin seçilmesi, iki model için sınıflandırma başarımını arttırmıştır. Buna dayanarak, bu konuda yapılacak çalışmaların umut vaat ettiği söylenebilir.

6.8. Sonuçların Geleneksel Yöntemlerle Karşılaştırılması

Bu tez çalışmasının konusu olan, sahne sınıflandırması için alt alanlara dayalı uzamsal bilginin kullanımını, bu bilgiyi dikkate almadan, resmin bütününden elde edilen öznitelikleri kullanarak yapılan sınıflandırmayla karşılaştırmak için Çizelge 6.5'de bir hata matrisi verilmiştir. Bu hata matrisini oluşturmak için, tez çalışmasında kullanılan veri kümesi içinde yer alan tüm resimler için, resmin bütününe ait HSV renk histogramları çıkarılmıştır. Bu histogramlar kullanılarak QDC (Quadratic Discriminant Classifier) ile, tezde kullanılan sahne sınıfları için sınıflandırma yapılmıştır. Bu yöntemin ortalama başarımı ise %41,8 olarak bulunmuştur. Bu değer, tez çalışmasında gerçekleştirilen her üç modelin başarımından da oldukça düşük bir değerdir. Bu sonuçtan da anlaşıldığı üzere, etkin bir sahne sınıflandırması için, resmin bütününden çıkarılan özniteliklere dayalı geleneksel yöntemler yeterli olmamaktadır. Resimdeki ayrıntıların dikkate alınmasını ve uzamsal bilginin modellenmesini sağlayan alt alan kullanımı sınıflandırma başarımını arttırmıştır.

Çizelge 6.7.QDC ile gerçekleştirilen sınıflandırma için hata matrisi

		Atanan		
		doğru	hatalı	başarım(%)
Doğru	kar	47	53	47
	bitki	42	58	42
	su	36	64	36
	gökyüzü	37	63	37
	dağ	55	45	55
	şehir	34	66	34

7. SONUÇLAR

Sayısal resim ve video arşivlerinin yaygınlaşmasıyla birlikte, bu arşivlerin etkin bir şekilde nasıl kullanılabileceği konusu gündeme gelmiştir. Bu konuda geliştirilecek başarılı uygulamalar, bu arşivlerin otomatik olarak değerlendirilmesini gerçekleştirerek hızlı ve etkili bir şekilde sonuca ulaşılmasını sağlayacaktır. Buradan hareketle, son yıllarda, görüntü işleme adı altında çeşitli çalışmalar yapılmaktadır. Bu çalışmalardan bir kısmı, resimlerdeki anlamsal bilgilerin çıkarılmasına yönelik olarak gerçekleştirilmektedir. Sahne sınıflandırması adı altında gerçekleştirilen bu çalışmalar ile, resimler, çeşitli sahne sınıflarına ayrılmaya çalışılmaktadır. Buradan elde edilecek sonuçlar ile arşivler üzerinde gerçekleştirilecek dizinler, özellikle, arama, tarama ve erişim süreçlerinin hızını ve başarımını önemli oranlarda arttıracaktır.

Bu konu üzerinde yapılan eski çalışmalarda, genellikle, renk, doku, kenar gibi öznitelikler, resmin bütününden çıkarılarak kullanılmış, bu öznitelikler ile çoğunlukla istatistiksel sınıflandırma yöntemleri denenmiştir. Ancak, resimdeki uzamsal bilgi ve uzamsal ilişkiler üzerinde durulmamıştır. Bir resimdeki anlamsal bilginin yorumlanmasında, uzamsal yerleşimin çok önemli olduğu bilinmektedir. Bu tespitten yararlanarak, sahne sınıflandırması probleminin çözümü için uzamsal konumdan ve uzamsal ilişkilerden yararlanılabilir.

Bu tez çalışmasında, sahne sınıflandırması, uzamsal bilgiye dayalı resim modelleri kullanılarak gerçekleştirilmiştir. Uzamsal bilgi, resmin alt alanlara bölünmesiyle elde edilmiştir. Resimler, ızgaraya bölme işlemi ile belli sayıda ve eşit büyüklükte alt alanlara ayrılmıştır. Tüm alt alanlar için, HSV renk uzayında renk histogramları hesaplanmıştır. Bu histogramlar kullanılarak, alt alanlar öbeklendirilmiştir. Böylelikle, renk özniteliğine göre benzer olan alt alanlar aynı öbek içinde toplanmış ve bundan sonraki süreçlerde buldukları öbek numarasıyla modellenmişlerdir. Öbek atamaları belirlenen alt alanlardan oluşan resimlerin sınıflandırılması için üç değişik model ortaya konmuştur.

Birinci modelde, alt alanlar, resimde buldukları konumdan bağımsız olarak ele alınmış, resimler, alt alanların atandığı öbeklerin listesi ile temsil edilmiştir. Bir resimde hangi öbeklere ait alt alanların bulunduğu belirlenerek buna ait histogramlar oluşturulmuştur. Öbeklerin sınıflardaki dağılımı, çok-terimli dağılım modeli ile örtüştüğünden, bu modelin parametreleri (bir öbeğin bir sınıfta bulunma olasılığını gösteren parametreler), en-büyük olabilirlik kestirimi ile hesaplanmıştır. Bunun için, örnek resimlerden elde edilen değerler kullanılmıştır. Verilen bir resmin hangi sınıfa ait olduğuna karar vermek için ise, örneklerden hesaplanan olasılık değerleri kullanılarak, Bayes karar kuralı ile sonsal olasılıklar bulunmuştur.

İkinci modelde, birinci modelden farklı olarak, alt alanların resimde buldukları konum da dikkate alınmıştır. Böylece resimler, hangi konumda, hangi öbeğe atanan alt alanın bulunduğunu belirten bir gösterimle ifade edilmiştir. Öbeklerin, sınıflara göre, konumlardaki dağılımı, çok-terimli dağılım ile öğrenilerek, bir öbeğin, belirli bir sınıf için, belirli bir resim konumunda bulunma olasılığı, en-büyük olabilirlik kestirimi ile bulunmuştur. Bu olasılıklar kullanılarak, verilen bir resmin hangi sınıfa ait olduğunu bulmak için, Bayes karar kuralı ile sonsal olasılıklar hesaplanmıştır.

Son model olan üçüncü modelde ise, alt alanların ikili alt-üst komşuluk ilişkileri kullanılmıştır. Bu modelde, her sınıf için, hangi öbeklere atanan alt alanların birbirleriyle komşu oldukları tespit edilmeye çalışılarak, önceki modellerde olduğu gibi, olası tüm komşuluk ilişkileri için olasılıklar hesaplanmıştır. Sınıflandırılmak istenen resimdeki öbek komşulukları tespit edilmiş, ve bu komşuluklara ilişkin olasılıklar kullanılarak sınıflandırma yapılmıştır.

Oluşturulan modeller, kar, bitki örtüsü, su, gökyüzü, dağ ve şehir sahnelerinin sınıflandırılması için kullanılmıştır. Modellerin başarımını sınamak için çeşitli deneyler gerçekleştirilmiştir. Deneyler kapsamında kullanılan veri kümesi, TRECVID2005 yarışması verilerinden rastgele seçilmiştir. Modellerin başarımına ilişkin hata matrisleri oluşturulmuştur. Buna göre, en iyi sınıflandırma, bitki ve şehir sınıfları için gerçekleşmiştir. Diğer sınıflarda, renk benzerliğinden dolayı birbirine karışmalar tespit

edilmiştir. Sonuç olarak, tez kapsamında ortaya konan modellerin, uzamsal bilgiyi dikkate almayan modellere göre daha başarılı olduğu gözlenmiştir.

KAYNAKLAR

- [1] NASA, "TERRA: The EOS flagship", <http://terra.nasa.gov>.
- [2] A. Vailaya, M. Figueiredo, A. Jain, H. J. Zhang. A Bayesian Framework for Semantic Classification of Outdoor Vacation Images. IEEE Trans. Image Processing, Vol. 10, No. 1, pp. 157-172, Jan. 2001.
- [3] A. Vailaya, A. Jain, H. J. Zhang. On Image Classification: City Images vs. Landscapes. Pattern Recognition Journal, Pattern Recognition, Vol. 31, No. 12, pp. 1921-1935, 1998.
- [4] A. B. Torralba, K. P. Murphy, W. T. Freeman, M. A. Rubin. Context-based Vision System for Place and Object Recognition. In Proc. ICCV, Oct. 2003.
- [5] M. Szummer, R. W. Picard. Indoor-outdoor Image Classification. IEEE Intl. Workshop on Content-based Access of Image and Video Databases, 1998.
- [6] Li Fei-Fei, Pietro Perona. A Bayesian Hierarchical Model for Learning Natural Scene Categories. CVPR, Vol. 2, pp. 524-531, 2005.
- [7] D. Blei, A. Ng, M. Jordan. Latent Dirichlet Allocation. Journal of Machine Learning Research, 3:993–1022, 2003.
- [8] P. Quelhas, F. Monay, M. Odobez, D. Gatica-Perez, T. Tuytelaars, L. Van Gool. Modeling Scenes with Local Descriptors and Latent Aspects. In Proc. ICCV, 2005.
- [9] L. W. Renninger, J. Malik. When is Scene Classification Just Texture Recognition? Vision Research, 44, 2301-2311, 2004.
- [10] Anil K. Jain, R. Duijn, J. Mao. Statistical Pattern Recognition : A Review. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 1, 2000.
- [11] P. E. Ross. Flash of Genius. Forbes, pp. 98-104, 1998.
- [12] S. Watanabe. Pattern Recognition: Human and Mechanical. New York: Wiley, 1985.
- [13] D. Judd, P. Mckinley, A. K. Jain. Large-Scale Parallel Data Clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 8, pp. 871-876, 1998.
- [14] H. Frigui, R. Krishnapuram. A Robust Competitive Clustering Algorithm with Applications in Computer Vision. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 5, pp. 450-465, 1999.

- [15] H. M. Abbas, M. M. Fahmy. Neural Networks for Maximum Likelihood Clustering. Signal Processing, Vol. 36, No. 1, pp. 111-126, 1994.
- [16] "TREC Video Retrieval Evaluation Home Page", <http://www-nlpir.nist.gov/projects/trecvid/> , 2006.
- [17] William H. Press, Brian P. Flannery, Saul A. Teukolsky, William T. Vetterling. Numerical Recipes in C: The art of scientific computing. ISBN 0-521-43108-5, Ch.14, pp. 617-622, 1992.
- [18] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 60, 2, pp. 91-110, 2004.
- [19] D. G. Lowe. Object Recognition from Local Scale-Invariant Features. International Conference on Computer Vision, pp. 1150-1157, 1999.
- [20] Gonzalez, Woods. Digital Image Processing (2nd ed.). Prentice Hall, ISBN 0201180758, 2002.
- [21] J. Canny. A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, No. 6, 1986.
- [22] R. O. Duda, P. E. Hart, D. G. Stark. Pattern Classification (2nd ed.). John Wiley & Sons, ISBN 0-471-05669-3, 2001.
- [23] J. R. Jensen. Introductory Image Processing: A Remote Sensing Perspective (2nd ed.). Prentice Hall, ISBN 0132058405, 1996.
- [24] T. Aach, A. Kaup, R. Mester. On Texture Analysis: Local Energy Transforms Versus Quadrature Filters. Signal Processing, Vol. 45, pp. 173-181, 1995.

ÖZGEÇMİŞ

Adı Soyadı : Ajda KARABULUT

Doğum Yeri : Karapınar

Doğum Yılı : 22.08.1980

Medeni Hali : Evli

Eğitim ve Akademik Durumu:

Lise 1998-1999 Balıkesir Sırrı Yırcalı Anadolu Lisesi

Lisans 1999-2003 Hacettepe Üniversitesi Bilgisayar Mühendisliği Bölümü

Yabancı Dil: İngilizce

İş Tecrübesi:

2003-2006 Hacettepe Üniversitesi Bilgisayar Mühendisliği Bölümü,

Araştırma Görevlisi