

YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YAPAY SİNİR AĞLARI İLE YÜZ SEZİMİ VE TAKİBİ

Elektrik-Elektronik Müh. Serkan ESER

F.B.E. Elektronik ve Haberleşme Anabilim Dalı Elektronik Programında

Hazırlanan

YÜKSEK LİSANS TEZİ

Tez Danışmanı: Prof.Dr.Tülay YILDIRIM

İSTANBUL,2006

İÇİNDEKİLER

	Sayfa
SİMGE LİSTESİ	iv
KISALTMA LİSTESİ.....	v
ŞEKİL LİSTESİ.....	vi
ÇİZELGE LİSTESİ	vii
ÖNSÖZ	viii
ÖZET	ix
ABSTRACT	x
1. GİRİŞ	1
1.1 Özellik Tabanlı (Feature Based) Yöntemler	2
1.2 Görüntü Tabanlı (Image Based) Yöntemler.....	2
1.3 Yüz Belirleme ve İzleme Problemine Getirilmiş Çözümler	4
2. GÖRÜNTÜ İŞLEME	7
2.1 Renk ve Algılama	7
2.2 Renk Uzayları.....	8
2.2.1 RGB	8
2.2.2 Normalize RGB	9
2.2.3 CMY	10
2.2.4 HSV	10
2.2.5 YCbCr	12
2.2.6 Karşıt Renk Uzayları	12
2.3 Matematiksel Morfoloji	13
2.3.1 Temel Küme İşlemleri	14
2.3.2 Morfolojik Görüntü İşleme	14
2.3.2.1 Aşındırma.....	15
2.3.2.2 Yayma	15
2.3.2.3 Açma ve Kapama.....	16
2.3.3 Bağlantılı Bileşenler Etiketleme (CCL, Connected Component Labeling).....	17
2.4 Ten Bölgelerinin Bulunması	18
2.5 Ten Sınıflayıcılar	19
2.5.1 Kapalı Sınırlı Ten Sınıflayıcılar	19
2.5.2 Parametrik Olmayan Sınıflayıcılar (Non-parametric Classifiers).....	19
2.5.3 Parametrik Sınıflayıcılar (Parametric Classifiers).....	19
2.5.4 Adaptif ve Durağan Sınıflayıcılar	21
2.5.5 Ten Sınıflayıcı Tasarımı	22
2.6 CAMSHIFT (Continuously Adaptive Mean Shift) Takip Algoritması.....	27
2.7 Tezde Önerilen Yöntem	27

2.8	Ten Bölgelerinden Özellik Çıkarımı	31
2.8.1	Sıkılık (Compactness).....	31
2.8.2	Katılık (Solidity).....	32
2.8.3	Görünüş Oranı (Aspect Ratio).....	32
2.8.4	Yönelim (Orientation).....	32
3.	ŞABLON EŞLEME.....	34
4.	ÖRNEK TANIMA.....	38
4.1	Sınıflama ve Kümeleme.....	38
4.1.1	Bayes Sınıflayıcı.....	38
4.1.2	En Yakın Orta Değer Sınıflayıcı (Nearest Mean Classifiers)	40
4.1.2.1	Öklit ve Mahalanobis Uzaklığı.....	40
4.1.3	K-Means (K-Ortalama) Algoritması.....	41
4.2	Özellik Çıkarımı	42
4.2.1	Ana Bileşen Analizi (PCA, Pricipal Component Analysis).....	43
4.2.2	Doğrusal Ayırtaç Analizi (LDA, Lineer Discriminant Analysis)	44
5.	BOYUT İNDİRGEME YOLUYLA YÜZ SEZİMİ	45
5.1	Özyüz (Eigenface) Metodu	45
6.	YAPAY SİNİR AĞLARI	51
6.1	Yapay Sinir Ağı Hücresi	52
6.2	Yapay Sinir Ağları Modelleri.....	54
6.2.1	Yapay Sinir Ağlarının Eğitilmesi	54
6.2.2	Geriye Yayılım (Backpropagation) Algoritması	55
7.	ÖNİŞLEME.....	58
7.1	En Uygun Düzlem Uydurma (Best Planar Fitting)	58
7.2	Histogram Eşitleme (Histogram Equalization)	60
7.3	Standart Sapma ile Normalizasyon.....	62
7.4	Hız Analizi	62
8.	YAPAY SİNİR AĞLARIYLA YÜZ İZLEME UYGULAMASI.....	63
8.1	Eğitim Setinin Hazırlanması	63
8.2	Doğrusal Alt Uzay Metotları ve YSA.....	64
8.2.1	PCA ve YSA	64
8.3	YSA	66
8.3.1	Algoritmanın Hızlandırılması	69
8.3.1.1	Uzamsal ve Frekans Domenleri.....	70
8.3.1.2	Aktivasyon Fonksiyonunun Örneklenmesi.....	75
8.4	Yüz Takibi.....	76
9.	DENEYSEL SONUÇLAR.....	78
10.	SONUÇ VE ÖNERİLER	84
	ÖZGEÇMİŞ.....	86

SİMGE LİSTESİ

Σ	Kovaryan matrisi
σ	Varyans değeri
ρ	Korelasyon katsayısı
η	Öğrenme adımı
w	Sinaptik ağırlıklar
W	Dönüşüm matrisi
λ	Özdeğerler
D	Özdeğerler matrisi

KISALTMA LİSTESİ

RGB	Red Green Blue
nRGB	Normalizde Red Green Blue
HSV	Hue Saturation Value
YCbCr	Luma Color blue Color red
PCA	Principal Component Analysis
LDA	Lineer Discriminant Analysis
CCL	Connected Component Labeling
CIE	Commission Internationale de L'éclairage

ŞEKİL LİSTESİ

Şekil 2.1 RGB uzayı	8
Şekil 2.2 HSV uzayının koni biçimindeki yapısı	11
Şekil 2.3 Spektrum ile HS değerleri arasındaki ilişki	11
Şekil 2.4 Aşındırma	15
Şekil 2.5 Yayma.....	16
Şekil 2.6 Açma ve Kapama	17
Şekil 2.7 Tüm örneklerin gerçek ve model dağılımları	21
Şekil 2.8 Normalize RGB uzayında renk dağılımları	24
Şekil 2.9 HSV uzayında renk dağılımları	25
Şekil 2.10 YCbCr uzayında renk dağılımları	25
Şekil 2.11 Ten sınıflayıcı sonuçları	26
Şekil 2.12 Karşılaştırmalı test sonuçları	28
Şekil 2.13 Ten bölgelerinin belirlenmesi	31
Şekil 3.1 Şablon yüz	35
Şekil 3.2 Şablon eşleme	36
Şekil 3.3 Şablon eşleme örnekleri	37
Şekil 4.1 Ana Bileşen Analizi	43
Şekil 4.2 Ana Bileşen Analizinin dezavantajı	44
Şekil 4.3 Doğrusal Ayırtaç Analizi.....	44
Şekil 5.1 Özdeğerler ve özyüzler.....	48
Şekil 5.2 Yüz uzayına izdüşüm	49
Şekil 5.3 Yüz ve yüz olmayan örneklerin yapılandırma hataları	50
Şekil 6.1 Temel sinir hücresi modeli	52
Şekil 6.2 Aktivasyon fonksiyonları	53
Şekil 6.3 İleri beslemeli yapı.....	54
Şekil 6.4 Geri beslemeli yapı	54
Şekil 7.1 Örnek yüzeyler ve en uygun düzlemler.....	60
Şekil 7.2 Örnek yüz ve sonuç görüntüsü.....	60
Şekil 7.3 Histogram eşitleme.....	61
Şekil 7.4 Histogram dönüşüm eğrisi.....	62
Şekil 8.1 Kanonik yüz.....	63
Şekil 8.2 Kullanılan yüz setlerinden örnekler	63
Şekil 8.3 Hazırlanan setten örnekler	64
Şekil 8.4 DFFS ve DIFS.....	65
Şekil 8.5 K-means ile gruplanmış örnekler.....	66
Şekil 8.6 Toplu eğitim.....	67
Şekil 8.7 İteratif eğitim	68
Şekil 8.8 Duyarlılık özgüllük eğrisi	68
Şekil 8.9 Çoklu yüz sezimi.....	69
Şekil 8.10 Görüntünün nokta nokta taranması	74
Şekil 8.11 Hız kazancı	75
Şekil 8.12 Örneklenmiş tanjant sigmoid fonksiyonu.....	75
Şekil 8.13 Yüzün belirlenmesi ve takip edilmesi	77
Şekil 9.1 PICs yüzsetinden örnekler	78
Şekil 9.2 CMU-130 yüzsetinden örnekler.....	78
Şekil 9.3 Oluşturulan setten örnekler.....	79
Şekil 9.4 CMU-130 sonuçları.....	82
Şekil 9.5 Oluşturulan setten sonuçlar.....	83

ÇİZELGE LİSTESİ

Çizelge 7.1 Önişlem süreleri	62
Çizelge 8.1 Aktivasyon fonksiyonu hesabı için geçen süreler.....	76
Çizelge 9.1 Yüz bulma yöntemleri ve sonuçlar.....	80

ÖNSÖZ

Danışmanın Sayın Prof.Dr.Tülay YILDIRIM’a tez çalışmam boyunca beni yönlendirdiği, bilgi ve görüşlerini bana aktardığı ve sabrı için teşekkür ederim.

Hayatım boyunca desteğini hissettiğim sevgili anneme teşekkür ederim.

ÖZET

Son yıllarda bilgisayarların işlem gücünün artması ve kamera sistemlerinin fiyatlarının düşmesi nedeniyle daha kolay elde edilebilir olması görsel yolla insan bilgisayar iletişimini daha popüler ve uygulanabilir bir hale getirmiştir.

Bu çalışmada video görüntülerinde insan yüzlerinin bulunması ve gerçek zamanlı izlenmesi problem olarak ele alınmıştır. Bilgisayar aracılığıyla dijital görüntülerde insan yüzlerinin ve yüzle ilgili özelliklerin bulunması insan bilgisayar arayüzlerinin oluşturulmasında, güvenlik ve izleme sistemlerinde, içerik tabanlı sorgulama sistemlerinde, video konferans sistemlerinde ve otomatik yüz tanıma sistemleri gibi birçok uygulamada kullanım alanı bulmaktadır.

Yapılan çalışmada kamerayla alınan video görüntülerinde insan tenine ait olduğu düşünülen bölgelerde veya resmin tamamında nöral tabanlı bir filtreyle insan yüzleri aranmış ve renk tabanlı bir algoritmayla takip edilmiştir. Nöral tabanlı filtre olarak çok katmanlı algılayıcı yapısında bir yapay sinir ağı kullanılmıştır.

Anahtar Kelimeler : Yüz İzleme, Yapay Sinir Ağları, Ten Sınıflayıcı

ABSTRACT

Especially lately, since computers became stronger and cheaper and camera equipments as well, the field of human computer communication has been a populer area to deal with.

In this work, it has been worked on finding and tracking human faces over video sequences. Finding human faces in digital images plays an important role in security and surveillance systems, building human computer interfaces, video communication systems and automatic face recognition systems.

In this thesis, human faces, which taken by a web cam, is found and tracked over sequences. A neural based filter is used to find human faces and a skin classifier is used for tracking. The neural based filter is a Multilayer Perceptron.

Key words: Face Tracking, Artificial Neural Networks, Skin Classifier

1. GİRİŞ

Takip işlemi kendisinden önce bir belirleme yani tanıma işlemine ihtiyaç duyar. Takip edilecek olan nesne veya örnek önce bulunmalı ve yeri belirlendikten sonra takip edilmelidir. Bu çalışmada hedef nesne olarak insan yüzleri seçilmiştir.

Yüz belirleme temel olarak dijital görüntülerdeki insan yüzlerinin arka plandan ayrılarak yerlerinin belirlenmesidir. Makina öğrenmesi açısından birini insan yüzlerinin ve diğerinde yüz olmayan örneklerin oluşturduğu iki sınıflı bir tanıma problemi olarak görülebilir. İnsanlar için oldukça kolay olan bu işlem insan yüzünün dinamik bir nesne olması yani görüntüsünde, ifadesinde, duruşunda çok fazla değişikliğin olması ve değişen ışık koşulları gibi nedenlerden dolayı bilgisayar görmesi açısından problemlili bir konudur. Başarılı bir yüz belirleme sisteminden değişen ışık koşulları, arka plan, kameraya olan uzaklık, yüzün pozisyonu, yüzün ifadesi ve sakal,bıyık,gözlük gibi etkenlerden etkilenmeksizin başarılı sonuçlar vermesi beklenmektedir.Şu an için insan yüzlerinin dış faktörlerle ilgili kısıtlamalar getirilmeden hızlı ve başarılı bir şekilde bulunması aktif bir çalışma konusudur.

Bilgisayar aracılığıyla dijital görüntülerde insan yüzlerinin ve yüzle ilgili özelliklerin bulunması insan bilgisayar arayüzlerinin oluşturulmasında, güvenlik ve izleme sistemlerinde, içerik tabanlı sorgulama sistemlerinde, video konferans sistemlerinde ve otomatik yüz tanıma sistemleri gibi birçok yerde kullanım alanına sahiptir.

Yüz tanıma algoritmaları arka plandan ayrılmış ve aynı boyutlardaki insan yüzleri üzerinde çalışırlar.Bu sebeple ilk önce gerçek dünyadan alınan görüntülerde insan yüzleri bulunmalı ve sonrasında tanıma işlemine geçilmelidir.Motor nöron rahatsızlıkları sebebiyle hareket edemeyen felçli hastaların yüz ifadelerindeki değişim bilgisayarla iletişim kurmak için kullanılabilir.Hastaların göz kırpmaya sıklıklarına göre bilgisayar başka komutları işletmektedir.Birçok uygulamada ilk önce yüzün yeri belirlendikten sonra gözler bulunmaktadır (Morris vd.,2002).Yüz bulmanın kullanıldığı bir diğer alan ise içerik tabanlı sorgulama sistemleridir.Yüz bulma algoritmaları sayesinde “İçinde insan bulunan görüntüleri getir.” şeklindeki bir sorguya sistem yanıt verebilir.

Yüz bulma algoritmaları temel olarak süre kısıtlamasının olmadığı durağan görüntü çerçeveleri veya gerçek zamanlı video görüntüleri üzerinde çalışmaktadır.

Bugüne kadar insan yüzlerinin bulunması için kenar belirleme gibi düşük seviyeli analiz metodlarından yüksek seviyeli örnek sınıflama metodlarına kadar birçok teknik önerilmiştir. Önerilen yöntemler problemi ele alış biçimlerine göre özellik ve görüntü temelli yöntemler

olmak üzere iki ana başlık altında toplanabilir.

1.1 Özellik Tabanlı (Feature Based) Yöntemler

Bu tür yöntemler renk, doku, kenar ve hareket gibi bilgileri kullanarak yüze ait göz, ağız gibi özellikleri arayan ve bu sayede yüz bulmayı hedefleyen düşük seviyeli analiz yöntemleridir. Genel olarak basit olmaları, hızlı sonuç vermeleri ve yüzün duruşundan görüntü temelli yöntemler kadar etkilenmemelerine karşın güvenilirlikleri daha azdır. Ancak belli kısıtlamaların getirildiği durumlarda başarılı ve hızlı sonuçlar vermektedirler.

1.2 Görüntü Tabanlı (Image Based) Yöntemler

İnsan yüzü gibi dinamik nesneler etkin olarak özellik çıkarmaya uygun değildirler. Temel olarak istatistiksel sınıflayıcılar, yapay sinir ağları gibi örnek temelli öğrenim yöntemlerinin ve şablon eşlemenin kullanıldığı yöntemlerdir. İnsan yüzleri iki boyutlu matrisler şeklinde bir bütün olarak ele alındığından bu tür yöntemler görüntü tabanlı olarak adlandırılmaktadır.

İstatistiksel sınıflayıcılar ve yapay sinir ağları bir olay hakkında karar verebilmek için olayı temsil eden yeterli sayıda örneği kullanarak örneklerden olay hakkında sonuçlar çıkaran yöntemlerdir. Olay bir nesne tanıma problemiye sınıflayıcıya çok sayıda nesne ile ilgili örnek (pozitif örnekler) ve ilgi duyulan nesne olmayan bir çok örnek (negatif örnekler) gösterilmelidir. Yüz bulma problemi ele alındığında insan yüzleri pozitif örnekleri oluştururken yüz olmayan her örnek negatif örnekleri oluşturur. Her ikisi beraber ise eğitim setini oluşturur.

Bir sınıflayıcıya gösterilen örneklerin kalitesi sınıflayıcının başarısını direkt olarak etkilemektedir. İnsan yüzlerinin çok fazla esnekliğe sahip dinamik nesneler oluşundan ve yüz olmayan örneklerin kendilerine has belirleyici özelliklerinin olmayışından yani grup içi dağılımlarının (within class variability) yüksek olmasından dolayı klasik görüntü işleme algoritmalarıyla özellik çıkarmaya uygun değildirler. Bu sebeple birçok görüntü temelli çalışmada yüz ve yüz olmayan örnekler yalnızca görüntü işleme metodlarıyla iyileştirilip (önişleme) bir bütün olarak sınıflayıcılara sunulup sınıflayıcıların kendilerinin örnekler hakkında ayırd edici özellikleri çıkararak bir sonuca varması hedeflenmiştir. İstatistiksel sınıflayıcılar ve yapay sinir ağları eğitimleri sırasında farklı özellikler çıkarır ve ayırd edici özellikleri seçerek kendi model parametreleri içine sıkıştırırlar (Bradski, 2005).

Görüntü tabanlı yöntemlerin çoğunda küçük boyutlardaki bir pencere belli oranlarda küçültülmüş görüntü çerçeveleri üzerinde gezdirilerek (window scanning) insan yüzleri aranır. İnsan yüzlerini belirlemek için resimden alınan parçalar öğrenilen veya insan eliyle hazırlanan modellerle karşılaştırılır.

Görüntü tabanlı yaklaşımlar yüzün rotasyonuna ve pozuna karşı oldukça duyarlıdır. Bu sebeple yapılan çalışmaların çoğunda yüzün yalnız önden cepheden görünümüne duyarlı tek bir sınıflayıcı kullanılmıştır. İnsan yüzlerini değişik açılardan tespit edebilmek için herbiri yüzün belli bir açıdan görüntüsüne duyarlı birden çok sınıflayıcı eğitilmelidir.

Görüntü tabanlı yöntemlerde kendi içinde genel olarak dört gruba ayrılabilir.

- **Şablon Eşleme (Template Matching):** En basit ve eski resim tabanlı yöntem olup temelde gri seviyeli insan yüzlerinin ortalaması alınarak oluşturulmuş bir şablon yüz, resim üzerinde gezdirilerek insan yüzleri bulunmaya çalışılır. Önceden oluşturulmuş sabit bir yüz şablonu (predefined face template) yerine esnekliğe sahip deforme olabilen (deformable face template) şablonların kullanılması başarıyı arttırmıştır.

Şablonla resimden alınan pencereler arasındaki benzerliğin ölçüsü olarak korelasyona bakılır.

- **Doğrusal Alt Uzak Metodları (Linear Subspace Methods) :** Gri seviyeli yüz resimleri çok boyutlu tüm resim uzayı içinde bir alt uzayı oluştururlar. Temelde bütün sınıflayıcılar ile bu alt uzay temsil edilmeye çalışılır. PCA (Principal Component Analysis), LDA (Linear Discriminant Analysis) ve FA (Factor Analysis) insan yüzlerinin oluşturduğu alt uzayı temsil etmek için kullanılan yöntemlerdir.

- **İstatistiksel Yaklaşımlar (Statistical Approaches) :** Destek vektör makinaları ve Bayes sınıflayıcı örnek olarak verilebilir. Özellikle destek vektör makinalarının iki sınıflı problemlerin çözümünde başarı sağladığının görülmesinden sonra çok sayıda yüz bulma algoritmasında kullanılmıştır.

- **Yapay Sinir Ağları (Artificial Neural Networks):** Yapay sinir ağları genel olarak matematiksel modeli oluşturulamayan veya modellemenin zor olduğu uygulamalarda başarılı sonuçlar veren, karmaşık doğrusal olmayan giriş çıkış ilişkilerini kendine gösterilen örneklerden yararlanarak kuran, insan beyninden esinlenerek tasarlanmış yapılardır.

Yapay sinir ağları özellikle süre sınırlamasının olmadığı birçok uygulamada kullanılmış ve başarılı sonuçlar elde edilmiştir.

Genel olarak görüntü tabanlı yöntemler özellik tabanlı yöntemlere göre daha başarılı olmalarına karşın çok fazla işlem yükü gerektirdiklerinden oldukça yavaşlardır. Bu yüzden birçok gerçek zamanlı uygulamada her iki yöntem beraber kullanılmaktadır.

1.3 Yüz Belirleme ve İzleme Problemine Getirilmiş Çözümler

Bugüne kadar durağan resimlerde veya video görüntülerinde insan yüzlerinin bulunması ve izlenmesi için özellik, görüntü veya her iki yaklaşımında kullanıldığı birçok çalışma yapılmıştır.

(Sung ve Poggio, 1996)' da daha sonra birçok çalışmada kullanılan bir önileme ve eğitim tekniği önerilmiştir. Küçük boyutlardaki her iki sınıfı oluşturan örneklerden kendilerine en iyi uyan parlaklık düzlemi çıkarıldıktan sonra histogram eşitleme yöntemi uygulanarak örnekler önilemden geçirilmiştir. Yaklaşık 46.000 yüz ve yüz olmayan örnek toplanmıştır. Yüz olmayan örnekleri etkin bir biçimde seçebilmek için önceden elle oluşturulan örneklerle eğitilmiş sınıflayıcının hatalarının toplanarak kendisine tekrar gösterildiği Bootstrap adlı bir algoritma önerilmiştir. Tüm görüntü uzayı içerisinde yüz olan ve olmayan örneklerin dağılımlarını modellemek için çok boyutlu Gauss kümeleri (Gaussian clusters) oluşturulmuştur. 6 tanesi yüz ve diğer 6'sı yüz olmayan örnekleri temsil eden 12 tane küme eliptik k-means kümeleme algoritması kullanılarak oluşturulmuştur ve her küme merkez noktası ve kovaryans matrisiyle ifade edilmiştir. Resimden alınan her bir görüntü parçası ile her bir küme arasındaki uzaklık iki farklı uzaklık yöntemiyle (mahalanobis, öklit) hesaplandıktan sonra elde edilen 12 tane uzaklık çiftini gösteren özellik vektörü çok katmanlı bir yapay sinir ağına sunularak karar verilmiştir.

(Cai vd., 1996)' da insan tenine ait olan bölgeleri çıkarmak için parametrik olmayan bir renk sınıflayıcı kullanılmıştır. Sonrasında insan teni olduğu düşünülen bölgelerin boyutlarına ve içlerinde boşluk olup olmadığına bakılarak (göz, ağız) yüz olmaya aday olanlar seçilmiştir. Seçilen bu bölgeler şalon eşleme ile kontrol edilerek insan yüzlerini ön cepheden ve profilden bulan bir sistem önerilmiştir.

(Yang ve Waibel., 1996)' da değişik kişi ve ışık koşullarına uyum sağlayabilen adaptif bir ten sınıflayıcı kullanarak ten bölgeleri belirlenmiş ve hareketli en büyük bölge yüz olarak kabul edilmiştir. Görüntü çerçeveleri arasında hareket kestirimi yapılarak insan yüzleri bütün çerçeve yerine daha küçük bir alanda aranarak takip edilmiştir.

(Mottaleb vd., 2002) ' da ilk olarak resimlerdeki ışık etkisini tolere etmek için geliştirilen bir

ışık dengeleme tekniğiyle resimdeki renk bileşenleri ayarlanmış ve daha sonra bir ten sınıflayıcı yardımıyla ten bölgeleri bulunmuştur. Ten olarak sınıflanan bölgelerde göz, ağız ve sınır haritaları oluşturulduktan sonra Hough transformuyla yüz ovali bulunmuştur.

(McKenna ve Gong, 1996)'da görüntü çerçevelerinde hareket eden kenarlar bulunarak insan yüzleri bu kenarlar etrafında oluşturulan sınır çerçevelerinde aranmıştır. Bu sayede sistem gerçek zamanlı çalışır duruma getirilmiş ve sınıflayıcının başarısı artırılmıştır. Sınıflayıcı olarak çok katmanlı bir yapay sinir ağı kullanılmış ve ağa sunulacak örnekler üzerindeki önileme süresini kısaltmak için önileme algoritması olarak standart sapma ile normalizasyon kullanılmıştır.

(Rowley vd., 1998)'da insan yüzlerini bulmak için herbiri lokal görüş alanına sahip (retinally connected) nöronlardan oluşan bir veya birbirine paralel daha fazla sayıda çok katmanlı ağ kullanılmış ve ağların çıkışları "and" veya "or" işlemlerine tabi tutularak sistemin performansı artırılmıştır. Önileme aşamasında ve yüz olmayan örneklerin toplanmasında Sung'ın önerdiği yöntemler kullanılmıştır. Önerilen yöntem oldukça başarılı sonuçlar vermesine karşın gerçek zamanlı çalışmaktan uzaktır.

(Yacoub vd., 1999)'da çok katmanlı yapay sinir ağlarının daha hızlı sonuç vermesi için bir yöntem önerilmiştir. Resimle ağın gizli katmanı arasındaki işlemler frekans uzayında yapılmış ve buna uygun önileme algoritması seçilerek küçük boyutlardaki video görüntülerinde (182*144) gerçek zamanlıya yakın yüz takibi yapılmıştır.

(Viola ve Jones, 2001)'da başarılı ve gerçek zamanlı çalışan bir yöntem önerilmiştir. Oldukça basit, hızlı ve sınıflama yetenekleri zayıf olan çok sayıda Haar-like özelliklerini kullanan sınıflayıcı birbirine seri olarak kullanılmıştır. Bu sayede diğer birçok görüntü tabanlı yöntemin aksine resmin her noktası kompleks ve uzun sürede sonuç veren bir sınıflayıcıyla taranmak yerine yüz olması muhtemel olmayan yerler birbirini takip eden basit sınıflayıcılar tarafından daha ilk aşamalarda elenmiştir. Bu şekilde birçok zayıf sınıflayıcının ard arda kullanılması reddetme temelli (rejection based) yöntemleri oluşturmaktadır.

Bu çalışmada ise video görüntülerinde insan yüzlerinin bulunması için çok katmanlı bir yapay sinir ağı kullanılmıştır. Ağın işlem hızını arttırmak için resimle ağın giriş katmanı arasında yapılan konvolüsyon işlemi aynı zamanda frekans uzayında da yapılmıştır. İnsan yüzlerinin çok sayıda, giderek azalan boyutlardaki görüntü çerçevelerinin her noktasında aranması, örneklerin özellik çıkarılmaksızın çok boyutlu ham bilginin ağa sunulması, özellik çıkarılması durumunda ise özellik çıkarma yöntemlerinin ve ağı oluşturan yapay sinir birimlerinin

aktivasyon fonksiyonlarının hesaplanmasının uzun zaman almasından dolayı nöral tabanlı bir filtreyle insan yüzlerinin gerçek zamanlı olarak bulunması bilgisayarların bugünkü işlem gücüne rağmen oldukça zordur. Bu sebeple video görüntülerinde insan teni olduğu düşünülen bölgelerde veya resmin tamamında yüzün yeri belirlendikten sonra renk tabanlı bir algoritmayla gerçek zamanlı olarak takip edilmiştir.

2. GÖRÜNTÜ İŞLEME

2.1 Renk ve Algılama

Renk fiziksel değil algısal bir olgudur ve retinayı uyaran elektromanyetik radyasyonun spektral karakteristiğiyle ilişkilidir. İnsan gözü yaklaşık olarak 370 nm ile 670 nm arasındaki dalga boyuna sahip radyasyona duyarlıdır ve elektromanyetik tayfın bu dar bölgesi ışık olarak adlandırılır.

Spektral güç dağılımı ışığı oluşturan frekans bileşenlerini, yani herbir dalga boyundaki radyasyonun genliğini verir. Var olan bütün renkler spektral güç dağılımı ölçülerek hesaplanabilmesine karşın ortaya gereğinden fazla renk çıkmasından dolayı tercih edilmezler. İnsan retinasında renkli görmeden sorumlu reseptörler, yaklaşık olarak mavi, yeşil ve kırmızı ışığa karşı düşen alçak, orta ve yüksek dalga boylarında duyarlı band geçiren filtre özelliği göstermelerinden dolayı insan retinası renkleri örnekler. Işığın spektral dağılımdaki her değişiklik bu sebepten dolayı ayrı bir renk olarak algılanamaz, yani birbirine yakın spektral özellikleri gösteren elektromanyetik radyasyonlar aynı renge gruplanır. Bu sebeple renklerin spektral olarak ifade edilmeleri çok uygun olmadığından renkler renk uzayları kullanılarak ifade edilirler (Ford, 1998).

Renk uzayları renkleri tanımlamak için kullanılan matematiksel modellerdir. Grassman'ın kanununa göre diğer ikisi üçüncü rengi oluşturmamak şartıyla yalnız üç ana renkle diğer birçok renk oluşturulabilmektedir, bu yüzden birçok renk uzayı üç boyutludur ve renkler bu uzaylar içinde sayısal değerlerle ifade edilirler. Renklerin bu şekilde sayısallaştırılması renklerin insanlar veya makineler arasında tanımlanabilmesini, taşınabilirliği sağlamıştır.

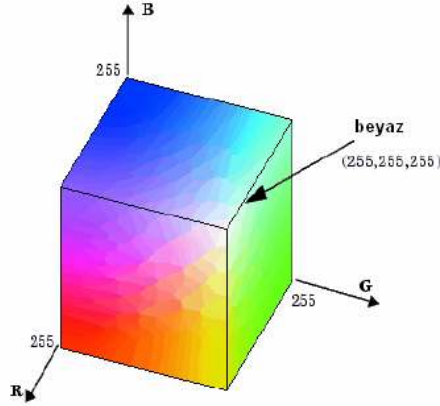
Değişik amaçlarla üretilmiş ve aralarında doğrusal veya doğrusal olmayan dönüşümlerin yapılabildiği çok sayıda renk uzayı bulunmaktadır. Yapılacak olan işe göre doğru uzayın seçilmesi oldukça önemlidir.

Renk uzayları cihaz bağımlı ve cihaz bağımsız olmak üzere ikiye ayrılırlar. Cihaz bağımlı renk uzaylarında üretilen renkler kullanılan cihazın teknik özelliklerine bağlı iken, cihaz bağımsız renk uzayları CIE (Commission Internationale de L'Eclairage :Uluslararası Aydınlatma Komisyonu) tarafından cihazların sınırlayıcı etkilerini göz ardı ederek insan gözünün görebildiği tüm renkleri ifade etmek için geliştirilmiş uzaylardır. Cihaz bağımlı uzaylara örnek olarak RGB, CMYK, YCbCr, HSV uzayları, cihaz bağımsız uzaylara ise CIE XYZ ve CIE Lab uzayları örnek olarak verilebilir.

2.2 Renk Uzayları

2.2.1 RGB

RGB renk uzayı koordinat eksenleri kırmızı, yeşil ve mavi olan üç boyutlu bir uzaydır ve toplamalı renk karışımı yöntemiyle bir birim küpün içinde renkleri tanımlayacak şekilde tasarlanmıştır. Televizyon, bilgisayar monitörleri, kameralar ve tarayıcılar gibi cihazlar bu renk uzayını kullanırlar.



Şekil 2.1 RGB uzayı

Eğer görünen spektrum retinadaki reseptörler tarafından algılandığına benzer bir şekilde algılanırsa, yani insan görmesinin ilk aşaması taklit edilirse spektrumla ilgili bilgi saklanabilir, işlenebilir ve görselleştirilebilir. Bu yüzden kameralar gibi çoğu görüntüleme cihazları retinadaki renk hücrelerine benzer şekilde yapay reseptörler kullanırlar ve bu reseptörler genellikle RGB olarak isimlendirilirler.

Herbir reseptör uyarıldığında, çıkışı gelen ışığın spektral dağılımı ve reseptörün duyarlılık fonksiyonunun çarpımlarının görünün dalga boyu sınırları içerisindeki toplamına eşittir. (2.1), bu dönüşüm integrallerini göstermektedir.

$$\begin{aligned}
 R &= \int_{380}^{780} S(\lambda) R(\lambda) d\lambda \\
 G &= \int_{380}^{780} S(\lambda) G(\lambda) d\lambda \\
 B &= \int_{380}^{780} S(\lambda) B(\lambda) d\lambda
 \end{aligned} \tag{2.1}$$

Reseptörlerin duyarlılık fonksiyonları aslında band geçiren bir filtrenin frekans yanıtı olarak

görülebilir. Spektral güç dağılımından üç boyutlu RGB vektörüne dönüşüm aslında etkili fakat kayıplı bir sıkıştırma tekniğidir. Kaybın nedeni değişik renklerdeki spektral dağılımların algısal olarak aynı etkiyi yaratabilmelerinden kaynaklanmaktadır.

(2.1) ‘deki denklemlerden görüldüğü gibi RGB değerleri cihazın reseptörlerinin duyarlık fonksiyonlarına bağlıdır, bu sebeple RGB renk uzayı cihaz bağımlı olarak isimlendirilir. RGB uzayı temel bir uzay olup diğer uzaylara doğrusal veya doğrusal olmayan şekillerde dönüştürülebilir.

Temel bir uzay olmasına ve çok sık kullanım alanı olmasına rağmen RGB uzayının bazı dezavantajları aşağıda sıralanmıştır.

- RGB renk uzayının bileşenleri arasında, reseptörlerin duyarlılık fonksiyonlarının iç içe girmesinden dolayı, korelasyonun yüksek olması
- Uzayın algısal olmayışı, yani insanların verilen RGB değerlerinin ifade ettiği rengi canlandırmada zorluk çekmeleri
- Algısal olarak birörnek (uniform) olmaması, yani iki renk arasındaki algısal farkla uzay içerisindeki farkın (öklit uzaklığı) arasındaki korelasyonun düşük olması
- Parlaklık bilgisiyre renk bilgisinin beraber kodlanmış olmalarıdır.

2.2.2 Normalize RGB

RGB uzayında herbir nokta renk bilgisiyre beraber parlaklık bilgisini de içerdiğinden bir cisim üzerine düşen ışığın spektral karakteristiği değişmeksizin yalnız şiddeti değiştiğinde bile cismi ifade eden piksellerin RGB uzayındaki yeri önemli miktarda değişir, yani cisim aynı kalmasına karşın cismin renginin değişmiş olduğu sanılır. Işık etkisiyle olan bu değişimi azaltmak için herbir renk bileşeni toplam renk bilgisine bölünerek normalize değerleri bulunur.

$$\begin{aligned} r &= \frac{R}{R + G + B} \\ g &= \frac{G}{R + G + B} \\ b &= \frac{B}{R + G + B} \end{aligned} \tag{2.2}$$

$$r + g + b = 1$$

2.2.3 CMY

CMY renk uzayı çıkarmalı bir renk uzayıdır ve RGB uzayının ikincil renkleri olan turkuaz, siyah ve sarı bu uzayın ana renklerini yani eksenlerini oluştururlar. Yazıcılar ve çiziciler gibi baskı cihazları bu uzayı kullanırlar. Herbir ana renk birer band durduran filtre görevi görerek yalnız bir rengi emer ve bu sayede değişik renkler oluşturulur. Her üç rengin karışımı siyah rengi oluşturmasına karşın pratik nedenlerden dolayı baskı cihazlarına ayrıca birde siyah renk eklenir.

CMY ile RGB uzayları arasındaki dönüşüm (2.3) ' de verilmiştir.

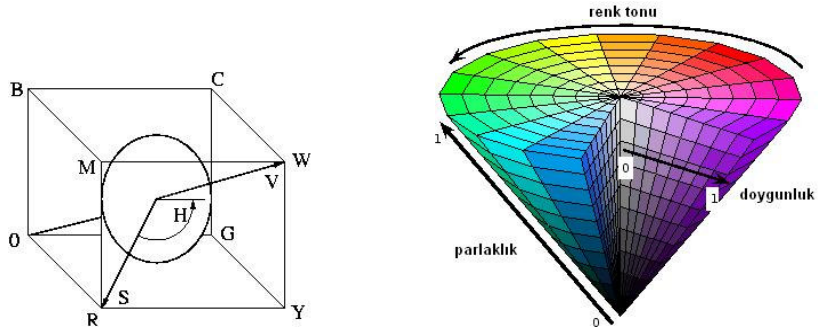
$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.3)$$

(2.3) eşitliğinde beyaz renk 1 rakamı ile temsil edilmiştir.

2.2.4 HSV

İlk olarak Newton renkleri, parlaklık değerlerini ihmal ederek, renk tonu (hue) ve doygunluk (saturation) özelliklerine göre bir daire üzerine yerleştirdi. Daha sonraları insan beyninin renkleri renk tonu, doygunluk veya saflık ve parlaklık özelliklerine göre organize etmeye eğilimli olduğu anlaşıldı. İnsanlar tarafından yapılan açık turkuaz gibi bir renk tanımlaması bu gerçeği göstermektedir.

HSV renk uzayı kartezyen koordinatlardaki RGB değerlerinin doğrusal olmayan bir biçimde yeniden düzenlenmesiyle oluşturulur. RGB renk uzayı içinde HSV değerleri ana diagonal etrafında belirlenmektedir. Şekil 2.2'de soldaki şekilde görüldüğü gibi RGB uzayındaki bir noktanın renk tonu (H) ana diagonalı dik bir daire üzerinde bulunduğu yerin açısı olarak, doygunluk (S) değeri ana diagonaldan dik uzaklığı olarak ve son olarak parlaklık (V) değeri de noktanın diagonal üzerine izdüşümünün RGB uzayının orjiniinden uzaklığı olarak hesaplanır. Şekil 2.2'de sağdaki şekilde HSV uzayının koni şeklindeki yapısı gösterilmiştir, HSV uzayı ayrıca çember ve silindir şeklinde de gösterilebilmektedir.



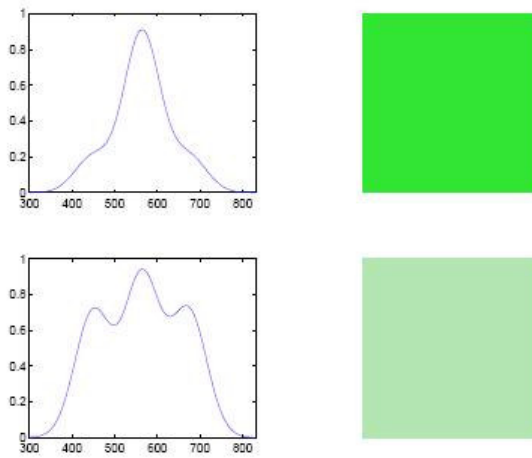
Şekil 2.2 HSV uzayının koni biçimindeki yapısı

Renk tonu, baskın rengi tanımlayan bir özelliktir. Açısal bir düzlemde yer aldığı için 0^0 ile 360^0 arasında değerler almasına karşın işlem kolaylığı açısından $[0,1]$ veya $[0,255]$ aralıklarına çekilmektedir.

Doygunluk, rengin beyazdan veya gri tonlardan ne kadar uzak olduğunun yani saflığının bir ölçüsüdür. Spektrumu tek bileşenden oluşan renkler (monochromatic) en yüksek saflığa sahip renklerdir.

Parlaklık, Işığın genliğinin bir ölçüsüdür.

HSV değerleriyle ışığın spektral dağılımı arasında tam bir ilişki kurmak mümkün olmamakla birlikte kabaca H değeri ışığın baskın dalga boyunun, S değeri ise ışığın baskın dalga boyundan ne kadar saçıldığının bir ölçüsüdür. Şekil 2.3’ de her iki grafikte işaretlerin baskın bileşenleri spektrumun yeşil renge karşılık gelen bölgesinde olduğundan her iki renkte yeşildir, fakat üstteki grafikte gösterilen işaretin dağılımı alttaki işarete göre daha az olduğundan doygunluk değeride daha yüksektir.



Şekil 2.3 Spektrum ile HS değerleri arasındaki ilişki

RGB uzayından HSV uzayına dönüşüm ise (2.5)'den başlayan işlemler gerçekleştirilir.

$$H = \begin{cases} \theta & ,eger \quad B \leq G \\ 360 - \theta & ,eger \quad B > G \end{cases} \quad (2.4)$$

Olmak üzere açısai renk tonu, doygunluk ve parlaklık değeri sırasıyla (2.5), (2.6) ve (2.7) denklemleriyle bulunur. RGB değeriinin [0,1] aralığına normalize edildiği kabul edilmiştir.

$$\theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R-G) + (R-B)]}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right\} \quad (2.5)$$

$$S = 1 - \frac{3}{(R+G+B)} [\min(R, G, B)] \quad (2.6)$$

$$V = \frac{R+G+B}{3} \quad (2.7)$$

2.2.5 YCbCr

YCbCr uzayı sayısal televizyon ve video uygulamalarında sıklıkla kullanılmaktadır.Y parlaklık Cb ve Cr ise renk bilgilerini içerir.Cb mavi renk bileşeniyle referans değeri arasındaki farkı, Cr ise kırmızı renk bileşeniyle referans değeri arasındaki farkı temsil eder.

YCbCr ile RGB uzayları arasındaki dönüşüm (2.8)'daki gibidir.

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ Cb &= B - Y \\ Cr &= R - Y \end{aligned} \quad (2.8)$$

2.2.6 Karşıt Renk Uzayları

18. yüzyıl sonlarında Alman fizyolog Ewald Hering karşıt renkler teorisini (opponent colours theory) ortaya atmıştır.Hering üç renklilik teorisinin (trichromatic theory) aksine , yaptığı deneyler sonucunda, insan retinasında siyah ve beyaz, mavi ve sarı, kırmızı ve yeşil arasındaki farklara duyarlı üç tip reseptör olduğunu ve üç karşıt bilginin renkli görmeyi sağladığını söylemiştir.20. yüzyıl başlarında bu iki teori tartışılmasına karşın daha sonraları retinadaki reseptörlerin Hering'in iddia ettiğinin tersine,Young ve Helmholtz tarafından Maxwell'in deneylerine dayanarak öne sürülmüş olan üç renklilik teorisine uygun olarak alçak, orta ve yüksek dalga boylarına duyarlı üç tipte oldukları anlaşılmıştır.Bu üç tip

reseptör çoğu zaman duyarlı oldukları dalga boyuna göre S (Short), M (Middle) ve L (Long) harfleriyle gösterilmektedir. Yaklaşık olarak S reseptörleri maviye, M reseptörleri yeşile ve L reseptörleride kırmızı renge karşılık gelmektedir. 1964 yılında retinadaki renkli görmeden sorumlu konik hücrelerin görüntülenebilmiştir.

Günümüzdeki modern karşıt renk teorisi birbirine zıt bu iki teorisin beraber nasıl işlediğini açıklayabilmektedir. Buna göre renkli görmenin ilk bölümünü oluşturan reseptörler Young ve Helmholtz'un önerdiği gibi üç renklidir, buna karşın üç renklilik teorisinin aksine reseptörlerden alınan sinyaller beyne direkt olarak yollanmaz, retinadaki sinir hücreleri tarafından karşıt sinyallere kodlanırlar. Üç tip reseptörün çıkışları parlaklık bilgisini oluşturmak için toplanır (L+M+S) renk bilgisini oluşturmak için ise kırmızı-yeşil (L-M) ve mavi-sarı (S-L+M) farkları alınarak beyne iletilir.

LMS sinyalinden karşıt sinyale dönüşüm reseptörlerden alınan üç kanallı bilgideki korelasyonu azaltır ve bu sayede daha etkin bir sinyal iletimi sağlanmış olur.

Şu an için bu teoriyi temel alan birçok karşıt renk teorisi ve bilgisayar görmesi alanında uygulamaları vardır. RGB uzayından karşıt renk uzayına basit bir dönüşüm (2.9)' de gösterilmiştir.

$$\begin{aligned} RG &= R - G \\ YeB &= 2B - R - G \\ WhBl &= R + G + B \end{aligned} \quad (2.9)$$

Bir diğer dönüşüm ise reseptörlerin logaritmik cevapları gözönüne alınarak (Fleck vd, 1996) tarafından önerilmiştir.

R_g ve B_y değerlerini $[0, 255]$ aralığında ifade edebilmek için direkt logaritma yerine (2.10)'daki eşitlik kullanılmıştır.

$$\begin{aligned} R_g &= \log(R) - \log(G) \\ B_y &= \log(B) - \left(\frac{\log(R) + \log(G)}{2} \right) \\ I &= \log(G) \end{aligned} \quad (2.10)$$

n , 0 ile 1 arasında keyfi olarak seçilebilen bir değerdir.

2.3 Matematiksel Morfoloji

Matematiksel morfoloji, görüntü işleme uygulamalarında kullanılan, basit küme işlemlerine

dayanan birçok yöntem sunar.

Morfolojik işlemler çoğunlukla ikili resimlerde (binary image), kenar bulma, gürültü giderme, görüntü iyileştirme, bölütleme, inceltme ve budama gibi amaçlarla kullanılırlar.

2.3.1 Temel Küme İşlemleri

A ve B 'nin iki boyutlu uzayda (Z^2) birer küme gösterdiğini kabul edelim. A ve B kümelerinin her elemanı iki değer içermektedir.

B kümesinin x ile ötelenmiş hali (2.11)'deki gibidir ve B_x ile gösterilir.

$$B_x = \{c : c = b + x, \quad b \in B\} \quad (2.11)$$

B kümesinin yansıması ise $\hat{B}$ ile gösterilir.

$$\hat{B} = \{x : x = -b, \quad b \in B\} \quad (2.12)$$

Eğer A kümesi B kümesinin bir alt kümesiye (2.13)'deki gösterim doğru olur .

$$A \subseteq B \quad (2.13)$$

C iki kümenin bileşimini gösteren bir küme ise (2.14)'daki gibi gösterilir.

$$C = A \cup B \quad (2.14)$$

D iki kümenin ortak elemanlarından oluşan bir küme ise (2.15)'deki gibi gösterilir.

$$D = A \cap B \quad (2.15)$$

2.3.2 Morfolojik Görüntü İşleme

İkili görüntülerde genel olarak beyaz pikseller (1 veya 255 sayıları ile temsil edilirler) ön plan olarak kabul edilirken siyah pikseller (0 ile temsil edilirler) arka plan olarak kabul edilir. Resimdeki ön planı oluşturan piksellerin koordinatları bir kümeyi (A kümesi) ,yapı elemanın (structuring element) koordinatları ise diğer kümeyi (B kümesi) oluşturur. Yapı elemanı boyutları resimden küçük olan ve yapılan işlemin etkisini değiştiren iki boyutlu bir matris olarak düşünülebilir.

Aslında uygulamada yapılan bütün morfolojik işlemler görüntü üzerinde, ön planı oluşturan piksellerin resimdeki dağılımı ve yapısal eleman kullanılarak yapılmasına ve herhangi bir

küme oluşturulmamasına karşın, yapılan işlemi matematiksel olarak ifade edebilmek için küme teorisinden bahsedilmiştir.

Temel morfolojik işlemler aşındırma (erosion) ve yayma (dilation) olup diğer morfolojik işlemler bu iki işlem kullanılarak yapılır.

2.3.2.1 Aşındırma

Aşındırma işleminin temel iki amacı görüntüde, belirlenen sayıdan az sayıda piksele sahip nesneleri veya gürültüyü yok etmek ve birbirine değen nesneler varsa bu nesneleri incelterek birbirinden ayırmaktır.

Yapılan işlem matematiksel olarak (2.16)' de gösterilmiştir.

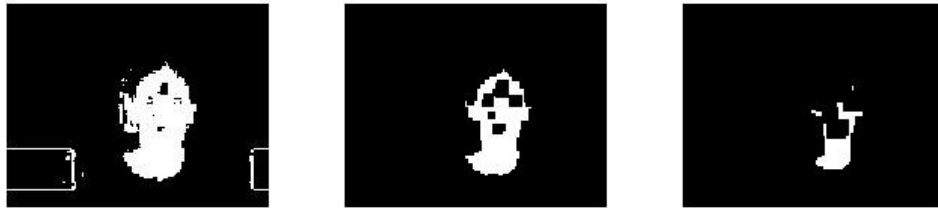
$$A \ominus B = \left\{ x : B_x^\wedge \subseteq A \right\} \quad (2.16)$$

Küme teorisi yönünden bakıldığında (2.16)'daki gibi ifade edilen işlem esasında konvolüsyona benzemektedir. Yapılan işlem aslında yapı elemanının simetriği alınıp resim üzerinde gezdirilmesi ve resminden alınan parçaya tam olarak uyup uymadığına bakılmasıdır.

Genellikle yapı elemanları kare matrisler olduklarından yansıtılmış halleri kendilerine eşittir.

$$B_x^\wedge = B \quad (2.17)$$

Şekil 2.4' de en solda orjinal görüntü, ortada 5*5 boyutlarında bütün elemanları 1 olan bir yapı elemanı ile elde edilen, en sağda ise aynı yapı elemanın boyutları 11*11 yapıldığında elde edilen sonuçlar gösterilmiştir.



Şekil 2.4 Aşındırma

2.3.2.2 Yayma

Yayma temel olarak görüntüdeki nesne veya nesneleri büyütme ve içlerindeki boşlukları

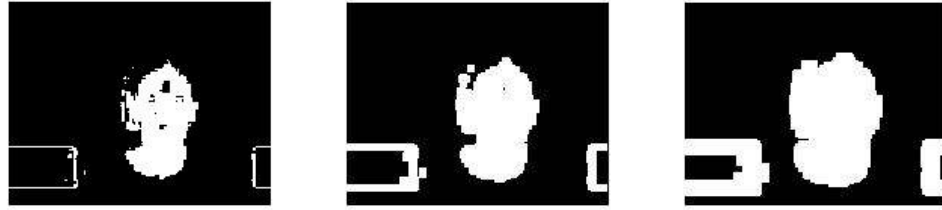
doldurmak için kullanılır.

Yapılan işlem matematiksel olarak (2.18)' gösterilmiştir.

$$A \oplus B = \left\{ x : \hat{B}_x \cap A \neq \emptyset \right\} \quad (2.18)$$

(2.18)'dan görüldüğü gibi yapı elemanı ile görüntüdeki nesneye ait pikseller karşılaştırıldığında eşit olan en az bir değer çıkması durumunda, yani ötelenmiş ve simetriği alınmış B kümesiyle A kümesinin kesişiminin boş küme olmadığı durumda, sonuç değeri 1'dir.

Şekil 2.5'de sırasıyla 5*5 ve 11*11 boyutlarında tüm elemanları 1 olan yapı elemanları kullanılarak elde edilen sonuçlar gösterilmiştir.



Şekil 2.5 Yayma

2.3.2.3 Açma ve Kapama

Açma ve kapama, yayma ve aşındırma işlemlerinin birbiri ardına uygulanmasıdır. Açma işlemi, görüntüye önce aşındırma sonrasında yayma işleminin uygulanması, kapama ise önce yayma sonrasında aşındırma işleminin uygulanması işlemleridir. Genel olarak açma işlemi görüntüdeki küçük nesneleri yok ederken, kapama işlemi küçük boşlukları doldurur.

Açma işlemi $A \circ B$ ile gösterilir ve (2.19)'de tanımlanmıştır.

$$A \circ B = (A \ominus B) \oplus B \quad (2.19)$$

Kapama işlemi ise $A \bullet B$ ile gösterilir ve (2.20)'de tanımlanmıştır.

$$A \bullet B = (A \oplus B) \ominus B \quad (2.20)$$

Şekil 2.6'da en solda orjinal görüntü olmak üzere, ortada açma ve sağda kapama işlemlerinin sonuçları gösterilmiştir. Kullanılan yapı elemanı 5*5 boyutlarında olup bütün elemanları 1'dir.



Şekil 2.6 Açma ve kapama

2.3.3 Bağlantılı Bileşenler Etiketleme (CCL, Connected Component Labeling)

Bağlantılı bileşenler etiketleme algoritması birbirine komşu olan piksellere aynı parlaklık değerini atayarak, birbirine komşu olmayan nesnelerin etiketlenmesini ve bu nesneler üzerinde şekil ve boyut analizi gibi işlemlerin yapılabilmesini sağlar. Etiketleme işlemi pikseller arasındaki 4 komşuluk (4 neighbors) veya 8 komşuluk (8 neighbors) ilişkisine bakılarak yapılır. Bir pikselin 4 komşuluğunu yatay ve dikeydeki komşuları oluştururken, bunlara diagonal komşuların eklenmesiyle 8 komşuluk elde edilir. P bir pikseli göstermek üzere, 4 ve 8 komşuluklarının oluşturduğu kümeler (2.21)'de gösterilmiştir.

$$N_4(p) = \{(x+1, y), (x-1, y), (x, y+1), (x, y-1)\}$$

$$N_8(p) = N_4 \cup \{(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)\} \quad (2.21)$$

CCL algoritması sıralı (sequential) ve tekrarlamalı (recursive) olmak üzere iki tiptir. Sıralı CCL algoritmasında görüntü en sol üst köşenden taranmaya başlanıp bağlantılı bileşenler aranmasına karşın tekrarlamalı CCL algoritmasında görüntü sıralı bir şekilde taranmaz. Tekrarlamalı bir CCL algoritması temel olarak aşağıdaki adımları izler.

- Orijinal görüntüdeki tüm 1 değerleri -1 yapılarak yeni bir görüntü oluşturulur.
- Görüntü sol üst köşenden başlayarak taranmaya başlar ve ilk -1 değerine ulaşıldığında etiket numarası bir artırılarak piksele atanır ve o piksel referans kabul edilir.
- Sonrasında yeni değer atanan pikselin (referans) komşularına (4 veya 8 komşuluk) sırayla bakılır. Eğer komşulardan birinin değeri -1, yani değer atanmamış ise o piksele referans pikselin değeri atanır, referans pikselin komşularının taranmasına sonra dönmek üzere ara verilir ve yeni piksel referans olarak seçilir. Aynı işlemler bütün resim taranmaya kadar iteratif olarak devam ettirilir.

2.4 Ten Bölgelerinin Bulunması

İnsan tenine ait renk bilgisi, insan bilgisayar arayüzü (HCI, Human Computer Interface) uygulamalarında önemli bir kullanım alanına sahiptir.

El işaretlerini kullanarak bilgisayarla iletişim kurma uygulamalarında (gesture recognition) görüntüdeki ten bölgelerinin insan ellerine ait olduğu kabul edilerek ten bölgelerinden özellik çıkarılıp bir sınıflayıcıya sunulmaktadır. Yüz sezimi probleminde ise ancak belli kısıtlamalar getirilmesi durumunda insan tenine ait bölgelerle ilgili basit yapısal özellikler çıkarılarak insan yüzleri bulunabilmesine karşın, ten bilgisi kullanılmasının asıl amacı kompleks ve uzun işlem süresi gerektiren bir sınıflayıcıyla resmin bütünüünün taranması yerine yalnız ten bölgelerinin taranarak işlem süresinin ve sınıflayıcının hatalarının azaltılmasıdır.

İnsan tenine ait renk bilgisi kullanmanın kendi içinde avantaj ve dezavantajları vardır. Renk bilgisi oldukça kısa sürede işlenebilen ve yönelimden bağımsız olmasına karşın, ışık koşullarının kontrol edilemediği durumlarda renk bilgisiyle çalışmak zorlaşmaktadır. İnsan gözü değişen ışık koşulları altında renkleri algılamakta oldukça başarılı olmasına karşın, bu durum makina görmesi açısından oldukça problemli bir konudur. İnsanlar tarafından, belli sınırlar içinde, aydınlatma seviyesinin değişmesine karşın bir cismin renginin değişmeden algılanmasına renk sabitliği (color constancy) denir. İnsan beyninin bu özelliği nasıl sağladığı tam olarak bilinmemesine rağmen, beyne reseptörlerden alınan ham bilginin gönderilmeyip, bunun yerine reseptörlerden alınan bilgilerin karşılaştırılarak gönderildiği bilinmektedir. Karşıt renk teorisine göre beyne parlaklık ve iki farklı renk bilgisi üç farklı kanaldan gönderilmektedir, aynı şekilde renk uzaylarının çoğunda parlaklık ve renk bilgisinin, RGB değerleri üzerinde lineer cebir işlemleri yapılarak, birbirinden ayrılması amaçlanmıştır. Parlaklık değerinden arındırılan renkler saf renkler (pure colours) olarak adlandırılır ve ışık şiddeti değişimlerine karşı mümkün olduğunca duyarsız olmaları istenir.

İnsan tenine rengini veren kandaki hemoglobin ve melanin pigmentidir. İnsan tenlerinin farklı renklerde görünmesinin temel nedeni tendeki melanin pigmenti oranlarının farklı olmasıdır. Tendeki pigment oranı arttıkça ten rengi koyulaşır ve saflığı azalır (Fleck, 1996). Renkli görüntülerde insan teninin bulunması için bir ten sınıflayıcı (skin classifier) oluşturulmalıdır. Bir ten sınıflayıcı oluşturmak için genel olarak değişik aydınlık koşulları altında, değişik ırklardan çok sayıda insan tenine ait örnek toplanır ve toplanan örnekler RGB uzayı yerine parlaklık bilgisinin renk bilgisinden ayrılabilirdiği bir uzayda temsil edilerek ten sınıflayıcı oluşturulur.

2.5 Ten Sınıflayıcılar

Ten sınıflayıcılar herhangi bir renk uzayı için tasarlanabilirler ve kendi içlerinde uyarlamalı (adaptive) olup olmadıklarına veya tasarım yöntemlerine göre alt gruplara ayrılabilirler.

2.5.1 Kapalı Sınırlı Ten Sınıflayıcılar

Bu tür sınıflayıcılarda, örnekler istenilen uzayda temsil edildikten sonra örneklerin tüm renk uzayı içinde oluşturdukları alt uzay karar sınırlarıyla temsil edilir. Karar sınırlarının örneklerin dağılımı yorumlanarak bir tasarımcı tarafından verilmesi ve genel olarak hızlı sonuç vermeleri avantajlarındandır. Bu tür sınıflayıcıların çıkışları iki değerlidir, yani bir pikseli direkt olarak ten veya ten olmayan sınıfa atar, bu ise bazı uygulamalarda istenmeyen bir durumdur.

2.5.2 Parametrik Olmayan Sınıflayıcılar (Non-parametric Classifiers)

Örneklerin dağılımı hakkında önceden herhangi bir öngörünün veya kabulün yapılmadığı sınıflayıcılardır. Eğitim setindeki her bir değer sıklığını (frequency), yani kaç örnekle temsil edildiğini gösteren bir histogram oluşturulur. Oluşturulan bu histogram, örneklerin olasılık dağılımlarını (probability distribution) göstermek için kullanılır. Histogramdaki bütün değerler histogram içindeki en büyük değere bölünerek normalize edilir ve bir tablo (look up table) oluşturularak saklanır. Oluşturulan tablolar genel olarak 1 veya 2 boyutludur ve tablodaki her değer kendisiyle aynı renk özelliklerine sahip bir pikselin ten olma ihtimalini verir.

Ten sınıflayıcının bir görüntüye uygulanması durumunda, sınıflayıcının çıkışı orjinal görüntüdeki her bir pikselin ten olma ihtimalini veren gri seviyeli bir ihtimal görüntüsüdür (probability image). Genellikle deneysel sonuçlardan elde edilen bir eşik değeri ile ihtimal görüntüsü ten ve ten olmayan bölgeleri gösteren iki seviyeli bir görüntüye dönüştürülür.

Parametrik olmayan sınıflayıcılar için örneklerin dağılımı üzerinde herhangi bir kabulün yapılmamış olması bir avantaj olmasına karşın, bu durum beraberinde sınıflayıcıyı örnek seçimine daha duyarlı kılar. Bu tür sınıflayıcıları eğitmek için çok sayıda örneğe ihtiyaç vardır.

2.5.3 Parametrik Sınıflayıcılar (Parametric Classifiers)

Parametrik ten sınıflayıcılar, örneklerin dağılımının önceden belirlenen bir model formunda olduğu kabul edilerek, model parametrelerinin örneklerin dağılımından çıkarıldığı sınıflayıcılardır. En çok kullanılan modeller 1 ve 2 boyutlu tek modlu normal gauss dağılımıdır (unimodal normal gaussian distribution).

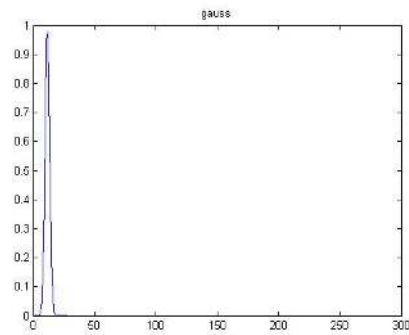
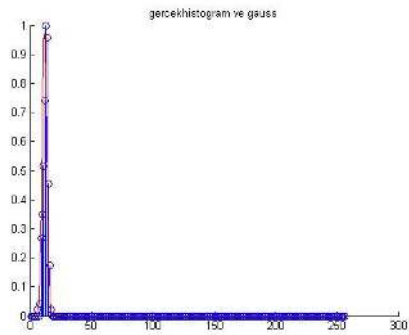
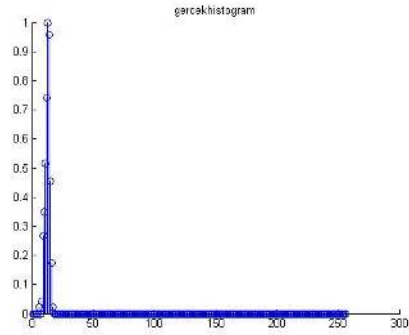
(Yang,1996) normalize r ve g değerlerini kullanarak, değişik ırklardan insanlara bile ait ten örneklerinin r-g düzleminde küçük bir bölgeye kümelendiğini ve 2 boyutlu normal gauss dağılımı formunda olduğunu gözlemledi.Yang' a göre insan tenlerinin farklı görünmesinin asıl nedeni ırksal ve kişisel farklılıklardan çok parlaklık etkisinden kaynaklanmaktadır.

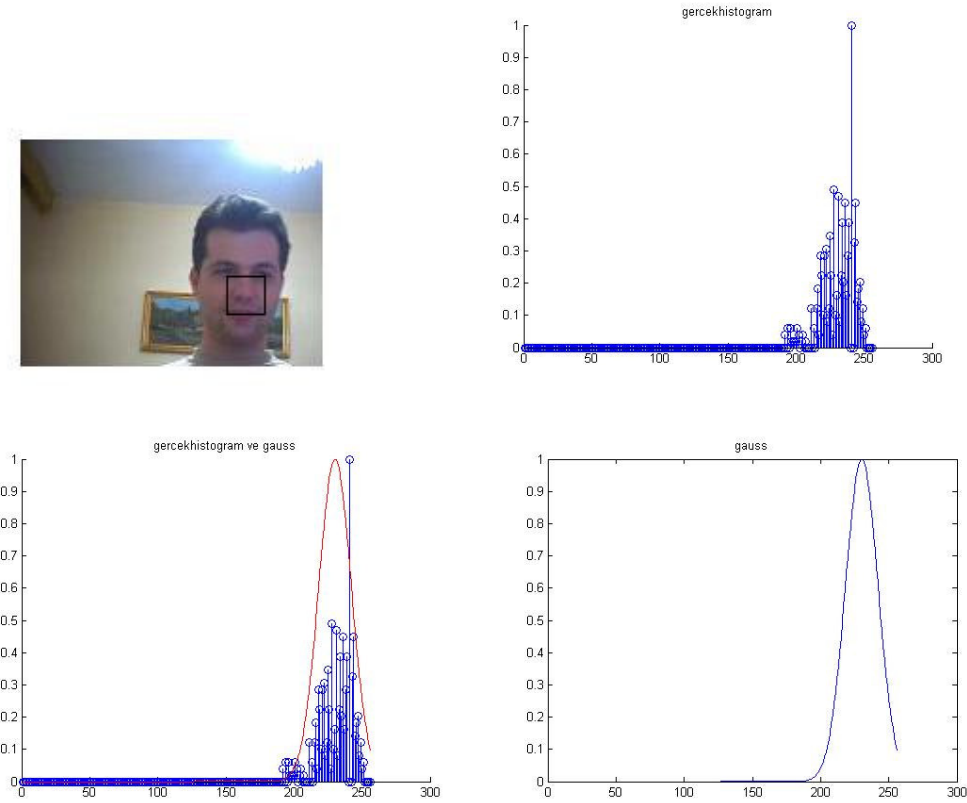
Gauss dağılımının model parametreleri örneklerin ortalama değeri (mean) ve kovaryans (covariance) matrisi hesaplanarak bulunur.2 boyutlu normal gauss dağılım (2.22)' de gösterilmiştir.

$$N(m, \Sigma) = \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left[-0.5(x-c)^T \Sigma^{-1}(x-c)\right] \quad (2.22)$$

Gauss dağılımı renk vektörü ile merkez (c) arasındaki mahalonabis uzaklığına bakılarak hesaplanır.

Şekil 2.7' de renkli görüntülerden işaretli kısım alınarak, piksellerin renk tonlarına (hue) göre sırasıyla gerçek dağılımı ve uydurulmuş gauss modelleri gösterilmiştir.Gerçek modellerle uydurulan modellerin benzerlik ölçütü olarak korelasyona bakılmış ve ilk resim için korelasyon değeri 0.7660, ikinci resim için ise 0.6430 olarak hesaplanmıştır.





Şekil 2.7 Ten örneklerinin gerçek ve model dağılımları

Parametrik sınıflayıcılar, parametrik olmayan sınıflayıcıların aksine birkaç parametreyle ifade edildiklerinden bellekte az yer kaplamalarına, daha kolay güncellenerek değişen aydınlatma koşullarına uyumlu hale getirilebilmelerine ve örneklerin dağılımını interpolate ederek daha az örnekle oluşturulabilmelerine karşın, gerçek dağılımın tam olarak gauss dağılımı göstermemesinden dolayı ten olmayan pikselleri ten olarak yanlış sınıflama oranları yüksektir ve mahalonabis uzaklığının hesaplanması işlem yükü getirdiğinden dolayı yavaştırlar.

2.5.4 Adaptif ve Durağan Sınıflayıcılar

Aslında tasarlanan bütün ten sınıflayıcılar aydınlık koşullarına karşı duyarlıdırlar ve ten sınıflayıcıların çoğu kontrollü aydınlatma koşulları altında çekilmiş fotoğraflar için tasarlanmıştır. Bu tür sınıflayıcılar mümkün olduğunca geniş bir eğitim setiyle ve karşılaşabilecekleri her durum için eğitilen genel (universal) sınıflayıcılardır ve ışık koşullarının kontrol edilemediği çoğu durumda başarısız olmaktadır. Bu sebeple adaptif ten sınıflayıcılar tasarlanmıştır. Adaptif ten sınıflayıcılar, video çerçevelerinde kendilerinin ten

olarak sınıfladığı piksellerin dağılımını kullanarak kendilerini güncellerler, yani sınıflayıcının çıkışı kendinden önceki çıkışlarına da bağlıdır. Bu sebeple FIR (Finite Impulse Response) filtre özelliği gösterirler.

Yang, yaptığı çalışmada adaptif parametrik bir ten sınıflayıcı kullanmıştır. Gauss model parametreleri, o anki model parametreleri ile bir önceki çerçeveden elde edilen parametrelerin ağırlıklı toplamı yapılarak hesaplanmıştır.

Adaptif sınıflayıcıların temel problemi değişen ışık koşullarına özelleşerek genel olma özelliklerini kaybetmeleri ve kendilerini tekrar toparlayamamalarıdır.

2.5.5 Ten Sınıflayıcı Tasarımı

Normalize RGB, HSV ve YCbCr uzaylarında, İkiizler ve Duygulu'nun (2004) TRECVID 2003 haber video arşivinden topladıkları ten seti kullanılarak, parametrik ve parametrik olmayan ten sınıflayıcılar oluşturularak başarıları karşılaştırılmıştır. Eğitim seti yaklaşık 31000 ten ve çok sayıda insan tenine ait olmayan örnek içermektedir.

Renkli görüntülerde ten bulma, temel olarak iki sınıflı bir sınıflandırma problemidir. Bir ten sınıflayıcı, her iki sınıfa da ait örnekler veya yalnız ten örnekleri kullanılarak oluşturulabilir.

Bayes yöntemiyle her iki sınıfa ait örnekler kullanılarak parametrik veya parametrik olmayan bir sınıflayıcı tasarlanabilir. Bir renk vektörünün ten olarak sınıflandırılması için, sınıflara bağlı olasılıkların (conditional probability) oranı belli bir eşik değerin üzerinde olmalıdır.

$$\frac{P(r|ten)}{P(r|\sim ten)} \geq \theta \quad (2.23)$$

θ eşik değerini, $P(r|ten)$ ve $P(r|\sim ten)$ ise sırasıyla ten ve ten olmayan örneklerin oluşturdukları, sınıflara bağlı olasılık dağılımlarını göstermektedir. Sınıflara bağlı olasılıklar, her bir sınıf için oluşturulan histogramlarla temel olarak aynı formda olup, histogramların ait oldukları sınıfı oluşturan toplam örnek sayısına bölünmesiyle elde edilir.

$$P(r|ten) = \frac{H_{ten}(r)}{\sum_r H_{ten}(r)} \quad (2.24)$$

$$P(r|\sim ten) = \frac{H_{\sim ten}(r)}{\sum_r H_{\sim ten}(r)} \quad (2.25)$$

Eşik değeri teorik olarak, öncül sınıf olasılıkları (priori probability), hatalı tespit (false detection) ve hatalı reddetme (false rejection) maliyetleri ile hesaplanır.

$$\theta = \frac{\lambda_{ht} P(\sim ten)}{\lambda_{hr} P(ten)} \quad (2.26)$$

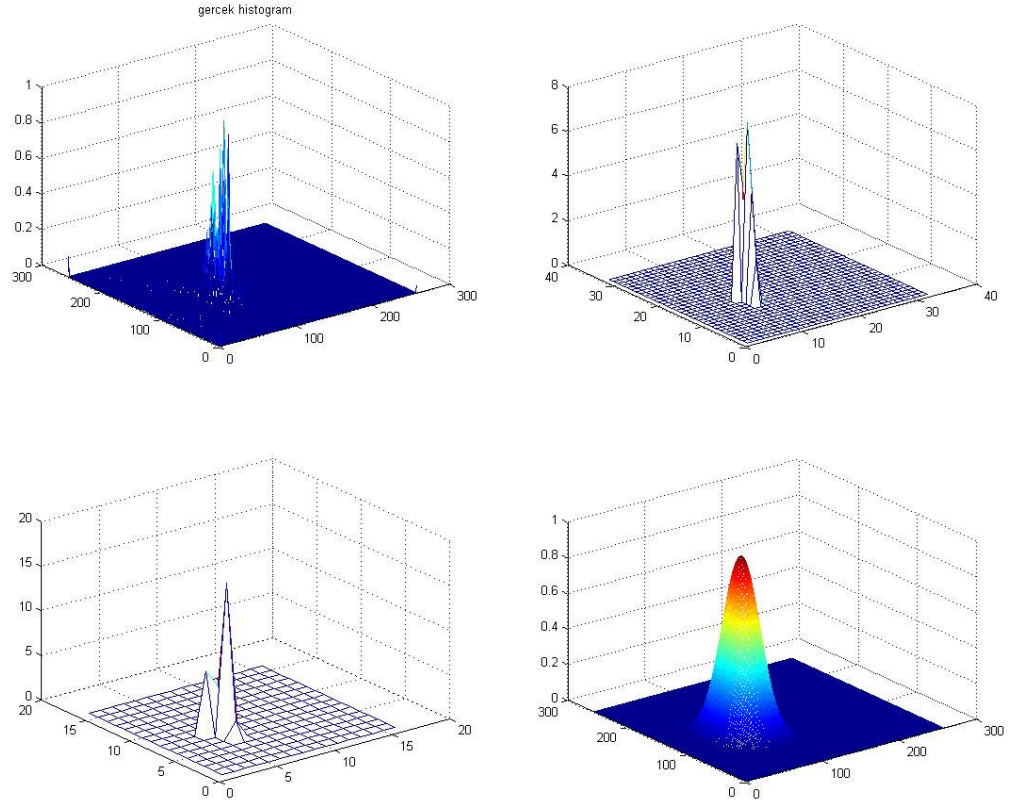
λ_{ht} ve λ_{hr} sırasıyla hatalı tespit ve hatalı reddetme maliyetlerini, $P(\sim ten)$ ve $P(ten)$ ise öncül sınıf olasılıklarını göstermektedir. Öncül sınıf olasılıkları bir sınıfa ait örnek sayısının tüm örnek sayısına oranı olup, ten sınıflama problemi için genelde eşit kabul edilir. Eşik değeri teorik olarak (2.26)'daki gibi hesaplanabilmesine karşın, genelde uygun eşik değeri deneysel olarak belirlenir (Bouzerdoun vd.,2000).

Ten olmayan örneklerin kendi içlerinde ortak bir karakteristik özelliklerinin olmaması ve insan teniyle aynı renk özelliklerine sahip birçok nesnenin bulunması, ten olmayan sınıfı temsil edecek örneklerin toplanmasını zorlaştırmaktadır. Bu sebeplerden ve ten sınıflayıcının hızlı cevap vermesi istendiğinden, yalnız insan tenine ait örnekler kullanılarak sınıflayıcı tasarımına gidilmiştir. İkizler ve Duygulu (2004) ten ve ten olmayan örnekleri kullanarak, her iki sınıf için de normal Gauss dağılımlarını hesaplamışlar ve yalnız ten örneklerini kullanarak oluşturdukları ten sınıflayıcı ile, Bayes yöntemini kullanarak oluşturdukları sınıflayıcıya nazaran, TRECVID 2003 video arşivinde daha başarılı sonuçlar elde ettiklerini bildirmişlerdir.

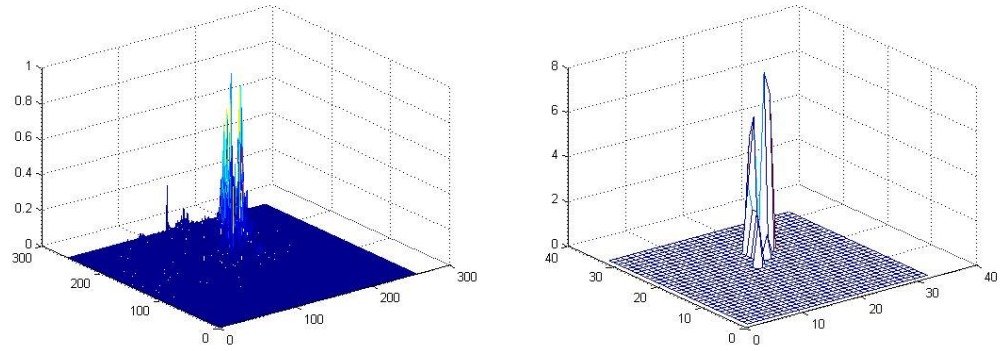
Gerçek bir dağılıma bir Gauss modeli uydurularak gerçek dağılım interpolate edilmiş, yani genelleştirilmiş olur. Bir diğer alan genişletme veya genelleştirme yöntemi ise ten sınıflayıcıyı oluşturan histogramı belli değerlere kuantalamaktır. Genel olarak tek boyutlu bir histogramda 0 ile 255 arasındaki değerleri temsil etmek için 256, iki boyutlu bir histogramda ise 256*256 seviye (level,bin) bulunmaktadır. Histogram kuantalama işlemi, sınıflayıcının örnek setine duyarlılığını azaltarak genelleme yeteneği kazandırmasının yanısıra sınıflayıcı için ayrılan bellek miktarında azaltmaktadır.

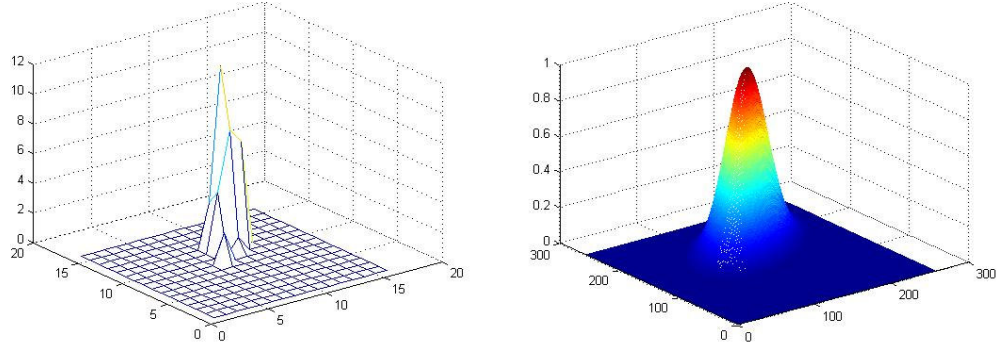
Bu sebeplerden dolayı eğitim setinin, normalize RGB, YCbCr ve HSV uzaylarında gerçek dağılımı, 32*32 ve 16*16 seviyelere kuantalanmış dağılımları ve Gauss dağılımı oluşturularak elde edilen sınıflayıcılar karşılaştırılmıştır.

Şekil 2.8, 2.9 ve 2.10'de sırasıyla normalize RGB, HSV ve YCbCr uzaylarında oluşturulan gerçek, 32*32, 16*16 seviyeli ve Gauss dağılımları gösterilmiştir.

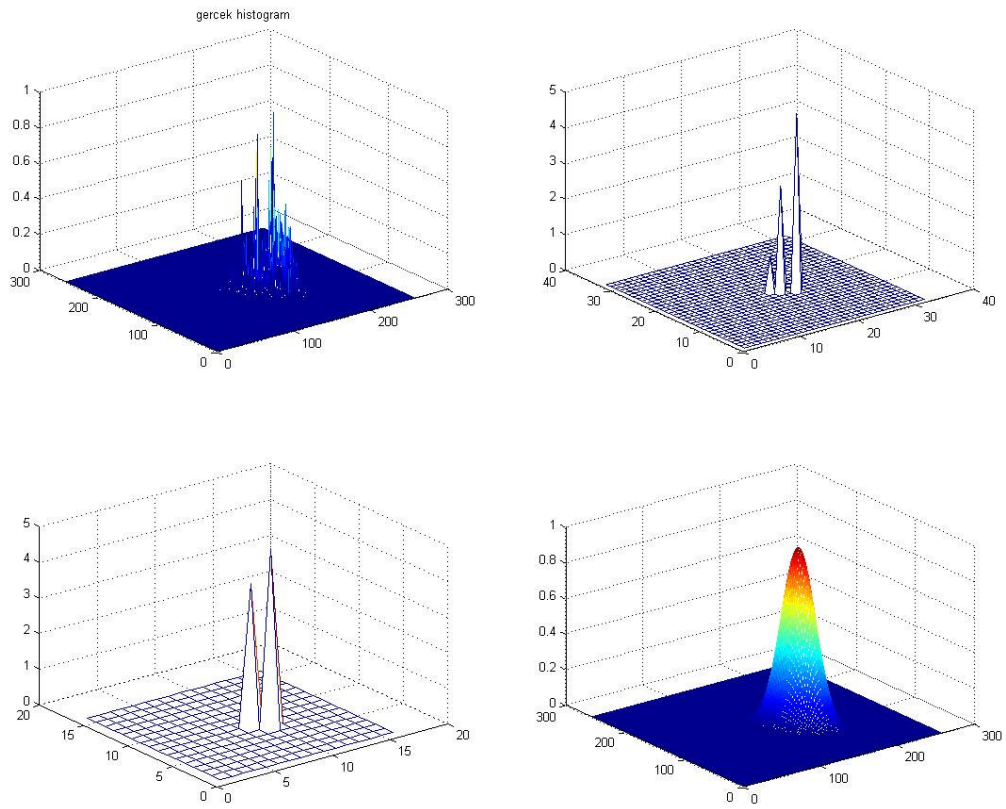


Şekil 2.8 Normalize RGB uzayında renk dağılımları





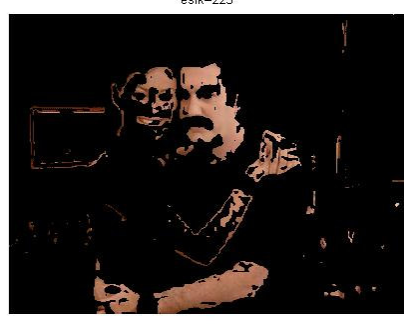
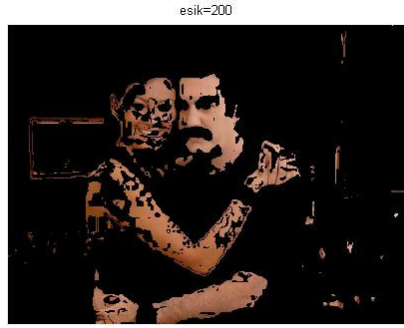
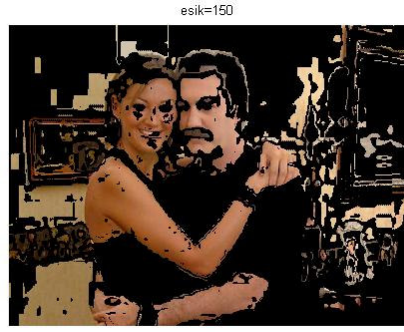
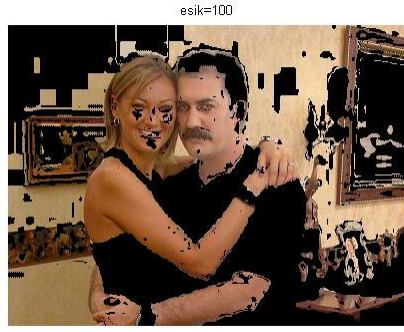
Şekil 2.9 HSV uzayında renk dağılımları



Şekil 2.10 YCbCr uzayında renk dağılımları

Oluşturulan sınıflayıcılar örnek resimler üzerinde denendiğinde en iyi sonucun YCbCr uzayında 16*16 ve 32*32 seviyeli histogramlardan elde edildiği görülmüştür.

Şekil 2.11’de sırasıyla 100, 150, 200 ve 225 eşik değerleri için 2 farklı görüntü ile elde edilen sonuç görüntüleri gösterilmiştir.



Şekil 2.11 Ten sınıflayıcı sonuçları

2.6 CAMSHIFT (Continuously Adaptive Mean Shift) Takip Algoritması

Ten sınıflayıcılar piksel tabanlı olduklarından çıkış görüntüleri oldukça gürültülüdür. Bu sebeple renkli görüntülerde insan tenine ait bölgelerin bulunması için sınıflayıcı çıkışındaki görüntü morfolojik işlemlerden geçirildikten sonra etiketleme algoritması kullanılarak gruplandırılır.

CAMSHIFT ise morfolojik işlemlerin yapılmadığı, bölgesel tabanlı ve kullanıcı girişine ihtiyaç duyan bir takip algoritmasıdır. Takip edilecek bölge mouse ile seçilir ve seçilen bu bölgeden oluşturulan bir histogram ile resimdeki her bir piksele bir değer atanır. Buradaki histogram genel sınıflayıcının aksine yalnız takip edilecek olan nesneye aittir. Sonuçta oluşan ihtimal görüntüsü (probability image) resmin her bir pikselinin seçilen bölgeye ne kadar benzediğini gösteren bir görüntüdür. Seçilen bölgenin koordinatları referans alınarak CAMSHIFT algoritması ile seçilen nesne takip edilir. Takip işlemi standart mean shift algoritmasının aksine boyutları değişen bir pencerenin ihtimal görüntüsü üzerinde gezdirilerek mouse ile işaretlenen noktaya en yakın lokal tepenin (peak, mode) bulunması ile yapılır.

2.7 Tezde Önerilen Yöntem

Işık koşullarının çok fazla kontrol edilemediği durumlarda video görüntülerinde özellikle adaptif olmayan renk sınıflayıcılar başarısız olmaktadır. Adaptif olanlar ise zamanla özelleşerek genel olma özelliklerini kaybedebilmektedirler. Genel renk sınıflayıcı ile CAMSHIFT algoritması birleştirilerek CAMSHIFT için gerekli olan kullanıcı girişi ortadan kaldırılmış ve daha iyi bir sonuç elde edilmiştir. Genel sınıflayıcı çıkışında insan teni olarak işaretlenen bölgeler CAMSHIFT algoritması ile referans kabul edilerek renk tabanlı bir takip işlemi yapılabilir.

Genel sınıflayıcı olarak YCbCr uzayında oluşturulan sınıflayıcı ile Fleck vd.'nin (1996), karşıt renkler teorisinden esinlenilerek tasarladığı renk sınıflayıcı karşılaştırılmış ve YCbCr uzayında tasarlanan 32*32 seviyeli sınıflayıcı ile yakın sonuçlar verdiği gözlemlenmiştir. Şekil 2.12'de örnek resimler üzerinde, tasarlanan sınıflayıcının ürettiği ihtimal görüntüsü, sonuç görüntüsü ve Fleck vd. (1996) tarafından tasarlanan sınıflayıcının ürettiği sonuç görüntüsü gösterilmiştir. İhtimal görüntüsü her pikselin yüz olma ihtimalini gösteren [0,255] aralığında değer alan bir görüntüdür, değeri 100 ve üzerinde olan pikseller ten olarak seçilmiştir.



ycbcr ten



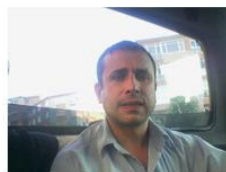
log opponent ten



ycbcr ten



log opponent ten



ycbcr ten



log opponent ten





Şekil 2.12 Karşılaştırmalı test sonuçları

Fleck tarafından tasarlanan sınıflayıcı kapalı karar sınırlarına sahip, ışık koşullarının kısmen kontrol edilebildiği, durağan görüntüler için tasarlanmış olduğundan sınıflayıcının video görüntülerinde başarısını arttırmak için karar sınırları üzerinde ve RGB değerlerinden karşıt renk uzayına dönüşümde değişiklik yapılarak çalışmada her iki sınıflayıcıda kullanılmıştır.

Orijinal sınıflayıcının dönüşüm ifadeleri (2.27) gösterilmiştir.

$$R_g = L(R) - L(G)$$

$$B_y = L(B) - \left(\frac{L(R) + L(G)}{2} \right)$$

$$I = L(G)$$

$$L(x) = 105 * \log(x + 1 + n) \quad n \in [0 \ 1] \quad (2.27)$$

$$Hue = \tan^{-1} \left(\frac{R_g}{B_y} \right)$$

$$Saturation = \sqrt{R_g^2 + B_y^2}$$

Bir pikselin ten olarak sınıflandırılması için uzay içerisinde (2.28)'de belirtilen iki kapalı karar bölgesinden birinde olmalıdır.

$$(110 < hue < 150 \text{ ve } 20 < saturation < 60) \text{ veya } (130 < hue < 170 \text{ ve } 30 < saturation < 130) \quad (2.28)$$

(2.27)'de gösterilen dönüşüm ifadelerinde RGB değerlerine 60 rakamı eklenmiş ve renk uzayı içinde aydınlık seviyesi yüksek pikselleri de ten olarak sınıflayan yeni bir karar bölgesi eklenmiştir. RGB değerlerine eklenen rakam ve oluşturulan yeni karar bölgesine deneysel

sonular neticesinde karar verilmiřtir.Kapalı karar sınırlarına sahip sınıflayıcılar daha ncede bahsedildiđi gibi zerlerinde bu tr iřlemler yapmaya izin verirler.Yeni eklenen kapalı blgenin sınırları (2.29)’da gsterilmiřtir.

$$10\langle saturation\langle 20 \text{ ve } 110\langle hue\langle 150 \quad (2.29)$$

řekil 2.11’de bir video erevesinde yalnız zerinde deđiřiklik yapılan genel sınıflayıcı ve genel sınıflayıcıyla CAMSHIFT algoritması birleřtirildiđinde alınan sonular gsterilmiřtir, sonu grntsnde en iteki dikdrtgen genel sınıflayıcının ten olarak sınıfladıđı alanı, dıřtaki dikdrtgen ise CAMSHIFT ile beraber kullanıldıklarında elde edilen sonucu gstermektedir.

Genel sınıflayıcı piksel tabanlı olduđundan ıkıřındaki grnt grltldr ve morfolojik iřlemlere ihtiya duyar.Genel sınıflayıcı ıkıřındaki grnt erozyona uđratıldıktan sonra grntde kalan paralar kenarları belirlenerek etiketlenir.Dzensiz řekillerdeki blgeler iin kenar erveleri oluřturularak ilgi alanları bulunur ve ilgi alanları birbirleriyle karřılařtırılarak en byk alan bulunur.En byk alan dikdrtgen formunda olduđundan aynı zamanda ten olmayan pikselleride iermektedir bu yzden genel sınıflayıcı ile en byk alan iindeki ten pikselleri bulunarak yalnızca renk tonundan oluřan (hue) tek boyutlu bir histogram oluřturulur.Yeni oluřturulan histogram veya sınıflayıcı genel deđil zel bir sınıflayıcıdır.Bradski’nin (1998) nerdiđi gibi oluřturulan histogram 16 seviyeye kuantalanır.Oluřturulan bu yeni sınıflayıcı ile resmin tamamı taranarak yeni ihtimal grnts oluřturulur ve genel sınıflayıcı tarafından en byk olarak seilen alanın merkezi CAMSHIFT algoritması iin merkez noktası seilerek grntdeki ten blgeleri bulunmaya alıřılır.

Bradski’nin nerdiđi standart CAMSHIFT algoritmasında grntye zel ten sınıflayıcı bir kere oluřturulduktan sonra saklanmaktadır, yani durađandır.Iřık kořullarındaki deđiřimi tolere etmek iin bir diđer yntem adaptif bir sınıflayıcı tasarlamak, yani sınıflayıcıyı kendi ıkıřı ile gncellemektir.Fakat ıřık kořullarının hızlı deđiřtiđi durumlarda ten sınıflayıcı kendi ıkıřı ile gncellendiđinde bařarısının dřtđ grlmřtr, bu yzden grntye zel ten sınıflayıcı genel ten sınıflayıcı ile gncellenmiřtir.İlk erveden sonraki ervelerde zel sınıflayıcının ten olarak belirlediđi blgede genel sınıflayıcının ten olarak pikseller kullanılarak yeni sınıflayıcılar oluřturulmaya devam edilmiřtir.Deđiřen ıřık kořulları yznden genel sınıflayıcının bařarısız olduđu ervelerde ise nceki ervelerde oluřturulan sınıflayıcılar kullanılmıřtır.

řekil 2.12’de yapılan iřlemler zetlenmiřtir.

Genel sınıflayıcı çıkışı



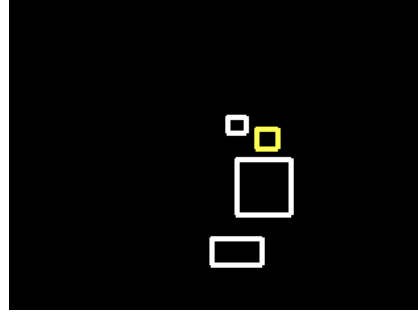
Erozyon



Etiketleme



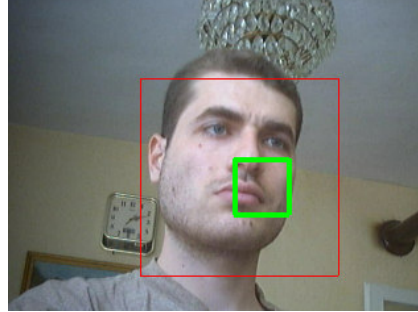
İlgi alanları



İhtimal görüntüsü



Sonuç görüntüsü



Şekil 2.13 Ten bölgelerinin belirlenmesi

2.8 Ten Bölgelerinden Özellik Çıkarımı

Renkli görüntülerden elde edilen ve insan tenine ait olduğu tahmin edilen pikselleri ön plan olarak gösteren ikili görüntüler üzerinde morfolojik işlemler uygulandıktan sonra bağlı bileşenler etiketlenir ve insan tenine ait olduğu düşünülen bölgeler (blop) bulunur. Ten olarak sınıflanan bölgelerden basit özellikler çıkarmak mümkündür.

2.8.1 Sıkılık (Compactness)

Sıkılık, (C) bir nesnenin dairesel formdan ne kadar deforme olduğunun bir ölçüsüdür ve

nesnenin alanının çevre uzunluğunun karesine oranı olarak ifade edilir.

$$C = \frac{A}{P^2} \quad (2.30)$$

Görüntüdeki bir nesnenin alanı (A) kendisini oluşturan tüm piksellerin, çevresi (P) ise yalnız sınır piksellerinin sayılarının toplamı olarak ifade edilir.

$$A = \sum_x \sum_y I(x, y) \quad (2.31)$$

İnsan yüzünün şekli oval olduğundan yüksek yoğunluk değerine sahiptir.

2.8.2 Katılık (Solidity)

Katılık, (S) cismin alanının kendisini çevreleyen sınır çerçevesinin (bounding box) alanına oranı olarak ifade edilir.

D_x ve D_y sınır çerçevesinin boyutlarını göstermek üzere;

$$S = \frac{A}{D_x D_y} \quad (2.32)$$

2.8.3 Görünüş Oranı (Aspect Ratio)

Görünüş oranı, (Ar) görüntüdeki cismi çevreleyen sınır çerçevesinin kenar uzunluklarının birbirine oranı olarak ifade edilir.

$$Ar = \frac{D_y}{D_x} \quad (2.33)$$

İnsan yüzünün ideal olarak uzunluğunun genişliğine oranının altın oran (golden ratio) olan 1.61 değerine eşit olduğu kabul edilmesine karşın, bu değer altındaki ve üstündeki görünüş oranına sahip birçok nesne yüz olarak kabul edilmektedir.

2.8.4 Yönelim (Orientation)

Bir nesnenin yönelimi momentleri hesaplanarak bulunabilir.

Sıfırıncı moment M_{00} ile gösterilmek üzere;

$$M_{00} = \sum_x \sum_y I(x, y) \quad (2.34)$$

Birinci dereceden momentler sırasıyla M_{10} ve M_{01} ile gösterilmek üzere;

$$M_{10} = \sum_x \sum_y xI(x, y) \quad (2.35)$$

$$M_{01} = \sum_x \sum_y yI(x, y) \quad (2.36)$$

İkinci dereceden momentler sırasıyla M_{20} ve M_{02} ile gösterilmek üzere;

$$M_{20} = \sum_x \sum_y x^2 I(x, y) \quad (2.37)$$

$$M_{02} = \sum_x \sum_y y^2 I(x, y) \quad (2.38)$$

θ nesnenin büyük ekseninin (major axis) görüntünün yatay eksenine ile yaptığı açığı ve x_c ile y_c merkez noktalarını göstermektedir.

$$\theta = \frac{1}{2} \arctan \left(\frac{2 \left(\frac{M_{11}}{M_{00}} - x_c y_c \right)}{\left(\frac{M_{20}}{M_{00}} - x_c^2 \right) - \left(\frac{M_{02}}{M_{00}} - y_c^2 \right)} \right) \quad (2.39)$$

Kuchi vd. (2002) görüntülerde etkiletilenen ten bölgelerinin sırasıyla yoğunluk, katılık ve görünüş oranlarına bakarak durağan ve video görüntülerinde insan yüzlerini aramışlardır. Yaptıkları denemeler sonucunda herbir özellik için birer eşik değeri belirlemişler ve uygun eşik değerlerini sağlayan ten bölgelerini yüz olarak sınıflamışlardır.

İnsan yüzlerini bulmak için ten bölgelerinden özellik çıkarmak basit ve hızlı sonuç veren bir yöntem olmasına karşın, ancak belirli kısıtlamalar altında doğru sonuç veren güvenilirliği az bir yöntemdir.

3. ŞABLON EŞLEME

Şablon eşleme, genel olarak görüntüdeki bir nesnenin önceden oluşturulmuş bir şablon kullanılarak tespit edilmesidir. Standart bir şablon, görüntüde aranan nesneyi temsil eden önceden tanınlanmış (predefined template) 2 boyutlu bir matristir. Yüz sezimi problemi için şablon insan yüzlerinin ortalaması alınarak bulunur. Birçok çalışmada durağan bir şablon yerine belli bir esnekliğe sahip deforme olabilen şablonlar (deformable template) kullanılarak daha başarılı sonuçlar elde edilmesine karşın, yapılan çalışmada yalnız durağan şablon modeli kullanılmıştır.

Görüntülerde insan yüzlerini bulabilmek için şablon görüntü üzerinde gezdirilerek görüntüden alınan parçalarla şablon arasındaki benzerliğe bakılır, benzerliğin ölçüsü korelasyon katsayısı (correlation coefficient) ile hesaplanır. Değişik boyutlardaki insan yüzlerini bulabilmek için görüntünün veya şablonun boyutları değiştirilerek birden fazla tarama işlemi yapılır.

Korelasyon iki çok boyutlu rastgele değişkenin (multivariate random variable) birbirlerine ne ölçüde doğrusal olarak bağlı olduklarının, yani beraber değiştiklerinin bir ölçüsüdür ve iki değişken arasındaki kovaryansın değişkenlerin standart sapmalarının çarpımına oranı olarak ifade edilir.

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (3.1)$$

Kovaryans değeri iki değişken arasındaki ilişki hakkında bilgi vermesine karşın, normalize bir değer olmadığından, ilişkinin gücü hakkında bir bilgi vermez. Kovaryans değerinin değişkenlerin standart sapmalarına bölünerek normalize edilmesiyle korelasyon katsayısı elde edilir.

Korelasyon katsayısı -1 ile 1 arasında değerler alır. Değişkenler arasındaki doğrusal ilişki arttıkça korelasyon değeri -1 veya 1 değerine yaklaşır, korelasyonun işareti ilişkinin yönünü gösterir, değişkenlerden biri artarken diğeri azalma eğilimindeyse aralarındaki korelasyon negatif işaretlidir. Değişkenlerin birbirinden bağımsız olmaları durumunda aralarındaki korelasyon sıfır olmasına karşın, korelasyonun değişkenler arasındaki yalnız doğrusal ilişkinin bir ölçüsü olmasından dolayı, bu ifadenin tersi doğru değildir, yani aralarında ilişki olan iki değişken arasındaki korelasyon sıfır olabilir.

X ve Y birer değişkeni temsil etmek üzere, eğer aralarında tam olarak bir doğrusal ilişki varsa korelasyon katsayısı ilişkinin yönüne göre 1 veya -1 değerini alır.

$$Y = a + bX$$

$$\begin{aligned} \rho_{xy} &= 1 \quad \text{eger } b > 0 \\ \rho_{xy} &= -1 \quad \text{eger } b < 0 \end{aligned} \quad (3.2)$$

X ve Y çok boyutlu değişkenleri aynı zamanda birer vektör gösterirler. Görüntüden alınan parça ve şablon vektörel veya matrissel formda ifade edilebilirler.

P görüntüden alınan parçayı, S şablonu, m ve n şablonun boyutlarını göstermek üzere resimden alınan bir parçayla şablon arasındaki korelasyon (3.3)'deki gibi hesaplanır.

$$\rho_{RS} = \frac{\sum_m \sum_n (P_{mn} - \mu_P)(S_{mn} - \mu_S)}{\sqrt{\sum_m \sum_n (P_{mn} - \mu_P)^2} \sqrt{\sum_m \sum_n (S_{mn} - \mu_S)^2}} \quad (3.3)$$

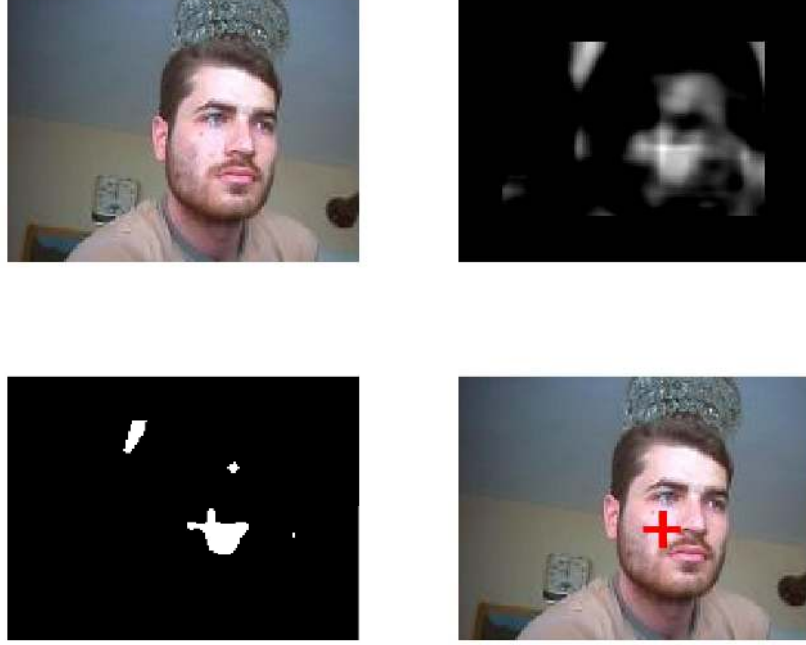
Önceden tanımlanmış bir şablonla tüm yüzleri temsil etmenin en iyi yolu, sakal, bıyık ve gözlük içermeyen kadın ve erkek resimlerinin aritmetik ortalamasını almaktır. Harvard yüz setinden 60 adet yüzün ortalaması alınarak yüze ait şablon elde edilmiştir.



Şekil 3.1 Şablon yüz

İnsan yüzü gibi sınıf içi dağılımı fazla ve sınıflar arası dağılımın az olduğu bir nesnenin tek bir şablonla temsil edilmesi oldukça zordur. Resimle şablon arasında yalnız korelasyona bakıldığında yüz olmayan bir çok bölgenin yüz olarak sınıflandırıldığı gözlemlenmiştir, bu hatalı tespitleri elimine etmek için ayrıca bir ten sınıflayıcı kullanılmıştır. Ten sınıflayıcıyla şablon eşleme sonucunda yüz merkezi olarak sınıflanan bölgelerin insan tenine ait olup olmadığına bakılmıştır.

Şekil 3.2' de orjinal görüntü, her bir pikselin yüzün merkezi olma olasılığını gösteren ihtimal görüntüsü, ihtimal görüntüsünün belli bir eşik değeri üzerinde olan noktalarını önplan olarak gösteren ikili görüntü ve şablon eşleme sonrasında ten sınıflayıcı kullanılarak elde edilen sonuç görüntüsü gösterilmiştir. Korelasyon katsayısı için eşik değeri 0.3 olarak seçilmiştir.

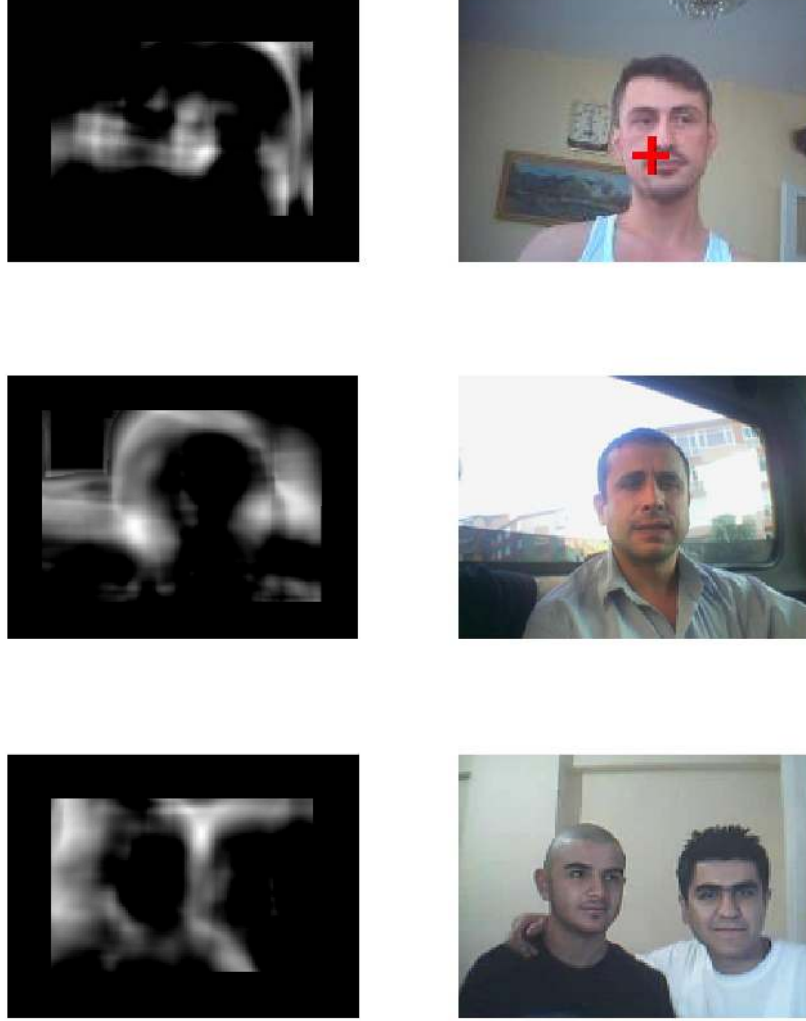


Şekil 3.2 Şablon eşleme

Şablon eşleme yöntemi fazla işlem yükü getirdiğinden oldukça yavaştır ve yüzün yalnız merkezi değil merkez etrafındaki noktalara da duyarlı olduğu gözlemlenmiştir. Bu sebeple görüntüdeki her pikseli incelemek yerine, görüntüyü pikseller arasında mesafe bırakarak incelemek mümkündür. Pikseller arasında 3 piksellik mesafe bırakılması durumunda sonucun değişmediği gözlemlenmiştir ve bu sayede algoritma yaklaşık 8 kat hızlandırılmıştır. Piksel atlayarak tarama işleminde ikili görüntü örneklenmektedir.

Şekil 3.3'de yüz merkezi olabilecek noktaları gösteren ihtimal ve ten sınıflayıcı sonrasında elde edilen sonuç görüntüleri gösterilmiştir.





Şekil 3.3 Şablon eşleme örnekleri

Şablon eşleme algoritmasını hızlandırmanın diğer bir yolu renkli görüntülerde ten olan bölgeleri belirleyerek yalnız bu bölgelerde insan yüzünü aramaktır, fakat görüntülerde insan teniyle aynı renkte birçok nesne olabileceğinden işlem süresi gerçek zamanlı bir sisteme kıyasla hala çok fazladır ve tek bir şablon ile geniş bir dağılımın temsil edilmesi mümkün olmadığından şablon eşlemenin yüz sezimi için uygun bir yöntem olmadığına karar verilmiştir.

4. ÖRNEK TANIMA

Örnek tanıma, önceden belirlenen veya karşılaşılmış örneklerden çıkarılmış istatistiksel sonuçlar yardımıyla, bir örneğin karakteristik özelliklerine göre sınıflanması veya tanımlanmasıdır. Tipik bir örnek tanıma sistemi girişindeki ham bilgiyi alır, ham bilgi ile ilgili özellikleri çıkarır ve daha önceden eğitilmiş bir sınıflayıcı ile test örneklerini değerlendirir. Bir örnek tanıma sistemi hakkında karar vereceği örnek veya nesnelerle ilgili bir hafızaya sahip olmalıdır. Hafıza, kullanılan sınıflayıcıya göre değişik şekillerde temsil edilir. Hafıza, çok sayıdaki örnekle, örneklerden çıkarılmış özelliklerle veya örneklerin ortalama değer ve dağılım modelleri gibi istatistiksel değerlerle temsil edilebilir.

4.1 Sınıflama ve Kümeleme

Bu bölümde tezde yararlanılan sınıflayıcılar ve kümeleme algoritması ile ilgili bilgi verilecektir.

Bir sınıflayıcı daha önceden etiketlenerek gruplara ayrılmış örneklere ihtiyaç duyar ve kendisine sunulan yeni örneği önceden belirlenmiş sınıflardan birine atar. Sınıflayıcıyı eğitmek için etiketlenmiş örnekler eğitim setini oluşturur. Bu eğitim yöntemi eğitici öğrenme (supervised learning) olarak isimlendirilir. Eğitici olmayan öğrenmede ise sisteme herhangi bir etiketlenmiş örnek seti göstermeksizin, sistemin örnekler arasında benzerlik ilişkisi kurarak, örnekleri sınıflaması veya kümelemesi istenir.

4.1.1 Bayes Sınıflayıcı

Bayes sınıflayıcı, örnek tanımda oldukça geniş kullanım alanına sahip, ihtimal dağılımı bilgisini kullanarak bir örneği ait olma ihtimali en yüksek sınıfa atayan istatistiksel bir sınıflama yöntemidir.

Bayes teoremi koşullu olasılık tanımı kullanılarak gösterilebilir. Koşullu olasılık (conditional probability) birbiriyle ilişkisi olan iki olaydan birinin durumuna göre diğerinin olma ihtimalini verir.

$$P(A|B) = \frac{P(A,B)}{P(B)} \quad (4.1)$$

$P(A|B)$, aslında B 'nin seçilen değerine karşılık A 'nin seçilen değerinin olma ihtimalini verir.

$$P(A = a|B = b) = \frac{P(A = a, B = b)}{P(A = a)} \quad (4.2)$$

B 'nin A 'ya göre olma olasılığı ise (4.3)'teki gibidir.

$$P(B|A) = \frac{P(A, B)}{P(A)} \quad (4.3)$$

(4.1) ve (4.3) ifadeleri kullanılarak Bayes kuralı elde edilir.

$$P(A|B)P(A) = P(B|A)P(B) \quad (4.4)$$

$$P(A|B) = \frac{P(B|A)P(B)}{P(A)} \quad (4.5)$$

Sınıflayıcı giriş vektörü x 'i herbir sınıf w ile gösterilmek üzere c adet sınıflatan birine atar.Örneğin seçilen sınıfa ait olma olasılığı, sınıfa bağlı olasılık (class conditional probability) $P(x|w)$ 'ye, öncül sınıf olasılığı (priori probability) $P(w)$ 'ye ve x 'in marjinal olasılığı (marginal probabiliy) $P(x)$ 'e bağlıdır.

$$P(w_i|x) = \frac{P(x|w_i)P(w_i)}{P(x)} = \frac{P(x|w_i)P(w_i)}{\sum_{i=1}^c P(x|w_i)P(w_i)} \quad (4.6)$$

(4.6)'deki ifadede paydadaki ifade sabit olup normalizasyonu sağlar (normalizing constant), bu sebeble ihmal edilebilir.Bu durumda sınıfa bağlı olasılığa ve sınıf olasılığına bakmak yeterlidir (bütün hata maliyetleri eşit kabul edilmek üzere).İki adet sınıf için örneğin sınıflardan birine ait olma şartı (4.7)'deki gibidir.Sınıf sayısının artması durumunda herbir sınıfa ait olma ihtimali diğer bütün sınıflara ait olma ihtimalleriyle karşılaştırılmalıdır.

$$P(x|w_1)P(w_1) > P(x|w_2)P(w_2) \quad ise \quad x \in w_1 \quad (4.7)$$

Sınıf olasılıkları $P(w)$ 'leri tahmin etmek oldukça zordur ve genelde eşit kabul edilirler.Bu durumda Bayes kuralı iki sınıflı bir problem için (4.8)'deki ifadeye indirgenmiş olur.

$$P(x|w_1) > P(x|w_2) \quad ise \quad x \in w_1 \quad (4.8)$$

Sınıfa bağlı olasılık dağılımları, sınıflara ait örneklerden oluşturulan histogramlardan elde edilir ve çoğu durumda dağılımın çok değişkenli normal Gauss dağılımı (multivariate normal distribution) formunda olduğu kabul edilerek, Gauss dağılımının parametreleri gerçek

dağılımdan elde edilir.

$$P(x|w_i) = \frac{\exp\left(-0.5(x - \mu_i)^T \Sigma^{-1}(x - \mu_i)\right)}{(2\pi)^{d/2} |\Sigma|^{1/2}} \quad (4.9)$$

4.1.2 En Yakın Orta Değer Sınıflayıcı (Nearest Mean Classifiers)

Sınıflayıcı tasarımı için en basit ve algısal yöntem örnekler arasındaki benzerliğe bakmaktır, yani benzer örnekler aynı sınıfa atanmalıdır. Benzerliğin ölçüsü olarak örnekler arasındaki uzaklığa bakılır.

En yakın orta değer sınıflama, aslında en yakın komşu kuralına (nearest neighbor rule) bir benzeridir.

En yakın komşu kuralında, bilinmeyen bir örnek kendisine en yakın örneğin sınıfına atanır. Benzerliğin ölçüsü olarak öklit uzaklığına bakılır. Daha iyi bir sınıflama en yakın k tane örneğin sınıfına bakılarak yapılır (k nearest neighbor). En yakın komşu sınıflamada örneklerin uzay içindeki dağılımları hakkında herhangi bir model çıkarımı yapılmaz, algoritma yalnız var olan eğitim setini kullanır. Bu sebepten k en yakın komşu sınıflama, sınıfların karmaşık yapıya sahip, hatta örtüştükleri (overlap) durumlarda dahi etkili olabilmesine ve yeni örneklerle kolaylıkla güncellenebilmesine karşın, test örneğinin bütün veritabanındaki bütün örneklerle arasındaki uzaklığa bakılmasından dolayı işlem süresi oldukça uzundur.

En yakın ortalama sınıflamada ise örneklerle ilgili çıkarım yapılarak her bir sınıf tek bir prototip, yani o sınıfın ortalama değeri ile temsil edilir. Bu yöntemde sınıfları temsil etmekte kullanılan örnekler ve örnekler arasındaki uzaklığın ölçülme yöntemi önem taşımaktadır. K en yakın komşu sınıflama da olduğu gibi burdada çoğunlukla öklit uzaklığına bakılmasına karşın, k en yakın komşu sınıflamanın aksine örneklerin dağılımları hakkında çıkarım yapılarak daha etkin bir uzaklık ölçümü yapmak mümkündür.

4.1.2.1 Öklit ve Mahalanobis Uzaklığı

Mahalanobis uzaklığı, öklit uzaklığının tersine değişkenler arasındaki korelasyonu dikkate alan bir uzaklık ölçüm yöntemidir. İki nokta arasındaki mahalanobis uzaklığı, örneklerin dağılım yönleri ve dağılımın büyüklüğü göz önüne alınarak hesaplanır. İki boyutlu bir uzayda bir noktadan eşit öklit uzaklığındaki noktalar bir daire çizerken, eşit mahalanobis uzaklığına sahip noktalar elips çizer. İki nokta arasındaki mahalanobis uzaklığı, Σ kovaryans matrisini göstermek üzere, (4.10)'deki gibi tanımlanır.

$$d_m^2 = [x - y]^T \Sigma^{-1} [x - y] \quad (4.10)$$

Kovaryans matrisinin diagonal olması, yani örnek dağılımını temsil eden elipsin koordinat eksenleriyle aynı yönlü olması durumunda hesaplanan uzaklık normalize öklit (normalized Euclidean) uzaklığı olarak isimlendirilir. Bu durumda değişkenler arasındaki korelasyon sıfırdır.

$$d_n^2 = \sum_i \frac{(x_i - y_i)^2}{\sigma_i^2} \quad (4.11)$$

Kovaryans matrisinin birim matris olması durumunda ise mahalanobis uzaklığı öklit uzaklığına eşit olur. Bu durum değişkenler arasında korelasyonun sıfır ve varyanslarının 1 olma durumudur.

$$d^2 = \sum_i (x_i - y_i)^2 \quad (4.12)$$

4.1.3 K-Means (K-Ortalama) Algoritması

K-means algoritması temel olarak bir uzayda bulunan noktalar ile seçilen merkezler arasındaki mesafeye göre benzerlik grupları oluşturarak kümeleme işlemi gerçekleştirir. Bir kümeyi oluşturan örneklerin, başka kümelerle ait örneklerden daha çok birbirlerine benzemeleri, yani küme içindeki varyansın mümkün olduğunca düşük ve kümelerin birbirinden mümkün olduğunca uzak olmaları istenir. Kümeleme stratejisi karesel hata (squared error) kriterine bağlıdır. Algoritmanın amacı belirlenmiş K adet kümeyi toplam karesel hatayı minimize edecek şekilde belirlemektir.

Toplam örnek sayısı n ve K tane sınıf $\{C_1, C_2, \dots, C_k\}$ olmak üzere, her bir kümenin merkezi kümeyi oluşturan elemanların aritmetik ortalaması alınarak hesaplanır.

$$m^{(k)} = \frac{1}{n_k} \sum_{i=1}^{n_k} x_i^{(k)} \quad (4.13)$$

Herhangi bir sınıf için karesel hata, sınıf içindeki her bir elemanın merkezden olan öklit uzaklığının karesinin toplamıdır. Sınıfa ait bu karesel hata toplamı sınıf içi varyans (within cluster variation) olarak isimlendirilir.

$$e_k^2 = \sum_{i=1}^{n_k} (x_i^{(k)} - m^{(k)})^T (x_i^{(k)} - m^{(k)}) \quad (4.14)$$

Bütün sınıfların sınıf içi varyanslarının toplamını minimize etmek K means algoritmasının amacıdır.

$$E_K^2 = \sum_{k=1}^K e_k^2 \quad (4.15)$$

K- means algoritması temel olarak dört adımdan oluşur.

- 1) Oluşturulacak olan küme sayısı ve K adet küme merkezi belirlenir.
- 2) Uzaydaki her bir nokta bütün küme merkezleriyle karşılaştırılarak kendilerine en yakın merkezin ait olduğu kümeye atanırlar.
- 3) Yeni oluşan durumda her bir kümenin merkezi yeniden hesaplanır.
- 4) Adım 2 ve 3 durdurma kriterinin belirlenen bir değerine kadar devam ettirilir. Durdurma kriteri sınıf değiştiren örnek sayısı veya toplam iterasyon sayısı ile ilgili olabilir.

Tezde kullanılan bir diğer sınıflayıcı olan yapay sinir ağı ayrı bir başlık altında incelenmiştir.

4.2 Özellik Çıkarımı

Birçok örnek tanıma uygulamasında örnekler sınıflayıcıya direkt olarak sunulmazlar. Örneğin, bir nesne tanıma probleminde fiziksel nesne bir kamerayla algılanarak nesnenin sayısal temsili elde edilmelidir. Kamera vasıtasıyla algılanan gerçek dünyadaki nesne, bilgisayar ortamında artık sayısal bir görüntüdür ve bu görüntü ham görüntü (raw image) olarak isimlendirilir. Gerçek nesnenin sayısal karşılığı aynı zamanda çok boyutlu bir uzayda bir vektör gösterir. Elde edilen ham görüntü veya örnek çoğu durumda çok fazla ölçüm içerdiğinden, yani çok boyutlu olduğundan örneklerin sınıflayıcıya bu halleriyle verilmesi problem oluşturabilir. Birinci problem, çok boyutlu bilgiyi işlemek oldukça fazla işlem yükü getirir, ikinci diğer önemli problem ise örnekteki ölçüm sayısı arttıkça, yani örneğin boyutu arttıkça, örnekleri temsil etmek için oluşturulacak eğitim seti üstel olarak artan sayıda örnekten oluşmalıdır (curse of dimensionality).

Bu yüzden örneklerden kendilerini en iyi temsil eden özellikler çıkarılır. Özellik çıkarımı, aslında örneklerin daha iyi temsil edildikleri daha düşük boyuttaki bir uzaya haritalanmalarıdır. Örneklerin, aralarında korelasyonun az olduğu mümkün olduğunca az özellik ile temsil edilmeleri istenir. Bu sayede daha az işlem yükü oluşur ve daha basit sınıflayıcılar tasarlanabilir.

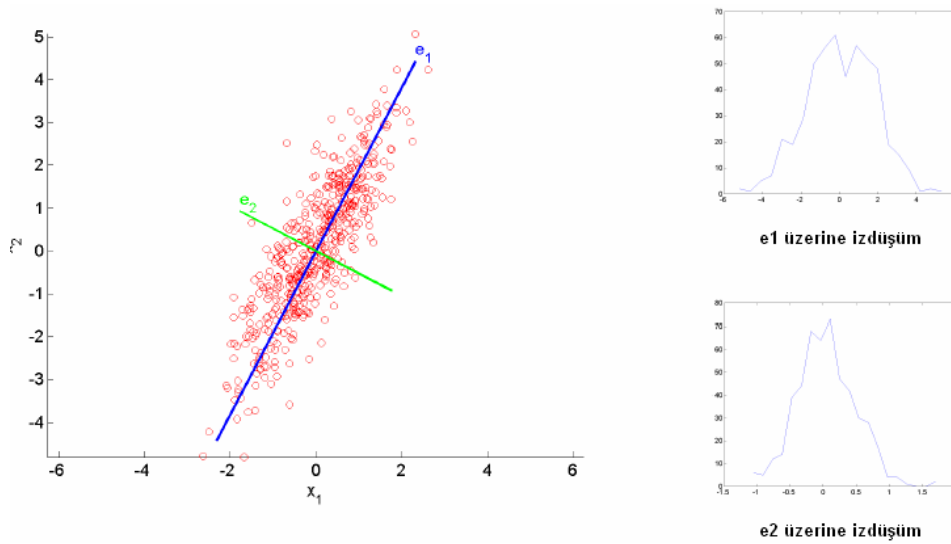
Örnekleri temsil eden özellikler, örneklerin kendilerinden direkt olarak elde edilebildiği gibi

değişik boyut indirgeme metodlarıyla da elde edilebilir. Örnek tanımda örneklerin boyutlarını düşürerek özellik çıkaran iki temel yöntem Ana Bileşen Analizi (PCA, Principal Component Analysis) ve Doğrusal Ayırtaç Analizi (LDA, Linear Discriminant Analysis) 'dır. Her iki boyut indirgeme tekniğinde de örneklerin uzay içinde birbirlerine göre olan yerleri değiştirilmez, yalnızca örnekleri daha iyi bir bakış açısıyla değerlendirebilmek için yeni eksen takımları oluşturulur.

4.2.1 Ana Bileşen Analizi (PCA, Pricipal Component Analysis)

PCA'nın avantajı genelleme yeteneğinden gelir. PCA örneklerin uzay içindeki dağılımını, yani varyansı kullanarak örneklerin indirgenmiş bir uzayda ifade edilmesini sağlar. Yöntem en iyi izdüşümü (projection) belirleyerek örneklerin varyansın maksimum olduğu ve aralarında korelasyonun olmadığı indirgenmiş bir uzayda temsil edilmesini sağlar. En yüksek varyansı elde etmek için örnekler, dağılımın yönünü ifade eden özvektörler üzerine düşürülerek, bir alt uzayda ifade edilirler.

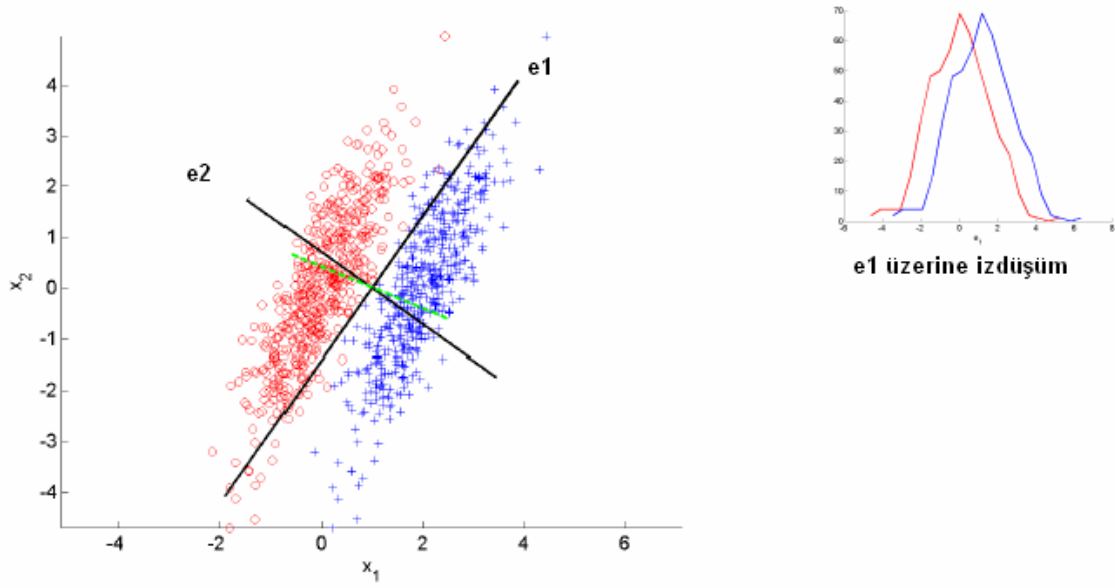
Şekil 4.1'de iki boyutlu bir uzayda örnek bir dağılımın özvektörleri ve örneklerin bu vektörler üzerine izdüşürüldükten sonra elde edilen histogramlar gösterilmiştir.



Şekil 4.1 Ana Bileşen Analizi

PCA'nın dezavantajı ise eğitici bir yöntem olması, yani örneklerin etiket bilgilerini dikkate almamasıdır. Bu durum kendisinden sonra kullanılan sınıflayıcı için problem olabilir. Şekil 4.2'de iki farklı sınıfa ait örneklerin yalnız ilk özvektör üzerine düşürüldüğünde sınıfların iç

içe girdikleri görülmektedir. Bu durumda sınıfları birbirinden ayırmak güçleşmektedir.

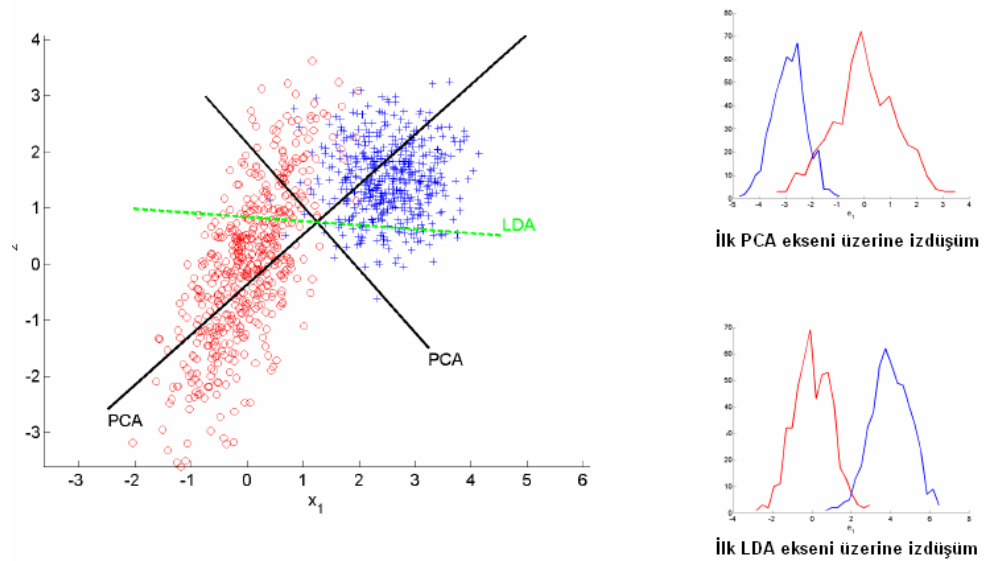


Şekil 4.2 Ana Bileşen Analizinin dezavantajı

4.2.2 Doğrusal Ayırtaç Analizi (LDA, Linear Discriminant Analysis)

PCA örnekler hakkında genelleme yaparak özellik çıkarmasına karşın, LDA indirgenmiş bir uzayda sınıfları birbirinden en iyi şekilde ayırmayı hedefler.

Şekil 4.2’de PCA eksenleri (özvektörler) ve LDA eksenini ile örneklerin eksenler üzerine izdüşümleri sonucu oluşan histogramlar gösterilmiştir.



Şekil 4.3 Doğrusal Ayırtaç Analizi

5. BOYUT İNDİRGEME YOLUYLA YÜZ SEZİMİ

Bu bölümde genelleştirme özelliğine sahip temel bir boyut indirgeme yöntemi olan Ana Bileşen Analizinin yüz seziminde nasıl kullanıldığı anlatılmıştır.

5.1 Özyüz (Eigenface) Metodu

Bir görüntü, boyutları kendi piksel sayısına eşit bir görüntü uzayında (image space) bir nokta veya vektör ile temsil edilir. Örneğin, $N_x \times N_y$ boyutlarındaki bir gri seviyeli görüntü P boyutlu bir uzayda bir nokta gösterir. P , N_x ile N_y 'nin çarpımına eşittir ve görüntünün herbir boyuttaki değeri 0 ile 255 arasındadır.

İki boyutlu matris formundaki bir görüntü tek bir sütun matrisi, yani bir vektör olarak ifade edilebilir. İnsan yüzlerine ait birçok örnek toplanıp, vektörel formda ifade edildiklerinde görüntü uzayı içerisinde belli bir bölgede toplanırlar. Bunun sebebi bütün insan yüzlerinin tüm farklılıklara rağmen temel yapılarının aynı olmasıdır, yani uzaydaki dağılımları tam olarak aynı olmasada, insan yüzünü belirleyen göz, burun ve ağzın yerleri arasında bir korelasyon vardır. Bu yüzden özyüz analizi için korelasyon çıkış noktasıdır.

Özyüz metodu Ana Bileşenler Analizinin görüntüler üzerinde uygulanmasıdır. İncelenen yüz ile ilgili özellikler, bütün eğitim setinden elde edilen kovaryans matrisinin özvektörleri üzerine izdüşümü yapılarak elde edilir. Eksenlerini özvektörlerin oluşturduğu görüntü uzayı içerisindeki bu alt uzay yüz uzayı (face space) olarak isimlendirilir. Sınıflama işlemi bu alt uzay içinde gerçekleştirilir. En yakın orta sınıflayıcı ve Bayes sınıflayıcı en çok kullanılan sınıflayıcı modelleridir. Turk ve Pentland (1991) yüz tanıma problemi için standart özyüz yöntemini önermişler ve en yakın orta değer sınıflayıcısını kullanmışlardır.

n boyutlu uzayda, N adet örnekten oluşan bir eğitim seti $\{x_1, x_2, \dots, x_N\}$ için n boyutlu uzaydan m boyutlu uzaya ($m < n$) haritalama ortonormal bir dönüşüm matrisi W ile yapılır.

$$y_k = W^T x_k \quad (5.1)$$

$W \in R^{n \times m}$ ve indirgenmiş uzaydaki özellik vektörü $y_k \in R^m$ 'dir. Toplam saçılım matrisi (scatter matrix) S_T ile gösterilirse;

$$S_T = \sum_{k=1}^N (x_k - \mu)(x_k - \mu)^T \quad (5.2)$$

Bütün yüz vektörlerinin aritmetik ortalaması alınarak ortalama yüz (mean face) elde edilir. Ortalama yüz $n \times 1$ boyutunda bir vektördür.

$$\mu = \frac{1}{N} \sum_{k=1}^N x_k \quad (5.3)$$

Lineer dönüşümden sonra elde edilen özellik vektörlerinin $\{y_1, y_2, \dots, y_N\}$ saçılım matrisi $W^T S W$ 'ye eşittir. Ana Bileşen Analizi ile indirgenmiş uzaydaki toplam dağılım matrisinin determinantını maksimum yapan en uygun dönüşüm matrisi bulunur.

$$W_{opt} = \arg \max_w |W^T S_T W| \quad (5.4)$$

(5.4)'deki ifadeyi maksimum yapan dönüşüm matrisi sütunlarını kovaryans matrisinin özvektörlerin oluşturduğu özvektör matrisidir.

$$\Sigma W = W D \quad (5.5)$$

Σ kovaryans D ise özdeğer matrisidir.

Eğitim setindeki herhangi bir yüzün ortalama yüzden uzaklığını gösteren görüntü, ortalama çıkarılmış görüntü (mean subtracted image) olarak isimlendirilir.

$$\bar{x} = x - \mu \quad (5.6)$$

Eğitim setindeki bütün vektörlerin ortalama vektörden uzaklıkları $n \times N$ boyutlarındaki bir fark matrisi ile gösterilir.

$$A = \begin{bmatrix} \bar{x}_1 & \bar{x}_2 & \dots & \bar{x}_N \end{bmatrix} \quad (5.7)$$

Kovaryans matrisi ise fark matrisinin kendisinin devriği ile çarpımına eşittir.

$$\Sigma = A A^T = \frac{1}{N} \sum_{i=1}^M (x_k - \mu)(x_k - \mu)^T \quad (5.8)$$

Özvektörler kovaryans matrisi normal olarak X 'den elde edilir, fakat kovaryans matrisinin boyutlarının fazla oluşundan dolayı özellik vektörlerinin hesaplanması oldukça fazla işlem

yükü getirir. Yüze ait bir örnek, n elemanlı bir vektör gösterdiğinde, kovaryans matrisi $n \times n$ boyutlarında bir matrisi gösterir. Çoğunlukla eğitim setindeki örnek sayısı, örneklerin boyut sayısından az olduğundan ($N \ll n$), gerçek kovaryans matrisinin özvektörleri $N \times N$ boyutlu bir matrisin özvektörlerinden hesaplanabilir.

$$Y = A^T . A = \frac{1}{N} \sum_{i=1}^{M_i} (x_k - \mu)^T (x_k - \mu) \quad (5.9)$$

Y matrisinin özvektörleri ve özdeğerleri sırasıyla v ve λ ile gösterilirse;

$$Y.v_i = \lambda_i.v_i \quad (5.10)$$

(5.10)'da Y matrisi yerine konulursa;

$$A^T A v_i = \lambda_i v_i \quad (5.11)$$

Her iki taraf fark matrisi A ile çarpılır.

$$A A^T A v_i = A \lambda_i v_i \quad (5.12)$$

λ_i değeri skaler bir değeri gösterdiğinden ifadenin sağında, başa alınabilir ve ifadede $\sum$ yerine konur.

$$X A v_i = \lambda_i A v_i \quad (5.13)$$

$A v_i$ ifadesi u_i ile gösterilirse;

$$X u_i = \lambda_i u_i \quad (5.14)$$

(5.14)' de görüldüğü gibi u_i $\sum$ 'in bir özvektörüdür ve Y matrisinin özvektöründen elde edilebilir.

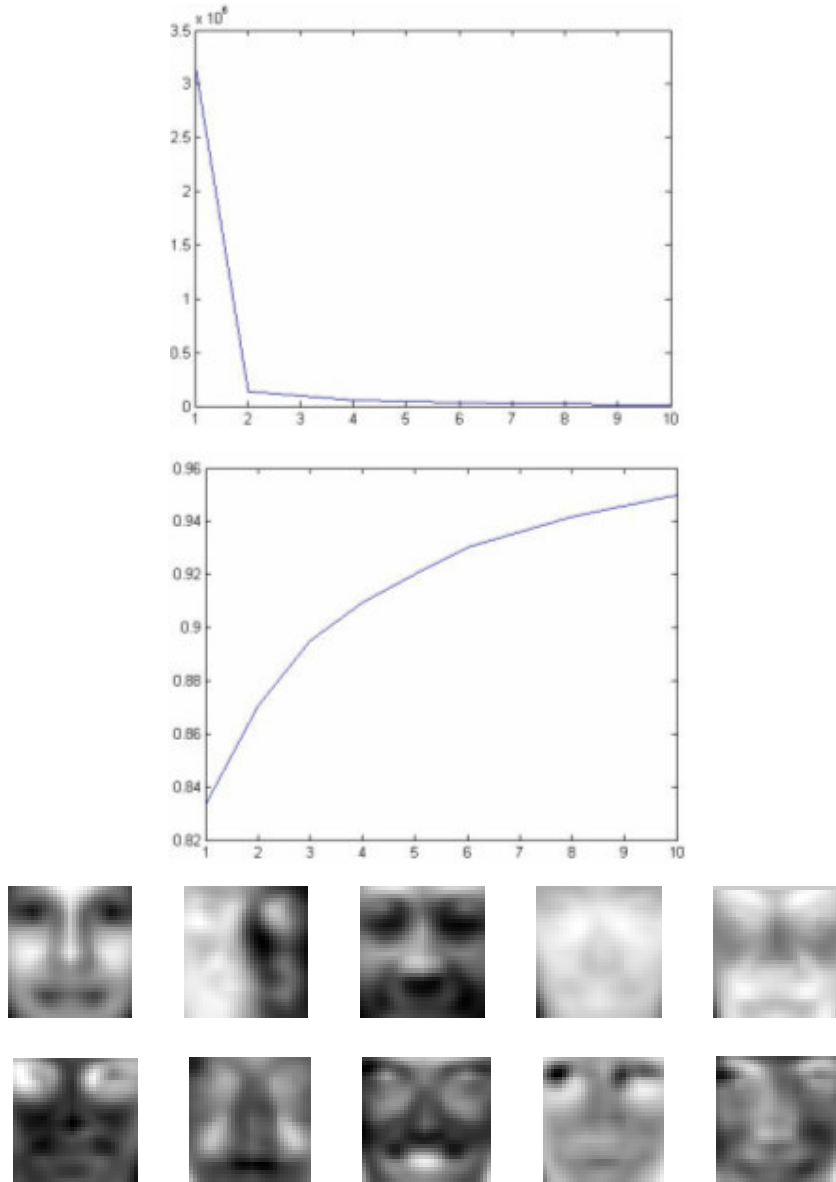
$$u_i = A v_i \quad (5.15)$$

Bu sayede işlem yükü hafifletilmiş ve özvektörler bu sayede etkin bir biçimde hesaplanmış olur.

Özvektörler, özdeğerlerine göre sıralanarak dönüşüm matrisi W elde edilir. Özdeğerler, özvektörlerin büyüklüklerini, yani varyansı verirler. Dönüşüm matrisi $n \times n$ boyutlarında olmasına karşın, örnek sayısının boyut sayısından az olması durumunda özdeğeri sıfırdan

farklı N tane özvektör vardır. Bu sebeple dönüşüm matrisi $n \times N$ boyutuna indirgenir. Dönüşüm işleminde N adet özvektör veya özyüz kullanmak yerine örneklerin varyansına katkısı az olan küçük değerli özvektörleri ihmal ederek m tane ($m \ll n$) vektör kullanılır. En yüksek özdeğere sahip özvektör, en yüksek varyansı temsil eder ve birinci özvektör olarak isimlendirilir, ayrıca birinci özvektör en yüksek genelleme özelliğine sahiptir.

Örnekler görüntü uzayında piksel değerleri ile temsil edilirken, yüz uzayında özvektörler üzerindeki değerlerini gösteren ağırlıklarla temsil edilirler. Şekil 5.1’de Harvard verisetinden alınan 300 örnekten elde edilen ilk 10 özdeğer, özdeğerlerin kümülatif toplamı ve özyüzler gösterilmiştir.

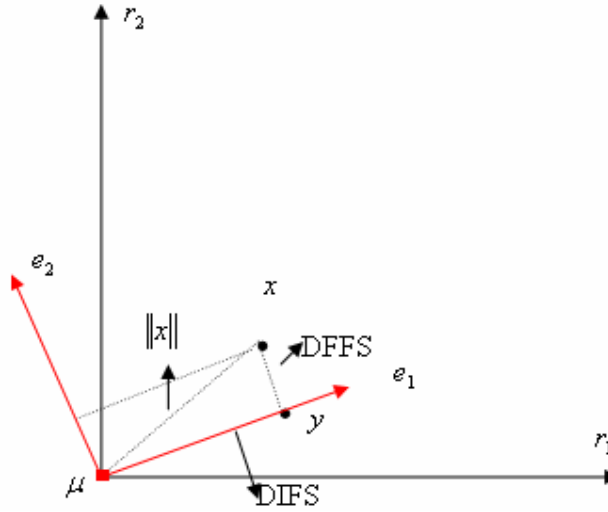


Şekil 5.1 Özdeğerler ve özyüzler

Standart özyüz yönteminde yüz sezimi için yeniden yapılandırma hatasına (reconstruction error, residual) bakılır. Yeniden yapılandırma hatası aslında örneğin yüz uzayından olan öklit uzaklığıdır (DFFS, Distance From Face Space). Bir diğer uzaklık ise izdüşümü alınan örnekle ortalama yüz arasındaki özellik (yüz) uzayı içindeki uzaklıktır (DIFS, Distance In Face Space). DIFS öklit veya mahalanobis uzaklığına bakılarak hesaplanabilir.

x ortalama yüz çıkarılmış örneği, μ ortalama yüzü, y ortalama çıkarılmış örneğin yüz uzayına izdüşümünü, D özdeğerler matrisini, λ özdeğerleri ve $\hat{x}$ 'de, m boyutlu alt uzay için yapılandırılan örneği göstermek üzere, (5.16)'deki gibi hesaplanır. Şekil 5.2'de e_1 ve e_2 yüz uzayının r_1 ve r_2 'de görüntü uzayının eksenleridir.

$$\hat{x} = Wy + \mu \quad (5.16)$$



Şekil 5.2 Yüz uzayına izdüşüm

Yapılandırma hatası ϵ^2 (5.17)'deki gibi hesaplanır.

$$\epsilon^2(x) = \left\| x - \hat{x} \right\|^2 = \left\| x \right\|^2 - \sum_{i=1}^m y_i^2 = \sum_{i=m+1}^n y_i^2 \quad (5.17)$$

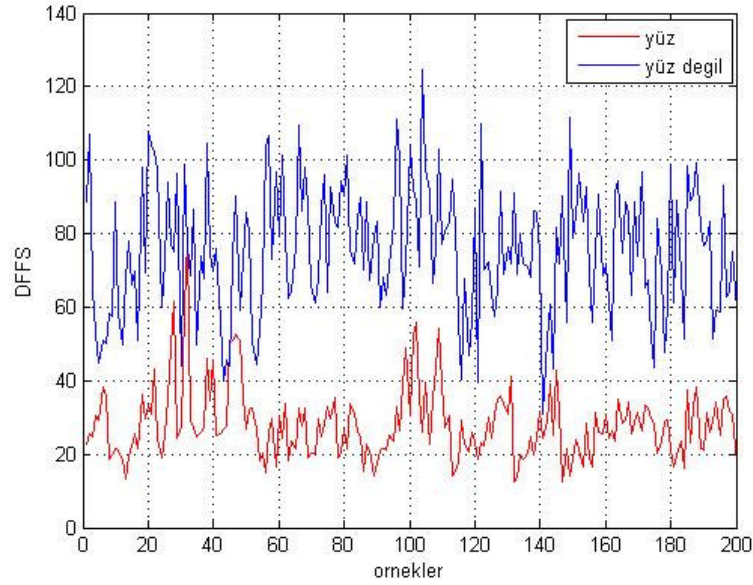
Uzay içindeki uzaklık (DIFS) ise öklit veya mahalanobis uzaklığına bakılarak hesaplanabilir. Öklit ve mahalanobis uzaklıklarının kareleri sırasıyla (5.18) ve (5.19)'deki gibi hesaplanır.

$$d_o^2 = y^T y = \sum_{i=1}^m y_i^2 \quad (5.18)$$

Yüz uzayında mahalanobis uzaklığı, izdüşümü alınan örnekler arasında korelasyon sıfır olduğundan, normalize öklit uzaklığına indirgenir.

$$d_m^2 = y^T D^{-1} y = \sum_{i=1}^m \frac{y_i^2}{\lambda_i} \quad (5.19)$$

Şekil 5.3’de 200 adet yüz ve yüze ait olmayan örneğin 20 boyutlu yüz uzayı için oluşan yapılandırma hataları gösterilmiştir. Örnekler histogram eşitleme yöntemiyle önışlemeden geçirildikten sonra 255’e bölünerek normalize edilmiştir.



Şekil 5.3 Yüz ve yüz olmayan örneklerin yapılandırma hataları

6. YAPAY SİNİR AĞLARI

Yaşayan organizmaların sinir hücreleri, öğrenme yeteneklerinden dolayı, makine öğrenmesi uygulamaları için bir araştırma konusu olmuştur. Beynin nöronları organize etme yeteneği böylece kesin hesaplamaları gerçekleştirmesi (örüntü tanıma, algılama gibi) bugünün en hızlı sayısal bilgisayarlarından daha hızlıdır. Diğer taraftan bir sinir hücresinin tepki hızı günümüzün bilgisayarlarına göre oldukça yavaş olmakla birlikte duyuşal bilgileri son derecede hızlı değerlendirebilmektedir. Bu nedenle insan beyni; öğrenme, birleştirme, uyarılma ve genelleştirme yeteneği nedeniyle son derece karmaşık, doğrusal olmayan ve paralel dağılmış bir bilgi işleme sistemi olarak tanımlanabilir.

Beynin bilgi işleme yöntemine uygun olarak YSA, bir öğrenme sürecinden sonra bilgiyi toplama, hücreler arasındaki bağlantı ağırlıkları ile bu bilgiyi saklama ve genelleme yeteneğine sahip paralel dağılmış bir işlemcidir. Öğrenme süreci, arzu edilen amaca ulaşmak için YSA ağırlıklarının yenilenmesini sağlayan öğrenme algoritmalarını ihtiva eder.

Yapay sinir ağlarının temel özellikleri aşağıdaki gibidir;

Doğrusal olmama; YSA' nın temel işlem elemanı olan hücre doğrusal değildir. Dolayısıyla hücrelerin birleşmesinden meydana gelen YSA da doğrusal değildir ve bu özellik bütün ağa yayılmış durumdadır. Bu özelliği ile YSA, doğrusal olmayan karmaşık problemlerin çözümünde en önemli araç olmuştur.

Öğrenme; YSA' nın arzu edilen davranışı gösterebilmesi için amaca uygun olarak ayarlanması gerekir. Bu, hücreler arasında doğru bağlantıların yapılması ve bağlantıların uygun ağırlıklara sahip olması gerektiğini ifade eder. YSA' nın karmaşık yapısı nedeniyle bağlantılar ve ağırlıklar önceden ayarlı olarak verilemez yada tasarlanamaz. Bu nedenle YSA, istenen davranışı gösterecek şekilde ilgilendiği problemde aldığı eğitim örneklerini kullanarak problemi öğrenmelidir.

Genelleme; YSA, ilgilendiği problemi öğrendikten sonra eğitim sırasında karşılaşmadığı test örnekleri için de arzu edilen tepkiyi üretebilir.

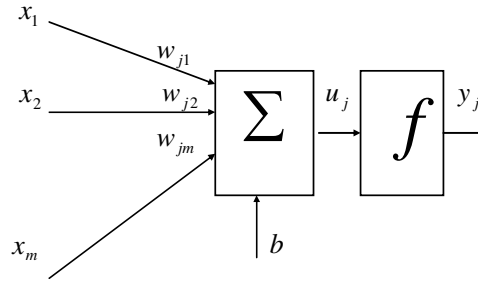
Uyarlanabilirlik; YSA, ilgilendiği problemdeki değişikliklere göre ağırlıklarını ayarlar. Yani, belirli bir problemi çözmek amacıyla eğitilen YSA, problemdeki değişimlere göre tekrar eğitilebilir, değişimler devamlı ise gerçek zamanda da eğitime devam edilebilir. Bu özelliği ile YSA, uyarlamalı örnek tanıma ve sinyal işleme gibi alanlarda etkin olarak kullanılır.

Hata Toleransı ; YSA, çok sayıda hücrenin çeşitli şekillerde bağlanmasından oluştuğundan

paralel dağılmış bir yapıya sahiptir ve ağına sahip olduğu bilgi, ağıdaki bütün bağlantılar üzerine dağılmış durumdadır. Bu nedenle, eğitilmiş bir YSA'nın bazı bağlantılarının hatta bazı hücrelerinin etkisiz hale gelmesi, ağına doğru bilgi üretmesini önemli ölçüde etkilemez. Bu nedenle, geleneksel yöntemlere göre hatayı tolere etme yetenekleri son derece yüksektir.

6.1 Yapay Sinir Ağı Hücresi

Temel bir yapay sinir ağı hücresi biyolojik sinir hücresine göre çok daha basit bir yapıya sahiptir. Yapay sinir ağı hücresinde temel olarak dış ortamdan ya da diğer nöronlardan alınan veriler yani girişler, ağırlıklar, toplama fonksiyonu, aktivasyon fonksiyonu ve çıkışlar bulunmaktadır. Dış ortamdan alınan veri ağırlıklar aracılığıyla nörona bağlanır ve bu ağırlıklar ilgili girişin etkisini belirler. Toplama fonksiyonu ise net girişi hesaplar, net giriş, girişlerle bu girişlerle ilgili ağırlıkların çarpımının bir sonucudur. Aktivasyon fonksiyonu işlem süresince net çıkışını hesaplar ve bu işlem aynı zamanda nöron çıkışını verir. Genelde aktivasyon fonksiyonu doğrusal olmayan (nonlinear) bir fonksiyondur. Şekil 6.1'de temel nöron modeli gösterilmiştir, b bir sabittir, bias veya aktivasyon fonksiyonunun eşik değeri olarak adlandırılır.



Şekil 6.1 Temel sinir hücresi modeli

x vektörü nöronun girişleridir. Bütün girişler eş zamanlı olarak ağa sunulur ve belirli bir ağırlık sabitleriyle çarpılarak ağırlıklandırılırlar. Hücre bağlantı ağırlıkları hücrenin ağırlık vektörü w_j ile giriş vektörü x 'in çarpımıdır.

$$u_j = w_j^T x + b_j = \sum_{i=1}^m x_i w_{ji} + b_j \quad (6.1)$$

(6.1)'de hesaplanan değer hücrenin aktivasyon değeridir. Ağırlıklı toplam sonucunda oluşan aktivasyon değeri $f(\cdot)$ ile gösterilen aktivasyon fonksiyonuna giriş değeri olarak verilir ve işlem biriminin çıkışı belirlenir.

$$y_j = f(u_j) \quad (6.2)$$

Oluşturulan bu model, gerçek sinir hücresinin çok sadeleştirilmiş bir hali olmasına karşın bir çok makina görmesi uygulamasında başarıyla kullanılmıştır.

İşlem birimi çıkışı, biyolojik nöronlardaki aksonlar boyunca iletilen sinyale karşılık gelir. Gerçek nöronda, nöronun uyarımı arttıkça akson boyunca iletilen sinyalin artan frekansına karşın, matematiksel nöron modelinde çıkış genliği değişmektedir. Aktivasyon fonksiyonu nöronun karakteristiğini temsil etmektedir ve değişik formlarda aktivasyon fonksiyonları kullanılmaktadır. En yaygın kullanıma sahip aktivasyon fonksiyonları sigmoid ve tanjant sigmoid fonksiyonlarıdır.

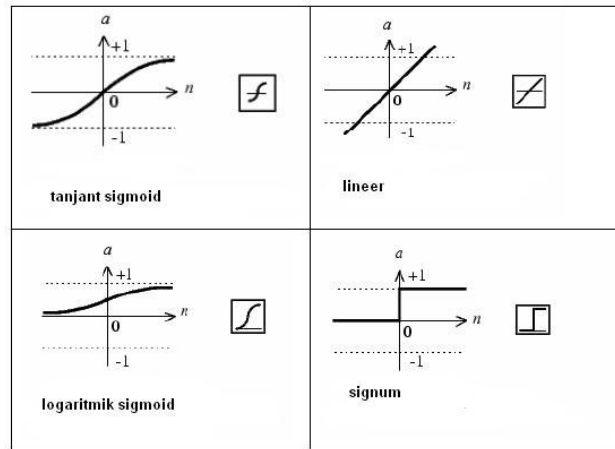
Her iki fonksiyonda sınırsız aralıkta değişen değerler sınırlı bir aralığa sıkıştırılmıştır ve türevleri alınabilen fonksiyonlardır. Sigmoid fonksiyonu 0 ile 1 arasında değerler alırken, bu aralık tanjant sigmoid için -1 ile 1 arasındadır. Sigmoid ve tanjant sigmoid fonksiyonlarının ifadeleri sırayla (6.3) ve (6.4)'de gösterilmiştir.

$$f(v) = \frac{1}{1 + \exp(-av)} \quad (6.3)$$

$$f(v) = \frac{1 - \exp(-av)}{1 + \exp(-av)} \quad (6.4)$$

a eğim parametresini göstermektedir.

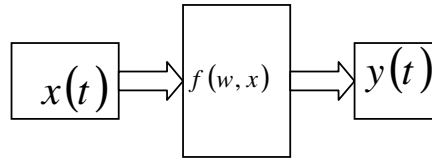
Diğer yaygın aktivasyon fonksiyonları da şekil 6.2'de gösterilmiştir.



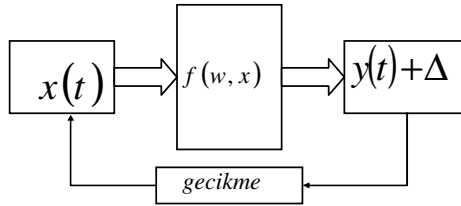
Şekil 6.2 Aktivasyon fonksiyonları

6.2 Yapay Sinir Ağları Modelleri

En basit ve en genel yapay sinir ağları tek yönlü sinyal akışını kullanırlar. Yapay sinir ağı modelleri temel olarak iki grupta toplanmaktadır. İleri beslemeli yapay sinir ağları ve geri beslemeli yapay sinir ağları. İleri beslemeli yapay sinir ağlarında gecikmeler yoktur, işlem girişlerden çıkışlara doğru ilerler. Çıkış değerleri öğreticiden alınan istenen çıkış değeriyle karşılaştırılarak bir hata sinyali elde edilerek ağ ağırlıkları güncellenir. Geri beslemeli yapay sinir ağlarında ise tıpkı kontrol uygulamalarında olduğu gibi gecikmeler söz konusudur. Geri beslemeli nöral ağ, çıkışlar girişlere bağlanarak ileri beslemeli bir ağdan elde edilir. İleri ve geri beslemeli ağ yapıları Şekil 6.3 ve Şekil 6.4’de gösterilmiştir.



Şekil 6.3 İleri Beslemeli Yapı



Şekil 6.4 Geri Beslemeli Yapı

6.2.1 Yapay Sinir Ağlarının Eğitilmesi

İnsan beyni doğumdan sonraki gelişme sürecinde çevresinden duyu organlarıyla algıladığı davranışları yorumlar ve bu bilgileri diğer davranışlarında kullanır. Yaşadıkça beyin gelişir ve tecrübelenir. Artık olaylar karşısında nasıl tepki göstereceğini çoğu zaman bilmektedir. Fakat hiç karşılaşmadığı bir olay karşısında yine tecrübesiz kalabilir. Yapay sinir ağlarının öğrenme sürecinde de, tıpkı dış ortamdan gözle veya vücudun diğer organlarıyla uyarıların alınması gibi dış ortamdan girişler alınır, bu girişlerin beyin merkezine iletilerek burada değerlendirilip tepki verilmesi gibi yapay sinir ağında da aktivasyon fonksiyonundan geçirilerek bir tepki çıkışı üretilir. Öğrenme eğitici veya eğitici olarak yapılabilir.

Eğitici (supervised) öğrenmede, her örnekler ve arzu edilen değerler bir eğitici tarafından

sağlanır. Arzu edilen çıkış ile sinir ağı çıkışı arasındaki fark hata ölçüsüdür ve ağ parametrelerini güncellemekte kullanılır.

Eğiticişiz öğrenmede ise, eğiticişli öğrenmedeki gibi arzu edilen çıkışlar bilinmemektedir. Bu yüzden kesin bir hata bilgisini ağın davranışını değıştirmekte kullanmak mümkün değıildir. Cevabın doğıruluğı veya yanlışlığı hakkında bilgi sahibi olunmadığı için öğrenme, girişlerin verdiğı cevaplar gözlenerek başarıya ulaşılır.

Çeşitli öğrenme algoritmalarıyla ağın ağırlıkları değıştirilerek, hata azaltılıp gerçek çıkışa yaklaşılmaya çalışılır. Ağırlıkların sürekli yenilenip istenilen sonuca ulaşılana kadar geçen zamana öğrenme adı verilir. Yapay sinir ağı öğrendikten sonra daha önce verilmeyen girişler verilir, ağın geçmiş tecrübesine dayanarak yeni örnekler hakkında karar vermesi beklenir.

6.2.2 Geriye Yayılım (Backpropagation) Algoritması

Geriye yayılım algoritması en genel çok katmanlı ağların eğiticişli öğrenme metodudur.Çok katmanlı ağ en az bir saklı katmana sahip ileri beslemeli bir ağ modelidir.Çıkış katmanı ile gizli katman arasındaki ağırlıklar, ağın çıkışı ve arzu edilen çıkış arasındaki fark kullanılarak hesaplanabilmesine karşın, giriş ile gizli katman veya birden fazla gizli katman varsa gizli katmanlar arasındaki ağırlıkların güncellenmesi bir problem olarak ortaya çıkar.Gizli katman veya katmanlar için arzu edilen değıerler bulunmadığından direkt olarak hata hesaplanamaz, bu durumda ağın çıkış hatası gizli katman veya katmanlara yansıtılır.Geriye yayılım algoritması her saklı katman için etkin bir hata hesaplama yöntemi sunar ve girişten saklı katmana veya saklı katmanlar arasındaki ağırlıkların güncellenebilmesini sağlar.

Ağın eğitilmesi ileri yönlü hesaplamadan, yani giriş katmanına örnekler sunulduktan sonra sonuçların elde edilmesiyle başlar.Ağın gerçek çıkışı ile arzu edilen değıer karşılaştırılarak hata fonksiyonu hesaplanır ve ağ çıkışı istenilen değıere yaklaşıncaya kadar hata fonksiyonu minimize edilmeye çalışılır.

Eğitme hatası ağın gerçek çıkışıyla arzu edilen çıkış arasındaki farkın karesi olarak verilir. y_k çıkış katmanındaki k .nöron çıkışı, d_k istenilen hedef değıeri, c saklı katmandaki nöron sayısını ifade etmek üzere p örneğinin sebep olduğı E^p hatası;

$$J(w) = E^p = \frac{1}{2} \sum_{k=1}^c (d_k - y_k)^2 \quad (6.5)$$

bağıntısıyla bulunur.Geriye yayılım algoritması ağırlığın bir fonksiyonu olan hata

fonksiyonunun eğim düşümünü (gradient descent) bulmaya dayanır.

$$\Delta w_{ji} = -\eta \left(\partial J / \partial w_{ji} \right) \quad (6.6)$$

η öğrenme oranı, ağırlık hataya karşı duyarlılığını ifade eder. Ağırlık güncellemesi, ağırlıkların eski değerleriyle değişim miktarları toplanarak yapılır.

$$w(m+1) = w(m) + \Delta w(m) \quad (6.7)$$

Tek bir gizli katmana sahip üç katmanlı bir ağ ele alınacak olursa, w_{ji} vektörü giriş ile saklı katman arasındaki bağlantıları, w_{kj} saklı katman ile çıkış katmanı arasındaki bağlantıları, u_k çıkış katmanındaki k .birimin aktivasyon değerini göstermek üzere zincir kuralı uygulanırsa;

$$\frac{\partial J}{\partial w_{kj}} = \frac{\partial J}{\partial u_k} \frac{\partial u_k}{\partial w_{kj}} = -\delta_k \frac{\partial u_k}{\partial w_{kj}} \quad (6.8)$$

$$\delta_k = -\frac{\partial J}{\partial u_k} \quad (6.9)$$

Sonuçları elde edilir. Aktivasyon fonksiyonunun türevi alınabilen bir fonksiyon olduğu kabul edilirse;

$$\delta_k = -\frac{\partial J}{\partial u_k} = \frac{\partial J}{\partial y_k} \frac{\partial y_k}{\partial u_k} = (d_k - y_k) f'(u_k) \quad (6.10)$$

$\frac{\partial u_k}{\partial w_{kj}} = y_j$ eşitliği kullanılırsa saklı katmanla çıkış katmanı arasındaki ağırlık güncelleme veya

öğrenme kuralı;

$$\Delta w_{kj} = \eta \delta_k y_j = \eta (d_k - y_k) f'(u_k) y_j \quad (6.11)$$

olarak belirlenir. Giriş ile saklı katman arasındaki ağırlıkların güncellenmesi ise daha karmaşık bir işlemdir. Çıkıştaki toplam hataya w_{ji} yani ara katman ağırlıklarının etkisi ()'deki bağıntı ile bulunabilir.

$$\frac{\partial J}{\partial w_{ji}} = \frac{\partial J}{\partial y_j} \frac{\partial y_j}{\partial u_j} \frac{\partial u_j}{\partial w_{ji}} \quad (6.12)$$

İfadenin sağdaki ilk terimi için;

$$\begin{aligned}
\frac{\partial J}{\partial y_j} &= \frac{\partial}{\partial y_j} \left(\frac{1}{2} \sum_{k=1}^c (d_k - y_k)^2 \right) \\
&= - \sum_{k=1}^c (d_k - y_k) \frac{\partial y_k}{\partial y_j} \\
&= - \sum_{k=1}^c (d_k - y_k) \frac{\partial y_k}{\partial u_k} \frac{\partial u_k}{\partial y_j} \\
&= - \sum_{k=1}^c (d_k - y_k) f'(u_k) w_{kj}
\end{aligned} \tag{6.13}$$

eşitliği kullanılarak;

$$\delta_j = f'(u_j) \sum_{k=1}^c w_{kj} \delta_k \tag{6.14}$$

elde edilir. Buradan ağırlık güncellemeleri;

$$\Delta w_{ji} = \eta x_i \delta_j = \eta \left(\sum_{k=1}^c w_{kj} \delta_k \right) f'(u_k) x_i \tag{6.15}$$

(6.11) ve (6.15) bağıntıları geriye yayılım algoritmasını tanımlar.

7. ÖNİŞLEME

Önişlemenin amacı görüntülerin iyileştirilmesidir. Görüntü iyileştirme dijital görüntülerin, insanların daha iyi algılayabilmesi veya görüntü işleme cihazlarının performanslarının artırılması için görünümünün değiştirilmesidir.

Görüntü iyileştirme algoritmalarının uygulamanın içeriğine uygun ve gerçek zamanlı uygulamalarda mümkün olduğunca basit ve hızlı cevap vermeleri istenir.

7.1 En Uygun Düzlem Uydurma (Best Planar Fitting)

Işık kaynağının durumuna göre insan yüzünün homojen olarak aydınlatılamadığı durumlarda yüz görüntüsü üzerinde istenmeyen gölgeler oluşmaktadır. Bu durumda insan yüzüne en çok benzeyen düzlem belli bir katsayı ile çarpılıp yüz görüntüsünden çıkarılarak istenmeyen gölgelenmeler giderilmeye çalışılır.

En uygun düzlem uydurma, bir yüzey için o yüzeye en çok benzeyen düzlemin En Küçük Kareler Yöntemi (Least Square Method) ile bulunmasıdır.

x ve y bağımlı z ise bağımsız değişkeni göstermek üzere bir yüzey $\{(x_1, y_1, z_1), (x_2, y_2, z_2), \dots, (x_n, y_n, z_n)\}$ gibi bir set ile düzlem ise 2 boyutlu bir fonksiyon ile gösterilir.

$$z = f(a, b, c) = a + bx + cy \quad (7.1)$$

En uygun düzlemi bulmak için gerçek değerlerle düzlem arasındaki karesel hata toplamını gösteren bir hata fonksiyonu tanımlanır. n adet noktadan oluşan 2 boyutlu ayrık bir işaret olan yüzey için toplam hata fonksiyonu (7.2)'de tanımlanmıştır.

$$R^3(a, b, c) = \sum_{i=1}^n [z_i - f(a, b, c)]^2 = \sum_{i=1}^n [z_i - (a + bx_i + cy_i)]^2 \quad (7.2)$$

x_i, y_i ve z_i bilinen a, b, c ise düzlemi belirleyen aranan değerlerdir. Toplam hata fonksiyonu a, b ve c 'nin yalnız bir değeri için minimum değerini alır ve bu noktada a, b ve c 'ye göre kısmi türevleri sıfırdır.

$$\frac{\partial R^3}{\partial a} = 0, \quad \frac{\partial R^3}{\partial b} = 0, \quad \frac{\partial R^3}{\partial c} = 0 \quad (7.3)$$

Hata fonksiyonu, kareli ifade açılarak yeniden düzenlenir ve a, b ve c 'ye göre türevleri

alınarak sıfıra eşitlenir.

$$R^3 = \sum_{i=1}^n z_i^2 - 2z_i(a + bx_i + cy_i) + (a + bx_i + cy_i)^2 \quad (7.4)$$

$$\frac{\partial R^3}{\partial a} = -2 \sum_{i=1}^n z_i - (a + bx_i + cy_i) = 0 \quad (7.5)$$

$$\frac{\partial R^3}{\partial b} = -2 \sum_{i=1}^n x_i z_i - (a + bx_i + cy_i)x_i = 0 \quad (7.6)$$

$$\frac{\partial R^3}{\partial c} = -2 \sum_{i=1}^n y_i z_i - (a + bx_i + cy_i)y_i = 0 \quad (7.7)$$

Bağımsız değişkenler bir tarafta toplanarak denklemler yeniden düzenlenirse 3 adet lineer denklem elde edilir.

$$na + b \sum_{i=1}^n x_i + c \sum_{i=1}^n y_i = \sum_{i=1}^n z_i \quad (7.8)$$

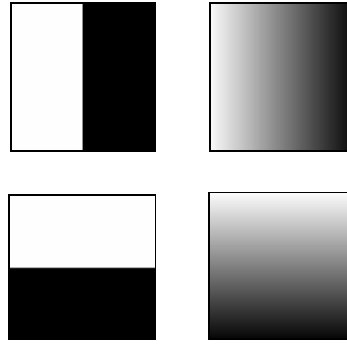
$$a \sum_{i=1}^n x_i + b \sum_{i=1}^n x_i^2 + c \sum_{i=1}^n x_i y_i = \sum_{i=1}^n x_i z_i \quad (7.9)$$

$$a \sum_{i=1}^n y_i + b \sum_{i=1}^n x_i y_i + c \sum_{i=1}^n y_i^2 = \sum_{i=1}^n y_i z_i \quad (7.10)$$

Denklemler matris formunda ifade edilirse;

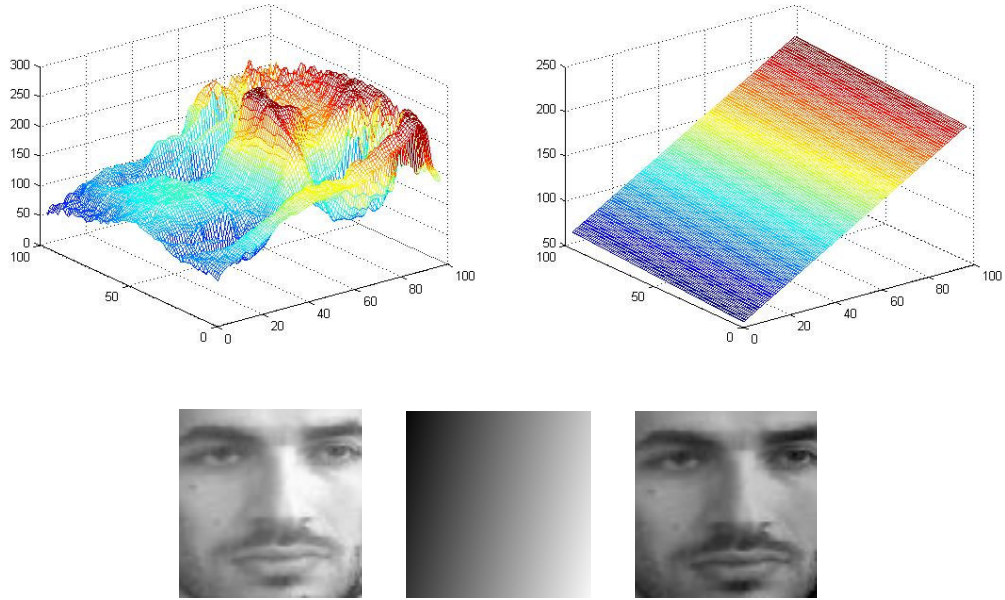
$$\begin{bmatrix} n & \sum_{i=1}^n x_i & \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i & \sum_{i=1}^n x_i y_i & \sum_{i=1}^n y_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n z_i \\ \sum_{i=1}^n x_i z_i \\ \sum_{i=1}^n y_i z_i \end{bmatrix} \quad (7.11)$$

İfadenin her iki tarafı en soldaki matrisin tersi ile çarpılarak a, b ve c değerleri hesaplanır. Şekil 7.1’de örnek yüzeyler ve en uygun düzlemler gösterilmiştir.



Şekil 7.1 Örnek yüzeyler ve en uygun düzlemler

Şekil 7.2' de ise bir yönü daha fazla aydınlatılmış yüz, en uygun düzlem işaretleri ile görüntüleri ve sonuç görüntüsü gösterilmiştir. En uygun düzlem 0.3 ile çarpılarak yüz görüntüsünden çıkarılmıştır.



Şekil 7.2 Örnek yüz ve sonuç görüntüsü

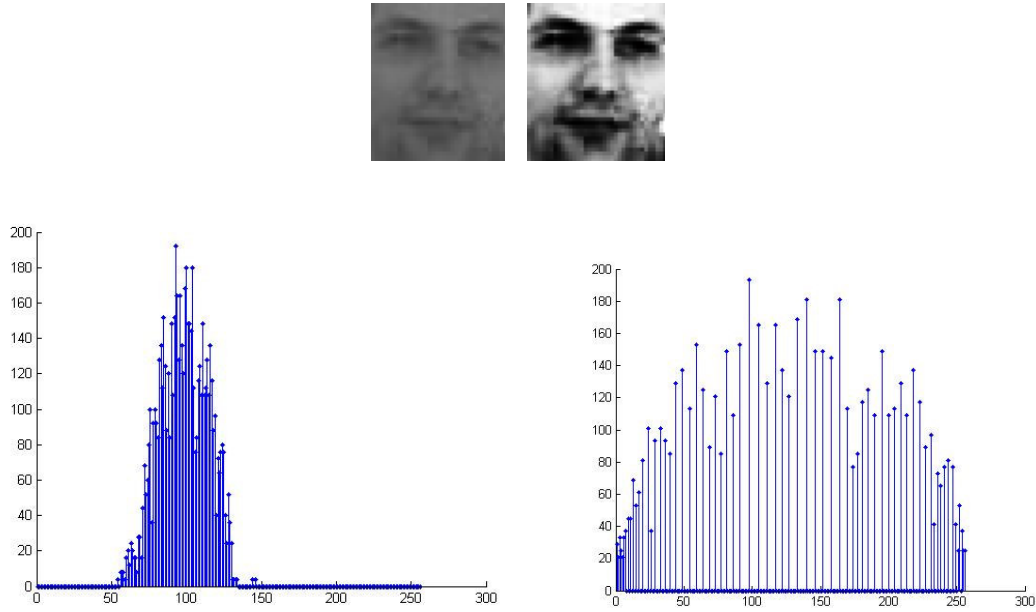
7.2 Histogram Eşitleme (Histogram Equalization)

Histogram eşitleme görüntülerdeki gri seviye aralığını, yani kontrastı değiştirmenin basit ve etkin bir yoludur. Giriş ile çıkış görüntüleri arasındaki dönüşüm bir tablo veya nokta diyagramı ile önceden belirlenebildiği gibi giriş görüntüsünün histogramı üzerinde çalışılarak da bulunabilir. İkinci durumda çıkış görüntüsünde mümkün olduğunca bütün gri seviye değerlerinin kullanılması ve her seviyede eşit sayıda piksel olması istenir.

Rowley vd.(1996) yapay sinir ağı için yüzün hangi bölümünün daha önemli olduğunu bulmak için hassasiyet analizi yapmışlardır.Yapılan analizde daha önceden ağı eğitimi için kullanılan örnekler boyutlarında ve yönelimlerinde küçük değişiklik yapılarak tekrar kullanılmıştır. 20×20 boyutlardaki insan yüzleri 4×4 boyutlarındaki alt bölümlere ayrılarak sırayla herbir alt görüntü parçası rastgele görüntü ile değiştirilmiş ve ağ bu her durum için test edilerek hata değerleri kaydedilmiştir.Gürültü ile değiştirilmesi durumunda ağı daha fazla hatalı sonuç vermesine sebep olan yerler ağ için daha önemli yerlerdir ve yapılan analiz sonucunda ağ için en önemli yerlerin gözler ve sonrasında burun ile ağız bölgeleri olduğu görülmüştür.

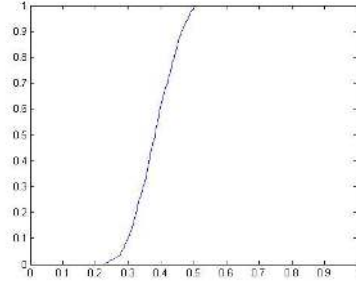
Histogram eşitleme ile farklı aydınlatma koşullarından ve kamera karakteristiklerinden dolayı gri seviyenin dar bir aralığındaki görüntülerin kontrastı açılarak ağ için önemli göz ve ağız gibi bölgelerin belirginleştirilmesi sağlanır.Örnekler ağa sunulmadan önce 255 değerine bölünüp $[0,1]$ aralığına çekilerek normalize edilmelidir.

Şekil 7.3’de giriş ve modifiye edilmiş görüntüler ile bu görüntülere ait histogramlar gösterilmiştir.



Şekil 7.3 Histogram eşitleme

Şekil 7.3’deki görüntü için çıkış görüntüsünde 256 seviye olması istenmiştir, bu durumda $[0,1]$ aralığında, giriş ile çıkış arasındaki ilişkiyi gösteren dönüşüm eğrisi Şekil 7.4’de gösterilmiştir.



Şekil 7.4 Histogram dönüşüm eğrisi

7.3 Standart Sapma ile Normalizasyon

Eğitim setindeki örneklerin değişen ışık koşullarına bağımlılığını azaltarak normalize etmenin bir diğer yolu ise sıfır ortalamalı normalize vektörler oluşturmaktır. Bunun için her bir örnekten ortalama değeri çıkarıldıktan sonra standart sapma değerine bölünür.

X eğitim setindeki bir örneği, $\bar{x}$ ortalama değeri, 1_m bütün elemanları 1 olan bir matrisi göstermek üzere sıfır ortalamalı normalize edilmiş örnek (7.12)'deki bağıntı ile bulunur.

$$\tilde{X} = \frac{X - \bar{x}1_m}{\sigma_x} = \frac{X - \bar{x}1_m}{\|X - \bar{x}1_m\|} \quad (7.12)$$

7.4 Hız Analizi

Özellikle zamanın önemli olduğu sistemlerde önişleme algoritmalarının hızı önem kazanmaktadır. Çizelge 7.1'de Matlab programında PIII 500Mhz makina ile 120×160 boyutlarındaki bir görüntüden 25×25 boyutlarındaki görüntü parçalarının alınıp önişlemeden geçirilmesi durumunda geçen süreler verilmiştir.

Çizelge 7.1 Önişlem için gereken süreler

Önişleme Algoritması	Geçen Süre (s)
En Uygun Düzlem Uydurma	10.05
Histogram Eşitleme	413.5
Standart Sapma	7.2

8. YAPAY SİNİR AĞLARIYLA YÜZ İZLEME UYGULAMASI

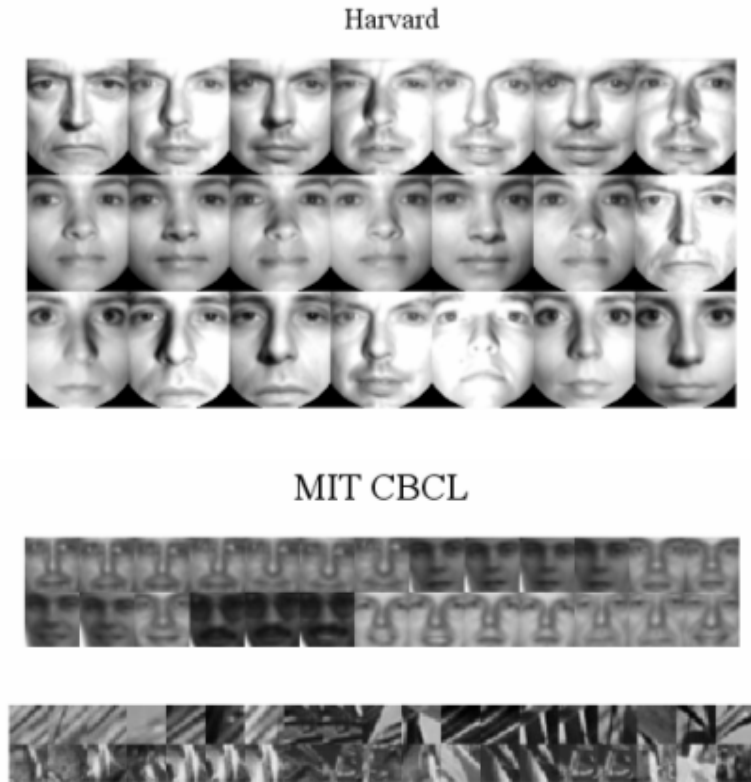
8.1 Eğitim Setinin Hazırlanması

İnsan yüzlerini temsil etmenin en iyi yolu yalnız göz, burun, ağız gibi en belirleyici özelliklerini içeren kanonik formda örnekler oluşturmaktır. Harvard yüz veritabanından alınan örnek bir yüz Şekil 7.1’ de gösterilmiştir.



Şekil 8.1 Kanonik yüz

Yapılan çalışmada Harvard ve MIT CBCL yüz veritabanları kullanılmıştır. Harvard yüz setinde farklı ışık koşullarında 5 setten oluşan toplam 651 adet yüz ve MIT CBCL yüz setinde de 2429 yüz ve 4548 yüz olmayan örnek bulunmaktadır. Şekil 8.2’de veritabanlarında alınan örnek görüntüler gösterilmiştir.



Şekil 8.2 Kullanılan yüz setlerinden örnekler

İnsan yüzlerini ve yüz olmayan örnekleri ağa öğretmek için çok sayıda temsil gücü yüksek örneğe ihtiyaç vardır. Ağın başarısı direkt kendisine sunulan örneklerin kalitesiyle alakalıdır. Veri setindeki örnekler kendilerinden özellik çıkarılıp çıkarılmamasına bakılmaksızın önışlemeden geçirilerek iyileştirilmelidir. Yapılan çalışmada önışleme algoritmaları olarak histogram eşitleme ve standart sapma ile normalizasyon kullanılmıştır. Ayrıca ağın, yüzün rotasyonuna karşı duyarlılığını azaltmak için yüze ait bazı örnekler -10° ile 10° arasında döndürülerek eğitim setine eklenmiştir. Şekil 8.3'de Harvard yüz veri setinden oluşturulmuş yüz o örnekleri gösterilmiştir.



Şekil 8.3 Yüze ait örnekler

8.2 Doğrusal Alt Uzay Metotları ve YSA

İnsan yüzleri kendi içlerinde varyansı yüksek bir sınıf oluştururlar. İnsan yüzleri sakal, bıyık, gözlük gibi etkenlerden dolayı oldukça fazla esnekliğe sahip nesneler olmalarından dolayı direkt olarak kendilerinin geometrik özelliklerine göre özellik çıkarmaya uygun değildirler. Diğer bir özellik çıkarma yöntemi ise istatistiksel veya bir eğitime dayanan özellik çıkarma yöntemleridir. Bu yöntemlerde örneklerden çıkarılacak özellikler bir eğitim seti incelendikten sonra çıkarılan sonuçlar ışığında belirlenir.

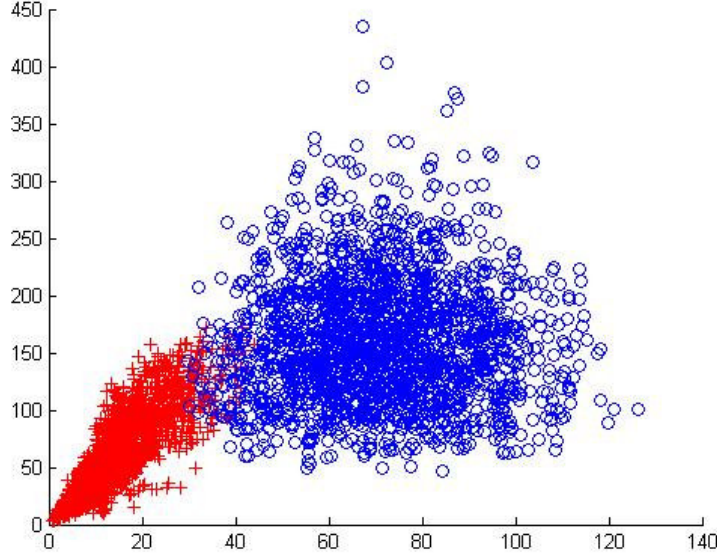
Doğrusal alt uzay metotları, örneklerin dağılımını inceleyerek özellik çıkaran yöntemlerdir ve temel olarak örneklerin boyutlarını azaltarak daha basit sınıflayıcılar tasarlanması için kullanılırlar.

8.2.1 PCA ve YSA

Örnekler PCA ile görüntü uzayından yüz uzayına haritalandığında, bu yeni uzayda sınıflama yapmak için örneğin yüz uzayından uzaklığına (DFFS, Distance From Face Space), örneğin uzay içindeki uzaklığına (DIFS, Distance In Face Space), ayrıca örneğin uzay içindeki konumuna, yani özellik vektörüne bakılabilir.

Yapılan çalışmada ilk olarak yüz ve yüz olmayan örnekler yüz uzayına haritalanıp uzaydan olan ve uzay içi uzaklıklarına göre 2 boyutlu özellik vektörleriyle temsil edilmiştir. Şekil 8.4'de 1000 adet yüz ve 1000 adet yüz olmayan örneğin uzay içi ve uzaydan olan uzaklıklarına göre dağılımları gösterilmiştir. Örneklerin yüz uzayına haritalanmasında

kullanılan dönüşüm matrisi Harvard yüz setinden oluşturulan 35*35 boyutlarında 300 adet yüz ile elde edilmiş ve ilk 20 özvektör kullanılmıştır. Yatay eksen uzaydan olan, dikey eksen ise uzay içindeki mahalanobis uzaklığını göstermektedir.



Şekil 8.4 DFFS ve DIFS

Oluşturulan bu 2 boyutlu özellik vektörleri ile 2 nöronlu ve tek gizli katmanlı bir ağ eğitilmiştir (DFFS+DIFS+YSA). Şekilden yüz örneklerinin DFFS ile DIFS değerleri arasında yüksek bir korelasyon olduğu, yüz olmayan örnekler için ise korelasyonun düşük olduğu, DIFS değerinin örnekleri ayırmada etkin bir özellik olmadığı ve her iki grubun iç içe geçtikleri (overlap) görülmektedir. Bu sebeple 2 boyutlu uzayda örtüşen bu örnekleri çok boyutlu uzayda ayırabilmek ve daha etkin özellik çıkarabilmek için her bir örnek, yalnızca ilk 20 özyüz kullanılarak, 20 elemanlı bir özellik vektörü ile temsil edilmiştir. Oluşturulan bu özellik vektörüne DFFS değeri de eklenerek 21 elemanlı özellik vektörleri oluşturulmuş ve bu özellik vektörleriyle 21 elemanlı tek gizli katmana sahip bir ağ eğitilmiştir (İndirgenmiş özellikler+DFFS+YSA). İlk 20 özvektör toplam varyansın büyük bir kısmını içermektedir.

$$\frac{\sum_{i=1}^{20} \lambda_i}{\sum_{i=1}^{1225} \lambda_i} = 0.9614 \quad (8.1)$$

λ_i büyükten küçüğe doğru sıralanmış özdeğerleri göstermektedir.

Son olarak denenen yöntemde ise yüz uzayında gruplar oluşturularak sınıflandırılmak istenen örneğin gruplara olan uzaklığına bakılmıştır (Sung ve Poggio, 1996). Yüz ve yüze ait olmayan örnekleri gruplamak için k-means algoritması kullanılmış ve 1000'er adet örnekten 6 yüz ve 6 yüz olmayan grup oluşturulmuştur. Şekil 8.5'de yüz ve yüze ait olmayan sınıfların ortalama değerleri gösterilmiştir.



Şekil 8.5 K-means ile gruplandırılmış örnekler

Oluşturulan gruplar, tek bir dönüşüm matrisi kullanılması durumunda (general PCA) yalnız ortalama değerleri ile temsil edilirler ve test edilecek örneğin her bir grup merkezine olan uzaklığı bir özellik vektörü oluşturur. Bir diğer yöntem ise her bir grubun kendi yüz uzayına haritalanmasıdır. Bu durumda grup sayısı kadar alt uzay oluşur ve test edilecek her bir örnek alt uzaya haritalanarak ortalama değerden uzaklığına bakılır. Bu yöntem özel PCA (individual PCA) olarak adlandırılır.

Her iki yöntemde denenerek özellik vektörleri çıkarılmış ve çok katmanlı ağlar eğitilmiştir.

8.3 YSA

YSA ile yüz seziminde en çok kullanılan yöntem, özellik çıkarma yöntemlerinin yetersiz kalmasından dolayı, örneklerin direkt olarak ağa sunulmasıdır. Ağın, eğitimi sırasında örnekleri en iyi temsil eden özellikleri kendisinin belirlemesi beklenmektedir.

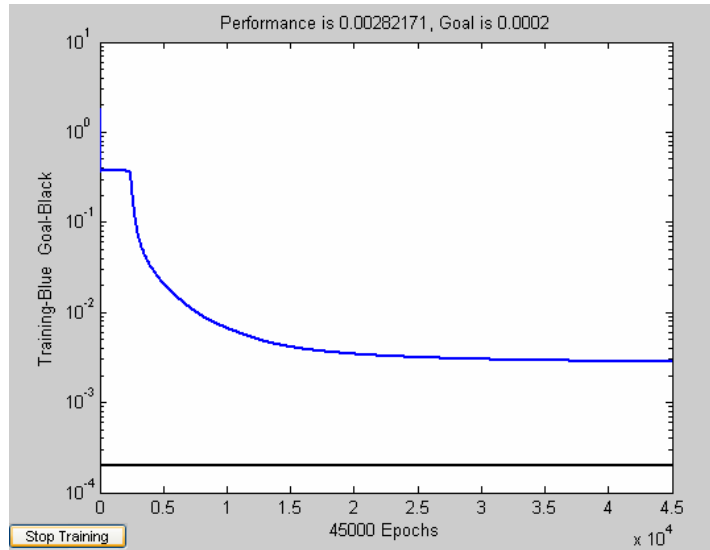
Ağın başarısını etkileyen temel etkenlerden biri ağın yapısı ve öğrenme parametreleridir. Bir diğer özellik ise ağın gizli katmanının girişine nasıl bağlı olduğudur, yani her bir nöron girişlerin hepsine bağlı olabildiği gibi (fully connected) yalnız bir kısmında (partially, retinally connected) bağlı olabilir.

Yapılan çalışmada 25 adet nörona sahip tek katmanlı ve tek çıkış nöronuna sahip, ara katman nöronlarının girişin tamamına veya bir bölümüne bağlı olduğu iki tipteki ağ iki farklı ön işleme algoritması kullanılarak eğitilmiştir. Kullanılan ön işleme algoritmaları standart sapma ile normalizasyon ve histogram eşitlemedir. Ayrıca ağın daha hızlı yanıt vermesi için

frekans uzayıda kullanılmıştır.Önişleme algoritmaları ve arakatman nöron görüş alanları değiştirildiğinde birbirine yakın sonuçlar elde edilmiştir.

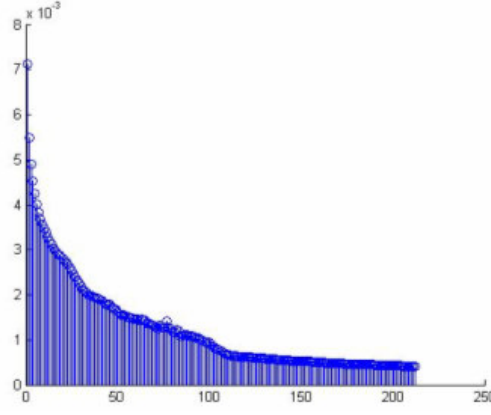
Normalde ara katman nöronlarının sayısı tek veya çift seçilebilmesine karşın, ara katmanla giriş arasındaki işlemler frekans uzayında yapıldığında, ara katman nöronlarının sayısında kısıtlama getirilir.YSA ile ara katman arasındaki işlem asında bir maskeleme işlemidir ve maskeleme işleminin konvolüsyon olarak ifade edilebilmesi için maskenin (burada sinaptik ağırlıklar) simetriği alınmalıdır.Bu yüzden ara katman nöronları tek sayıda olmalıdır.Ağa sunulan örnekler 25*25 boyutlarındadır ve ara katmandaki her bir nöronun 5*5 boyutunda bir görsel alana bakması istendiğinde ara katmandaki nöron sayısı 25 olmalıdır.Arakatmandaki nöron sayısı, ağı daha hızlı çalışması için, genel olarak 20 ile 30 arasındaki değerler seçilir.

Ağı örneklerle eğitmenin temel iki yöntemi ağı her bir örneğin ardından güncellemek (incremental) veya bütün örnekleri gösterdikten sonra (batch) ağı güncellemektir.Ağa örnekler bir bütün halinde gösterildiğinde ağı eğitilemediği görülmüştür.



Şekil 8.6 Toplu eğitim

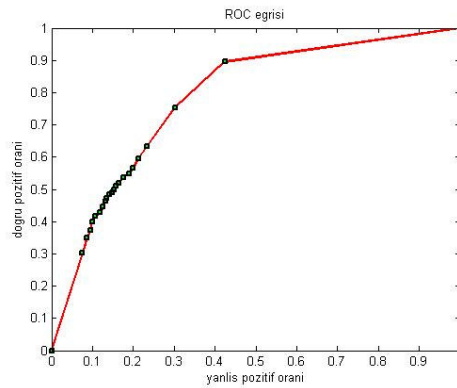
Ağın her bir örnek ardından güncellenerek eğitilmesi durumunda ağın ortalama karesel hatasını (MSE, mean square error) gösteren bir grafik Şekil 8.7’dedir.



Şekil 8.7 İteratif eğitim

Ağın yüz olmayan örnekleri yüz olarak sınıflama eğilimi olduğundan, yüz olmayan örneklerin sayısı yüz olan örneklerden fazla tutulmuştur. Ağa, MIT CBCL veritabanından 3000 yüz olmayan örnek ve Harvard veritabanından elde edilen yaklaşık 300 yüz örneği gösterilmiştir. Çıkış nöronunu 0.8 ve üzeri yapan değerler yüz olarak sınıflandırılmıştır. Örnek seti ağa yaklaşık 200 defa gösterilmiş ve bu değer üzerinde ağın eğitim setindeki örnekler için özelleştiği, yani ezberlediği (overfitting) gözlemlenmiştir. Bir diğer önemli parametre ise öğrenme adımı ve momentum katsayılarının seçilmesidir. En uygun değerler öğrenme katsayısı için 0.02 ve momentum için 0.8 olarak seçilmiştir. Öğrenme adımının seçilen değerden büyüdükçe ağın eğitilemediği ve küçüldükçe eğitimin çok uzadığı gözlemlenmiştir.

Eşik seviyesinin yüksek seçilmesinin nedeni yüze benzeyen birçok örneğin elenmesidir. Yüze ait ve ait olmayan 1000'er adet örnek kullanılarak duyarlılık özgüllük (ROC, Receiver Operating Characteristic Curve) eğrisi oluşturulmuştur. ROC eğrisi ağın doğru tespit oranıyla yanlış pozitif oranının nasıl değiştiğini göstermektedir. Yanlış pozitif yüz olmayan örneklerin yüz olarak sınıflanmasını ifade etmektedir.

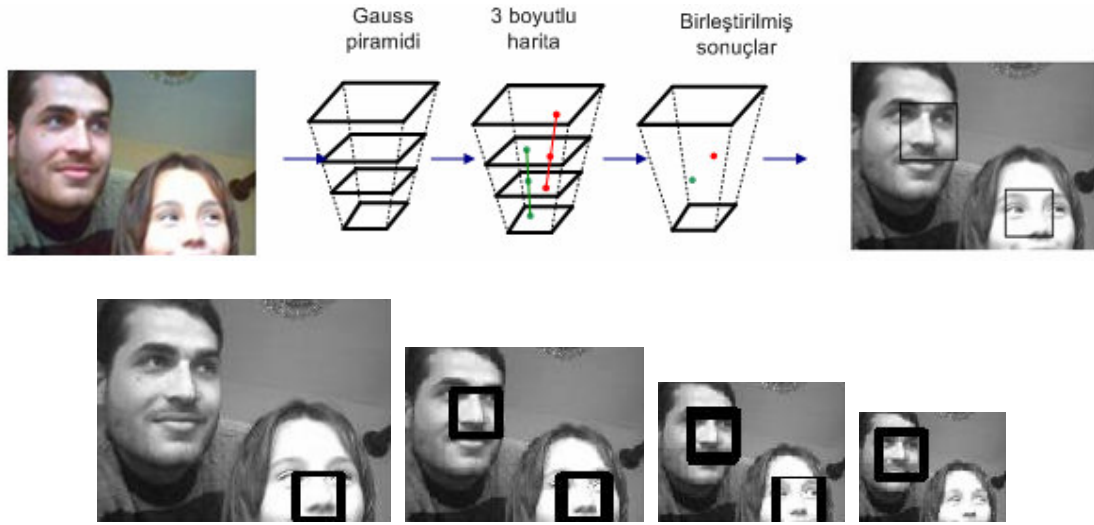


Şekil 8.8 Duyarlılık özgüllük eğrisi

Eğrideki her bir nokta bir karar eşiği için ağırlık doğru tespit oranı ile yanlış pozitif değeri çiftini gösterir. Eşik değeri $[-1 \ 1]$ aralığında 0.1 aralıklarla değiştirilmiştir.

Ağ 25×25 boyutlarındaki örneklerle eğitilmiştir ve daha büyük insan yüzlerini tespit edebilmek için her bir görüntü çerçevesi 1.2 oranında küçültülerek 4 adet görüntü çerçevesi oluşturulmuştur. Her bir çerçeve YSA ile tarandıktan sonra elde edilen sonuçlar bir sonişlem (postprocessing) aşamasında yorumlanarak insan yüzlerinin varlığı hakkında karar verilir.

Yüzün merkezi etrafında bir çok noktada yüz olarak sınıflandırılmasından dolayı (multiple detection) belli bir sıklık eşiğinin altında kalan yanlış sonuçlar elimine edilmiş ve birden fazla çerçevede belirlenen (örtüşen) yüzlerden yüz olma ihtimali en yüksek olan seçilmiştir. Şekil 8.9'da her bir çerçevede bulunan yüzler ve sonuç görüntüsü gösterilmiştir.



Şekil 8.9 Çoklu yüz sezimi

8.3.1 Algoritmanın Hızlandırılması

Yapay sinir ağları ile yüz bulmanın temel problemlerinden birisi, resmin nokta nokta taranması gerektiğinden ve aktivasyon fonksiyonlarının hesaplanmasının zorluğundan dolayı uzun işlem süresine ihtiyaç duymasıdır. Bu sebeple ağırlık gizli katmanıyla görüntü arasındaki nöral hesaplamalar frekans uzayında gerçekleştirilmiş ve aktivasyon fonksiyonunun değeri sürekli hesaplanmak yerine örneklenerek oluşturulan tablodaki (look up table) değerler okunarak algoritma hızlandırılmıştır.

Yapılan çalışmada programlama aracı olarak MATLAB ve C++ kullanılmıştır. C++ ile çalışırken Intel firması tarafından geliştirilen görüntü işleme kütüphanesi OpenCV (Open Source Computer Vision) ve hızlı Fourier dönüşümünü gerçekleştirmek içinde FFTW (

Frigo,M., Johnson,S.G ,1998) kütüphaneleri kullanılmıştır.

8.3.1.1 Uzamsal ve Frekans Domenleri

YSA ile insan yüzleri resmin kenarları hariç her noktasında aranmaktadır.Temelde ağın gizli katmanıyla resim arasında yapılan işlem bir maskeleme veya filtreleme işlemidir.Resimden alınan normalize edilmiş görüntü parçaları ile ara katman nöronları arasında skaler çarpım yapılır.Bu yönüyle ara katmandaki her bir nöron, nöral tabanlı bir filtre görevi görmektedir.

Simetrik bir maske veya çekirdek için maskeleme işlemiyle konvolüsyon işlemi birbirinin aynıdır, simetrik olmayan bir maske için ise bu iki işlemin denk olması için maskenin her iki boyutta simetriği alınır.Ara katman nöron ağırlıkları birer maske olarak görülebilir.

Uzamsal domendeki konvolüsyon işlemi frekans boyutunda noktasal çarpıma denktir.Noktasal çarpımın konvolüsyon işleminden daha az işlem yükü getirmektedir, ayrıca uzamsal domenle frekans domeni arasındaki dönüşümler hızlı Fourier dönüşümü kullanılarak hızlı bir şekilde yapılabilmektedir, bu sayede işlemler hızlandırılır (Yacoub, 1999).

P görüntüden alınan bir parçayı, Φ boyutları $n \times n$ ağırlık matrisini, g aktivasyon fonksiyonunu ve $b1$ eşik değerlerini gösteren bir vektör olmak üzere, görüntüden alınan bir parça için, ara katman nöronlarının çıkış değerleri hesaplanabilir.

$$h_i = g \left(\sum_{j=1}^n \sum_{k=1}^n \Phi_i(j,k)P(j,k) + b1(i) \right) \quad (8.2)$$

I görüntünün bütünü, r ve c görüntüden alınan parçaların merkez koordinatlarını ve n bir tek sayı olmak üzere (8.3)'deki ifade görüntünün bütünü için genişletilebilir.

$$h_i(r,c) = g \left(\sum_{j=-\frac{(n-1)}{2}}^{\frac{(n-1)}{2}} \sum_{k=-\frac{(n-1)}{2}}^{\frac{(n-1)}{2}} \Phi_i(j,k)I(r+j,c+k) + b1(i) \right) \quad (8.3)$$

Aynı ifade görüntünün bütünüyle ağırlık matrisinin konvolüsyon toplamı olarak ifade edilebilir.

$$h_i = g(\Phi_i * I + b1(i)) \quad (8.4)$$

Tek bir gizli katmana ve çıkış nöronuna sahip bir ağ için çıkış değeri ise gizli katman nöronlarının çıkışlarının kombinasyonu olarak ifade edilir.

$$O = g\left(\sum_{i=1}^m w2_i h_i + b2\right) \quad (8.5)$$

İfadede O çıkış nöronunun değerini, $w2$ gizli katmanla çıkış katmanı arasındaki ağırlıkları ve $b2$ eşik değerini göstermektedir.

Ara katmanla resim arasındaki konvolüsyon işlemi, görüntünün ve ağırlık matrislerinin Fourier dönüşümlerinin noktasal çarpımlarının ters Fourier dönüşümüne eşittir.

$$I * \Phi = F^{-1}(F(I).F(\Phi)) \quad (8.6)$$

Görüntüden alınan her bir parça ağa sunulmadan önce ön işlemden geçirilmektedir. Ön işleme yalnız resimden alınan parçayı ilgilendirdiğinden lokal bir işlemdir, buna karşın frekans boyutunda yapılan işlemler ise globaldir yani resmin bütününe ilgilendirmektedir. Bu durumda kullanılan ön işleme algoritması önem kazanır, histogram eşitleme gibi global olarak ifade edilemeyen ön işleme algoritmaları işlemlerin frekans boyutunda yapılmasını engeller. Yapılan çalışmada ön işleme algoritması olarak standart sapma ile normalizasyon kullanılmıştır.

$\left[\tilde{X} \right]_{rc}$ sıfır ortalamalı normalize bir görüntü parçasını göstermek üzere;

$$\left[\tilde{X} \right]_{rc} = \frac{[X]_{rc} - \bar{x}_{rc} 1_{nn}}{\| [X]_{rc} - \bar{x}_{rc} 1_{nn} \|} \quad (8.7)$$

İfadede $\bar{x}_{rc}$ görüntü parçasının ortalama değerini ve 1_{nn} tüm elemanları 1 olan bir matrisi göstermektedir. Normalize görüntü parçasıyla ağırlık matrisinin konvolüsyonu (8.8)'de gösterilmiştir.

$$\left[\tilde{X} \right]_{rc} * \Phi_i = \frac{\left([X]_{rc} - \bar{x}_{rc} 1_{nn} \right) * \Phi_i}{\| [X]_{rc} - \bar{x}_{rc} 1_{nn} \|} \quad (8.8)$$

Konvolüsyon işlemi lineer bir işlem olduğundan ifade yeniden düzenlenir.

$$\left[\tilde{X} \right]_{rc} * \Phi_i = \frac{\left([X]_{rc} * \Phi_i - \bar{x}_{rc} (1_{nn} * \Phi_i) \right)}{\| [X]_{rc} - \bar{x}_{rc} 1_{nn} \|} \quad (8.9)$$

Görüntüden alınan parçanın ortalama değeri, bütün elemanları 1 olan bir matrisle konvolüsyonu olarak ifade edilebilir ve (8.10)'da gösterildiği gibi hesaplanır.

$$\bar{x}_{rc} = \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n X_{rc}(j, k) = \frac{[X]_{rc} * 1_{nn}}{n^2} \quad (8.10)$$

Sıfır ortalamalı görüntü parçasının norm veya standart sapma değeri de konvolüsyon toplamı cinsinden ifade edilebilir.

$$\begin{aligned} \left\| [X]_{rc} - \bar{x}_{rc} 1_{nn} \right\| &= \sqrt{\sum_{i=1}^{n^2} \left(x_{rc}(i) - \bar{x}_{rc} \right)^2} \\ &= \sqrt{\sum_{i=1}^{n^2} x_{rc}(i)^2 - n^2 \bar{x}_{rc}^2} \end{aligned} \quad (8.11)$$

$$\sum_{i=1}^{n^2} x_{rc}(i)^2 = [X_{rc}^2] * 1_{nn} \quad (8.12)$$

(8.9),(8.10),(8.11) ve (8.12) ifadeleri birleştirildiğinde (8.13)'teki ifade elde edilir.

$$\left[\tilde{X} \right]_{rc} * \Phi_i = \frac{[X]_{rc} * \Phi_i - ([X]_{rc} * 1_{nn}) \left(\frac{1_{nn} * \Phi_i}{n^2} \right)}{\sqrt{\left[X_{rc}^2 \right] * 1_{nn} - \frac{([X_{rc} * 1_{nn}])^2}{n^2}}} \quad (8.13)$$

İfadede $\frac{1_{nn} * \Phi_i}{n^2}$ ağırlık matrisinin ortalama değeridir ve μ ile gösterilmek üzere ifade yeniden düzenlenirse;

$$\left[\tilde{X} \right]_{rc} * \Phi_i = \frac{[X]_{rc} * (\Phi_i - \mu_i 1_{nn})}{\sqrt{\left[X_{rc}^2 \right] * 1_{nn} - \frac{([X_{rc} * 1_{nn}])^2}{n^2}}} \quad (8.14)$$

Pay ifadesinden görüldüğü gibi sıfır ortalamalı görüntü parçasıyla ağırlık matrisinin konvolüsyonu, ham görüntü parçasıyla sıfır ortalamalı ağırlık matrisinin konvolüsyonuna eşittir.

$$\left([X]_{rc} - \bar{x} 1_{nn} \right) * \Phi_i = [X]_{rc} * (\Phi_i - \mu_i 1_{nn}) = [X]_{rc} * \tilde{\Phi}_i \quad (8.15)$$

Bu sayede önışleme işleminin bir kısmı çalışma zamanından önce yapılarak önışlem yükü

azaltılır.

Bir nöronun resmin bütünü için ürettiği çıkış matrisi ise görüntü üzerindeki bütün parçalar üzerinde işlem yapılarak bulunur.

$$h_i = \frac{I * (\Phi_i - \mu_i 1_{nn})}{\sqrt{I^2 * 1_{nn} - \frac{(I * 1_{nn})^2}{n^2}}} \quad (8.16)$$

Uzamsal domende tanımlı ifade frekans domeninde daha hızlı bir şekilde hesaplanabilir. Yapılan işlemler özetlenecek olursa;

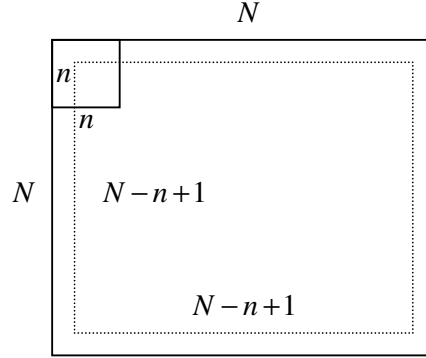
- Çalışma zamanı öncesinde sıfır ortalamalı ağırlık matrisi ve bütün elemanları 1 olan matrisin Fourier dönüşümleri alınır.
- Görüntünün ve kendisiyle noktasal çarpımının Fourier dönüşümleri alınır. $(F(I), F(I^2))$
- Görüntünün karesinin ve elemanları 1 olan matrisin Fourier dönüşümlerinin noktasal çarpımının ters dönüşümü alınır. $[I^2 * 1_{nn} = F^{-1}(F(I^2).F(1_{nn}))]$
- Görüntünün kendisinin ve elemanları 1 olan matrisin Fourier dönüşümlerinin noktasal çarpımının ters dönüşümü alınır. $[I * 1_{nn} = F^{-1}(F(I).F(1_{nn}))]$
- Ara katmandaki her bir nöron için merkezlenmiş ağırlık matrisiyle görüntünün Fourier dönüşümlerinin noktasal çarpımının ters dönüşümü alınır. $[I * \tilde{\Phi}_i = F^{-1}(F(I).F(\tilde{\Phi}_i))]$
- Son olarak bölme ve karekök işlemleri yapılarak ara katmandaki her bir nöron için çıkış matrisi bulunur.

Uzamsal domenden frekans domenine dönüşüm hızlı Fourier dönüşümü (FFT, fast Fourier transform) kullanılarak yapılır. Ayırık Fourier dönüşümünün (DFT, Discrete Fourier Transform) hızlı bir biçimde hesaplanmasına olanak tanıyan yöntemler hızlı Fourier dönüşümü olarak adlandırılır. N noktalı tek boyutlu bir işaret için ayırık Fourier dönüşümü (8.17)'de verilmiştir.

$$X(k) = \sum_n^{N-1} x(n) e^{-j(2\pi/N)kn} \quad (8.17)$$

$k = 0, 1, 2, \dots, N-1$ olmak üzere ayırık Fourier dönüşümü N^2 işlem adımı

gerektirmektedir.Hızlı Fourier dönüşümünde ise aynı işlem için $N \log N$ adım gereklidir.Resmin boyutları $N \times N$, görüntüden alınan parçanın boyutları $n \times n$ ve ara katman nöron sayısı m olmak üzere, görüntüyle gizli katman arasında yapılan işlemler klasik olarak uzamsal domende yapıldığında gereken işlem adım sayısı $(N - n + 1)^2 n^2 m$ 'dir.

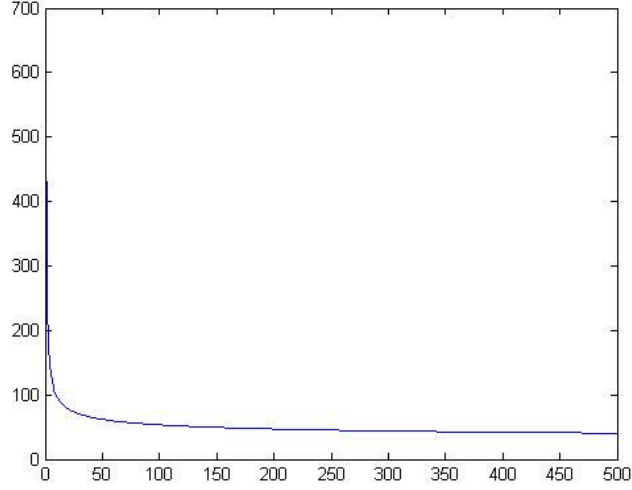


Şekil 8.10 Görüntünün nokta nokta taranması

İşlemler frekans uzayında yapıldığında ise arakatmandaki herbir nöron için m adet ters Fourier dönüşümü olmak üzere toplam $(m + 4)$ Fourier dönüşümü ve frekans uzayında mN^2 işlem yapılmaktadır.Uzamsal ve frekans domeninde yapılan işlem sayılarının oranı algortimanın aynı işlem için ne kadar hızlandığının ölçüsüdür.

$$r = \frac{(N - n + 1)^2 n^2 m}{(m + 4)N^2 \log N^2 + mN^2} \quad (8.18)$$

Yapılan çalışmada ara katman nöron sayısı 25 ve görüntüden alınan parça 25×25 boyutlarında seçilmiştir.Bu durumda değişik boyutlardaki görüntü çerçeveleri için işlem süresinin hangi oranda kısaldığı Şekil 8.11'de gösterilmiştir.



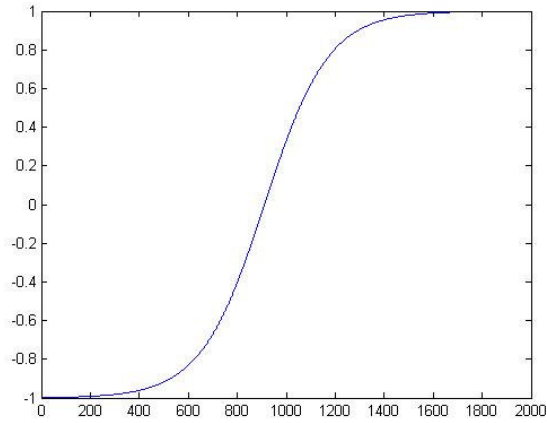
Şekil 8.11 Hız kazancı

8.3.1.2 Aktivasyon Fonksiyonunun Örneklenmesi

Yapılan çalışmada aktivasyon fonksiyonu olarak tanjant sigmoid fonksiyonu kullanılmıştır.

$$y = \frac{1 - e^{-x}}{1 + e^x} \quad (8.19)$$

Aktivasyon fonksiyonu sürekli hesaplanmak yerine fonksiyondan 2000 adet örnek alınarak bir tablo oluşturulmuş ve program içerisinde bu tablo kullanılmıştır.



Şekil 8.12 Örneklenmiş tanjant sigmoid fonksiyonu

Boyutları $N \times N$ olan bir görüntüden $n \times n$ boyutlarında görüntüler alınması durumunda ve ara katman nöron sayısı m olmak üzere program içerisinde aktivasyon fonksiyonu

$(N - n + 1)^2 (m + 1)$ defa çağırılmaktadır.

C++ ile yapılan analizde 120×160 ve 240×320 boyutlarındaki görüntüler için aktivasyon fonksiyonunun hesaplanması ve örneklenerek tablolaştırılması durumunda geçen süreler gösterilmiştir.

Çizelge 8.1 Aktivasyon fonksiyonu için geçen süreler

Görüntü Boyutu	Gerçek Fonksiyon için geçen süre (s)	Tablo için geçen süre(s)
120×240	0.8	0.009
240×320	2.9	0.03

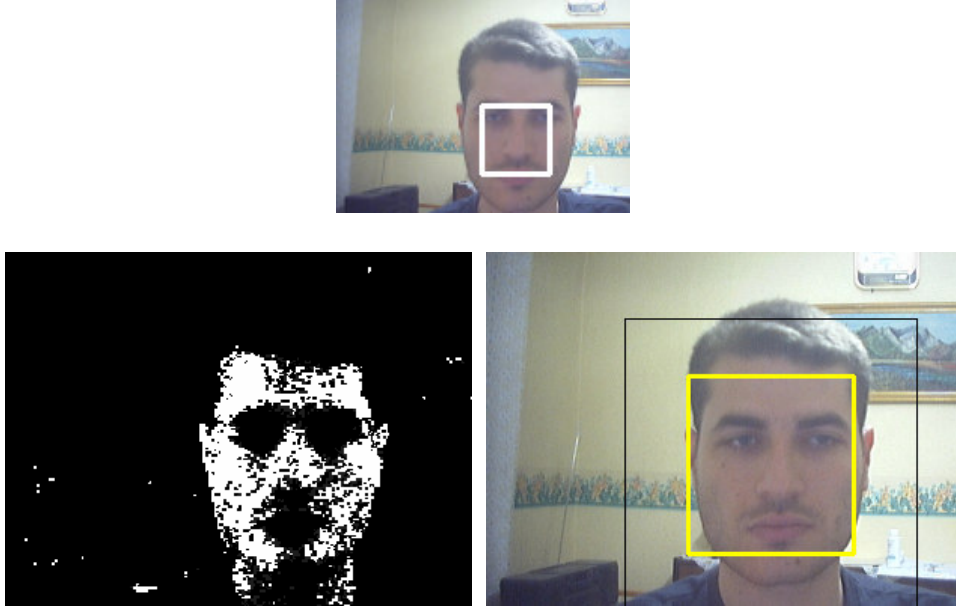
8.4 Yüz Takibi

Yapılan çalışmanın amacı insan yüzlerini en etkin ve hızlı bir şekilde bularak takip etmektir. Uygun ön işleme ve FFT algoritmaları kullanılarak küçük boyutlu görüntülerde (120×160) gerçek zamanlıya yakın takip işlemi yapılabilmesine karşın daha etkin bir yöntem ihtiyacı duyulmuştur. Bu sebeple iki farklı takip algoritması kullanılmıştır.

Birinci yöntemde kullanıcının bilgisayara yakın olduğu varsayılarak 120×160 boyutlarındaki görüntülerde insan yüzleri aranmış ve görüntüde birden fazla yüz bulunması durumunda en büyük yüz takip edilmiştir. YSA ile insan yüzü olarak belirlenen bölge önerilen takip algoritması ile izlenmiştir.

İkinci yöntemde ise kullanıcının kameradan daha uzak olduğu varsayılarak 320×240 boyutlarındaki renkli görüntülerde insan teni olduğu düşünülen yerler belirlenip sonrasında insan yüzü olup olmadığına bakılmış ve yüz olarak tespit edilen bölge takip edilmiştir.

Şekil 8.13'de YSA ile yüz belirlendikten sonra insan tenine ait ihtimal görüntüsü ve takip görüntüsü gösterilmiştir. Sağ alttaki şekilde içteki dikdörtgen yüz bölgesini, dıştaki dikdörtgen ise ihtimal görüntüsü işlem yapılan bölgeyi göstermektedir.



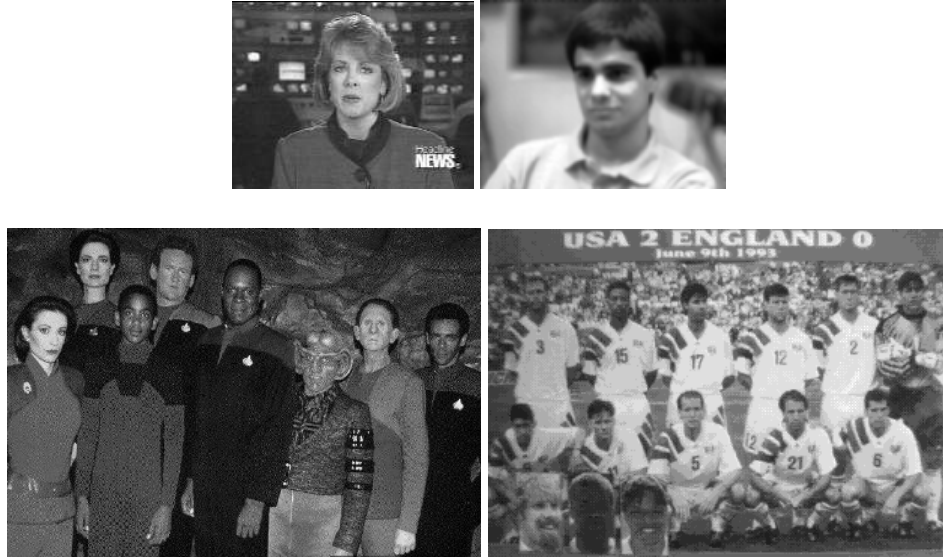
Şekil 8.13 Yüzün belirlenmesi ve takip edilmesi

9. DENEYSEL SONUÇLAR

Sonuçlar için üç farklı test veritabanı kullanılmıştır. Kullanılan iki veri tabanı genel kullanıma sahip PICs veritabanı ve CMU-130 (Rowley vd.,1996) veritabanlarıdır. Üçüncü veritabanı ise renkli görüntüler toplanarak oluşturulmuştur. Her üç veritabanından örnek görüntüler sırasıyla Şekil 9.1, Şekil 9.2 ve Şekil 9.3’de gösterilmiştir.



Şekil 9.1 PICs verisetinden örnekler



Şekil 9.2 CMU-130 verisetinden örnekler



Şekil 9.3 Oluşturulan setten örnekler

PICs veritabanından 70 adet görüntü alınarak veriseti oluşturulmuştur, CMU-130 setinde 560 tane yüz içeren 130 adet siyah beyaz görüntü bulunmaktadır, oluşturulan set ise 166 tane yüz içeren 50 adet renkli görüntüden oluşmaktadır.

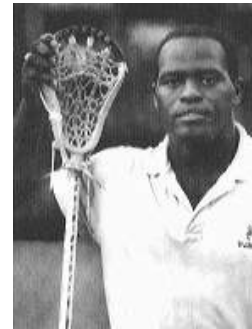
Aşağıdaki tabloda yöntemlerin kısaltmaları ve doğru sınıflama oranlarıyla, yanlış olarak yüz sınıflandırdıkları örnek sayısı (false positive) tablolanmıştır.

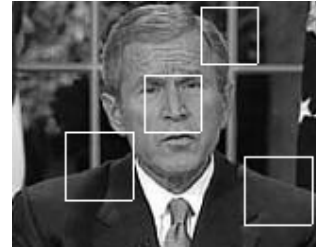
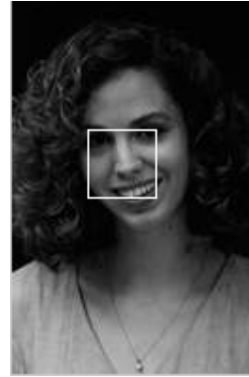
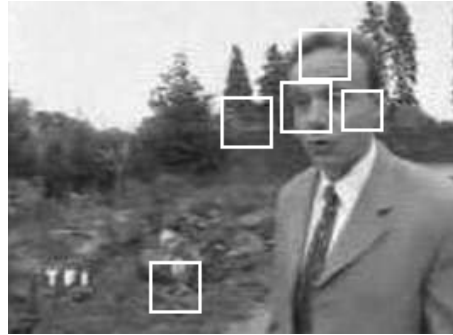
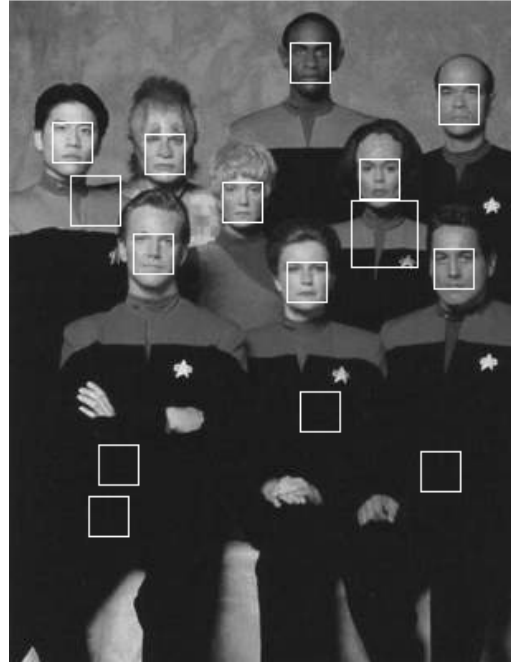
Çizelge 9.1 Yüz bulma yöntemleri ve sonuçlar

Yöntem	PICs	Oluşturulan Set		CMU-130
		Renk Bilgisi		
		Var	Yok	
DFFS+DIFS+YSA	%82.6 / 5	%67.3 / 15	%67.3 / 260	—
DFFS+indirgenmiş özellikler+YSA	%86,2 / 2	%64.2 / 10	%64.2 / 223	—
Genel PCA+YSA	%83.5 / 4	%68.6 / 12	%68.6 / 241	—
Özel PCA+YSA	%83.6 / 4	%67.3 / 11	%67.3 / 233	—
YSA	%100 / 0	%86,6 / 3	%86.6 / 110	% 68.2 / 667

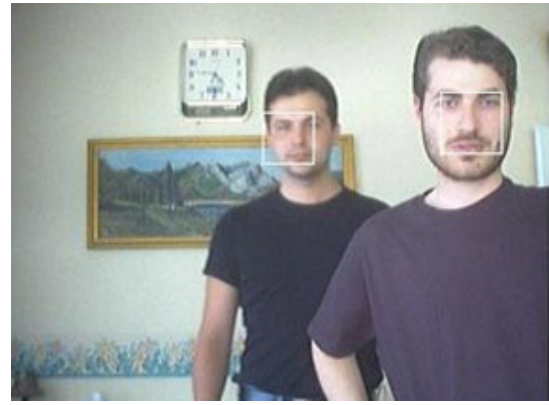
Şekil 9.4’de yalnız YSA sınıflayıcının CMU-130 ve oluşturulan set üzerinde renk bilgisi kullanılarak ve kullanılmaksızın elde edilen bazı sonuçlar gösterilmiştir.

Tablodaki sonuçlardan görüleceği gibi örneklerden boyut indirgeyerek özellik çıkarmaksızın örneklerin ağa direkt olarak sunulması durumunda daha başarılı sonuçlar elde edilmiştir.





Şekil 9.4 CMU-130 sonuçları





Şekil 9.5 Oluşturulan setten sonuçlar

10. SONUÇ VE ÖNERİLER

Yapılan çalışmada video görüntülerinde insan yüzlerinin etkin ve hızlı bir şekilde bulunarak takip edilmesi amaçlanmıştır, bu sebeple nöral ve renk tabanlı sınıflayıcılar tasarlanmıştır. Renk tabanlı sınıflayıcı insan tenine ait pikselleri diğerlerinden ayıran bir ten sınıflayıcı, nöral tabanlı sınıflayıcı ise insan yüzlerini yüz olmayan örneklerden ayıran çok katmanlı algılayıcı yapısında bir yapay sinir ağıdır.

Ten sınıflayıcı nöral tabanlı sınıflayıcıya göre daha basit ve hızlı cevap veren bir sınıflayıcı olması nedeniyle, daha kompleks ve uzun işlem süresine ihtiyaç duyan nöral tabanlı filtre ile yalnız insan tenine ait ve benzer bölgelerde insan yüzlerinin aranarak işlem süresinin kısaltılması, nöral tabanlı filtre ile görüntünün tamamının taranması durumunda ise yüz olarak sınıflanan bölgelerin insan teni olup olmadığını kontrol ederek ağın hatalarının ayıklanmasında ve insan yüzü belirlendikten sonra takip işleminde kullanılmıştır. Yapılan çalışmada uygun önileme algoritması seçimi, ağ ile görüntü arasındaki işlemlerin frekans uzayında yapılması ve nöronların aktivasyon fonksiyonlarının örneklenip bir tablo oluşturularak daha hızlı hesaplanması sonucu yalnız nöral tabanlı filtre ile küçük boyutlu görüntülerde (120×160) gerçek zamanlıya yakın yüz takibi yapılabilmesine karşın, aynı zamanda yüz bölgesi bir kere belirlendikten sonra her çerçevede nöral tabanlı filtre ile aranmak yerine ten sınıflayıcı kullanılarak takip edilmiştir.

Yapılan çalışmada örneklerden herhangi bir özellik çıkarılmaksızın, yalnız önilemden geçirilerek, parlaklık değerleri kullanıldığı gibi Ana Bileşen Analizi ile örneklerden özellik vektörleri çıkarılarak da çok katmanlı algılayıcılar eğitilmiştir. Örneklerden özellik çıkarılmadığı durumda, çok boyutlu ve yalnız parlaklık bilgisini içeren örnekler ağı sunulduğunda, eğitim süresi oldukça uzamasına karşın daha başarılı sonuçlar elde edilmiştir.

Tasarlanan sistemin hızlı çalışması istendiğinden tek katmanlı tek bir ağ kullanılmıştır. Özellikle süre kısıtlamasının olmadığı uygulamalarda birden fazla arakatmana sahip, birbirine paralel birden çok ağ kullanılarak sonuçları yorumlanıp daha etkin bir sınıflama yapılabilir.

Tasarlanan nöral tabanlı filtre insan yüzlerinin yalnız önden görünüşüne duyarlıdır, insan yüzlerini değişik açılardan tespit edebilmek için ise herbiri yüzün belli bir açıdan görüntüne duyarlı birden çok sınıflayıcı eğitilerek her bir görüntü birçok sınıflayıcı ile taranmalı veya görüntüden alınan parçaların pozunu tahmin edilerek (pose estimation) uygun olan sınıflayıcı belirlenip görüntüden alınan parçanın yüz olup olmadığına karar verilmelidir.

KAYNAKLAR

- Bouzerdoun, A. ve Chai, D., (2000), "A Bayesian Approach to Skin Color Classification in YCbCr Color Space", IEEE Region Ten Conference (TENCON' 2000), Kuala Lumpur, Malaysia, vol II, pp. 424-424.
- Bradski, G., (1998), "Computer Vision Face Tracking For Use in a Perceptual User Interface", Intel Technology Journal.
- Cai, J., Goshtasby, A., (1996), "Detecting Human Faces in Color Images" , Image Vis. Comput. 18, 63-75.
- Fleck, M., Forsth, A., (1996), "Finding Naked People", European Conference on Computer Vision , Volume II, pp. 592-602.
- Ford, A. ve Roberts, A., (1998), "Colour Space Conversions".
- Frigo, M., Johnson, S.G., (1998), "FFTW 1.3 User's Manual".
- İkizler, N., Duygulu, P., (2005), "Improvement of face detection algorithms for news videos", Signal Processing and Communications Applications Conference.Proceedings of the IEEE 13th, pp.308-311.
- Kuchi, P., Gabbur, P., Subbanna, B., Sumam, D., (2002), "Human Face Detection and Tracking using color modeling and connected component operators", IETE Journal Of Research, vol.48, pp. 289-293.
- McKenna, S. ve Gong, S., (1996), "Combined Motion and Model Based Face Tracking", British Machine Vision Conference, pages 755-765.
- Morris, T., (2002), "Blink Detection for Real Time Eye Tracking", Journal of Network and Computer Applications, Volume 25 , Issue 2 , Pages: 129 – 143.
- Mottaleb, A., (2002), "Face Detection in Color Images", IEEE Transactions on Pattern Analysis and Machine Intelligence Volume 24 , Issue 5 Pages: 696 – 706.
- Rowley, H., Baluja, S., Kanade,T., (1996), "Neural Network Based Face Detection", IEEE-Transactions on Pattern Analysis and Machine Intelligence (PAMI), Vol. 20, No. 1, January.
- Sung, K.K ve Poggio, T., (2002), " Example-based learning for viewbased face detection", IEEE Trans. PAMI, 20:39—51.
- Turk, M., Pentland, A., (1991), "Face Recognition using Eigenfaces", Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 586-591.
- Viola, P. ve Jones, M., (2001), "Real time Object Detection", Computer Vision and Image Understanding ,volume 98 , Issue 1, Pages: 182 - 210 .
- Yang, J., Waibel, A., (1996), "A Real time Face Tracker", IEEE Workshop on Applications of Computer Vision, pages 142—147.
- Yacoub, S.B., Fasel, B., Luetin, J., (1999),"Fast Face Detection using MLP and FFT", Proc. Second International Conf. On Audio and Video-based Biometric Person Authentication (AVBPA' 99).
- Yang, M., Kriesman, D., Ahuja,N., (1996), "Detecting Faces in Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(1), pp 34-58, 2002.

ÖZGEÇMİŞ

Doğum tarihi	19.10.1978	
Doğum yeri	Balıkesir	
Lise	1992-1996	Yüzüncü Yıl Endüstri Meslek ve Teknik Lisesi
Lisans	1997-2001	Pamukkale Üniversitesi Mühendislik Fak. Elektrik-Elektronik Mühendisliği Bölümü
Yüksek Lisans	2003-2006	Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Elektrik-Elektronik Müh. Anabilim Dalı, Elektronik Programı