

İmge Kümeleri Tabanlı Yüz Tanıma için Aydınlatma İyileştirme ve Poz Hizalama

Meltem Seyirt

YÜKSEK LİSANS TEZİ

Elektrik Elektronik Mühendisliği Anabilim Dalı

Ekim 2017



Illumination Improvement and Pose Alignment for Image Set Based Face Recognition

Meltem Seyirt

MASTER OF SCIENCE THESIS

Electrical and Electronics Engineering Department

October 2017

İmge Kümeleri Tabanlı Yüz Tanıma için Aydınlatma İyileştirme ve Poz Hizalama

Meltem Seyirt

Eskişehir Osmangazi Üniversitesi
Fen Bilimleri Enstitüsü
Lisansüstü Yönetmeliği Uyarınca
Elektrik Elektronik Mühendisliği Anabilim Dalı
Kontrol ve Kumanda Bilim Dalı
YÜKSEK LİSANS TEZİ
Olarak Hazırlanmıştır

Danışman: Yrd.Doç.Dr.Hasan Serhan Yavuz

"Bu Tez TÜBİTAK tarafından 114E014 no'lu proje çerçevesinde desteklenmiştir."

Ekim 2017

ONAY

Elektrik Elektronik Mühendisliği Anabilim Dalı YÜKSEK LİSANS öğrencisi Meltem Seyirt'in YÜKSEK LİSANS tezi olarak hazırladığı "İmge Kümeleri Tabanlı Yüz Tanıma için Aydınlatma İyileştirme ve Poz Hizalama" başlıklı bu çalışma, jürimizce lisansüstü yönetmeliğin ilgili maddeleri uyarınca değerlendirilerek oy birliğiyle kabul edilmiştir.

Danışman : Yrd. Doç. Dr. Hasan Serhan Yavuz

İkinci Danışman : Doç. Dr. Hakan Çevikalp

Yüksek Lisans Tez Savunma Jürisi:

Üye : Yrd. Doç. Dr. Hasan Serhan Yavuz

Üye : Prof. Dr. Rifat Edizkan

Üye : Yrd. Doç. Dr. Cihan Topal

Fen Bilimleri Enstitüsü Yönetim Kurulu'nun tarih ve sayılı kararıyla onaylanmıştır.

Prof.Dr. Hürriyet ERŞAHAN
Enstitü Müdürü

ETİK BEYAN

Eskişehir Osmangazi Üniversitesi Fen Bilimleri Enstitüsü tez yazım kılavuzuna göre, Yrd. Doç. Dr. Hasan Serhan Yavuz danışmanlığında hazırlamış olduğum “**İmge Kümeleri Tabanlı Yüz Tanıma için Aydınlatma İyileştirme ve Poz Hizalama**” başlıklı tezimin özgün bir çalışma olduğunu; tez çalışmamın tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı; tezimde verdiğim bilgileri, verileri akademik ve bilimsel etik ilke ve kurallara uygun olarak elde ettiğimi; tez çalışmamda yararlandığım eserlerin tümüne atıf yaptığımı ve kaynak gösterdiğimi ve bilgi, belge ve sonuçları bilimsel etik ilke ve kurallara göre sunduğumu beyan ederim. 20/10/2017

Meltem Seyirt

ÖZET

Otomatik yüz tanıma süreci son yıllarda popülerliğini arttırmış bir konudur. İmge tabanlı yaklaşımların hakim olduğu yüz tanıma süreci, bilgisayar ve kamera teknolojilerinin gelişimiyle yerini video tabanlı yaklaşımlara bırakmaktadır. Video tabanlı yüz tanıma uygulamalarında, özellikle kişilerin farklı aydınlatma veya cepheden, yandan görünüm vb. farklı pozlar içeren imge kümelerinin eşleştirilmesi zorluklar içermektedir. Bu tez çalışmasında, bu zorluklardan yola çıkılarak aydınlatma ve poz çeşitliliklerinin yüz tanıma sistemlerine etkilerinin azaltılmasına yönelik iyileştirmeler araştırılmıştır. Yüz tanıma algoritmalarına engel teşkil eden aydınlatma farklılıklarının etkilerinin en aza indirilmesi amaçlanmış ve bu etkilerin azaltılmasının ardından yüze ait poz bilgileri kullanılarak imgeler sınıflandırılmıştır. Pozlara göre ayrıştırılan yüzler, sınıf içi değişimlerinin azaltılması için popüler bazı hizalama algoritmaları kullanılarak hizalanmıştır. Böylece hizalama algoritmalarının yüz tanıma sistemleri üzerindeki etkileri araştırılmıştır. Yapılan deneyler sonucunda yüz imgelerindeki değişimlerin azaltılmasıyla birlikte video tabanlı yüz tanıma deneylerinin başarımlarında gelişmeler sağlandığı tespit edilmiştir.

Anahtar Kelimeler: yüz sezimi, yüz hizalama, yüz tanıma, imge kümeleri

SUMMARY

The automatic face recognition process has become a popular topic in recent years. The facial recognition process, where previously single-image based methods were more common, has left its place in video-based approaches by the development of camera. In video based recognition applications, it becomes more difficult to match the image sets of the same person whose frames captured under different illumination conditions or when the compared frames include different face poses such as frontal versus profile. In this thesis, we investigate how to reduce the effects of lighting and exposure variations on face recognition systems. Our aim is to minimize the effects of the lighting differences which are an obstacle to the face recognition algorithms. After this phase, we classified preprocessed face images according to the face pose and each face image is aligned by using some popular alignment algorithms to reduce intra-class variations. Also, effects of alignment algorithms on face recognition systems have been investigated. Experimental results demonstrate that set based recognition methods give higher correct recognition rates with the application of the proposed methodology schemes.

Keywords: face detection, face alignment, face recognition, image sets

TEŞEKKÜR

Beni bu konuda çalışmaya yönlendiren ve tez çalışmam boyunca bilgi ve deneyimlerini bana aktararak gösterdiği her türlü destek ve yardımlardan dolayı tez danışmanlarım Yrd. Doç. Dr. Hasan Serhan Yavuz ve Doç. Dr. Hakan Çevikalp'e,

Tüm hayatım boyunca desteklerini her zaman arkamda hissettiğim annem Semra Yalçın, babam Burhan Yalçın, kardeşim Tuncay Yalçın, dedem Fahrettin Yalçın, babaannem Perihan Yalçın, anneannem Sevim Ayer'e,

Destek ve yardımlarından dolayı eşim Cengizhan Seyirt'e,

Tanıştığım için kendimi çok şanslı hissettiğim hiçbir destek ve yardımı benden esirgemeyen kıymetli çalışma arkadaşlarım Merve Elmas, Golara Ghorban Dordinejad, Barış Can Çam ve Şahin Işık'a,

TÜBİTAK'a, 114E014 numaralı projede verdikleri destekten dolayı teşekkürlerimi içtenlikle sunarım.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	vi
SUMMARY	vii
İÇİNDEKİLER	ix
TEŞEKKÜR	viii
ŞEKİLLER DİZİNİ	xi
ÇİZELGELER DİZİNİ	xiii
SİMGELER VE KISALTMALAR DİZİNİ	xiv
1. GİRİŞ VE AMAÇ	1
2. LİTERATÜR ARAŞTIRMASI	4
3. MATERYAL VE YÖNTEM	9
3.1. Yüz Sezimi	9
3.1.1. Viola-Jones yüz sezimi	9
3.1.2. Zhu & Ramanan yüz sezimi	12
3.1.3. Face++ yüz sezimi	14
3.1.4. INTRAFACE yüz sezimi	17
3.2. Aydınlatma Farklılıklarının Giderilmesi	21
3.3. Hizalama için Poz Tahmini	30
3.3.1. k -ortalamalar gözetimsiz öğrenme yöntemi	30
3.3.2. Temel bileşen analizi yönteminin kullanımı	31
3.3.3. Negatif olmayan matris ayrışımı yönteminin kullanımı	33
3.3.4. Grafikte düzenlenmiş negatif olmayan matris ayrışımı yönteminin kullanımı	34
3.4. Yüz Hizalama	35
3.4.1. Nirengi noktaları kullanılarak hizalama	35
3.4.2. Robust Alignment by Sparse and Low-Rank decomposition (RASL) yöntemi ile hizalama	36
3.4.3. Scale Invariant Feature Transform (SIFT)-akış yöntemi ile hizalama	37
3.4.4. SIFT-akış yöntemi ile afin dönüşüm kullanılarak hizalama	43
3.5. Hizalanmış İmgeler Kullanılarak Yüz Tanıma	44
3.5.1. Afin zarf modelleme	45

İÇİNDEKİLER (devam)

3.5.2. İndirgenmiş afin zarf modelleme	46
3.5.3. Konveks zarf modelleme	48
3.5.4. İndirgenmiş konveks zarf modelleme	48
3.5.5. Karar ağaçları kullanarak imge kümesi tabanlı tanıma yöntemlerinde hız geliştirmesi	49
4. BULGULAR VE TARTIŞMA	52
4.1. Veri Tabanı	52
4.2. Poz Tahmini Deneyleri	54
4.3. Yüz Hizalama Deneyleri	59
4.4. Hizalanmış İmge Kümelerinin Yüz Tanıma Deneylerinde Kullanılması	67
5. SONUÇ VE ÖNERİLER	69
KAYNAKLAR DİZİNİ	70
EK AÇIKLAMALAR	77
Ek Açıklamalar - A	78
Ek Açıklamalar - B	79
Ek Açıklamalar - C	80
Ek Açıklamalar - D	81
Ek Açıklamalar - E	82

ŞEKİLLER DİZİNİ

<u>Şekil</u>	<u>Sayfa</u>
3.1 Dikdörtgensel bölge tabanlı Haar benzeri öz nitelikler.İki dikdörtgenli (A)ve (B)'de, üç dikdörtgenli (C)'de, dört dikdörtgenli (D)'de gösterilmiştir (Viola ve Jones, 2001).	10
3.2 (x, y) noktasındaki integral imge piksel değeri,orijinal imgede aynı noktanın sol ve üstünde kalan piksellerin toplanmasıyla elde edilen değerdir (Viola ve Jones, 2001).	10
3.3 Zayıf sınıflandırıcıların art arda eklenmesiyle oluşan nihai sınıflandırıcının şematik gösterimi (Viola ve Jones, 2001)	11
3.4 Bakış açısından kaynaklanan topolojik değişiklikleri kodlayan karışım ağaçlarından örnekler [3]	13
3.5 Face++ algoritma sistemi genel görünümü (Zhou vd., 2013)	15
3.6 Kullanılan ağların genel yapısı (Zhou vd., 2013)	16
3.7 Eğitimli İniş Metodunun gösterimi (Xiong ve De La Torre, 2013)	17
3.8 (a) El ile işaretlenmiş 66 nirengi noktası (b) Yüz sezici ile bulunan yüzdeki ortalama nirengi noktaları (Xiong ve De La Torre, 2013)	18
3.9 Eğitimde kullanılacak yedi poz grubu gösterimi	21
3.10 Uyarlamasız Histogram Eşitleme yönteminin uygulaması	22
3.11 Uyarlamalı Histogram Eşitleme yönteminin uygulaması	22
3.12 SSR yönteminin uygulaması	23
3.13 MSR yönteminin uygulaması	23
3.14 Homomorfik filtreleme yönteminin uygulaması	24
3.15 SSQ yönteminin uygulaması	24
3.16 MSQ yönteminin uygulaması	25
3.17 DCT yönteminin uygulaması	25
3.18 Dalgacık tabanlı normalleştirme yönteminin uygulaması	25
3.19 WD yönteminin uygulaması	26
3.20 NLM yönteminin uygulaması	26
3.21 ANL yönteminin uygulaması	27
3.22 SF yönteminin uygulaması	27
3.23 GF yönteminin uygulaması	28
3.24 DoG yönteminin uygulaması	28
3.25 WEB yönteminin uygulaması	29
3.26 MSW yönteminin uygulaması	29
3.27 LSSF yönteminin uygulaması	29

ŞEKİLLER DİZİNİ (devam)

3.28 Lowe'nin piramit şeması (Liu vd., 2011)	39
3.29 Solda: Ölçek uzay imgeleri kümesi,sağda: DoG imgeleri (Liu vd., 2011)	39
3.30 Kabadan inceye SIFT akış eşleşmesinin gösterimi	43
4.1 İmge kümeleri örnekleri	53
4.2 RASL uygulama ile (pozlar ayrıştırılarak) hizalanmış imge kümeleri örnekleri,Poz kategorisi: a) cephe (b) sağ profil (c) sol profil (d) aşağı doğru (e) yukarı doğru, Birinci sıra: giriş imgeleri, İkinci sıra: hizalanmış imgelerden örnekler	61
4.3 RASL: Hizalama öncesi (birinci imge) ve sonrası (ikinci imge) ortalama yüz imgeleri Poz kategorisi: a)cephe (b)sağ profil (c)sol profil (d)aşağı bakış (e)yukarı bakış	62
4.4 SIFT öznitelikleriyle hizalanmış imge kümeleri örnekleri. Satırlardaki imge kümeleri sırasıyla Senaryo 1, 2, 3 ve 4'e aittir.	63
4.5 SIFT öznitelikleriyle hizalama öncesi (birinci imge) ve hizalama sonrası (ikinci imge) ortalama yüz imgeleri (a)Senaryo-1, (b) Senaryo-2, (c)Senaryo-3, (d)Senaryo-4	64
4.6 SIFT-flow yöntemiyle (pozlar ayrıştırılarak) hizalanmış imge kümeleri örnekleri. Poz kategorisi: a)cephe (b)sağ profil (c)sol profil (d)aşağı bakış (e)yukarı bakış. Birinci sıra: giriş imgeleri, İkinci sıra: hizalanmış imgelerden örnekler	65
4.7 SIFT-flow (uygulama-2): Hizalama öncesi (birinci imge) ve sonrası (ikinci imge) ortalama yüz imgeleri. Poz kategorisi: a)cephe (b)sağ profil (c)sol profil (d)aşağı bakış (e)yukarı bakış	66
4.8 Nirengi noktaları için tasarlanan hedef şablonları (a) cephe, (b) sağa doğru, (c) sağ profil (d) aşağı doğru (e) yukarı doğru (f) sola doğru (g) sol profil.	66
4.9 Şablon hizalama yöntemiyle hizalanmış imge kümeleri örnekleri. Satırlardaki imge kümeleri sırasıyla Senaryo 1, 2, 3 ve 4'e aittir.	67
6.1 Viola-Jones yüz sezimi uygulaması	82
6.2 Zhu & Ramanan tarafından önerilen yüz sezici ile yüz sezimi sonuçları	83
6.3 Face++ algoritması ile alınan sonuçlar	84
6.4 Eğitimli İniş Metodu kullanılarak alınan sonuçlar	85
6.5 Video çekimi için tasarlanan senaryolar (a) Senaryo-1 (b) Senaryo-2 (c) Senaryo-3 (d) Senaryo-4	86

ÇİZELGELER DİZİNİ

<u>Çizelge</u>	<u>Sayfa</u>
4.1 Orijinal veri üzerinde poz sınıflandırma başarımlar oranları	55
4.2 Verilen yöntemler ile normalize edilmiş veri üzerinde k -ortalamlar poz sınıflandırma başarımlar oranları	56
4.3 Verilen yöntemler ile normalize edilmiş veri üzerinde PCA+ k -ortalamlar poz sınıflandırma başarımlar oranları	57
4.4 Verilen yöntemler ile normalize edilmiş veri üzerinde NMF+ k -ortalamlar poz sınıflandırma başarımlar oranları	58
4.5 Verilen yöntemler ile normalize edilmiş veri üzerinde GNMF+ k -ortalamlar poz sınıflandırma başarımlar oranları	59
4.6 ESOGU-285 Yüz Video Veri Tabanı Üzerindeki Sınıflandırma Başarımları(%) .	68

SİMGELER VE KISALTMALAR DİZİNİ

Kısaltma	Açıklama
ANL	Adaptive NL Means based Normalization
ASR	Adaptive single-scale retinex
AZM	Afin Zarf Modelleme
BHDT	Binary Hierarchical Descision Tree
CRNP	Collaboratively Regularized Nearest Points
DoG	Difference-of-Gaussian
DCT	Discrete Cosine Transform
DWT	Discrete Wavelet Transform
GF	GradientFaces
GNMF	Graph Regularized NN Matrix Factorization
HoG	Histogram of Gradient
HOMO	Homomorfik Filtreleme
KMC	K-Means Clustering
KZM	Konveks Zarf Modelleme
LoG	Laplacian-of-Gaussian
LSSF	The Large- and Small-scale Features
MMD	Manifold-to-Manifold distance
MRF	Markov Random Field
MSQ	Multi-scale Self Quotient Image
MSR	Multi Scale Retinex
MSW	Multi-scale WeberFaces
MSM	Mutual Subspace Method
NCuts	Normalized Cuts

Kısaltma	Açıklama
NLM	Non-local Means based Normalization
NMF	Nonnegative Matrix Factorization
PCA	Principle Component Analysis
RNP	Regularized Nearest Points
RASL	Robust Alignment by Sparse and Low-Rank decomposition
SIFT	Scale Invariant Feature Transform
SRN-ADML	Self-Regularized Non-Negative Coding and Adaptive Distance Metric Learning
SVD	Singular Value Decomposition
SSR	Single Scale Retinex
SSQ	Single-Scale Self Quotient Image
SANP	Sparse Approximation Nearest Points
SF	Steerable Filter
SSE	Sum of Squared Error
SDM	Supervised Descent Method
TT	Tan-Triggs normalization
WA	Wavelet-based image denoising
WD	Wavelet-based normalization
WEB	WeberFaces

1. GİRİŞ VE AMAÇ

Bilgisayar ortamında bir kişinin kimliğinin tespit edilebilmesi için çeşitli biyometrik özellikler kullanılmaktadır. Biyometri tabanlı kişi tanıma sistemi, bir bireyin fiziksel ya da davranışsal kimliğini ölçen ve varolan kayıtlarla karşılaştırarak tanımlama işlemi yapan otomatik bir sistemdir. Başka bir deyişle, kişinin tanıma işlemi personel kimlik kartları, manyetik kartlar, anahtarlar veya şifreler yerine kolay ve rahat bir doğrulama sunan parmak izi, yüz, iris, avuçiçi izi, DNA ve retina gibi biyometrik özellikler ile gerçekleştirilir. Yüz biyometrisi kişinin kimliği hakkında geniş bir spektrum sunması ile tanıma sistemleri tarafından en çok tercih edilen biyometrik özellik haline gelmiştir. Günümüzün gelişen teknolojisiyle birlikte yüz tanıma sistemi bir fotoğrafı ya da video görüntüsünü yüzün karakteristik özelliklerini tanımlayan bir koda dönüştürebilir. Bu sistem kişinin özel alanını ihlal etmeden kişiyi uzaktan tanımlamak için kullanılır.

Yüz tanıma sistemleri, bir insanın karşısındaki kişinin kimliğini tespit ederken kullandığı tanıma mekanizmasıyla benzer şekilde çalışmaktadır. İnsanın gelişmiş görsel yetenekleri sayesinde tanıma problemi ne kadar zor olursa olsun başarılı bir şekilde gerçekleştirilir. İnsan, aynı kişiye ait yüz imgesini farklı aydınlatma koşulları altında ya da farklı pozlarda rahatlıkla tanıyabilmektedir. Ancak bilgisayarla yapılan tanımda sistem daha genel bir perspektif sunmakta ve tanıma işlemi daha karmaşık bir problem haline gelmektedir. Bu yüzden yüz tanıma sistemlerinin geliştirilmesiyle ilgili çalışmaların kökeni uzun yıllar öncesine dayanmaktadır.

Yüz tanıma, askeri, güvenlik ve hukuki amaçların yanı sıra personel katılım/devam takibi gibi ticari amaçlarla kullanılabilen geniş bir uygulama sahasına sahiptir. Yüz tanıma görüntü işleme, bilgisayarla görü ve örüntü tanıma alanındaki birçok araştırmacı tarafından ilgi gören bir konudur. Ancak, popüler bir çalışma alanı olmasına rağmen problemin zorluğu sebebiyle başarıyı yüksek gürbüz bir yüz tanıma sistemi henüz oluşturulamamıştır. Aydınlatma farklılıkları, kamera açısı veya yüzün hareket ettirilmesi ile oluşan poz değişimleri (yüz dönmeleri vs.), yüzdeki kısmi kapanmalar, doğal görünümlerdeki farklılıklar (sakal, bıyık, yoğun makyaj vs.) yüz tanıma sistemlerinin başarımlarını olumsuz yönde etkilemektedir. Bu problemlerin tamamının üstesinden gelebilen gürbüz bir sistem oluşturulabilmesi için yüz tanıma sistemlerinde önileme adımlarına daha fazla önem verilmesi gerekmektedir.

Yüz tanıma veri toplama kolaylığı sunması sebebiyle yüz imgelerinden oluşan büyük bir verinin oluşmasına olanak sağlamaktadır. Bir video kamerası ile çekilen bir kişinin yüzüne ait videodan yüzlerce yüz imgesi elde edilebilmektedir. Buna ek olarak bilgisayarların işlem gücünün yükselmesiyle birlikte bu büyük verinin işlenebilirliğinin artması da bu konuya olan yönelimi arttırmıştır. Video kamerasıyla elde edilen görüntülerin çerçeveler bazında ayrıştırılması, kontrollü ortamlarda elde edilen tekil imgelere göre çok daha fazla çeşitlilik içermekte ve kişi hakkında daha fazla bilgi edinilmesini sağlamaktadır. Böylece kişinin eşsiz yüz biyometrisi farklı özellikleriyle incelenebilmektedir. Bu avantajlar göz önüne alınarak yüz tanıma alanında yapılan çalışmalar daha çok video tabanlı yöntemler üzerinde yoğunlaşmaya başlamıştır.

İmge kümeleri tabanlı tanıma yaklaşımları son yıllarda ilgi çeken bir araştırma konusudur. Bu yaklaşımda, klasik yüz tanıma sistemlerinden farklı olarak video görüntüleri kişiye ait bir imge kümesi olarak tanımlanmaktadır. Bu sistemde sorgu aracı imge kümeleri olmakta ve tanıma imge kümelerinin karşılaştırılmasıyla yapılmaktadır. Küme modellemeleri genellikle ayrıık imgelerin bir öznitelik uzayında noktalar biçiminde, imge kümelerinin ise küme içindeki noktaları içine alan bölgeler biçiminde temsil edilmesiyle gerçekleştirilir. Böylece, her bir kimlik çok boyutlu öznitelik uzayında bir bölge biçiminde modellenmiş olur. Sorgu kümesi sınıflandırılması, sorgu kümesi imge modeli ile eğitilen kimlik imge modellerinin geometrik uzaklıklarına bakılarak yapılır. En yakın uzaklıkta olan kimlik modeli, sorgulanan imge kümesi ile eşleştirilir. İmge kümeleri tabanlı yöntemlerin en büyük avantajı yüz imgelerine doğal koşullar altında ulaşabilme özelliğidir. Standart ve kontrollü ortamlar yerine video kamerasının yerleştirildiği herhangi bir noktadan kayıt yapılabilmekte ve bir çok veri elde edilebilmektedir.

Yüz tanıma sistemlerinde genellikle yüz sezici çıktıları direkt olarak kullanılmaktadır. Ancak, daha önceden de bahsedilen gerçek dünya ile ilgili zorluklar sınıf içi değişkenliği arttırdığından bu çıktıların güvenilirliğini azaltmaktadır. Özellikle kafanın yunuslama (pitch), sapma (yaw) ve yuvarlanma (roll) hareket eksenlerindeki değişimleri sınıf içi değişkenliği, sınıflar arası değişkenlikten daha belirgin hale getirmektedir. Bu durum, poz ve perspektif dönüşümlerin değişimlerinin yüksek olduğu veritabanlarında yüz sezicilerin çıktılarına müdahale etme zorunluluğunu gündeme getirmektedir.

Yüz tanıma sistemlerinin başarımı, bu istenmeyen sınıf içi farklılıkların ortadan kaldırılmasıyla artırılabilir. Bu önışleme adımı yüz imgelerinin hizalanması olarak adlandırılmaktadır. Yüz hizalama işleminde kafa hareketlerinin ve perspektif dönüşümlerinin küçük ölçülerde olduğu yüz imgeleri için sabit bir poz konumuna hizalama yapılması mantıklı görünmektedir. Ancak kafa hareketlerinin yüksek olduğu durumlarda bir

poz sınıfına ait yüz imgelerinin o sınıfa ait şablona hizalanması yüz tanıma sistemlerinin başarımlarını arttırıcı yönde etki yapması beklenmektedir. Ayrıca otomatik poz tahminini de içeren bu hizalama süreci yüz sezicilerinin çıktılarının daha kullanışlı hale gelmesini sağlamaktadır. Bu amaçla, kanonik pozlar ve pozlara uygun olan standart yüz şablonları belirlenip, yüz imgelerinin belirlenen yüz şablonlarına otomatik olarak eşleştirilmesi üzerine yeni çalışmalar yapılmıştır. En verimli çalışan modelleri belirlemek için geniş hacimli veritabanı üzerinde denemeler yapılarak sonuca varılmıştır.

Çalışmalarda çoğunlukla karşılaşıldığı kadarıyla imge kümeleri oluşturulurken yüz sezme algoritmalarının döndürdüğü çıktılar ya hiç hizalanmadan ya da insan eliyle hizalanarak kullanılmıştır. Gerçek zamanlı uygulamalarda, pozların otomatik olarak ayırt edilmesi ve her bir pozun kendi kategorisinde doğru olarak hizalanması gerekmektedir. Bu hizalama otomatik hale getirilerek yüz tanıma sistemleri içinde kullanımı daha elverişli hale gelmektedir. Yüz tanıma sistemleri için önemli bir önışleme adımı olan yüz hizalama adımı kullanılarak sistemlerin başarımlarının arttırılması beklenmektedir.

Veritabanlarında katılımcı sayısının yükseldiği durumlarda mevcut yöntemlerin tanıma başarımlarında düşmeler olması kaçınılmazdır. Bu problemin üstesinden gelebilmek için tez çalışmasında çeşitli yenilikler denenmiştir. İlk olarak, farklı pozların uygun biçimde otomatik olarak gruplandırılması ve hizalanması üzerine geliştirmeler yapılmıştır. Literatürde hız ve doğruluk açısından en uygun yüz sezme algoritması olduğu düşünülen algoritmalar kullanılmıştır. Yüz imgeleri elde edildikten sonra aydınlatma etkilerinin azaltılması için literatürde yer alan birçok yöntem denenmiş ve aydınlatma farklılıklarının giderildiği imgeler üzerinde poz sınıflandırma çalışmaları yapılmıştır. Hizalanan yüz imgelerinde oluşan imge kümeleri literatürde yer alan birçok imge kümesi tabanlı yöntemde kullanılarak tanıma başarımları karşılaştırılmıştır.

2. LİTERATÜR ARAŞTIRMASI

İmge kümeleri tabanlı yüz tanıma sistemleri örüntü tanıma, makine öğrenimi, nesne tanıma ve görüntü işleme alanlarının kesişim noktasında yer almaktadır. Klasik yüz tanıma sistemleri, veritabanı oluşturma, verilerden bilgilerin çıkarılması, sistemin eğitilmesi ve sistemin testi aşamalarından oluşmaktadır. İmge kümeleri tabanlı sistemler de bu aşamaların tamamını içermektedir. Bu aşamalar birbirlerine bağlı şekilde çalışmaktadır. Veritabanı oluşturma aşamasında belirlenen hedef doğrultusunda koşulları sağlayan veriler toplanarak standart bir formda biriktirilir. Biriktirilen bu veriler yine belirlenen hedef doğrultusunda işlenerek bilgi çıkarımı yapılır. Elde edilen bilgilerin kullanılarak test aşamasında belirleyici özelliği olan parametreler çıkarılır. Bu aşama ise sistemin eğitim aşamasıdır. Eğitilmiş sisteme verilen yeni verilerle ilgili sonuçların elde edildiği aşama ise test aşamasıdır. Tanıma işlemi test aşamasında gerçekleştirilmektedir.

Görünüş tabanlı yüz tanıma uygulamalarında, düşük çözünürlük, aydınlatma şiddeti farklılıkları, yüz ifade değişiklikleri ve benzeri zorlaştırıcı problemler içeren durumlar söz konusu olduğunda imge kümeleri içeren yaklaşımların daha başarılı tanıma oranları verdiği (Yamaguchi vd., 1998; Chang vd., 2007; Kim vd., 2007) rapor edilmiştir. İmge kümeleri ile yüz tanıma sistemlerinin başarılı sonuçlar elde etmesi üzerine son yıllarda birçok araştırmacı bu alan üzerine yoğunlaşmaya başlamıştır (Zhao vd., 2003; Hadid ve Pietikainen, 2004; Wang vd., 2009; Cevikalp ve Triggs, 2010; Cui vd., 2012; Chen vd., 2013b).

Video görüntüleri tekil imgelere göre çok daha fazla bilgi içermektedir. Bu da yüz tanıma sistemlerinde başarıyı arttırıcı yönde etki göstermektedir. Ayrıca büyük veri işleme kapasitesi günümüzdeki teknolojiyle çok daha kolay hale gelmektedir. Video görüntülerinden elde edilen imge dizilerine hesapsal yaklaşımlar, imge kümeleri ile yüz tanımanın temellerini oluşturmaktadır (Cevikalp ve Triggs, 2010) .

Video görüntülerinin içerdiği yüksek bilgiye rağmen kontrollü ortamlarda oluşturulan veritabanlarında daha yüksek tanıma oranları elde edilmektedir. Bunun sebebi kişilerin inisiyatifinden bağımsız olarak toplanabildikleri için mimik, yüz ifadesi, poz, aydınlatma ve kısmi kapanmalar konusunda çeşitlilikler içermesidir. İmge kümeleri kullanılan sistemlerinin klasik yöntemlerden farkı imge-imge karşılaştırması yerine imge-kümesi-imge kümesi karşılaştırması yapmasıdır. Bu şekilde kişilerin kimlikleri bir kişiye ait

videodan elde edilen imgeler kullanılarak oluşturulan imge kümesiyle yapılmaktadır (Cevikalp ve Triggs, 2010; Cui vd., 2012).

İmge kümeleri kullanılan sistemlerde kişi sınıfları imge kümelerinden öğrenilen alt uzaylarla temsil edilirler (Wang vd., 2009). Alt uzaylar öklid olmayan, Grassmann manifold ya da eğrisel Riemann manifold olarak bilinen biçimdedir ve imgeler bu alt uzaylarda birer nokta ile, imge kümeleri ise kümeye ait olan noktaların birleşiminden oluşan bölgelerle temsil edilir (Harandi vd., 2013). Kişilere ait imge kümelerinin sınıflandırılması için bazı benzerlik ölçütleri kullanılır ve karşılaştırma gerçekleştirilir. Benzerlik ölçüsü olarak doğrusal uzaylar kullanılmışsa açısız uzaklıklar, afin uzaylar kullanılmışsa geometrik uzaklıklar (bölgeler arası en yakın mesafe uzunluğu) kullanılmaktadır. İmge kümeleri tabanlı yöntemlerin farklılaştıkları nokta bu kümelerin modellenmesi ve kümeler arası uzaklık ölçüsüdür.

Yamaguchi ve diğerleri (Yamaguchi vd., 1998), müşterek alt uzay (mutual subspace) tabanlı yaklaşım kullanarak imge kümelerini doğrusal alt uzaylar biçiminde, kümeler arasındaki uzaklıkları ise alt uzaylar arasındaki kanonik açılar cinsinden kullanmıştır. Ohkawa vd. (2011) geliştirilmiş müşterek alt uzayları el figürlerini sınıflandırmada kullanmıştır. Hamm ve Lee (2008 a , 2008 b) kaynakçalarında, veri kümesi vektörler yerine doğrusal alt uzaylardan oluştuğu durumlarda verileri ayırt etmede Grassmann manifoldlarını kullanmıştır. Bu manifoldların kernel versiyonları ve değişik uzaklık ölçütleri çalışmaları Wang vd. (2009) ve Harandi vd. (2011) çalışmalarında denenmiştir. Kernel yöntemlerinde daha başarılı sonuçlar rapor edilmesine rağmen bu tip yöntemlerdeki en önemli problem, alt uzayları ayırt etmede kısıtlı bilgilerin (doğrusal alt uzay aralarındaki açılar) kullanılmasıdır. İmge kümelerini modellemede doğrusal alt uzay yerine afin alt uzay kullanımları da yapılmıştır. Fitzgibbon ve Zisserman (2003) ve Arandjelovic ve Zisserman (2005) çalışmalarında filmlerdeki temel karakterleri tanımada afin alt uzaylarla birlikte birleştirilmiş manifold uzaklığı (joint manifold distance) ve buna uygun bir topaklama yaklaşımı kullanılmıştır. Lokal doğrusal alt uzay modellerinin birleşiminden doğrusal olmayan manifoldlar oluştuğu düşünülürse, imge kümelerini ayırt etmede manifoldlar arası uzaklık ölçüleri de önemli yer tutmaktadır. Başka bir tip modelleme yaklaşımı, imge kümelerinin parametrik bir olasılık dağılım fonksiyonuyla betimlenip, dağılımlar arasında benzerliklerin ölçülmesine dayanır (Arandjelovic vd., 2005). Buna benzer olasılıksal öğrenme yöntemlerinin farklı biçimleri Wibowo vd. (2012, 2013) çalışmalarında denemiştir. Ancak parametrik olasılık dağılımı ile modelleme içeren yaklaşımların büyük dezavantajları vardır. Bu modeller, olasılık dağılımının uygun biçimde tahmin edilmesi için geniş hacimli eğitim verilerine ihtiyaç duymaktadır; zor bir parametre tahmin edici problemi çözmek zorundadır ; sorgu kümeleri kullanılarak yapılan deneylerde veritabanındaki eğitim kümeleri

arasında kuvvetli istatistiksel ilişkiler olmadığı durumlarda genelde başarılı sonuçlar vermez (Wang vd., 2012).

İmgeler doğrusal ve afin uzay yöntemleri ile modellenemediği gibi, bu amaçla afin uzaylarla karşılaştırıldığında daha kompakt bir model olan konveks zarflar da kullanılabilir. Afin ve konveks zarfların farklı türevleri literatürde bulunmaktadır. Örneğin Sparse Approximation Nearest Points (SANP) (Hu vd., 2012) yöntemi imge kümelerinin afin alt uzaylarla temsilini afin uzay kombinasyonlarında kullanılan örneklerin seyrekliğini kullanarak genişletmiştir. İki imge kümesinin SANP'ları ortalama imge ve kümelerin afin zarf modelleri kullanılarak tanımlanır ve ardından kümeler küme imgelerine benzetilerek eş zamanlı olarak kümelere en yakın noktalar aranır. SANP yöntemi iyi sonuçlar dönmesine rağmen, afin zarf modelleri dışında 3 tane daha parametre ayarlanması gerektiğinden oldukça karmaşıktır. Buna ek olarak $L1$ norm minimizasyonu içeren karmaşık bir eniyileme probleminin çözmesinden dolayı yavaş bir yöntemdir ve gerçek zamanlı uygulamalarda kullanılması mümkün değildir. Benzer bir şekilde, Yang vd. (2013) düzenlenmiş afin zarf modellerini kullanarak afin zarflar arası $L2$ norm uzaklık ölçütünü kullanmıştır. Bu diğer yöntemlerle karşılaştırıldığında çok daha basit bir eniyileme problemini çözmekte ve çok daha hızlı olmasına rağmen hala gerçek zamanlı uygulamalar için uygun değildir. Buna ek olarak bu yöntemlerin yeni genişlemiş halleri olan işbirlikçi temsilleri önerilmiştir (Wu vd., 2013; Zhu vd., 2013). Bu yöntemlerdeki temel farklılık galerideki tüm imge kümelerini tek bir afin/konveks zarf modeli ile modelleyerek test kümesini yeniden yapılandırma artanlarını kullanarak sınıflandırmasıdır. Buna rağmen bu yöntemler büyük ölçekli veriler için yeterli çözümleri üretememektedir. Ayrıca bunlara ek olarak seyrek modelleri kullanan imge kümesi tabanlı tanıma sistemleri de mevcuttur (Cui vd., 2014; Chen vd., 2013 a; Chen vd., 2013 b). Bahsedilen yöntemlerinin birçoğunun lineer olmayan yüz modellerini kestirmek için kernel versiyonları da mevcuttur.

Yakın zamanlarda, Hayat vd. (2014, 2015) çalışmalarında imge kümelerinin lineer olmayan geometrik yapısını derin öğrenme çerçevesinde elde etmiştir. Bu yöntemde her bir imge kümesine ait modellerin öğrenilmesi için Uyarlamalı Derin Ağ Şablonları eğitilerek bu modellerle test kümesi arasındaki en küçük yeniden oluşturma hatasına bakılarak sınıflandırma yapılmaktadır.

İmge kümeleri ile yüz tanıma çalışmalarında çoğunlukla karşılaşıldığı kadarıyla imge kümeleri oluşturulurken yüz sezme algoritmalarının döndürdüğü çıktılar ya hiç hizalanmadan ya da insan eliyle hizalanarak kullanılmıştır. Yüz tanıma başarımında yüz hizalamanın etkisi önemlidir. Kameradan yakalanan görüntülerde yüzün kameraya

uzaklığına göre değişim gösteren ölçekleme farklılıkları, yüzün dönme açısı veya eğiklikleri tanıma başarımlarını etkilemektedir. Yüksek tanıma oranları elde etmenin bir yolu kanonik bir yüz şablonu belirleyip, bireylerin yüzlerindeki göz, burun veya ağız gibi çeşitli nirengi noktalarının her birey için şablonda tasarlanan yerleşkelere ölçekleme, kaydırma ve döndürme gibi temel doğrusal konformal dönüşümler kullanılarak haritalanmasıdır. Yapılan bazı çalışmalarda bu konunun önemi literatürde yer alan çalışmalarda denenmiş ve rapor edilmiştir (Yavuz, 2013). Bu çalışmada nirengi noktaları işaretlemeleri el ile yapılmıştır. Yüzdeki önemli noktaların otomatik olarak doğru biçimde işaretlenmesi zor bir problemidir (Cevikalp vd., 2011; Yavuz vd., 2013). Otomatik yüz hizalama konusu, özellikle poz açıları farklılıkları olduğu durumlarda ya da yüzü kapatan aksesuarların (ya da gürültü içeren imgeler) olması durumlarında çözülmesi oldukça zor olan bir problemidir. Bu durumlarda, standart piksel tabanlı hizalama yaklaşımları yerine daha gürbüz olan matris ayrıştırması prensiplerine göre çalışan yaklaşımlar gerektirmektedir (Wright vd., 2009). Benzer şekilde, Peng vd. (2012) çalışmasında bir grup ardışıl doğrusal olarak ilişkili imgeleri (yüzün kısmen kapatıldığı imgeler de olabilir) düşük ranklı matris ayrıştırması prensiplerine göre geliştirerek daha yüksek doğru hizalama başarımları gerçekleştirmişlerdir. Peng vd. (2012) çalışmasında verilen yöntemin büyük hacimli yüz hizalama uygulamalarında daha hızlı bir biçimde çalışacak iyileştirmesi Wu vd. (2011) çalışmasında verilmiştir. Benzer yöntemi temel alan ancak yüz hizalamada iteratif konveks optimizasyon teknikleri ilave eden bir diğer yaklaşım Wu vd. (2012) çalışmasında sunulmuştur. Farklı teknikler kullanan yüz hizalama çalışmalarına da rastlamak mümkündür. Açık suret regresyonu yaklaşımı (Cao vd., 2014) yüz hizalamada bir vektör regresyon fonksiyonu öğrenerek yüz suretini modelleme temeline dayanır. Shen vd. (2013) çalışmalarında imge erişim (image retrieval) tekniklerine dayanan yüz hizalama yöntemi önermiştir. Ding ve Martinez (2010) çalışmalarında yüz dokusu analizinin yüz betimlemede yeterli bilgi içermediği için göz, ağız ve burun bölgelerinde öğrenilen özel konteks bilgilerinin de ilave edilmesinin tanıma başarımlarını arttıracakını belirtmiştir.

Bu yöntemlerin yanı sıra 2 boyutlu ve 3 boyutlu yaklaşımlar yapılarak yüz imgelerindeki pozluluğun normalleştirilmesini hedef alan bazı çalışmalar bulunmaktadır. 3 boyutlu normalleştirmeyi sağlayan yöntemlerde öncelikle 2 boyutlu yüz imgeleri 3 boyuttaki modellere hizalanmakta ve ardından 3 boyutlu dönüşümler kullanılarak poz cephe görünümüne dönüştürülmektedir (Blanz ve Vetter, 2003; Asthana vd., 2011). 2 boyutlu normalleştirme yöntemlerinde ise cephe görünümü Markov Random Field (MRF) gibi grafik modellerle temsil edilmektedir. Eğitim aşamasında her poza ait MRF modeller kullanılarak pozlar arası bağımlılıklar tespit edilmektedir (Ashraf vd., 2008; Rahimzadeh Arashloo ve Kittler, 2011).

Wang vd. (2014) iki video dizisi arasında hizalama yapmak için zaman kaydırmasına dayanan bir yöntem önermiştir. Yaklaşım, özellik tanımlayıcılarını kullanarak olası tüm çift eşlemeler için bir maliyet hesaplamıştır. Çerçeve i 'yi birinci sıradan çerçeve j ile ikinci sıralamadan eşleştirme maliyeti, iki çerçeve arasındaki eşleştirilmiş özellik tanımlayıcılarının sayısı ile ters orantılıdır. Yanlış olumlu özellik eşleşmeleri, maliyet matrisi üzerinde nispeten zayıf bir etkiye sahip olduklarından filtrelenmemektedir. Sıra hizalama algoritması, en uygun çerçevelerin tepe noktasını çıkarmak için en kısa yol algoritmasını kullanır.

Naseer vd. (2015), Wang vd. (2014) tarafından bu konu üzerine yapılan çalışmalar genişletmiştir. Bu çalışmalarda, maliyet matrisinin oluşturulması için Histogram of Gaussian (HoG) gibi özellikler tanımlayıcıları ve Konvolusyonel Sinir Ağı katmanındaki düğüm aktivasyonlarını içeren çeşitli yaklaşımlar araştırılmıştır.

3. MATERYAL VE YÖNTEM

Otomatik yüz hizalama süreci geleneksel olarak yüz sezimi, sezilen yüz bölgesi içinde göz, ağız gibi önemli noktaların bulunması ve yüz bölgesinin önceden belirlenen bir yüz şablonuna haritalanması biçimindeki üç aşama ile gerçekleştirilir. Bu geleneksel yaklaşımın yanı sıra bazı hizalama yöntemlerinde sezilen yüz bölgeleri kullanılarak noktasal haritalama yerine matris tabanlı yaklaşımlar kullanılmıştır. Tezin bu bölümünde öncelikle her iki yaklaşım için ortak olan yüz sezimi aşamasından bahsedilmiştir. Ardından sırasıyla noktasal ve matris tabanlı yaklaşımlar ile hizalama yapılan yöntemler incelenmiştir.

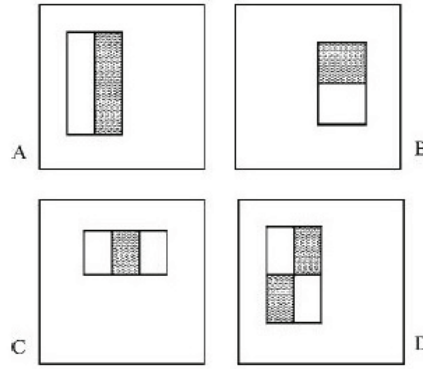
3.1 Yüz Sezimi

Yüz sezme amacıyla kullanılan çok farklı yöntemler mevcuttur (Yang vd., 2002). Genel olarak görünüş tabanlı yüz sezme yöntemleri yüz koordinatlarının bulunması problemini, yüz sınıfını (yüz imgeleri) arka plan sınıfından (yüz olmayan imgeler) ayırma amacı güden iki-sınıflı bir sınıflandırıcı tasarlama problemi olarak ele alır.

3.1.1 Viola-Jones yüz sezimi

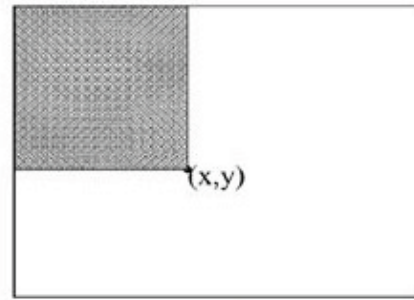
Yüz sezimi konusunda belki de en popüler çalışma, günümüzde de oldukça yaygın olarak kullanılan Viola ve Jones tarafından gerçekleştirilmiştir (Viola ve Jones, 2001). Bugün bile en iyi yöntemler arasında sayılan bu çalışmada imgeleri betimleme amacıyla Haar dalgacık benzeri öznitelikler kullanan araştırmacılar, integral imge adını verdikleri teknik sayesinde bu öznitelikleri çok hızlı bir şekilde hesaplamışlardır. Haar dalgacık benzeri öznitelikler imgeyi dikdörtgensel bölgeler bazında inceleyerek bu dikdörtgensel bölgelerin beyaz kısımlarındaki tüm piksellerin toplanıp siyah kısımdakiler de çıkarılarak elde edilen özniteliklerdir. Bu dikdörtgensel bölge öznitelikleri Şekil 3.1'de gösterildiği gibi iki dikdörtgenli, üç dikdörtgenli ve dört dikdörtgenli olabilmektedir.

Dikdörtgen özniteliklerin çıkarımı dikdörtgenlerin içerdiği piksel değerlerinin toplanmasını gerektirdiğinden bu işlem çok yüksek miktarda işlem yapılmasını gerektirmektedir. Bu da özellikle boyutları büyük imgelerde maliyeti yüksek bir işlemdir. Bu maliyetin azaltılması adına integral imge kavramı ortaya atılmıştır. İntegral imge, imgedeki her piksel için sol ve üstündeki piksellerin değerleri toplanarak oluşturulan yeni bir imgedir. İntegral imge Şekil 3.2'de gösterildiği gibi orijinal imgenin (x, y) noktasının sol



Şekil 3.1 Dikdörtgensel bölge tabanlı Haar benzeri öznelikler. İki dikdörtgenli (A)ve (B)'de, üç dikdörtgenli (C)'de, dört dikdörtgenli (D)'de gösterilmiştir (Viola ve Jones, 2001).

ve üstünde kalan piksel değerlerinin toplamından elde edilen değerin integral imgedeki aynı (x, y) noktasına aktarılmasıyla oluşmaktadır.



Şekil 3.2 (x, y) noktasındaki integral imge piksel değeri, orijinal imgede aynı noktanın sol ve üstünde kalan piksellerin toplanmasıyla elde edilen değerdir (Viola ve Jones, 2001).

Bir imgeye ait integral imge Denklem 3.1 'de verildiği şekilde hesaplanabilir. İntegral imgenin oluşturulması dikdörtgensel özneliklerin çıkarılmasındaki işlemlerin yükünü azaltmakta ve yüz seziminin gerçek zamanlı olarak yapılabilmesine olanak sağlamaktadır .

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (3.1)$$

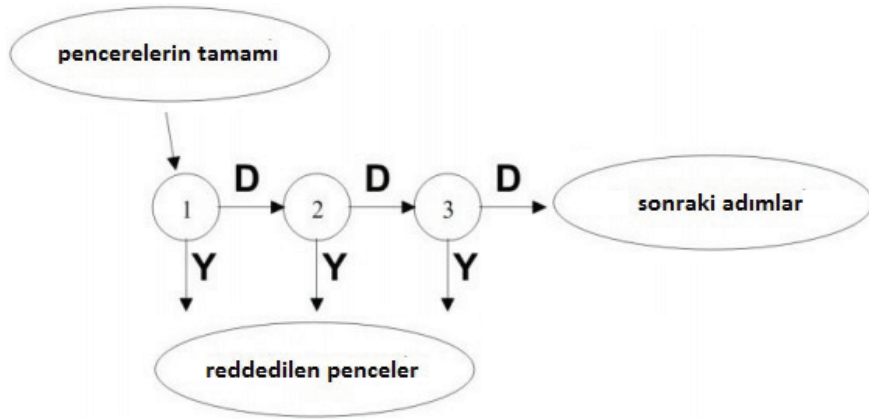
Dikdörtgensel özneliklerin kullanımı yüz bulma işleminde yüz ya da yüz değil şeklinde ayırt edecek bir sınıflandırıcı kullanılmasını gündeme getirmiştir. Bu sebeple,

AdaBoost algoritmasındaki her biri minimum sayıda yanlış sınıflandırılmış örnek veren eşik değerli bir fonksiyon olan zayıf sınıflandırıcılar ile pozitif ve negatif örnekleri en iyi şekilde ayıran öznelilikler çıkartılır. Her zayıf sınıflandırıcı bir öznelilik f , bir eşik değeri θ ve eşitsizlik işaretinin yönünü belirten bir polarite p 'den oluşacak şekilde $h(x, f, p, \theta)$ olarak tanımlanır ve sınıflandırıcı Denklem 3.2'deki şekilde ifade edilebilir.

$$h(x, f, p, \theta) = \begin{cases} 1, & pf(x) < p\theta \\ 0, & \text{aksihalde} \end{cases} \quad (3.2)$$

Bu durumda nihai sınıflandırıcı bir dizi sınıflandırıcının art arda eklenmesiyle oluşturulmuştur. Bu zayıf sınıflandırıcılar eğitim kümesindeki yanlış sınıflandırılmış örneklerin ağırlıklarını arttırarak bu ağırlıklarla en az hatayı veren yeni bir sınıflandırıcıya girdi olarak alınır. Bu işlem belirlenen yanlış tahmin miktarına ulaşana kadar devam ettirilir.

Zayıf sınıflandırıcıların art arda eklenmesiyle oluşturulan sınıflandırıcı yapısı Şekil 3.3'de gösterilmiştir. Birinci sınıflandırıcıdan dönen pozitif sonuçlar ikinci sınıflandırıcıya alınarak ilk sınıflandırıcıda pozitif olarak ayrılmış negatif örnekler ayıklanır. Bu işlem sonraki adımlarda da devam eder. Bu işlem daha önceden de belirtildiği gibi yanlış tahmin sayısına ulaştığında sınıflandırma işlemi tamamlanır. Her aşamadaki zayıf sınıflandırıcılar AdaBoost ile eğitilir.



Şekil 3.3 Zayıf sınıflandırıcıların art arda eklenmesiyle oluşan nihai sınıflandırıcının şematik gösterimi (Viola ve Jones, 2001)

Bu sınıflandırma yüzlere ait olmayan arka plan imgelerinin birçoğunun erken bir şekilde reddedilmesini ve yüz görüntülerinin çok hızlı bir şekilde bulunmasını sağlamıştır. Viola-Jones yöntemi kullanılarak yapılan yüz sezimi sonuçları Şekil 6.1’de verilmiştir.

Bu yöntem özellikle ön cephe görünümlü yüz içeren imgelerde yüksek başarı oranına sahip olmasına rağmen, poz ve aydınlatma farklılıkları içeren imgelerde yüzleri kaçırabilmektedir. Ancak, özellikle profil imgelerinde başarıyı yeterli seviyelerde gözlenemediği için daha zorlayıcı durumlarda yüzü sezebilen bir başka yüz sezicinin kullanılması ihtiyacı doğmuştur.

3.1.2 Zhu & Ramanan yüz sezimi

Yüz sezimi yöntemlerinin, aydınlatma ve poz farklılıklarının herhangi bir ön işleme adımı olmaksızın uygulanması hizalama yönteminin veritabanından bağımsız hale gelmesi açısından önemlidir. Bu sebeple, yüz sezimi için aydınlatma ve poz farklılıklarının etkisini en aza indirebilen yöntemlerin araştırılması gündeme gelmiştir. Yapılan araştırma sonucu yakın zamanda önerilen yüz sezicilerde, geleneksel yüz hizalama sürecindeki üç farklı aşama (yüz sezimi, nirengi noktaları yerleşimi, poz tahmini) tek bir aşamada çözebilecek biçimlerde tasarlandıkları görülmüştür. Doğru bir şekilde tespit edilen nirengi noktalarının konumları kullanarak hizalama işlemi kolaylıkla yapılabilir. İki imge ele alındığında tespit edilen nirengi noktaları kullanarak basit bir dönüşüm hesaplanarak ikinci imge ilkine göre hizalanabilmektedir. Hizalama işlemi için şablon bir imge yerine nirengi noktalarını içeren ve verilen poz sınıfı için sabit olan bir nirengi noktası yerleşimi dahi imgelerin tamamının hizalanmasını sağlayacaktır. Nirengi noktalarının yerleşimi konusunda ele alınan yöntemlerden biri de aydınlatma şiddeti farklılıkları ve poz açısı değişimlerine karşı daha gürbüz olduğu düşünüldüğü için Zhu vd. (2013) tarafından önerilen yöntemdir.

Zhu vd. (2013) tarafından önerilen yüz sezimi yöntemi yüzdeki elastik deformasyonları ve üç boyutlu yapıyı kodlayan bir algoritmadır. Bu yöntem ortak bir bölge havuzu ile karar ağaçlarının karışımını kullanmaktadır. Her nirengi noktasında bir bölge tanımlanmakta ve topolojik değişiklikleri modellemek için küresel karışımlar kullanılmaktadır. Bu karışım ağaçlarında bazıları Şekil 3.4’de verilmiştir.

V ortak bölge havuzu, m karışım ve $V_m \subseteq V$ olmak üzere her ağaç $T_m = (V_m, E_m)$ şeklinde tanımlanabilir. İmge I ve imgedeki piksel yerleşikleri $l_i = (x_i, y_i)$ olarak alındığında



Şekil 3.4 Bakış açısından kaynaklanan topolojik değişiklikleri kodlayan karışım ağaçlarından örnekler (Zhu vd., 2013)

skor fonksiyonu Denklem 3.3, 3.4 ve 3.5’de gösterildiği şekilde tanımlanabilir.

$$S(I, L, m) = App_m(I, L) + Shape_m(L) + \alpha^m \quad (3.3)$$

$$App_m(I, L) = \sum_{i \in V_m} w_i^m \cdot \phi(I, l_i) \quad (3.4)$$

$$Shape_m(L) = \sum_{ij \in E_m} = a_{ij}^m dx^2 + b_{ij}^m dx + c_{ij}^m dy^2 + d_{i,j0}^m dy \quad (3.5)$$

Skor fonksiyonundaki ilk terim görünüş puan fonksiyonudur. Bu fonksiyon belirli bir bölgede şablon w_i ile lokal eşleşmelerin yapılmasını sağlamaktadır. Denklem içerisinde imgedeki l_i yerleşkesinden çıkarılan HoG betimleyicileri ϕ olarak adlandırılmıştır. Burada tanımlanan görünüm skoru tüm parçalar için şablonlar ve şablonların iç çarpımları toplamıdır. Skor fonksiyonundaki ikinci terim ise parça çiftinin bağlı konumları değerlendiren deformasyon skorudur. Deformasyonu diğerleri üzerinde belirli uzaklıklar tercih eden bir yay olarak düşünebiliriz. Bu durumda yay çok deforme olursa, ikinci dereceden maliyet ödemesi yapılması gerekmektedir. Denklemdaki $[a, b, c, d]$ yayın geri kalan konumunu ve sertliğini tanımlar; α ise bias terimidir. Böylece skor fonksiyonu hem görünüm hem de deformasyonu tanımlamış olmaktadır.

Ele alınan parçaların en iyi skor yapılandırma ve en iyi skor karışımını bulma işlemi karışımlar ve parça konumları üzerindeki puan fonksiyonu maksimize edilerek yapılır (Denklem 3.6). Kenar yapısı bir ağaç olduğunda, bu maksimizasyon dinamik programlama ile verimli bir şekilde yapılabilir.

$$S^*(I) = \max_m [\max_L S(I, L, m)] \quad (3.6)$$

Modelin öğrenme aşamasında pozitif parça yerleri ve bakış açıları ile açıklamalı tam denetimli bir eğitim veri kümesi kullanılmıştır. Ağaç yapıları, Chow-Liu algoritması (Chow ve Liu, 1968) kullanılarak öğrenilmektedir. Denklem 3.7’de verilen skor fonksiyonu lineer olduğundan görünüm şablonu bir destek vektör makinesi (DVM) kullanılarak öğrenilebilir.

$$S(I, z) = \beta \cdot \phi(I, z) \quad (3.7)$$

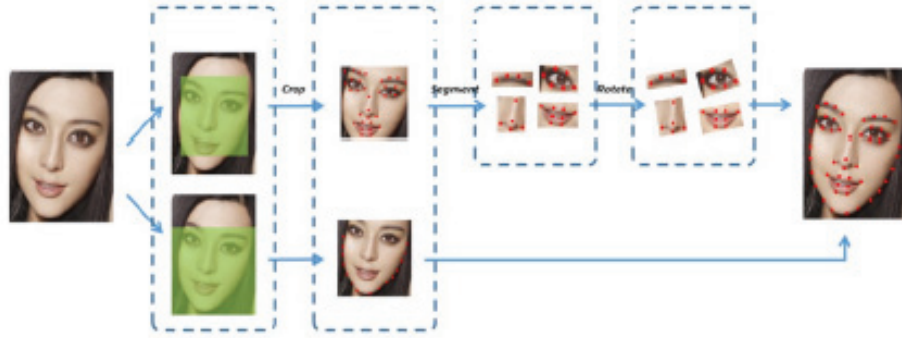
Burada esas nokta bakış açısına göre bu bölgelerin kaybolabilmesinin temel alınmasıdır. Bu sebeple bölgeler için farklı karışım şablonları hazırlanır. Bu durum düşük karmaşıklıkla çok sayıda model oluşturulmasını sağlar. Son olarak modelin tüm parametreleri (bölüm şablonları, elastik deformasyon modları ve görünüme dayalı topoloji) maksimum marjin tabanlı sınıflandırıcılar ile eğitilir.

Zhu vd. (2013) tarafından önerilen yüz sezimi, nirengi noktası yerleşimi ve poz tahmini adımlarını bir arada gerçekleştirebilen birleşik yöntemle alınan sonuçlar Şekil 6.2’de gösterildiği gibidir. Bu yöntem kullanılarak yapılan deneylerde yüzdeki nirengi noktalarının başarılı bir şekilde tespit edildiği görülmüştür. Ancak sistemin gerçek zamanlı olması planlandığından her yöntem için nirengi noktalarının yerleşimi sırasında geçen süre de önem taşımaktadır. Tüm yöntemler için yapılan süre testleri sonucunda nirengi noktalarının yerleşimi sırasında en fazla sürenin Zhu & Ramanan tarafından önerilen yöntem sırasında geçtiği sonucuna ulaşılmıştır. Bu da sistemin gerçek zamanlı koşullara uygunsuzluğunu ortaya koymuştur. Hem başarımlar hem de süre açısından en başarılı yöntemin bulunması için araştırmalara devam edilmiştir.

Hız sorununa ek olarak en başarılı yöntemlerden biri olan Zhu & Ramanan yüz sezicisinin bile nirengi noktalarının yerleşimi konusunda tam başarıma ulaşamadığı gözlenmiştir. Başarısız bir nirengi noktası yerleşimi imgelerin hizalanması konusunda büyük sorunlar yaratabilmektedir. Bu sebeple, en iyi başarımları elde edebilmek için nirengi noktalarından bağımsız olan bir çalışma gündeme gelmektedir. Yüzün konumunun tam doğrulukla bulunmasının sağlanması durumunda nirengi noktaları yerine farklı özniteliklerin kullanılmasıyla hizalama işlemi çok daha yüksek başarımlarla elde edebilmektedir. Nirengi noktası bağımlı hizalama sistemlere ek olarak nirengi noktasından bağımsız ve daha gülbüz özniteliklerin çıkarımıyla ilgili yöntemler ilerleyen bölümlerde anlatılacaktır.

3.1.3 Face++ yüz sezimi

Literatürde yer alan bir diğer başarılı yöntem ise Zhou vd. (2013) tarafından önerilen face++ yüz sezim yöntemidir. Açık kaynak kodlu bir yazılım olan bu yöntem derin katmanlı sinir ağı kullanarak nirengi noktalarının yerleşimini sağlamaktadır. Dört aşamalı katmanlı ağ kaskad yapısını kullanan bu yöntem kabadan inceye bir analiz sağlamaktadır. Kabadan inceye kaskad yapı ve geometrik hassaslaştırma yapılarının bir araya gelmesiyle nirengi noktası yerleşimi sağlanmıştır. Genel görünüm Şekil 3.5’de verildiği şekildedir.



Şekil 3.5 Face++ algoritma sistemi genel görünümü (Zhou vd., 2013)

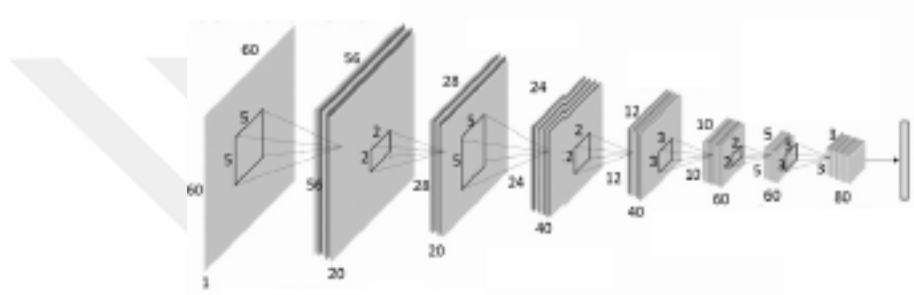
Sistem kaskad ağ seviyelerinden oluşmakta ve her bir seviye bir öncekinden daha ayrıntılı iyileştirmeler yapmaktadır. İlk aşamada yüz iki parçaya ayrılmıştır. Bu parçalardan biri yüzün iç kısmını içerirken diğeri çevresini içermektedir. İkinci seviyede ilk tahminler yapılmaktadır. Diğer aşamalarda ise yüze ait her bir eleman ayrı ayrı incelenerek iyileştirmeler yapılmaktadır. Üçüncü ve dördüncü seviye yüze ait bölümlerin yönlendirme iyileştirmelerinden sorumlu seviyelerdir. Sistemdeki yapay sinir ağı stokastik gradyan düşüş yöntemiyle eğitilir.

Yüz için, iç noktalar ve çevre noktaları olarak iki ayrı bölge belirlenmiştir. İç noktalar 51 noktadan oluşurken, çevre noktaları 17 nokta içermektedir. İlk seviyede, 2 yapay sinir ağı iç ve çevre noktalar için ayrı birer sınırlayıcı kutu bulunması için eğitilirler. Sistemin geri kalanı bu kutularla beslenirler.

İç noktaların eğitilmesinde bir başka 3 seviyeli yapay sinir ağı eğitilmiştir. İç noktalara ait sınırlayıcı kutunun bulunmasının ardından ikinci aşamada 51 iç nokta için başlangıç konumları belirlenir. Bu başlangıç konumlarına göre kaşlar, gözler ağız ve burun olmak üzere 6 yüz bileşeni hesaplanır. Üçüncü seviyede ise her bir yüz bileşenine ait nirengi noktaları doğru şekilde yerleştirilecek şekilde eğitilir. Son olarak yüz bileşenlerinin dönme açısı hesaplanarak döndürülmüş parçalar 4. seviyeye final görünümünü elde etmek için gönderilir.

Çevre noktaları için ise basit bir ağ kaskadı oluşturularak yerleşimleri hesaplanır. Çeneyi çevreleyen sınırlayıcı kutu ikinci seviyeye girdi olarak alınarak noktaların yerleşmeleri bulunur. Bu yerleşmelerin bulunmasında satır vektörleri kullanılır. Çevre noktaların bulunmasında üçüncü ve dördüncü seviye ağlar zaman kısıtlaması sebebiyle kullanılmamaktadır.

Bu kaskad yapının gerçekleştirilebilmesi için derin yapay sinir ağları kullanılmıştır. Bu sistem satır vektörlerini kullanarak istenilen nokta için regresyon uygulamaktadır. Bu derin yapı Şekil 3.6’da gösterildiği şekildedir.



Şekil 3.6 Kullanılan ağların genel yapısı (Zhou vd., 2013)

Girdi düğümünün arkasına üç katman yerleştirilmiştir. Her katman girdi olarak alınan çok kanallı imgeye çeşitli filtreler uygulayarak çıktıları oluşturmaktadır. t . katmandaki imge I^t , I katmanının girdi imgesi, B ve F ayarlanabilir parametreler olmak üzere çıktılar Denklem 3.8’de gösterildiği şekilde yazılabilir.

$$C_{i,j,k}^t = \left| \tanh \left(\sum_{x=0}^{h_t-1} \sum_{y=0}^{w_t-1} \sum_{z=0}^{c_t-1} I_{i-x,j-y,z}^{t-1} \cdot F_{x,y,k,z}^t + B_k \right) \right| \quad (3.8)$$

Standart uygulamaya göre hareket edilerek çıktıların mutlak değer fonksiyonları uygulanarak sisteme doğrusalsızlık sağlanır. Konvolüsyonun ardından örtüşmeyen havuzlama bölgeleri kullanılarak en büyük havuzlama Denklem 3.9’da verildiği şekilde uygulanır.

$$I_{i,j,k}^t = \max_{0 \leq x < d, 0 \leq y < d} (C_{i-d+x,j-d+y,k}^t) \quad (3.9)$$

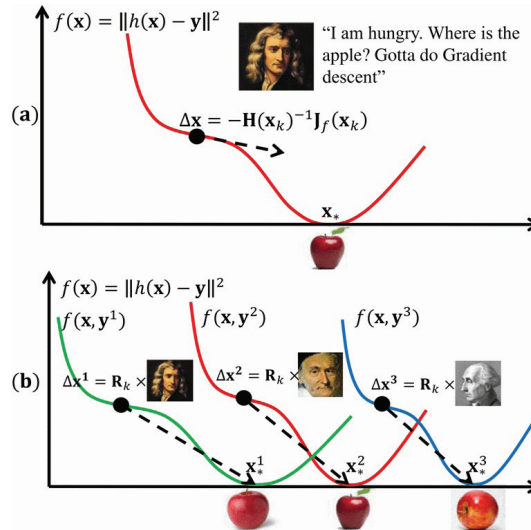
Son tahmin etme bir ya da iki bağlantılı katman tarafından üretilir. Parametreler $L2$ kaybı minimum olacak şekilde Denklem 3.10 'da gösterildiği şekilde ayarlanır.

$$\sum_{I^0} |layer_m \circ layer_{m-1} \circ \dots \circ layer_1(I^0) - labels(I^0)|_2^2 \quad (3.10)$$

Face++ yöntemiyle alınan yüz sezimi sonuçları Resim 6.3'de verilmiştir. Face++ yöntemi için hız açısından herhangi bir sorun olmamasına rağmen yeterince iyi sonuçlar alınamamıştır. Bazı durumlarda Viola-Jones yüz sezim yönteminin bulduğu yüzler bile yöntem tarafından kaçırılmıştır. Yüz bulunamayan imgelerin göz ardı edilmesi durumunda veritabanında büyük kayıplar yaşanacağından özellikle poz ve aydınlatma farklılıkları karşısında çok daha gürbüz bir sistemin kullanılması gerekmektedir.

3.1.4 INTRAFACE yüz sezimi

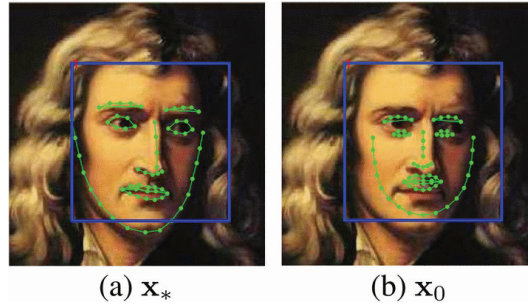
Yapılan araştırmalar sonucu yüzdeki nirengi noktalarının bulunabilmesi için kullanılabilecek bir diğer yöntem ise Xiong ve De La Torre (2013) tarafından önerilen INTRAFACE yöntemidir. Bu yöntem için Eğitmenli İniş Metodu (Supervised Descent Method (SDM)) kullanılmıştır (Xiong ve De La Torre, 2013). SDM doğrusal olmayan en küçük kareler problemlerini eniyilemek için kullanılan eğitmenli bir yöntemdir. Denetimli bir şekilde genel iniş haritaları öğrenerek, nirengi noktalarını başarılı bir şekilde tespit edebilmektedir. Eğitmenli iniş metodunun ana fikri Şekil 3.7'de gösterilmiştir.



Şekil 3.7 Eğitmenli İniş Metodunun gösterimi (Xiong ve De La Torre, 2013)

Şekil 3.7.a’da doğrusal olmayan en küçük kareler probleminin Newton’un metoduyla çözümünü gösterilmektedir. Burada $f(x)$ imge özniteliklerini içeren doğrusal olmayan fonksiyon, y bilinen bir vektör (şablon) ve x hareket parametreleri (döndürme, ölçek, katı olmayan hareket) olmak üzere geleneksel Newton yöntemiyle Hessian ve Jacobian matrisleri hesaplanmıştır. Şekil 3.7.b ise Eğitimli İniş Metodunun genel yapısını göstermektedir. Eğitim verisi farklı yerleşkelerdeki y^i bir dizi fonksiyon olan $\{f\{x, y^i\}\}$ oluşmakta ve minimum değer $\{x_*^i\}$ bilinmektedir. SDM bu eğitim verisini kullanarak lineer olmayan en küçük kareler fonksiyonlarının her birinin ortalamalarını minimize eden parametre güncellemeleri serisini öğrenir. Bu parametre güncellemeleri iki bölüme ayrılır; örneğe özgü bileşen ve genel iniş yönü $\mathbf{R}_k$. SDM eğitim sırasında genel iniş yönünü bulur. Eğitim sırasında ise verilen y ’nin bulunan genel iniş yönünde $(\mathbf{R}_k)$ y ’ye özgü bileşene haritalanmasıyla bulunur.

Verilen imge $\mathbf{d} \in \mathcal{R}^m$, $\mathbf{d}(x) \in \mathcal{R}^p$ imgedeki nirengi noktaları olmak üzere $\mathbf{h}$ lineer olmayan öznitelik çıkarma fonksiyonu ve $\mathbf{h}(\mathbf{d}(x)) \in \mathcal{R}^{128p}$ çıkarılmış SIFT öznitelikleri olarak tanımlanmaktadır. Eğitim süresinde nirengi noktalarını bilinen yerlerdedir ve bunlar fonksiyonda x_* olarak görülmektedir. Ayrıca, test senaryolarının çoğaltılması adına yüz sezimi çalıştırılarak nirengi noktalarının ilk yapılandırması tespit edilmekte ve bu x_0 ortalama nirengi noktaları olarak verilmektedir (Şekil 3.8).



Şekil 3.8 (a) El ile işaretlenmiş 66 nirengi noktası (b) Yüz sezici ile bulunan yüzdeki ortalama nirengi noktaları (Xiong ve De La Torre, 2013)

Bu düzenlemeler altında, yüzdeki nirengi noktalarının yerleşimi Denklem 3.11’de verilen fonksiyonun Δx üzerinden minimize edilmesiyle tespit edilmektedir.

$$f(x_0 + \Delta x) = \|\mathbf{h}(\mathbf{d}(x_0 + \Delta x)) - \phi_*\|_2^2 \quad (3.11)$$

Denklemden $\phi_* = \mathbf{h}(\mathbf{d}(x_*))$ elle işaretlenmiş olan nirengi noktalarındaki SIFT değerlerini temsil etmektedir. Eğitim kümesindeki imgelerde ϕ_* ve Δx bilinmektedir.

Burada türev süreci düşünülerek $\mathbf{h}$ 'nin iki kez türevlenebilir olduğu kabul edilmektedir. Newton'un yöntemine benzer şekilde ikinci dereceden Taylor seri genişlemesi kullanılmaktadır.

$$f(x_0 + \Delta x) \approx f(x_0) + \mathbf{J}_f(x_0)^\top \Delta x + \frac{1}{2} \Delta x^\top \mathbf{H}(x_0) \Delta x \quad (3.12)$$

Denklem 3.12'de verilen $\mathbf{J}_f(x_0)$ ve $\mathbf{H}(x_0)$, f fonksiyonunun x_0 üzerinden Jacobian ve Hessian matrisleridir. Formülü basitleştirmek için x_0 ihmal edildiğinde Denklem 3.12'nin Δx 'e göre türevinin sıfıra götürülmesi x değerinin ilk güncellemesini vermektedir.

$$\Delta x_1 = -\mathbf{H}^{-1} \mathbf{J}_f = -2\mathbf{H}^{-1} \mathbf{J}_h^\top (\phi_0 - \phi_*) \quad (3.13)$$

Denklem 3.13'de $\phi_0 = \mathbf{h}(\mathbf{d}(x_0))$ olmak üzere $\mathbf{J}_f = 2\mathbf{J}_h^\top (\phi_0 - \phi_*)$ 'in tanımlanması için zincir kuralı kullanılmıştır. Böylece ilk Newton adımı $\Delta \phi_0 = \phi_0 - \phi_*$ 'in $\mathbf{R}_0 = -2\mathbf{H}^{-1} \mathbf{J}$ matrisinin satır vektörlerine yansıtılmış hali olmaktadır. Bu aşamadan sonra $\mathbf{R}_0$ iniş yönü olarak kullanılacaktır. Yöntemde $\mathbf{R}_0$ eğitim kümesinden direkt olarak öğrenilmiştir. Böylece yöntem fonksiyonunun iki kez türevlenebilir olması koşulu ortadan kalkmıştır. Buna rağmen, test aşamasında ϕ_* bilinmemekte ancak eniyileme süresince sabit kalmaktadır. İniş yönünün test aşamasında kullanılabilmesi için, eğitimde ϕ_* bilgisinin kullanılması gerekmektedir. Bu sebeple, Denklem 3.13 ϕ_0 'in jenerik lineer birleşiminin sapma ($\mathbf{b}_0$) ifadesiyle toplamı şeklinde Denklem 3.14'de gösterildiği şekilde tekrar yazılmıştır.

$$\Delta x_1 = \mathbf{R}_0 \phi_0 + \mathbf{b}_0 \quad (3.14)$$

Böylece SDM eğitim kümesindeki örnekleri kullanarak $\mathbf{R}_0$ ve $\mathbf{b}_0$ değerlerini öğrenir. Ancak, f , x altında ikinci dereceden ise algoritma tek bir güncelleme adımında yakınsanabilir. Bu sebeple, ikinci dereceden olmayan fonksiyonların üstesinden gelebilmek için bir iniş yönü dizisi üretilmelidir. Denklem 3.15'de görüldüğü üzere $\phi_{k-1} = \mathbf{h}(\mathbf{d}(x_{k-1}))$ bir önceki nirengi noktası olan x_{k-1} 'den çıkarılan öznel vektörü olmak üzere bir imge için Newton'un metodu ile görüntüye özel gradyan yönlerinde bir güncelleme dizisi oluşturulmuştur.

$$x_k = x_{k-1} - 2\mathbf{H}^{-1} \mathbf{J}_h^\top (\phi_{k-1} - \phi_*) \quad (3.15)$$

SDM genel iniş yönlerini ($\mathbf{R}_k$) ve sapma ifadesini ($\mathbf{b}_k$) Denklem 3.16'daki fonksiyonu kullanarak eğitim kümesindeki imgelerin tamamı için birbirini izleyen x_k 'lar

x_* 'a yakınsayana kadar aramaya devam etmektedir.

$$x_k = x_{k-1} + \mathbf{R}_{k-1}\phi_{k-1} + \mathbf{b}_{k-1} \quad (3.16)$$

Bir sonraki adım sistemin $\mathbf{R}_k$ v e $\mathbf{b}_k$ p arametrelerinin eğitim kümesinden öğrenilme aşamasıdır. Verilen yüz imgeleri $\{\mathbf{d}^i\}$ ve elle işaretlenmiş nirengi noktaları $\{\mathbf{x}_*^i\}$ olmak üzere her imge için başlangıç değerleri x_0^i , $\mathbf{R}_0$ v e $\mathbf{b}_0$ k ullanılarak tahmin edilen ve en uygun nirengi noktası yerleşimi arasındaki beklenen kayıp minimize edilmektedir. Bu amaç için L2 kayıp fonksiyonu seçilmiş, $\Delta x^i = x_*^i - x_0^i$ ve $\dim_0^i = \mathbf{h}(\mathbf{d}^i(x_0^i))$ olmak üzere eniyileme yapılmıştır.

$$\arg \min_{\mathbf{R}_0, \mathbf{b}_0} \sum_{\mathbf{d}^i} \int p(\mathbf{x}_0^i) \|\Delta \mathbf{x}^i - \mathbf{R}_0 \phi_0^i - \mathbf{b}_0\|^2 d\mathbf{x}_0^i \quad (3.17)$$

Burada x_0^i normal dağılımlı olarak varsayılmış ve parametreleri yüz sezicinin varyansı olarak alınmıştır. Böylece integral Monte Carlo örneklemesiyle yaklaştırılmıştır.

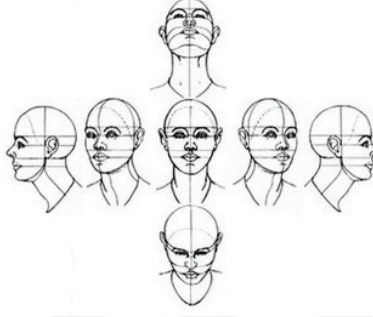
$$\arg \min_{\mathbf{R}_0, \mathbf{b}_0} \sum_{\mathbf{d}^i} \sum_{\mathbf{x}_0^i} \|\Delta x_*^i - \mathbf{R}_0 \phi_0^i - \mathbf{b}_0\|^2 \quad (3.18)$$

Denklem 3.18'den yola çıkarak her adımda Denklem 3.16'daki güncelleştirme önceden bulunan $\mathbf{R}_{k-1}$ ve $\mathbf{b}_{k-1}$ kullanılarak uygulanarak yeni bir veritabanı oluşturulur. Başka bir ifadeyle, öğrenilen $\mathbf{R}_{k-1}$ ve $\mathbf{b}_{k-1}$ kullanılarak şimdiki nirengi noktaları güncellenir. Ayrıca yeni bir en iyi parametre güncellemesi ve yeni öznitelik vektörü de bulunur. Böylece Denklem 3.19 kullanılarak yeni veritabanı için $\mathbf{R}_k$ ve $\mathbf{b}_k$ yeni doğrusal fonksiyon ile öğrenilir.

$$\arg \min_{\mathbf{R}_k, \mathbf{b}_k} \sum_{\mathbf{d}^i} \sum_{\mathbf{x}_k^i} \|\Delta x_*^{ki} - \mathbf{R}_k \phi_k^i - \mathbf{b}_k\|^2 \quad (3.19)$$

Kullanılan bu yöntemde daha önceden eğitilen modeller kullanılarak testler gerçekleştirilmiştir. Hız ve doğruluk açısından diğer yöntemlerden daha iyi olan INTRAFACE yöntemi kullanılarak alınan sonuçlar Şekil 6.4'de verilmektedir. Yöntem hız açısından oldukça başarılı olmasına rağmen profil yüzlerin nirengi noktalarının yerleşimi konusunda sorunlar bulunmaktadır. Daha önceden de bahsedildiği gibi nirengi noktalarının yerleşiminin yapılamaması ya da yanlış yapılması durumunda hizalama işlemi başarısız olmaktadır. Nirengi noktalarının yerleşimi konusunda yeterli başarıml elde edilemediğinden nirengi noktasından bağımsız olarak yüzün yalnızca konumu ve büyüklüğü hakkında bilgi veren daha basit ve hızlı bir yüz sezimi yöntemi kullanılmasına karar verilmiştir.

Bu sebeple Haar benzeri öznitelikleri kullanan sistemin çeşitli poz sınıflarıyla eğitilmesine karar verilmiştir. Yedi farklı poz sınıfı için oluşturulan pozitif ve negatif örnek kümeleri kullanılarak profil yüz imgelerinin sezimi için verimli hale gelecek bir yüz sezicisinin eğitilmesine karar verilmiştir. Bu eğitim için ayrılan yedi poz grubunun gösterimi Şekil 3.9’da gösterildiği gibidir.



Şekil 3.9 Eğitimde kullanılacak yedi poz grubu gösterimi

2. Aydınlatma Farklılıklarının Giderilmesi

Aydınlatma değişikliği etkilerini en aza indirmek için en sık uygulanan ön işleme yöntemi histogram eşitlemedir. Histogram eşitleme imgedeki gri seviye renk değerlerinin birbirine eşit piksel sayısına sahip bir histograma dönüştürme işlemidir. Bu yöntemde imgenin histogramının düz bir yapıya sahip olması hedeflenmektedir. Bu yöntem histogramı belli bir alanda toplanmış imgeler için daha iyi sonuçlar vermektedir.

Histogram eşitleme uyarlamalı olmayan ve uyarlamalı olmak üzere ikiye ayrılmaktadır. Uyarlamalı olmayan histogram eşitleme imgenin tümüne uygulanmaktadır. Uyarlamalı olmayan histogram eşitleme şu şekilde özetlenebilir: i . pikselin gri seviye değerinin bulunma olasılığı $p(i) = n_i/N$ olarak verildiğinde dönüştürülmüş yoğunluk değeri, i_{new} , Denklem 3.20’ de verilen şekilde hesaplanmaktadır.

$$i_{new} = \sum_{j=0}^i \frac{n_j}{N} = \sum_{j=0}^i p(j) \quad (3.20)$$

Burada n_i d eğer i . gri seviye değerindeki piksel sayısını belirtmektedir. Bu işlem piksellerin yoğunluk değerlerinin orijinal aralığından başka bir aralığa haritalanmasını sağlamaktadır. 8-bitlik aralıkta $[0, 255]$ bulunan bir imge $[0,1]$ uzayına haritalanmaktadır.

Uyarlamasız histogram eşitleme yöntemi uygulanmış örnek imgeler Şekil 3.10 'da verilmiştir.



Şekil 3.10 Uyarlamasız Histogram Eşitleme yönteminin uygulaması

Uyarlamalı histogram eşitleme yönteminde ise imge alt bölümlere ayrılarak histogram eşitleme işlemi bölgesel olarak yapılmaktadır. İmge ızgara şeklinde dikdörtgen bölgelere ayrılarak, her bir bölgeye Denklem 3.20'de verilen işlem uygulanır. İmge için kullanılacak bölge boyutları ve sayısı imgeye göre değişiklik göstermektedir. Bölgelere histogram eşitleme işlemi uygulanmasının ardından alt bölgeler bilineer interpolasyon yöntemi ile birleştirilerek iyileştirilmiş imge elde edilmektedir (Şekil 3.11).



Şekil 3.11 Uyarlamalı Histogram Eşitleme yönteminin uygulaması

En sık kullanılan yöntemlerden biri olan histogram eşitleme yönteminin yanı sıra yüz tanıma performansının artırılması için aydınlatma farklılıklarının etkisini ortadan kaldırmayı amaçlayan birçok yöntem mevcuttur. Bu yöntemlerden biri de tek ölçekli retineks algoritmasıdır (Single Scale Retinex (SSR)) (Hines vd., 2004). Bu yöntemde görüntünün yansıma ve parlaklık fonksiyonlarının özellikleri ile ilgili olarak ortak varsayımlar altında iyileştirme yapılmaktadır. İmgenin histogramının üst ve alt kısımlarını

kesilip geri kalan kısmını 8-bit aralığına yeniden ölçeklenmektedir. Yöntemin uygulanmış örnek imgeleri Şekil 3.12’de verilmiştir.



Şekil 3.12 SSR yönteminin uygulaması

Bir diğer yöntem ise SSR’ın genişletilmiş hali olan çok ölçekli retinex (Multi Scale Retinex (MSR)) algoritmasıdır (Jobson vd., 1997). Bu yöntem SSR yönteminde sorun oluşturan uygun Gauss ölçeğinin seçilmesi problemini çözerek birçok ölçekteki Gauss ile uygulanmış SSR çıktılarını birleştirmektedir. Yöntemin uygulanmış örnek imgeleri Şekil 3.13’de verilmiştir.



Şekil 3.13 MSR yönteminin uygulaması

Retineks teorisini kullanan bir diğer yöntem ise homomorfik filtrelemedir (Williams ve Madisetti, 1997). Diğer yöntemlere benzer şekilde imge yansıma ve parlaklık fonksiyonlarına ayrılarak filtrelenmektedir. Bu yöntemde normalizasyon işlemi frekans uzayında uygulanmaktadır. Bu uygulama için parametre ayarlaması yapılması gerektiğinde filtre tasarımının temel sorunu hala devam etmektedir. Yöntemin uygulanmış örnek imgeleri Şekil 3.14’de verilmiştir.



Şekil 3.14 Homomorfik filtreleme yönteminin uygulaması

Uyarlamalı yumuşatma ile tek ölçekli retinex tekniğine benzer şekilde, bir diğer fotometrik normalleşme tekniği de self-quotient (Single-scale Self Quotient Image (SSQ)) imgedir (Wang vd., 2004). Direkt olarak retineks teorisinden üretilmemiş olsa da, retineks tabanlı yöntemlerden özellikle çok ölçekli retineks ile bağlantılıdır. Bu yöntem sözde Lambertian modeli ve quotient imge konseptini kullanarak aydınlatma etkilerini en aza indirmeyi amaçlamaktadır. Tek ölçekli self-quotient imge yönteminin uygulanmış örnek imgeleri Şekil 3.15'de verilmiştir.



Şekil 3 .15 SSQ yönteminin uygulaması

Bu yöntem tek ölçekli olarak uygulanmakla beraber yazarlar tarafından yöntemin çok ölçekli versiyonu da önerilmiştir. Çok ölçekli self-quotient imge yönteminin (Multi-scale Self Quotient Image (MSQ)) (Chen vd., 2006) uygulanmış örnek imgeleri Şekil 3.16'da verilmiştir.

Bir diğer retineks tabanlı yöntem olan ayrık kosinüs dönüşümü (Discrete Cosine Transform (DCT)) tekniği aydınlatma farklılıklarını DCT'un düşük frekanslı katsayıları ile ilişkili olduğunu varsaymaktadır (Chen vd., 2006). Bu düşük frekanslı katsayıların atılarak görüntüye geri dönüşüm uygulandığında aydınlatmadan bağımsız yüz temsili elde edileceğini önermektedir. Yöntemin uygulanmış örnek imgeleri Şekil 3.17'de verilmiştir.



Şekil 3.16 MSQ yönteminin uygulaması



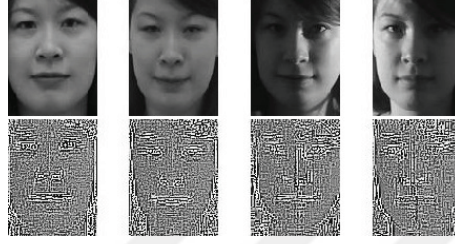
Şekil 3.17 DCT yönteminin uygulaması

Aydınlatma normalleştirme yöntemlerinden bir diğeri de dalgacık tabanlı normalleştirmedir (Du ve Ward, 2005). Bu yöntemde imge 2D ayrık dalgacık dönüşümüyle (Discrete Wavelet Transform (DWT)) dört alt-banta ayrılmaktadır. Bu bantlar yaklaşım katsayıları tarafından oluşturulan alçak-alçak bant ve ayrıntı katsayıları ile oluşturulan alçak-yüksek, yüksek-alçak ve yüksek-yüksek bantlardır. Bantlara ayırma işleminin ardından her bir banda histogram eşitleme uygulanarak imgenin kontrastı iyileştirilir. Ardından ayrıntı katsayıları 1'den büyük bir sayıyla çarpılarak imgenin kenarları vurgulanır. Tüm bantlar değişikliğe uğratıldıktan sonra tersine DWT uygulanarak normalleştirilmiş imge oluşturulur. Yöntemin uygulanmış örnek imgeleri Şekil 3.18'de verilmiştir.



Şekil 3.18 Dalgacık tabanlı normalleştirme yönteminin uygulaması

Dalgacık tabanlı yöntemlerden bir diğeri ise dalgacık tabanlı gürültü giderme (Wavelet-based Image Denoising (WD)) yöntemidir (Zhang vd., 2009 b). Görüntüden gürültü giderme modelinde, anahtar yüz özellikleri gürültü ile eşdeğer yüksek frekanslı olgulardır. Bu yöntemde WD yöntemi kullanılarak imgenin parlaklık bileşeni bulunmakta ve bu kullanılarak aydınlatmadan bağımsız yansıma bileşeni elde edilmektedir. Yöntemin uygulanmış örnek imgeleri Şekil 3.19’da verilmiştir.



Şekil 3.19 WD yönteminin uygulaması

Aydınlatma etkisinin azaltılması için kullanılabilecek bir diğ yöntem ise yerel olmayan ortalama temelli normalleşme tekniğidir. Non-local Means based Normalization (NLM) yönteminde parlaklık fonksiyonunu hesaplamak için yerel olmayan ortalamalar kullanılarak yansıma elde edilmektedir (Štrucve ve Pasevic, 2009). Buna ek olarak, NLM yönteminin uyarlamalı hali de yazarlar tarafından önerilmiştir. Yöntemin uygulanmış örnek imgeleri Şekil 3.20’de verilmiştir. Uyarlamalı yerel olmayan ortalama temelli normalleşme tekniğinde (Adaptive Non-local Means based Normalization (ANL)) ise parlaklık fonksiyonu yerel olmayan ortalamalar gürültü ayıklama yöntemiyle bulunmaktadır (Štrucve ve Pasevic, 2009). Buradaki uyarlamalı ifadesi ise imgenin yumuşatılması sırasında imgenin yerel kontrastının kullanılmasından gelmektedir. Yöntemin uygulanmış örnek imgeleri Şekil 3.21’de verilmiştir.



Şekil 3.20 NLM yönteminin uygulaması



Şekil 3.21 ANL yönteminin uygulaması

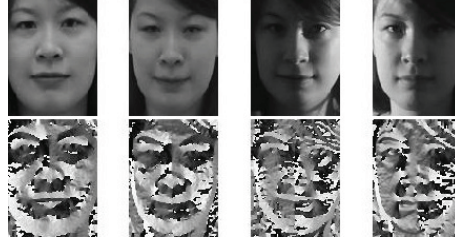
Yüz görüntülerinden aydınlatma kaynaklı görünüm varyasyonlarını kaldırmak için kullanılan bir diğer yöntem ise yönlendirilebilir filtreler (Steerable Filter (SF)) yöntemidir (Freeman ve Adelson, 1991). Bu yöntemde bir filtrenin küçük oryantasyon farklarıyla oluşturulmuş versiyonları imgeye uygulanmakta ve bunların yanıtları enterpolasyon yöntemi ile birleştirilmektedir. Bu filtrelerden kaç tane ve hangi oryantasyonlarda kullanılacağı yönlendirilebilir filtreler kavramı ile açıklanarak bir filtreler kümesi oluşturulmuştur. Yöntemin uygulanmış örnek imgeleri Şekil 3.22’de verilmiştir.



Şekil3.22 SF yönteminin uygulaması

Gradyan yüzler (GradientFaces (GF)) yöntemi ise yüz imgesindeki her bir pikselin imge gradyanlarının oryantasyonlarını kullanarak normalizasyon yapmayı hedeflemektedir (Zhang vd., 2009 a). Hesaplanan bu imge versiyonu aydınlatmadan bağımsız yeni imge olarak tanımlanmaktadır. Oluşturulan bu Gradyan yüz imgesinin aydınlatma değişikliklerine karşı gürbüz olduğu belirlenmiştir. Yöntemin uygulanmış örnek imgeleri Şekil 3.23’de verilmiştir.

Aydınlatma normalizasyonu için kullanılan bir diğer yöntem ise Gauss filtre farklarına dayalı Gauss farkı (Difference-of-Gaussian (DoG)) yöntemidir (Nilufar vd., 2012). Bu yöntemde imgeye bant geçiren filtre uygulanmakta ve imgenin aydınlatmadan



Şekil 3.23 GF yönteminin uygulaması

bağımsız normalleştirilmiş hali elde edilmektedir. Burada en önemli nokta imgeye filtreleme yapılmadan önce bir önileme yapılmasıdır. İmgeye gamma düzeltmesi ya da logaritmik dönüşüm uygulandığında beklenen sonuçların elde edilmesinin mümkün olduğu belirtilmektedir. Yöntemin uygulanmış örnek imgeleri Şekil 3.24’de verilmiştir.



Şekil 3.24 DoG yönteminin uygulaması

Normalizasyon yöntemlerinden bir diğeri ise Weber yüzleri (WeberFaces (WEB)) normalleştirme tekniğidir (Wang vd., 2011). Bu yöntemde işlevi değiştirilmiş Weber kontrastı içinde nispi gradyan hesapları yapılarak yüz imgesinin aydınlatmadan bağımsız gösterimi elde edilmektedir. Yöntemin uygulanmış örnek imgeleri Şekil 3.25’de verilmiştir. Bu yöntem tek ölçekli olarak uygulanmakla beraber bu yöntemin çok ölçekli versiyonu da yazarlar tarafından önerilmiştir. Çok ölçekli Weber yüzleri (MSW) normalleştirme yönteminde hesaplana nispi gradyanlar farklı komşuluklarla hesaplanmakta ve bunların lineer kombinasyonu aydınlatmadan bağımsız yüz imgesi olarak kullanılmaktadır (Wang vd., 2011). Yöntemin uygulanmış örnek imgeleri Şekil 3.26’da verilmiştir.

Aydınlatma değişikliklerinin normalleştirilmesi için kullanılan son yöntem ise büyük ve küçük ölçekli özellikler (The Large- and Small-scale Features (LSSF)) normalleştirme tekniğidir (Xie vd., 2011). Bu yöntemde yüz imgesi küçük ölçekli ve büyük ölçekli



Şekil 3.25 WEB yönteminin uygulaması



Şekil 3.26 MSW yönteminin uygulaması

özellikler açısından ikiye ayrılmaktadır. Ardından büyük ölçekli özelliklerin yer aldığı imgeye aydınlatma normalleştirmesi uygulanmaktadır. Küçük ölçekli özelliklerin yer aldığı imgeye ise yalnızca küçük bir düzeltme yapılmaktadır. Ardından bu iki imge birleştirilerek aydınlatma değişikliklerinde arınmış yeni imge elde edilmektedir. Yöntemin uygulanmış örnek imgeleri Şekil 3.27’de verilmiştir.



Şekil 3.27 LSSF yönteminin uygulaması

3.3 Hizalama için Poz Tahmini

Video imgelerinden sezilen yüz bölgelerinin belirlenen şablonlara hizalanabilmesi için öncelikle hangi poz sınıfına ait olduğu tespit edilmelidir. Poz tahmini için kullanılan yöntemin hızlı sonuçlar vermesi sistemin gerçek zamanlı çalışması açısından büyük öneme sahiptir. Bu sebeple eğitimsiz sistemlerin kullanımıyla poz tahminin yapılmasına karar verilmiştir. Böylece pozların yedi sınıfa ayırımı basit ve hızlı bir şekilde yapılabilmektedir.

3.3.1 k -ortalamalar gözetimsiz öğrenme yöntemi

MacQueen (1967) tarafından geliştirilen k -ortalamalar gözetimsiz öğrenme yöntemi verinin istenilen sayıda kümeye ayrıştırılmasına imkan vermektedir. Bu yöntemde her bir örnek yalnızca tek bir sınıfa dahil olabilmektedir. Bu da kümeleme işleminin daha katı olmasına sebep olmaktadır. Bu yöntemde eşit büyüklükteki kümeler elde edilmekte ve bu kümeyi temsil eden verinin kümenin merkezi olduğu kabul edilmektedir. k -ortalamalar kümeleme yönteminde kümeleme işlemi karesel hata kriteri (Sum of Squared Error) kullanılmaktadır. En düşük hataya sahip kümeleme sonucu kümeleme işleminin sonucu olarak alınmaktadır. Kümeleme işlemine k tane rastgele seçilen merkez ile başlanmaktadır. Bu merkezler $x \in C_i$ kümesinde yer alan bir örnek, m_i merkez noktası ve $dist$ Öklid uzaklık ölçütü olmak üzere Denklem 3.21'de verilen hata fonksiyonu minimize edilerek güncellenmektedir.

$$SSE = \sum_{i=1}^k \sum_{x \in C_i} dist^2(m_i, x) \quad (3.21)$$

Küme içi benzerliklerin en fazla, kümeler arasında benzerliğin en az olduğu durumda küme merkezleri söz konusu kümeyi temsil eden ölçüt olarak belirlenmektedir. Özet olarak k -ortalamalar yönteminin işlem adımları; (i) Küme sayısının ve merkezlerinin başlangıç değerlerinin rastgele belirlenmesi, (ii) Her bir değişken için belirlenen küme merkezlerinden olan mesafelerin hesaplanması, (iii) Her bir değişkenin en yakın küme merkezine göre bir kümeye atanması, (iv) Amaç fonksiyonunun en küçüklenmesi, (v) Küme merkezlerinin yeniden hesaplanması ve yeni kümelerin belirlenmesi, (vi) Küme üyeliklerinde herhangi bir değişim olmayana kadar analizin devam ettirilmesi şeklindedir.

k -ortalamalar yönteminin işlem yükü yönünden basit olması ve hızlı çalışması sebebiyle poz sınıflandırmasında kullanılmıştır. Poz sınıf sayısının şablon oluşturma işlemi için sabit tutulması sebebiyle eğitmenli yöntemlerin kullanılmasına ihtiyaç duyulmamıştır.

3.3.2 Temel bileşen analizi yönteminin kullanımı

Tek başına kullanıldığında k -ortalamalar yönteminin yüz imgelerinin poz sınıflarına ayırması konusunda yetersiz kalabileceği tarafımızca önceden yapılan çalışmalar göz önüne alınarak tespit edilmiştir. Bu sebeple veriyi doğrusal ilişkiler içeren çok sayıda veriden, birbirinden bağımsız az sayıda veriye dönüştüren temel bileşen analizi (Principle Component Analysis (PCA) (Jolliffe, 2002)) yöntemi incelenmiştir.

PCA veriyi çok daha az sayıda değişkenle en iyi şekilde ifade eden dönüşümü uygulamaktadır. Bu dönüşüm sonrasında elde edilen veri ilk veriye ait temel bileşenler olarak adlandırılmaktadır. Bu yöntemin en önemli uygulamalarında biri de eigenface yöntemidir. Bu yöntemde veri kümesi $\mathbf{X}$, M sayıdaki $N \times 1$ boyutunda vektörlerden oluştuğu varsayılarak Denklem 3.22 ve 3.23’de gösterilen şekilde düşünülmektedir.

$$\mathbf{X} = [\mathbf{x}^1 | \mathbf{x}^2 | \dots | \mathbf{x}^M] (N \times M) \quad (3.22)$$

$$\mathbf{X} = \begin{bmatrix} x_1^1 & x_1^2 & \dots & x_1^M \\ x_2^1 & x_2^2 & \dots & x_2^M \\ \vdots & \vdots & \ddots & \vdots \\ x_N^1 & x_N^2 & \dots & x_N^M \end{bmatrix} \quad (3.23)$$

Birden fazla değişkenin içeren veriler analiz edildiğinde değişkenlerin ölçütleri birbirinden farklılık gösterebilmektedir. Bu sebeple, öncelikle bu değişkenlerin standartlaştırılması gerekmektedir. Bu işlem verinin ortalaması $\mathbf{m}$ ’nin 0’a çekilmesiyle sağlanmaktadır. Ortalama bulma işlemi Denklem 3.24’de gösterildiği şekilde yapılmakta ve standartlaştırılmış veri $\mathbf{X}^T$ olmak üzere Denklem 3.25’de gösterilen şekilde değişkenlerin ölçütleri benzetilmektedir.

$$\mathbf{m} = \frac{1}{M} \sum_{i=1}^M \mathbf{x}^i \quad (3.24)$$

$$\mathbf{X}^T = \begin{bmatrix} x_1^1 - m_1 & x_1^2 - m_1 & \dots & x_1^M - m_1 \\ x_2^1 - m_2 & x_2^2 - m_2 & \dots & x_2^M - m_2 \\ \vdots & \vdots & \ddots & \vdots \\ x_N^1 - m_N & x_N^2 - m_N & \dots & x_N^M - m_N \end{bmatrix} \quad (3.25)$$

Değişkenlerin benzer ölçütlerle temsil edilmesinin ardından veri için kovaryans matrisi aşağıdaki şekilde hesaplanmaktadır.

$$\mathbf{C} = \mathbf{X}^T \mathbf{X} = \begin{bmatrix} (x_1^1 - m_1)^2 & (x_1^2 - m_1)(x_2^1 - m_2) & \dots & (x_M^1 - m_1)(x_N^1 - m_N) \\ (x_1^2 - m_2)(x_2^1 - m_1) & (x_2^2 - m_2)^2 & \dots & (x_M^2 - m_2)(x_N^2 - m_N) \\ \vdots & \vdots & \ddots & \vdots \\ (x_N^1 - m_N)(x_1^M - m_1) & (x_N^2 - m_N)(x_2^M - m_2) & \dots & (x_N^M - m_N)^2 \end{bmatrix}$$

Bu matris içinde köşegen üzerinde bulunan değerler verinin tek boyutta ortalama etrafındaki dağılımını göstermektedir. Bu değerler varyans olarak adlandırılmaktadır. Diğer değerler ise bir değişkenin diğer değişkenlere bağlı olarak nasıl değiştiğini göstermektedir. Bu değerler ise kovaryans olarak adlandırılmaktadır. Bu değişkenler arasında benzer bir ilişki varsa yani bir değişken artıyor ya da azalırken diğer değişken de aynı şekilde artıyor ya da azalıyorsa kovaryans değeri pozitif olmaktadır. Bunun tam tersi bir durumda ise kovaryans negatif olmaktadır. Değişkenler arasında herhangi bir ilişki bulunmadığı takdirde kovaryans sıfır olmaktadır.

Kovaryans matrisinin hesaplanmasının ardından bu matrise özdeğer-özvektör ayrıştırması yapılmaktadır. Kovaryans matrisi $\mathbf{C}$, $N \times N$ boyutunda bir matris, λ özdeğer, $\mathbf{v}$ sıfırdan farklı bir sütun vektörü olan özvektör olmak üzere Denklem 3.27 kullanılarak ayrıştırılır.

$$\mathbf{C}\mathbf{v} = \lambda\mathbf{v} \quad (3.27)$$

Özdeğerler büyükten küçüğe sıralandıktan sonra bu özdeğerlere karşılık gelen özvektörler kullanılarak veriyi temsil eden en iyi $\mathbf{W}$ izdüşüm matrisi Denklem 3.28'de verilen şekilde oluşturulur.

$$\mathbf{W} = [\mathbf{w}_1 \mathbf{w}_2 \dots \mathbf{w}_p] \quad (3.28)$$

Veri oluşturulan bu uzay üzerine haritalanarak veriye ait öznitelikler Denklem 3.29'da verilen şekilde belirlenmiş olur.

$$\mathbf{y}^i = \mathbf{W}^T \mathbf{x}^i, \quad i = 1, 2, \dots, M \quad (3.29)$$

Böylece veriye ait öznitelikler belirlenerek verinin boyutu indirgenmiş ve veri kolay ayırt edilebilir hale gelmiştir. Bu aşamanın ardından k -ortalamalar kümeleme yöntemi uygulandığında performansın geliştiği gözlenmiştir.

3.3.3 Negatif olmayan matris ayrışımı yönteminin kullanımı

Veriyi temsil etmek için kullanılan bir diğer yöntem ise Cai vd. (2008) tarafından önerilen Negatif olmayan Matris Ayrışımı yöntemidir. NMF çarpımları veriye en iyi şekilde yakınsayacak iki negatif olmayan matris bulmayı hedeflemektedir. Negatif olmama kısıtı verinin parça tabanlı temsil edilmesine sebep olmaktadır. Bu durum da yalnızca toplanabilir kombinasyonlara izin vermektedir. Bu sebeple NMF yöntemi Singular Value Decomposition (SVD) yönteminin gelişmiş hali olarak adlandırılmaktadır.

NMF yöntemi negatif elemana sahip olmayan matrislerin analizini sağlamaktadır. Bu yöntem sütunları birer örnek vektörü olmak üzere verilen bir veri matrisi $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n] \in \mathbf{R}^{M \times N}$ için $\mathbf{U} \in \mathbf{R}^{M \times K}$ ve $\mathbf{V} = [v_{jk}] \in \mathbf{R}^{N \times K}$ olmak üzere iki matris bulmayı hedeflemektedir. Bulunan bu iki matrisin çarpımlarının yaklaşık olarak verinin kendisi olması beklenmektedir (Denklem 3.30).

$$\mathbf{X} \approx \mathbf{UV}^\top \quad (3.30)$$

Bu maliyet fonksiyonunu eniyilemek için sık kullanılan iki yöntem mevcuttur. Bulardan ilki bu iki matris arasındaki Öklid uzaklıklarının karekökünün bulunmasıdır. Bu işlem Denklem 3.31 'de verilen şekilde gösterilmektedir.

$$O_1 = \|\mathbf{X} - \mathbf{UV}^\top\|^2 = \sum_{i,j} (x_{i,j} - \sum_{k=1}^K u_{ik}v_{jk})^2 \quad (3.31)$$

Bir diğer yöntem ise bu iki matris arasındaki ıraksamayı bulmaktır. Bu durumda maliyet fonksiyonu $\mathbf{Y} = [x_{ij}] = \mathbf{UV}^\top$ olmak üzere Denklem 3.32'de verilen şekilde tanımlanmaktadır.

$$O_2 = D(\mathbf{X}||\mathbf{UV}^\top) = \sum_{i,j} (x_{ij} \log \frac{x_{ij}}{y_{ij}} - x_{ij} + y_{ij}) \quad (3.32)$$

Her iki denklem için de maliyet fonksiyonları $\mathbf{U}$ ve $\mathbf{V}$ için aynı anda konveks değildir. Bu durumda bu fonksiyonlar için evrensel minimum bulunması mümkün değildir. Bu sebeple O_1 maliyet fonksiyonu Denklem 3.33'de verilen şekilde minimize edilmektedir.

$$u_{ik} = u_{ik} \frac{(\mathbf{XV})_{ik}}{(\mathbf{UV}^\top \mathbf{V})_{ik}}, \quad v_{jk} = v_{jk} \frac{(\mathbf{X}^\top \mathbf{U})_{jk}}{(\mathbf{VU}^\top \mathbf{U})_{jk}} \quad (3.33)$$

O_2 maliyet fonksiyonunun minimizasyon algoritması ise Denklem 3.34’de verilmiştir.

$$u_{ik} \leftarrow u_{ik} \frac{\sum_j (x_{ij} v_{jk} / \sum_k u_{ik} v_{jk})}{\sum_j v_{jk}}, v_{jk} \leftarrow v_{jk} \frac{\sum_i (x_{ij} u_{ik} / \sum_k u_{ik} v_{jk})}{\sum_i u_{ik}} \quad (3.34)$$

Gerçekte $K \ll M$ ve $K \ll N$ olduğundan NMF tekniği orijinal verinin sıkıştırılmış yakınsamasını bulmaktadır. Bu durumda bu yakınsama Denklem 3.35’de verilen şekilde gösterilmektedir.

$$x_j \approx \sum_{k=1}^K \mathbf{u}_k v_{jk} \quad (3.35)$$

Bu durumda x_j ’deki her bir veri vektörü $\mathbf{U}$ ’nun $\mathbf{V}$ ile ağırlıklandırılmış lineer kombinasyonları haline gelmektedir. Bir başka deyişle, $\mathbf{U}$, $\mathbf{X}$ verisinin temelini içeren bir lineer yakınsama olarak tanımlanabilmektedir. Bu yöntemde yer alan negatif olmama kısıtı diğer matris faktörizasyon yöntemlerine göre farklı temeller arasında ilave kombinasyonlar elde edilmesini sağlamaktadır.

3.3.4 Grafikte düzenlenmiş negatif olmayan matris ayrışımı yönteminin kullanımı

Negatif olmama kısıtıyla birlikte NMF parça tabanlı temsilleri öğrenmede oldukça başarılıdır. Ancak, bu yöntem bu öğrenme işleminin Öklid uzayında gerçekleştirmektedir. Bu durumda veri uzayının içsel geometrik yapısı ve ayırıcı yapısı göz ardı edilmektedir. Bu durum da gerçek dünya uygulamaları için önemli bir sorun teşkil etmektedir. Bu durumun çözümü için Grafikte Düzenlenmiş Negatif olmayan Matris Ayrışımı (Graph Regularized Nonnegative Matrix Factorization (GNMF)) yöntemi önerilmiştir (Cai vd., 2008). Bu yöntemde verinin manifold yapısı en yakın komşu grafiği oluşturularak kurulmaktadır. Oluşturulan grafiğin yapısının korunmasıyla, NMF algoritmasına üstünlük sağlamıştır. Yüz imgelerine ait poz bilgisinin elde edilebilmesi için tüm yöntemler denenmiş ve doğrulukları ilerleyen bölümlerde tablolar halinde verilmiştir.

Veriden seçilen iki nokta x_j ve x_l veri dağılımının içsel geometrisinde birbirine yakın noktalar ise bu noktaları temsil eden z_j ve z_l temsilleri de yeni bulunan temele (basis) göre birbirlerine yakın olacaktır. Bu durum literatürde yerel değişmezlik varsayımı olarak geçmekte ve birçok algoritmanın geliştirilmesinde rol oynamaktadır. Bu geometrik yapı en yakın komşu grafiği kullanılarak modellenenilmektedir. Bu iki nokta arasındaki yakınlık Denklem 3.36’da gösterilen şekilde tanımlanmaktadır.

$$\mathbf{W}_{jl} = \mathbf{x}_j^\top \mathbf{x}_l \quad (3.36)$$

3.4 Yüz Hizalama

Pozlar ayrıştırıldıktan sonra her bir poz kategorisi uygun şablona hizalanmalıdır. Hizalama işleminde yüz içerisindeki çeşitli nirengi noktaları referans alınarak şablonla eşleştirmede kullanılabileceği gibi imgeler birleştirilmiş vektörler yığını şeklinde de incelenebilir. Ayrıca SIFT öznitelikleri kullanılarak da hizalama yapılabilir.

3.4.1 Nirengi Noktaları Kullanılarak Hizalama

İmgelere uygulanan yüz sezimi algoritmasının çıktıları olan yüz imgelerinin hizalanmasında farklı yöntemler ön plana çıkmaktadır. İlk olarak yüz sezimi ile tespit edilen nirengi noktaları arasında bir afin dönüşüm hesaplanarak hizalama yapılmıştır. INTRAFACE yöntemi çıktıların en başarılı olduğu tespit edildiğinden bu yöntem sonucu dönen nirengi noktaları kullanılmıştır. Homojen koordinatlarda yazılmak koşuluyla referans imgeye ait nirengi noktaları $\bar{\mathbf{p}} = (x, y, 1)$, hedef imgeye ait nirengi noktaları $\bar{\mathbf{p}}' = (x', y', 1)$ şeklinde tanımlansın. Uygulanan afin dönüşüm Denklem 3.37'de gösterildiği şekilde gösterilmektedir.

$$\begin{pmatrix} x_1 & y_1 & 1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ x_N & y_N & 1 \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \end{pmatrix} = \begin{pmatrix} x'_1 \\ \cdot \\ \cdot \\ x'_N \end{pmatrix} \quad (3.37)$$

$$\begin{pmatrix} x_1 & y_1 & 1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ x_N & y_N & 1 \end{pmatrix} \begin{pmatrix} a_{21} \\ a_{22} \\ a_{23} \end{pmatrix} = \begin{pmatrix} y'_1 \\ \cdot \\ \cdot \\ y'_N \end{pmatrix} \quad (3.38)$$

Bu dönüşümün parametreleri, gözlemler olarak ele alınan noktaların eski koordinatları ile bilinmeyen seçilen yeni koordinatları arasında kurulan matematik ilişkilerden faydalanmak suretiyle en küçük kareler çözümü yapılarak hesaplanır. Hesaplanan dönüşüm hedef imgeye uygulanarak referans imgeyle olan poz farklılığının giderilmesi amaçlanmaktadır. Böylece yüz tanıma sisteminin karşılaştığı sınıf içi varyansın yüksek olma zorluğunun önüne geçilmesi hedeflenmiştir.

3.4.2 Robust Alignment by Sparse and Low-Rank decomposition (RASL) yöntemi ile hizalama

Yüz hizalama işlemi sırasında kullanılan bir diğer yöntem Learned-Miller'in katılma algoritmasını temel alan RASL yöntemidir (Peng vd., 2011). Bu yöntemde, imgeler birleştirilmiş vektörler yığını halinde işlenerek, piksel değerlerinin entropilerinin toplamı en aza indirilmektedir. Yığın, bir kişiye ait tüm imgelerin satır vektörleri şeklinde art arda eklenmesiyle oluşturulmaktadır. Oluşturulan yığının, her bir satırı yaklaşık olarak sabit olacak şekilde optimize edilmektedir. Bu bakış açısından yola çıkarak, mükemmel şekilde hizalanmış yığın matrisinin rankının '1' olması beklenmektedir. Ancak gerçek imgelerde özellikle aydınlatma ve poz farklılıklarından dolayı hizalanmış yığın matrisi, bilinmeyen bir ranka sahip olmaktadır ve rankın '1' olması mümkün olmamaktadır. Bu durumda problem, yığın matrisinin rankının en aza indirilmesi biçiminde ele alınmaktadır. $I_1^0, \dots, I_n^0$ doğrusal bağlantılı imgeler olması durumunda yığın matrisi Denklem 3.39'da gösterildiği gibi yazılabilmektedir.

$$A = [vec(I_1^0) | \dots | vec(I_n^0)] \quad (3.39)$$

İmgelerin kameraya göre pozları ve konumları bir kişiye ya da sahneye ait imgeleri birbirinden çok farklı hale getirebilmektedir. Bu yanlış hizalama bilgisayarla görü alanındaki bir çok uygulama için büyük sorunlar yaratmaktadır. Bu sebeple yanlış hizalamalar bir alan dönüşümü olarak tanımlanmıştır. Böylece I_1 ve I_2 yanlış hizalanmış imgeler olmak üzere bunlar arasındaki terslenebilir dönüşüm $\tau: \mathcal{R} \rightarrow \mathcal{R}^2$ de olmak üzere Denklem 3.40'da gösterildiği şekilde yazılabilir.

$$I_2(x, y) = (I_1 \circ \tau)(x, y) \doteq I_1(\tau(x, y)) \quad (3.40)$$

Denklem 3.39 ve 3.40'da verilen modellerin bir araya getirilmesiyle imgelerin yanlış hizalaması formüle edildiğinde Denklem 3.41'de verilen matris oluşmaktadır. Bu matriste $I_1, I_2, \dots, I_n$ aynı duruma ait yanlış hizalanmış imgeler ve $I_1 \circ \tau_1, I_2 \circ \tau_2, \dots, I_n \circ \tau_n$ ise dönüşüme uğratarak hizalanmış imgeler olmak üzere $I_j^0 = I_j \circ \tau_j$ olarak tanımlanabilmektedir.

$$D \circ \tau \doteq [vec(I_1^0) | \dots | vec(I_n^0)] \in \mathcal{R}^{m \times n} \quad (3.41)$$

Buradan yola çıkarak imge hizalama problemi Denklem 3.42'de verilen eniyileme problemi haline dönüşmektedir.

$$\min_{A, \tau} rank(A) \text{ s.t. } D \circ \tau = A \quad (3.42)$$

Hizalanmış imgeler kısmi kapanmalar sebebiyle kolayla bozulmaya uğrayabilmektedir. Bu sebeple imgedeki tüm piksellere oranla küçük bir parça olan bu

hataların tanımlanması gerekmektedir. Bu hata e_j olarak tanımlandığında, $\{J_j \circ \tau_j - e_j\}_{j=1}^n$ mükemmel şekilde hizalanmış ve bozulmalardan arındırılmış imgeler haline gelmektedir. Böylece formül $E = [vec(e_1)|...|vec(e_n)]$ olmak üzere Denklem 3.43 'de gösterildiği şekilde güncellenmektedir.

$$\min_{A,E,\tau} rank(A) s.t. D \circ \tau = A + E, ||E||_0 \leq k \quad (3.43)$$

Burada L0 normu E matrisi içindeki sıfır olmayan elemanları saymaktadır. k ise tüm imgelerdeki bozulmaya uğraması beklenen piksellerin maksimum değerini temsil etmektedir. Bu problemi Lagrange formda yazmak daha uygun olacağından formül Denklem 3.44'de verilen şekli almıştır.

$$\min_{A,E,\tau} rank(A) + \gamma ||E||_0 s.t. D \circ \tau = A + E \quad (3.44)$$

Burada verilen γ parametresi hatanın seyreklik karşı çözümün rankını değiştirmektedir. Denklem 3.41'de verilen eniyileme problemi direkt olarak çözülebilecek bir problem değildir. Bu sebeple yapılan bazı değişiklikler sonucu bu eniyileme problemi Denklem 3.45'de verildiği şekle dönüştürülmüştür.

$$\min_{A,E,\tau} ||A||_* + \lambda ||E||_1 s.t. D \circ \tau = A + E \quad (3.45)$$

Burada $||A||_*$, A matrisinin nükleer normunu; $||E||_1$, E matrisinin L1 normunu ve γ pozitif ağırlık faktörünü temsil etmektedir. Böylece minimum ranka ulaşan imge matrisi kullanılarak hizalanmış imgeler elde edilmektedir.

3.4.3 Scale Invariant Feature Transform (SIFT)-akış yöntemi ile hizalama

Yüz hizalama için kullanılan bir diğer yöntem SIFT-akış yöntemidir. Bu yöntemde optik akış algoritması temel alınarak piksel tabanlı bağımlılık yerine SIFT betimleyicilerin bağımlılığı üzerine çalışılmıştır. İmgede sık SIFT betimleyicileri çıkarılarak, bu betimleyicilerin akış vektörleri doğrultusunda eşleşmesi için bir amaç fonksiyonu tanımlanmaktadır ve problem amaç fonksiyonunun eniyilenmesi ile çözümlenmektedir. Amaç fonksiyonu optik akış algoritmasından farklı olarak çift katmanlı kanı yayılma algoritması ile eniyilenir. Bunun ardından SIFT-akışı için coarse-to-fine eşleme şeması, hızı ve eşleştirme sonucu iyileştirmek için kullanılmaktadır.

SIFT betimleyicileri imgeye ait yerel gradyan bilgisini tanımlayan ve öznitelik çıkarımı sağlayan bir ifadedir. Her bir piksel değeri 4×4 'lük komşuluklarıyla oluşturulan

hücrede 8 farklı yönelme için 128 boyutlu bir vektör oluşturularak bir piksel için SIFT temsili oluşturulmaktadır.

SIFT öznelikleri 4 adımda çıkarılmaktadır. Bunlar ölçek uzayı uç noktalarının bulunması, anahtar noktalarının belirlenmesi, yönelimlerin atanması ve anahtar nokta tanımlayıcılarının tespitidir. İlk adım olan ölçek uzayı uç noktalarının bulunması için ölçek ve yönelimden bağımsız DoG fonksiyonu kullanılmaktadır. DoG hesaplanması için öncelikle imge farklı ölçeklerdeki σ ölçek parametresi kullanılarak Gauss çekirdekleri Denklem 3.46 'daki şekilde hesaplanır.

$$G(x, y, \sigma) = \frac{1}{2\phi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (3.46)$$

Ardından Denklem 3.47 'de verilen şekilde DoG hesaplanması yapılmaktadır.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (3.47)$$

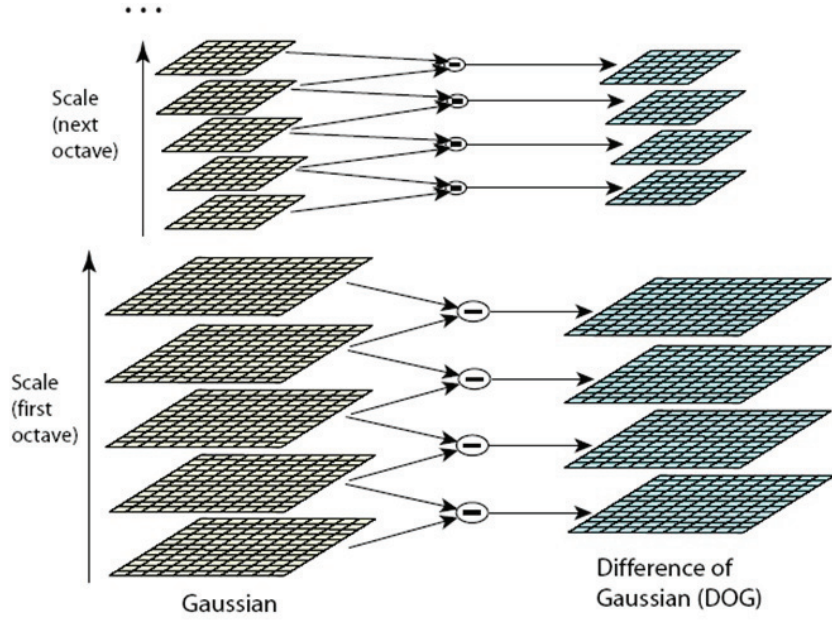
Denklem 3.47'de kullanılan k çarpım faktörü, $L(x, y,)$ imgenin ölçek uzayını tanımlamakta ve G ile giriş imgesi $I(x, y)$ 'nin konvolusyonu ile oluşmaktadır. DoG yapısı Şekil 3.29 'da gösterildiği şekilde oluşturulur. Girdi olarak alınan imge tekrarlamalı olarak farklı Gauss fonksiyonları ile konvole edilerek imgeye ait farklı oktavlarda ölçek uzayı oluşturulur. Ardından komşu Gauss imgeleri birbirlerinden çıkarılarak DoG imgeleri oluşturulur. Bu işlemin ardından tüm Gauss imgelerin örnek sayıları 2 faktörüyle azaltılarak işlem tekrarlanır.

Bu fonksiyon basit bir imge çıkarma işlemiyle hesaplanabilecek verimli bir fonksiyondur. Buna ek olarak DoG tarafından çalışılan normalleştirilmiş ölçeğe sahip Laplacian-of-Gaussian (LoG), $\sigma^2 \nabla^2 G$ fonksiyonuna benzer bir yaklaşım sunmaktadır.

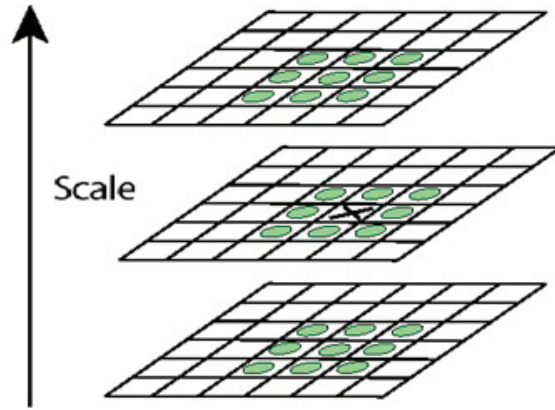
Doğru ölçek değişmezliğini sağlamak için Laplas fonksiyonunun σ^2 faktörüyle normalleştirilmesi gerekmektedir. $\sigma^2 \nabla^2 G'$ nun maksima ve minima değerlerinin imgeye

ait diğer fonksiyonlarla karşılaştırıldığında (gradyan, Hessian, Harris köşe bulma) en kararlı öznelikleri döndürdüğü tespit edilmiştir.

$D(x, y, \sigma)$ fonksiyonunun maksima ve minimasının bulunması için her örnek pikseli üst ve alt ortamlarda 9ar komşu, aynı ortamda 8 komşu olmak üzere 26 komşu pikseliyle Şekil 3.28'de verildiği şekilde karşılaştırılır. Bu karşılaştırma sonucunda bu piksel en büyük veya en küçük değere sahipse uçdeğer olarak atanır.



Şekil 3.28 Lowe'nin piramit şeması (Liu vd., 2011)



Şekil 3.29 Solda: Ölçek uzay imgeleri kümesi,sağda: DoG imgeleri (Liu vd., 2011)

Uçdeğerler genellikle piksellerin üzerinde değil de aralarında bulunurlar. Bu da bulunan uçdeğerlerin yaklaşık değerler olarak dönmesine sebep olmaktadır. Bu yaklaşık değerler gürültüye karşı dayanıksızdır ve nadiren köşe noktalarında konumlanmaktadır. Bu sebeple noktaların daha başarılı eşleşmeler yapması için geliştirilmesi gerekmektedir. Noktaların yaklaşık değerlerinin belirlenmesini ardından gerçek değerlerin bulunması adına ölçek uzayında enterpolasyon uygulanmaktadır. Bu sebeple bulunan noktaların gerçek konumlarını belirlemek amacıyla yerel örnek noktasına bir üç boyutlu, ikinci dereceden

fonksiyon uydurma yöntemi uygulanmaktadır. Bu yöntemle Taylor açılımı kullanarak $D(x, y, \sigma)$ fonksiyonunu Denklem 3.48’de verildiği şekilde yazılır.

$$D(x) = D + \frac{\partial D^\top}{\partial x} + \frac{1}{2}x^\top \frac{\partial^2 D}{\partial x^2} \quad (3.48)$$

Burada D ve D ’nin türevi yerel örnek noktası $x = (x, y, \sigma)^\top$ ’de değerlendirilmiştir. Bu eşitliğin x ’e göre türevi alınır ve sıfıra eşitlenerek, aranan uçdeğerler hesaplanabilir ve ara piksellerin gerçek konumları bulunabilir. Bu durumda fonksiyon Denklem 3.49’da verilen şekle dönüşür.

$$\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x} \quad (3.49)$$

Fonksiyonun uç noktadaki değeri, $D(\hat{x})$, kararsız uç noktadaki değerlerin elenmesi için kullanışlı bir yöntemdir. Böylece Denklem 3.49’un Denklem 3.48’de yerine konulmasıyla Denklem 3.50’deki şeklini alır.

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^\top}{\partial x} \hat{x} \quad (3.50)$$

İmge $[0, 1]$ aralığında kabul edildiği takdirde tüm $|D(\hat{x})| < 0,03$ değerindeki uçdeğerler yoksayılarak kararsız olan uçdeğerler elenmiş olunur.

Tam kararlılık için düşük kontrastlı uçdeğerlerin elenmesi yeterli değildir. DoG fonksiyonu köşeler boyunca güçlü cevaplar oluşturur ve bu da gürültüye karşı duyarlılığı artırır. Kenar üzerindeki noktalar birbirine benzerdir ve herhangi bir ayırt edicilik söz konusu değildir. Bu sebeple ayırt ediciliği yüksek olan köşe noktaları seçmek gerekmektedir. Bu seçimin yapılmasında Hessian matrisi kullanılmaktadır. DoG fonksiyonunda zor ayırt edilebilen bir zirve, kenar üzerinde yüksek temel eğrilik geçişine sahiptir ancak bu değer dikey yönde küçüktür. Temel eğriler, anahtar noktanın ölçek ve konumunda, Denklem 3.51’de verilen 2×2 ’lik Hessian matrisi, H , kullanılarak hesaplanabilir.

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (3.51)$$

H matrisinin özdeğerleri D ’nin temel eğrilik geçişine orantılıdır. Bu sebeple özdeğerlerin yerine yalnızca özdeğer oranları hesaplanabilir. Hessian matrisi simetrik matris olduğundan, matrisin izi ve determinantı α genliği büyük olan özdeğer ve β genliği küçük olan özdeğer Denklem 3.52’de verildiği şekilde hesaplanabilir.

$$\begin{aligned} Tr(\mathbf{H}) &= D_{xx} + D_{yy} = \alpha + \beta \\ Det(\mathbf{H}) &= D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \end{aligned} \quad (3.52)$$

Özdeğerlerin toplamını H 'nin izinden, çarpımını ise belirteninden Denklem 3.53'deki şekilde hesaplayabiliriz. Böylece ölçeği bilinen kararlı anahtar noktalar bulunmuş ve ölçekten bağımsızlık sağlanmıştır.

$$\frac{Tr(H)^2}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r\beta + \beta)^2}{r\beta^2} = \frac{(r + 1)^2}{r} \quad (3.53)$$

Anahtar noktalara yerel imge gradyan yönelimi ataması yapılarak döndürmeden bağımsızlık sağlanabilmektedir. Anahtar noktaların ölçeği yumuşatılmış Gauss imgesi L 'nin seçilmesi için kullanılmaktadır. Seçilen bu ölçekte her imge örneği $L(x, y)$, gradyan büyüklüğü $m(x, y)$ ve yönelim $\theta(x, y)$ olmak üzere piksel farklılıkları kullanılarak Denklem 3.54'de verilen şekilde yazılabilir. Anahtar nokta etrafındaki tüm piksellerin eğim yönelimleri kullanılarak, bir oryantasyon histogramı oluşturulur. Histogram 360 derecenin bölündüğü, 36 basamaktan oluşur. Her nokta, eğim genliğine göre ağırlıklandırılarak bu histograma eklenir.

$$m(x, y) = \sqrt{(L(x+1, y) + L(x-1, y))^2 + (L(x, y+1) + L(x, y-1))^2}$$

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (3.54)$$

Histogramdaki tepe noktaları yerel gradyanların baskın yönelimini vermektedir. Histogramdaki en büyük tepe noktası tespit edilir ve ardından bu tepe noktasının büyük bir bölümünü içeren diğer tepe noktaları da kullanılarak anahtar noktanın yönelimi tespit edilir. Bu sebepten dolayı aynı yerleşkeye ve ölçeğe sahip noktada farklı yönelimleri olan birden fazla anahtar noktası oluşabilir. Bu da eşleşmenin kararlılığını arttıran bir etkidir.

Bu tüm imgeye uygulandığında ise SIFT imgesi oluşturulmaktadır. Bu aşamadan sonra elimizde bulunan bu iki SIFT imgesinin arasındaki yoğun benzerliği bulma (dense correspondence) aşamasına geçilmektedir. Önerilen SIFT akış yönteminde optik akış algoritmasına benzer şekilde SIFT özniteliklerinin akış vektörleri doğrultusunda eşleşmesi sağlanmıştır. Böylece enerji fonksiyonu, $\mathbf{p} = (x, y)$ imgenin koordinatları, $\mathbf{w}(\mathbf{p}) = (u(\mathbf{p}), v(\mathbf{p}))$ $\mathbf{p}$ 'de tanımlana akış vektörleri, s_1 v e s_2 kullanılacak olan SIFT imgeleri ve ϵ konumsal komşuluk olmak üzere aşağıdaki Denklem 3.55'de gösterildiği şekilde tanımlanmaktadır.

$$E(w) = \sum_{\mathbf{p}} \min(\|s_1(p) - s_2(\mathbf{p} + \mathbf{w}(\mathbf{p}))\|_1, t) + \quad (3.55)$$

$$\sum_{\mathbf{p}} \eta(|u(\mathbf{p})| + |v(\mathbf{p})|) + \quad (3.56)$$

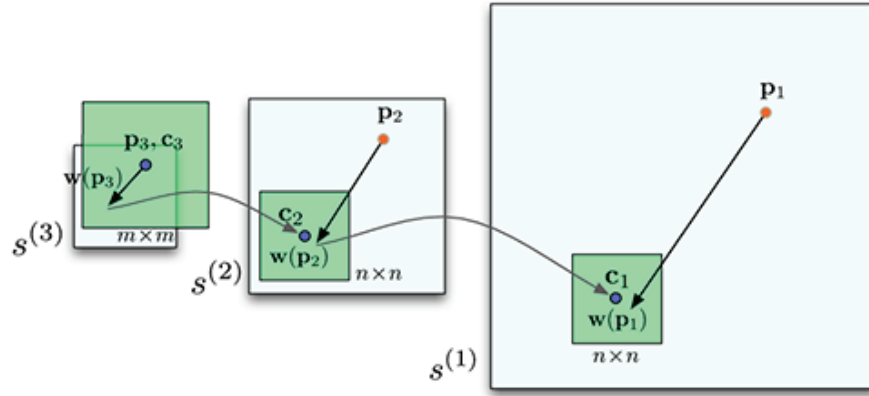
$$\sum_{(\mathbf{p}, \mathbf{q}) \in \epsilon} \min(\alpha|u(\mathbf{p}) - u(\mathbf{q})|, d) + \min(\alpha|v(\mathbf{p}) - v(\mathbf{q})|, d) \quad (3.57)$$

Bu formülasyon üç ayrı terimi içermektedir. Eşitlik 3.55 veri terimi olarak adlandırılmakta ve SIFT özniteliklerini $\mathbf{w}(\mathbf{p})$ a kış vektörü doğrultusunda eşleşmesini sağlamaktadır. Eşitlik 3.55 küçük yerdeğiştirme terimi olarak adlandırılmakta ve akış vektörlerinin olabildiğince küçük olması kısıtını koymaktadır. Eşitlik 3.56 ise akıcılık terimi olarak tanımlanmakta ve komşu piksellerin akış vektörlerinin benzer olması kısıtını koymaktadır. Tanımlanan bu amaç fonksiyonunda t ve d eşik değerleri olmak üzere veri ve akıcılık terimlerinde $L1$ normu kullanılmıştır.

Amaç fonksiyonu optik akış algoritmasından farklı olarak çift katmanlı kanı yayılma algoritması ile eniyilenmektedir. Bunun ardından SIFT-akışı için kabadan inceye eşleme şeması, hızı ve eşleştirme sonucu iyileştirmek için kullanılmaktadır. Kabadan inceye eşleşme şemasında temel amaç, imgenin koordinat sistemindeki kaba bir seviyedeki akışı tespit ettikten sonra bu akışı yavaş yavaş yayarak sonunda kabadan inceye bir akış tanımlamaktır. Bu işlem Şekil 3.30'da gösterilmiştir. Kabadan inceye eşleşme şeması olağan eşleşmeye göre önemli derecede hızlı bir eşleşme sağlamaktadır. Buna ek olarak başarımı çok daha düşük enerji seviyelerinde sağlamaktadır. Böylece kabadan inceye eşleşmenin eniyileme problemine çok daha iyi bir çözüm getirdiği söylenebilmektedir.

Büyük veritabanlarında çalışılması durumunda birbirinden çok farklı imgeler bulunması olası olduğundan benzer yerel yapılar gösteren imgelerin ayıklanması durumu ortaya çıkmaktadır. Bu sebeple sorgu imgesine yakın imgeler tanımlanmış ve N yakın komşusuyla arasında SIFT akış uygulanmıştır. Bu sebeple SIFT öznitelikleri kullanılarak konumsal histogram eşleme kullanılmıştır. Rastgele seçilen 5000 SIFT özniteliği üzerinde k -ortalama kümeleme algoritması çalıştırılarak 500 görsel kelime içeren bir bag-of-words modeli oluşturulmuştur. Oluşturulan kelimelerin histogramları iki seviyeli konumsal piramit kullanılarak tespit edilmiş ve iki imge arasındaki benzerliğin ölçülmesi için histogram çakışması uygulanmıştır.

SIFT-akış yöntemi hız ve hizalama açısından oldukça başarılı bir yöntemdir. Özellikle ortalama yüz imgeleri referans alındığında çok başarılı hizalamalar yaptığı gözlemlenmektedir. Ancak bu yöntemde referans imgeye göre dönüşüme uğratılmış imge



Şekil 3.30 Kabadan inceye SIFT akış eşleşmesinin gösterimi

piksel tabanlı dönüşüme tabi tutulduğundan birçok bozulmaya uğramaktadır. Pozların ayrı ayrı hizalanması işleminde dahi birçok deformasyon sebep olan yöntem yüz tanıma deneylerinde tanıma oranını azaltıcı yönde etki yapması beklenmektedir.

3.4.4 SIFT-akış yöntemi ile afin dönüşüm kullanılarak hizalama

İmgelerde oluşan deformasyon sorununun çözümü için SIFT-akış yönteminde önerilenden farklı olarak akış vektörleri afin dönüşüm hesaplamakta kullanılmıştır. Homojen koordinatlarda yazılmak koşuluyla referans imge $\mathbf{p} = (x, y, 1)$, hedef imge $\mathbf{p}' = (x', y', 1)$ şeklinde tanımlansın. Hedef imge, akış vektörleri kullanılarak $x' = x + u(\mathbf{p}), y' = y + v(\mathbf{p})$ şeklinde yazılabilir. Bu durumda imgelerdeki her bir pikselin yerleri kullanılarak daha önceden tanımlanan imgeye ait afin dönüşüm hesaplanmıştır. Tanımlanan bu denklem üstbelirlenmiş bir denklemdir ve çözümü bulunmamaktadır. Bu denklem daha genel bir biçimden Denklem 3.58 'de gösterilen şekilde tanımlanır.

$$y = \mathbf{H}x \quad (3.58)$$

Bu durumda x 'in çözümü hatanın enerjisinin (Denklem 3.59) minimize edilmesiyle elde edilmektedir.

$$J(x) = ||y - \mathbf{H}x||_2^2 \quad (3.59)$$

$J(x)$ 'in genişletilmesiyle Denklem 3.60, 3.61 ve 3.62 oluşmaktadır.

$$J(x) = (y - \mathbf{H}x)^\top (y - \mathbf{H}x) \quad (3.60)$$

$$= y^\top y - y^\top \mathbf{H}x - x^\top \mathbf{H}^\top y + x^\top \mathbf{H}^\top \mathbf{H}x \quad (3.61)$$

$$= y^\top y - 2y^\top \mathbf{H}x + x^\top \mathbf{H}^\top \mathbf{H}x \quad (3.62)$$

Bu denklemde görüldüğü üzere Denklem 3.61'deki terimleri tamamı skalerdir. Ayrıca $x^\top \mathbf{H}^\top y$ terimi de $y^\top \mathbf{H}x$ skaler teriminin transpozudur. Bu terimin türevinin alınmasıyla Denklem 3.63'ü oluşturmaktadır.

$$\frac{\partial}{\partial x} J(x) = -2\mathbf{H}^\top y + 2\mathbf{H}^\top \mathbf{H}x \quad (3.63)$$

Bu denklemin sıfıra eşitlenmesiyle Denklem 3.64 oluşmaktadır.

$$\frac{\partial}{\partial x} J(x) = 0 \Rightarrow \mathbf{H}^\top \mathbf{H}x = \mathbf{H}^\top y \quad (3.64)$$

Bu durumda $\mathbf{H}^\top \mathbf{H}$ terslenebilir olarak kabul edildiğinde çözüm Denklem 3.65'de verilen şekilde olmaktadır.

$$x = (\mathbf{H}^\top \mathbf{H})^{-1} \mathbf{H}^\top y \quad (3.65)$$

Bu durumda enküçük kareler çözümü Denklem 3.66'da gösterildiği şekilde bulunmaktadır.

$$\min_x \|y - \mathbf{H}x\|_2^2 \Rightarrow x = (\mathbf{H}^\top \mathbf{H})^{-1} \mathbf{H}^\top y \quad (3.66)$$

3.5 Hizalanmış İmgeler Kullanılarak Yüz Tanıma

Hizalama başarımlarının ölçülmesi için hizalanmış imgeler kullanılarak yüz tanıma performansları test edilmelidir. Bu sebeple, imge kümeleri tabanlı yüz tanıma sistemleri seçilmiştir. Çoğul imgeler söz konusu olduğunda, yüz tanıma problemi bilinmeyen bir kişiye ait çoğul imgelerden oluşan bir imge kümesini alıp, eğitim kümesindeki etiketli imge kümeleri arasında bu imge kümesine en çok benzeyeni bulma olarak tanımlanır. İmge kümeleri ile yüz tanımda, tekil soru imgeleri yerine aynı bireye ait birçok imge, imge kümesi olarak ele alınır; uygun metrik altında kümelerin modellenmesi ve kümeler arasındaki uzaklık ölçütlerine göre sorgulama yapılır.

Yüz tanımda imge kümelerinin kullanılmasının çeşitli faydaları vardır. Öncelikle imge kümelerini kullanan yüz tanıma yöntemlerinin tek bir imge kullanan yöntemlerden daha başarılı olması beklenir. Çünkü, imge kümeleri aynı kişiye ait yüz imgeleri arasındaki olası farklılık bilgilerini içermektedir ve bu bilgiyi kullanan yüz tanıma sisteminin güvenilirliği daha fazla olacaktır. Ayrıca, kontrollü ortamlarda çekilen imgelerle eğitilen ve tek bir test imgesi kullanan sistemlerde kişilerin bazı kurallara uyması beklenir (kameraya önden bakmak, belli bir uzaklıkta durmak ve çekim esnasında kıpırdamamak gibi). Oysa güvenlik amaçlı uygulamalarda kullanılan kameralardan kişilere herhangi bir yaptırım

olmadan imge kümeleri kolaylıkla elde edilir. Bu yüzden imge kümelerini kullanan yöntemlerle çalışmak daha doğal ve kolaydır.

İmge kümeleri tabanlı yüz tanıma sistemlerinden Cevikalp ve Triggs (2011) tarafından önerilen ve her bir imge kümesini bu kümesine ait örneklerin oluşturduğu basit geometrik konveks yapıları (affine ve konveks zarflar – affine/convex hulls) kullanarak modellenen imge kümesi tabanlı yüz tanıma sistemi kullanılmıştır. Bu yöntemde modellerin arasındaki benzerlikleri ölçmek için de bu geometrik modeller arasındaki uzaklıkları kullanılmaktadır. Doğrusal konveks modellerin uygun olmadığı durumlarda ise kernel hilesi kullanılmakta ve bu sayede doğrusal olmayan karmaşık imge manifoldlarının doğru bir şekilde modellenmesi sağlanmaktadır.

Yüz imge kümelerini oluşturan imgeleri $\mathbf{x}_{ci} \in R^d$ ile gösterelim. Burada $c = 1, 2, \dots, C$ imge kümelerini (dolayısıyla kişileri) $i = 1, 2, \dots, n_c$ ise c . imge kümesindeki örnek sayısını belirtmektedir. Önerilen yöntemde her bir imge kümesi konveks bir modelle (afin ve konveks zarflar) yakınsamıştır. Tanıma sırasında bilinmeyen kümeyi sınıflandırmak için bu konveks modeller arasındaki uzaklıklara bakılır ve test imge kümesi en yakın uzaklığı veren imge kümesine ait kişiye atanır. Konveks kümeler arasındaki uzaklığı bulmak için kullanılan strateji aşağıdaki şekilde özetlenebilir:

H_i ve H_j doğrusal olarak ayrıştırılabilen iki konveks model olsun. Bu iki model arasındaki uzaklık, bu modeller üzerindeki herhangi iki nokta arasındaki uzaklıkların en küçüğü ile ifade edilir. Bu uzaklık ölçümü Denklem 3.67’de verildiği şekildedir.

$$D(H_i, H_j) = \min_{\mathbf{x} \in H_i, \mathbf{y} \in H_j} \|\mathbf{x} - \mathbf{y}\| \quad (3.67)$$

3.5.1 Afın zarf modelleme

İlk olarak imge kümelerinin bu kümeleri oluşturan örneklerin oluşturduğu afin zarflarla modellendiğini varsayalım. Afın zarflar imge örneklerinin oluşturduğu en küçük afin altuzay olup, Denklem 3.68’de gösterilen formülle bulunurlar.

$$H_c^{aff} = \left\{ \mathbf{x} = \sum_{k=1}^{n_c} \alpha_{ck} \mathbf{x}_{ck} \mid \sum_{k=1}^{n_c} \alpha_{ck} = 1 \right\}, c = 1, 2, \dots, C \quad (3.68)$$

Normal altuzayların aksine afin zarfların orijinden geçme şartı olmadığından, afin zarflar yüz örneklerini içine alacak şekilde kaydırılmış altuzaylar olarak görülebilirler. Yukarıdaki formülden de görüldüğü üzere afin zarflar örneklerin afin kombinasyonlarını kullanarak sonsuz sayıda yeni örnekler oluşturmakta ve orijinal örneklerin altuzaydaki

konumlarını kısıtlamamaktadırlar. Bu yüzden imge kümelerini yakınsamak için oldukça gevşek sayılabilecek modellerdir.

Doğrusal olarak ayrıştırılabilen durumlarda affine zarflar üzerinden herhangi bir referans noktası μ_c seçip, mesela affine zarfı oluşturan örneklerden herhangi biri yada onların ortalaması gibi, affine zarfları aşağıdaki gibi yazmak mümkündür.

$$H_c^{aff} = \{\mathbf{x} = \mu_c + \mathbf{U}_c \mathbf{v}_c | \mathbf{c}_c \in \mathcal{R}^l\}, c = 1, 2, \dots, C \quad (3.69)$$

Denklem 3.69'da verilen $\mathbf{U}_c$ afin altuzayı geren birimdik vektörlerinin oluşturduğu matrisi, $\mathbf{v}_c$ ise afin zarf içindeki örneklerin birimdik vektörlere göre koordinatlarını yazmak için gerekli katsayıları içeren vektörü ifade etmektedir. Numerik olarak $\mathbf{U}_c$ matrisi $\{\mathbf{x}_{c1} - \mu_c, \dots, \mathbf{x}_{cnc} - \mu_c\}$ fark vektörlerine tekil değer ayrıştırması (Singular Value Decomposition) uygulanarak bulunur. Genellikle sıfıra yakın çok küçük özdeğerlere karşılık gelen özvektörler verilerdeki gürültüye karşılık geldiğinden birimdik vektörlere dahil edilmez. Elimizde doğrusal olarak ayrıştırılabilen iki afin zarf olduğunu ($\{\mathbf{U}_i \mathbf{v}_i + \mu_i\}$ ve $\{\mathbf{U}_j \mathbf{v}_j + \mu_j\}$) varsayalım. Bu durumda bu afin zarfların üzerilerindeki en yakın iki nokta Denklem 3.70'de verilen eniyileme problemi çözülerek bulunur.

$$\min_{\mathbf{v}_i, \mathbf{v}_j} \|(\mathbf{U}_i \mathbf{v}_i + \mu_i) - (\mathbf{U}_j \mathbf{v}_j + \mu_j)\|^2 \quad (3.70)$$

Denklem 3.70'deki eniyileme probleminde $\mathbf{U} = (\mathbf{U}_i \quad -\mathbf{U}_j)$ ve $\mathbf{v} = (\mathbf{v}_i \quad \mathbf{v}_j)^\top$ olarak tanımlanırsa, Denklem 3.71'deki eniyileme problemi elde edilir.

$$\min_{\mathbf{v}_j} \|\mathbf{U} \mathbf{v} + (\mu_i - \mu_j)\|^2 \quad (3.71)$$

Bu eniyileme probleminin eniyileyen $\mathbf{v}$ vektörü $\mathbf{v} = \mathbf{U}(\mathbf{U}^\top \mathbf{U})^{-1} \mathbf{U}^\top$ olacaktır. Bu durumda iki afin zarf arasındaki uzaklık Denklem 3.72 'de verilen formülle bulunur.

$$D(H_i^{aff}, H_j^{aff}) = \|(\mathbf{I} - \mathbf{P})(\mu_i - \mu_j)\| \quad (3.72)$$

Burada $\mathbf{P}$ matrisi her iki afin zarfı geren uzayın birimdik izdüşüm matrisini ifade eder ve nümerik olarak $\mathbf{P} = \mathbf{U}(\mathbf{U}^\top \mathbf{U})^{-1} \mathbf{U}^\top$ eşitliğiyle bulunur. $(\mathbf{I} - \mathbf{P})$ matrisi ise bu uzayın sıfır altuzayının birimdik izdüşüm matrisidir.

3.5.2 İndirgenmiş afin zarf modelleme

Afin zarfların doğrusal olarak ayrıştırılamadığı durumlarda (afin zarfların birbirleriyle kesiştiği durumlar), iki afin zarf arasındaki uzaklık her zaman sıfır olacağından

bir önceki bölümde verilen formüller işe yaramayacaktır. Afin zarflar birkaç sebepten dolayı kesişebilirler: İlk olarak imge kümelerini oluşturan imgeler arasında aykırı örnekler olabilir (yanlış yada kötü çekilmiş imgeler aykırı örnek olarak görülebilirler). Bu durumda aykırı değerlerin olduğu imge kümelerine karşılık gelen affine zarflar çok büyük olacaktır. Bu problemi çözmek için daha gürbüz afin zarf modelleme yöntemleri kullanılmalı veya aykırı değerler tespit edilip imge kümelerinden atılmalıdır. İkinci olarak yüz imgelerinden elde edilen öznitelikler kötü seçilmişse, bu öznitelikler imgeleri ayrıştırmak için yeterli olmayabilirler. Bu durumda ya daha iyi öznitelikler çıkarma yoluna gidilmeli yada kernel hilesi kullanılarak veriler doğrusal olarak ayrıştırılabilecek bir uzaya çıkarılmalıdır. Son olarak başta söylediğimiz gibi afin zarflar imge kümelerinin giriş uzayında gömülü olduğu bölgelerin kestirimi için oldukça gevşek modellerdir ve imge kümelerinin kesiştiği yada öznitelik uzayında birbirine çok yakın olduğu durumlarda bunlara karşılık gelen afin zarflarda kesişebilir.

Afin zarfları daha sıkılaştırmak için Denklem 3.72’de verilen α_{ck} katsayıları sınırlandırılır ve indirgenmiş afin zarflar Denklem 3.73’de verildiği şekilde elde edilir.

$$H_c = \left\{ \mathbf{x} = \sum_{k=1}^{n_c} \alpha_{ck} \mathbf{x}_{ck} \mid \sum_{k=1}^{n_c} \alpha_{ck} = 1, K \leq \alpha_{ck} \leq U \right\} \quad (3.73)$$

Bu eşitlikte (L, U) en alt ve en üst sınırları belirleyen parametrelerdir. Gevşek afin zarflarda bu sınırlar aktif değildir ((L, U) = (-∞, ∞)). Eğer L = 0 olarak seçilirse indirgenmiş afin zarflar konveks zarflara karşılık gelir. Benzer şekilde alt sınır $-\infty < L < 0$ ve üst sınır $U = \infty$ olarak seçilirse, buna karşılık gelen indirgenmiş afin zarf konveks zarflardan daha geniş olmasına rağmen normal afin zarflara oranla daha küçüktür. İndirgenmiş afin zarflar içindeki noktalar $\mathbf{x} = \mathbf{X}_c \alpha_c$ olarak gösterilebilir. Burada $\mathbf{X}_c$, sütun vektörleri imge kümesindeki yüz imgelerine karşılık gelen özellik vektörlerinin içeren matrisi, α_c ise α_{ck} katsayılarını içeren vektörü ifade etmektedir. İndirgenmiş afin zarfların üstündeki en yakın iki nokta Denklem 3.74’deki karesel eniyileme problemi çözülerek bulunabilir.

$$(\alpha_i^*, \alpha_j^*) = \arg \min_{\alpha_i, \alpha_j} \|\mathbf{X}_i \alpha_i - \mathbf{X}_j \alpha_j\|^2 \quad (3.74)$$

Yukarıdaki problem konveks bir problem olup, evrensel uygun çözümü vardır. Uygun α_i^*, α_j^* vektörlerinin bulunmasının ardından, indirgenmiş afin zarflar arasındaki uzaklık $D(H_i, H_j) = \|\mathbf{X}_i \alpha_i^* - \mathbf{X}_j \alpha_j^*\|$ formülü ile bulunur.

Denklem 3.52’de verilen optimizasyon problemi imge kümelerini oluşturan yüz örneklerinin iç çarpımları şeklinde yazılabilir. Bu durum bize, afin zarfların doğrusal olarak

ayrıştırılmadığı durumlarda kernel hilesinin kullanılmasına olanak verir. Kernel hilesini kullanarak, yani örnekler arası iç çarpımları $\langle \mathbf{x}_{ck}, \mathbf{x}_{c'l} \rangle$ kernel fonksiyonları $k(\mathbf{x}_{ck}, \mathbf{x}_{c'l}) = \langle \phi(\mathbf{x}_{ck}), \phi(\mathbf{x}_{c'l}) \rangle$ ile değiştirerek, örnekler giriş uzayı $\mathcal{R}^d$ 'den $\phi : \mathcal{R}^d \rightarrow \tau$ fonksiyonu kullanılarak çok daha büyük boyutlu bir uzaya aktarılmış olurlar. Böylece doğrusal olmayan karmaşık topolojilere sahip imge kümeleri arasındaki uzaklıklar daha doğru şekilde yakınsanabilir.

3.5.3 Konveks zarf modelleme

Afin zarflara alternatif olarak imge kümelerini modellemek amacıyla konveks zarflar kullanılabilir. Konveks zarflar imge örneklerinin konveks bileşimlerinden meydana gelir ve afin zarflardan tek farkları bileşen katsayılarının negatif olmamasıdır. Bir başka deyişle indirgenmiş afin zarflarda $L = 0$ alınır konveks zarflar elde edilmiş olur. Bu formülasyon Denklem 3.75 'de verildiği şekildedir. Konveks zarflar imge kümelerinin gömülü olduğu en küçük konveks yapılar olduklarından oldukça sıkı sayılabilecek yapılardır.

$$H_c^{con} = \left\{ \mathbf{x} = \sum_{k=1}^{n_c} \alpha_{ck} \mathbf{x}_{ck} \mid \sum_{k=1}^{n_c} \alpha_{ck} = 1, \alpha_{ck} \geq 0 \right\}, c = 1, .2, \dots, C \quad (3.75)$$

Denklem 3.52'de verilen optimizasyon probleminde $L = 0$ ve $U = 1$ olarak seçildiğinde iki konveks zarf üzerindeki en yakın noktalar belirlenerek iki konveks zarf arasındaki uzaklık bulunmuş olur. İki konveks zarf üzerindeki en yakın noktaları bulma problemi, hard-margin DVM (Destek Vektör Makineleri) sınıflandırıcılarında sınıfları en iyi ayıran hiper-düzlemi bulma problemi ile aynıdır. Bu sebeple konveks zarflar arasındaki uzaklıkları bulmak için DVM algoritması da kullanılabilir. DVM algoritması kullanıldığında iki konveks zarf arasındaki uzaklık $2/\|\mathbf{w}\|$ olacaktır; burda $\mathbf{w}$ iki imge kümesini en iyi şekilde ayıran hiper-düzlemin normal vektörüdür.

Aykırı değerlerin etkisini azaltmak için indirgenmiş afin zarflarda olduğu gibi, üst sınır birden küçük bir sayıya atanır. Kernel hilesi indirgenmiş afin zarflarda olduğu gibi konveks zarflarda da kullanılabilir ve bu şekilde doğrusal olmayan topolojiye sahip imge kümeleri arasındaki uzaklıklar daha doğru bir biçimde yakınsanabilir.

3.5.4 İndirgenmiş konveks zarf modelleme (İKZM)

İmge kümeleri, kernel hilesi içeren SVDD (Support Vector Data Description) yöntemi kullanarak modellenenebilir. Eğitim verileri için her bir imge kümesini sınırlandıran konveks bölgeler belirlendikten sonra az sayıda elemanların (yalnızca destek vektörlerin) kullanılması, imge kümelerini modellemede yeterlidir. Bu durum konunun giriş kısmında

belirtilen veri depolama ve yüksek hafıza gereksinimi problemini önemli ölçüde çözer çünkü sınıfların modelini bilmek için az sayıda veri örneği kullanmak yeterlidir.

SVDD yönteminin Gauss kernel ile birlikte kullanıldığında 2 parametresi vardır; tavan parametresi ve Gauss kernel genişliği. Bu parametreler arasında Gauss kernel genişliği parametresi, veri sınırlarını belirlemede daha etkindir. Gauss kernel genişliği küçüldükçe daha fazla destek vektör döndürmektedir. Sınıfları modellemede tüm verileri kullanmak yerine yalnızca destek vektörleri kullanmak yeterlidir. Sınıf sınırlarını belirleyecek olan destek vektör sayısı da tavan ve kernel parametrelerinin ayarlanmasıyla belirlenebilir.

Küçük hacimli veritabanı üzerindeki deney sonuçlarında oldukça yüksek tanıma oranları elde edilirken, veritabanı hacmi büyütüldüğünde tanıma oranlarının küme tabanlı tüm yöntemlerde düşmesi beklenmektedir. Ayrıca, küçük hacimli veritabanında doğrusal yöntemler ile kernel yöntemler birbirlerine yakın tanıma oranları verirken, büyük veritabanı için tanıma oranlarındaki farklar büyümüştür. Bunun en önemli sebebi, az sınıf sayısı olması durumunda afin uzaylar veya konveks zarflarla modellenen imge kümelerinin oluşturduğu bölgelerin birbirlerine fazla karışmamasıdır. Sınıf sayısı çoğaldıkça, sınıfı modelleyen imge kümelerinin kapsadığı bölgeler, örnek olarak temsil edildiği üzere iç içe geçmekte (kesişmekte) ve bu kümelerin ayrıştırılması zorlaşmaktadır.

Geniş hacimli veritabanları için öngörülen bir diğer problem de sorgu kümesi sınıflandırma hızlarıdır. Sınıflandırma için geçen süre imge kümesinin büyüklüğüne bağlıdır, büyük hacimli kümelerin sınıflandırılması daha uzun sürmektedir. İKZM yöntemi sınıflandırma hızları açısından kazanımlar getirmesi beklenmektedir. Veri örneklerinin sayılarının azaltılması sadece depolama ile ilgili problemleri çözmenin yanı sıra, daha az örnek kullanılması yüzünden uzaklık hesabındaki minimizasyon probleminin de daha kısa sürede çözülmesini sağlayarak sınıflandırma sürelerini de kısaltması beklenmektedir.

3.5.5 Karar ağaçları kullanarak imge kümesi tabanlı tanıma yöntemlerinde hız geliştirmesi

İmge kümelerinin temsil eden temel veri örneklerinin bulunmasının ardından test kümesi ile galeride yer alan tüm imge kümeleri arasındaki uzaklıklara bakılarak tanıma gerçekleştirilmektedir. Bu uzaklık hesaplama işlemi galeri kümeleri ile test kümesi arasındaki uzaklıklar bağımsız olduğundan paralel olarak da gerçekleştirilebilir. Birebir karşılaştırma söz konusu olduğunda test kümesi ile galeri kümelerinin bir kısmının karşılaştırılması gereksiz olmaktadır. Örneğin test kümesi bir erkek katılımcıya aitken kadın

katılımcılara ait kümelerle karşılaştırma gereksizdir. Buna benzer şekilde obje sınıflandırma uygulamaları için köpek sınıfına ait bir imge kümesi ile tamamen farklı sınıflar olan araba, uçak vb. sınıflarla karşılaştırma gereksiz olmaktadır. Bu anlamsız karşılatırmaları engellemek ve test süresini kısaltmak için ikili hiyerarşik karar ağaçları (Binary Hierarchical Descision Tree (BHDT)) kullanılabilir. Bu yöntem galerideki imge kümelerini her grupta birer imge kümesi kalacak şekilde iki gruba ayırır. Bu da karşılatırma sayısını C' den $\log_2 C'$ ye düşürmektedir. Böylece test süresinde belirgin bir azalma meydana gelmektedir (C' - veritabanındaki katılımcı sayısı).

Bu sistem içerisinde hem test hızı hem de başarımlar oluşturulan ağacın yapısına bağlıdır. Her bir düğümde yer alan imge kümelerinin dengeli olarak ayrıldığı bir ağaç yapısının oluşturulması gerekmektedir. Bu sebeple imge kümelerinin ikiye ayırmak için Shi ve Malik, 2000 tarafından önerilen Normalized Cuts (NCuts) yöntemi kullanılmıştır. Bu kümeleme yöntemi veri örneklerinin sonsuz boyuttaki uzaya haritalayarak veriyi sonsuz boyuttaki uzayda en büyük uzaklıkta ayıracak bir hiperdüzlem kullanarak veriyi ikiye böler. Bir başka deyişle, veriyi arasındaki marjin en büyük olacak şekilde dengeli iki gruba ayırır. Bu sistem imge kümeleri kullanılarak uygulandığında ayrılması gereken imge örnekleri yerine imge kümeleri haline gelmiştir. Bu durumda veri örnekleri yerine veri kümeleri konulması gerekmektedir. Bu sebeple benzerlik matrisini oluşturmak için konveks küme zarf uzaklıkları kullanılmıştır. Bu durumda benzerlik matrisi $\mathbf{W}$ 'nin kenarları olan w_{ij} Denklem 3.76'de verildiği şekilde hesaplanmaktadır.

$$w_{ij} = \begin{cases} \exp(-d(H_i^{convex})/t), & i \neq j \\ 0, & \text{aksihalde} \end{cases} \quad (3.76)$$

Bu denklemde verilen t değeri Gauss kernel fonksiyonunu genişliğidir ve kullanıcı tarafından ayarlanmaktadır. Bunun ardından imge kümelerini iki gruba ayırmak için Denklem 3.77'de verilen genel özdeğer problemi çözülmüştür.

$$\mathbf{L} = \lambda \mathbf{D} \mathbf{a} \quad (3.77)$$

Denklem 3.72'de bahsedilen $\mathbf{L} = \mathbf{D} - \mathbf{W}$ Laplas matrisi ve $\mathbf{D}$ ise $\mathbf{W}$ matrisinin satırlarının toplamını içeren köşegen matrisidir. Bu matrislerin boyutu imge kümelerindeki tüm imgelerin oluşturduğu $n \times n$ 'lik matrise göre çok daha küçük boyutlara sahip $C \times C$ boyutunda bir matristir. Son olarak Denklem 3.72'in en küçük ikinci özdeğerine ($\mathbf{a}^*$) karşılık gelen özvektör kullanılarak Denklem 3.78 'de verilen şekilde imge kümeleri ikiye kümeye ayrılır.

$$\begin{cases} y_i = -1, & \text{ise } \mathbf{a}^*_i \geq 0 \\ y_i = 1, & \text{ise } \mathbf{a}^*_i < 0 \end{cases} \quad (3.78)$$

İmge kümeleri her düğümden ikiye ayrıldıktan sonra test kümesinin sınıflandırılması için iki yöntem vardır. İlk yöntemde test kümesi ilke bölütlenmiş eğitim kümeleri arasındaki konveks zarf uzaklıklarına bakılmaktadır. Ardından test kümesi en yakın düğüme atanarak ait olduğu sınıf tespit edilir. Bir diğer yöntem ise eğitim aşamasında her bir düğümden lineer olmayan karar sınırlarını belirlemektir. Bu yöntemde test kümesine en yakın düğüm oy çokluğu ile karar alma ilkesiyle bulunmaktadır. Bu prosedür yalnızca kernel fonksiyonu değerlendirmesine ihtiyaç duyduğu için ilk yöntemle göre çok daha hızlı sonuçlar dönmektedir. Bu sebeple ikinci yöntemin kullanılması çok daha uygundur. Her iki yöntem de karşılaştırma sayılarını azaltarak büyük ölçekli verilerdeki hız problemine çözüm getirmektedir.



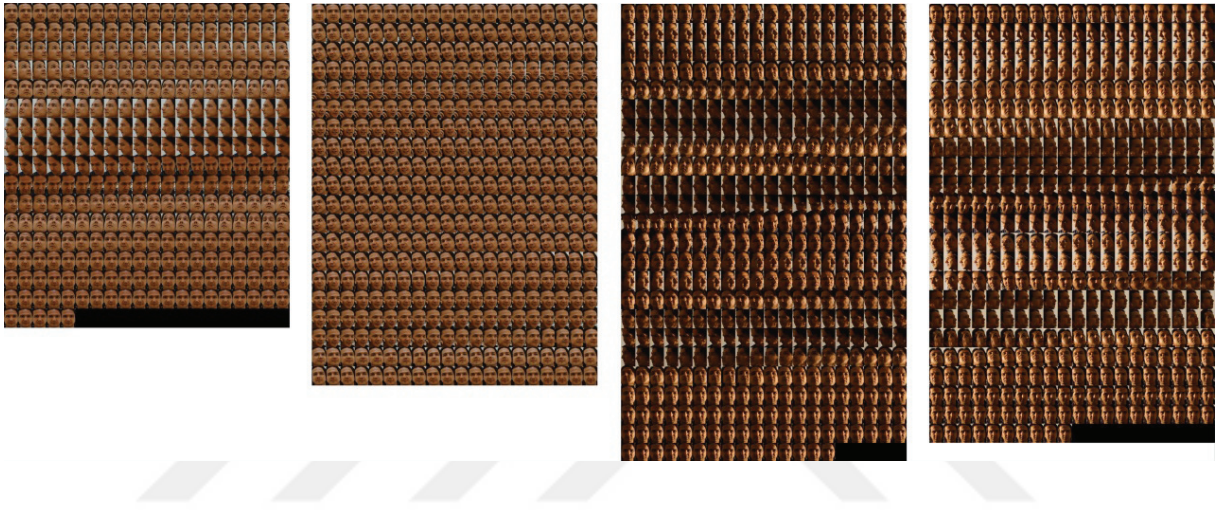
4. BULGULAR VE TARTIŞMA

4.1 Veri Tabanı

Deneysel çalışmalar için, tez çalışmasının desteklendiği proje kapsamında 285 kişinin katılımıyla oluşturulan ESOGU-285 video veritabanı kullanılmıştır (Yalcin vd., 2015). Veritabanı, yüz ve mimik çeşitliliklerini içeren kısa videolardan (yaklaşık 1 dk) oluşacak şekilde tasarlanmıştır. Katılımcılardan, aralarında en az 3 hafta zaman aralığı olacak şekilde 2 farklı oturum için çekim yaptırmaları istenmiştir. Her bir oturumda sırasıyla 4 farklı senaryo altında video çekimleri gerçekleştirilmiştir. Birinci senaryo, Honda/UCSD veritabanının tasarımına benzetilerek doğal ışık ortamında kişinin yüzünü yavaşça sağa, sola ve aşağı yukarı hareket ettirerek yüzün çeşitli açılardan görüntülendiği senaryodur. İkinci senaryo, yine doğal ışık ortamında ancak bu kez konuşma ve mimiklerin yoğun olarak canlandırıldığı bir senaryodur. Bu senaryoda kişilerden günlük hayatta cep telefonu ile konuşuyormuş gibi davranmaları istenmiştir. Buradaki temel amaç kişinin tercih edeceği gülme, şaşırma, kızgınlık gibi mimik çeşitliliklerini görüntülemektir. Üçüncü ve dördüncü senaryolar birinci senaryonun benzeridir. Üçüncü senaryoda katılımcının sol tarafında kalan projeksiyon lambası; dördüncü senaryoda katılımcının sağ tarafında kalan projeksiyon lambası açılarak görüntüler kaydedilmiştir. Projeksiyon lambalarının sadece bir taraftan açık olması, yüzün diğer yarısında büyük gölgeler oluşturmakta ve yüz tanıma problemini zorlaştırmaktadır. Bu senaryolar, aydınlatma değişiklikleri altında yüz imgelerini inceleyebilmek için tasarlanmıştır.

285 kişi için her iki oturumda 4'er senaryo (toplam 8 senaryo) çekimlerinden oluşturulan veritabanı 2280 videodan oluşmaktadır. Videoların uzunluğu katılımcıların yüzlerini hareket ettirme hızlarına bağlı olarak değişmektedir. Katılımcılar süre veya hareket hızı konusunda kısıtlandırılmamıştır. Video kareleri cinsinden incelendiğinde en kısa videonun 100 kare (frame), en uzun videonun 1360 kareden oluştuğu tespit edilmiştir. Tüm veritabanında toplam 764006 kare mevcuttur, bu sayı toplam video sayısına bölündüğünde videoların ortalama 335 kareden oluştuğu düşünülebilir.

Yüz imgelerinin sezimi ve hizalanmasının ardından literatürdeki diğer çalışmalarda izlenen deneysel çalışmalarla uyumlu olabilmek için imgeler gri seviyelere dönüştürülmüş ve imgelere histogram eşitleme uygulanmıştır. Benzer biçimde, imge boyutları da hafıza yetersizliği sorunu yaşanmaması için 40×30 olacak şekilde küçültülmüştür. Başka bir ilave ön işleme yapılmamıştır. Histogram eşitleme aydınlatma farklılıklarının oluşturduğu algı problemini kısmen ortadan kaldırmaktadır. Seçilen katılımcının birinci oturumdaki senaryolarda (soldan sağa senaryo 1, 2, 3 ve 4 sıralaması ile verilmiştir) çekilen videolarından oluşturulan imge kümeleri örnekleri Şekil 4.1’de gösterilmiştir.



Şekil 4.1 İmge kümeleri örnekleri

4.2 Poz Tahmini Deneyleri

Tez çalışmasının temel amacı, küme tabanlı yüz tanıma yöntemlerinde kullanılmak üzere poz kategorilerine ayrıştırılıp, uygun biçimlerde hizalanan yüz imgeleri oluşturmaktır. Buradaki temel hedef titiz bir veri ön işleme gerçekleştirmek ve bu sayede tanıma yöntemlerinin başarımlarında iyileşmeler gözlemlemektir. Ön işleme basamağının ilk adımı sezilen yüz imgeleri kullanılarak aydınlatma farklılıklarının etkisinin ortadan kaldırılmasıdır. Literatür araştırması sonucu bulunan yöntemlerin tümü veritabanından elde edilen yüz imgelerine uygulanmış ve bu yöntemlerin başarımları ve etkileri pozların sınıflandırması aşamasında elde edilmiştir.

Pozların kategorilere ayrıştırılmasındaki temel hedef daha başarılı hizalamalar elde edebilmektir. Bu amaçla, yedi poz kategorisi belirlenmiştir. Bu kategoriler yüzün (1) cepheden (2) sağ profilden, (3) sol profilden, (4) aşağıya baktığı durumda, (5) yukarıya baktığı durumda, (6) cephe ile sağ profil arasına baktığı durumda ve (7) cephe ile sol profil arasına baktığı durumda görüşleridir. Bu sebeple poz sınıfları verinin kümeleme işlemi küme sayısı $k = 7$ olarak yapılmış ve 4 yöntem arasında performans karşılaştırması yapılmıştır. Kümeleme işlemi için aşağıda verilen 4 yöntem kullanılmıştır.

- Klasik k -ortalamlar kümeleme yöntemi(KMC)
- PCA + k ortalamlar
- NMF yöntemi kullanılarak kümeleme yöntemi.
- GNMF yöntemi kullanılarak kümeleme yöntemi.

Poz sınıflandırma için kullanılacak yöntemler öncelikle manuel olarak yüz bölgesi kesilmiş ve herhangi bir ön işleme uygulanmamış hali üzerinde test edilmiştir. Bu testler için veritabanından rastgele seçilen 50 kişinin tüm videoları kullanılmıştır. Yüz imgeleri manuel olarak kesilerek poz sınıflarına ayrılmıştır. Deneylerde eğitim test verilerinin oranı 1:3 olarak belirlenmiştir. KMC, PCA, NMF ve GNMF yöntemlerinin işlenmemiş veri üzerindeki başarımları yüzdeleri Çizelge 4.1’de verilmiştir.

Çizelgede de görüldüğü gibi NMF ve GNMF yöntemleri kümeleme başarımları üzerinde ufak bir geliştirme sağlamıştır. Bu geliştirme poz sınıflandırma işlemi açısından yeterli değildir. Ancak orjinal veride aydınlatma değişikliklerinin etkisinin çok fazla olması

Çizelge 4.1 Orijinal veri üzerinde poz sınıflandırma başarımları

Metot	Başarımları(%)
KMC	49.22
PCA	52.47
NMF	53.78
GNMF	54.15

sebebiyle bu geliştirmenin küçük bir oranda olduğu düşünülmektedir. Aydınlatma etkilerinin ortadan kaldırılması sonucu ayrıştırılmaya daha uygun bir hale gelerek pozlar arasındaki sınıf içi benzerlik azalacaktır. Bu sebeple aydınlatma etkilerinin azaltılması poz sınıflarının tespit edilmesinde önemli bir adımdır. Bu noktadan yola çıkarak aynı deneyler aydınlatma etkilerinin ortadan kaldırmak için kullanılan yöntemler üzerinde denenmiş ve başarımları tespit edilmiştir. Bu yöntemler şu şekilde sıralanabilir:

- Uyarlamasız histogram eşitleme
- Uyarlamalı histogram eşitleme
- Single Scale Retinex (SSR)
- Multi Scale Retinex (MSR)
- Homomorfik Filtreleme (HOMO)
- Single-scale Self Quotient Image (SSQ)
- Multi-scale Self Quotient Image (MSQ)
- Discrete Cosine Transform (DCT)
- Wavelet-based normalization (WA)
- Wavelet-based image denoising (WD)
- Non-local Means based Normalization (NLM)
- Adaptive Non-local Means based Normalization (ANL)
- Steerable Filter (SF)
- GradientFaces (GF)
- Difference-of-Gaussian (DoG)

- Tan-Triggs normalization (TT)
- WeberFaces (WEB)
- Multi-scale WeberFaces (MSW)
- The Large- and Small-scale Features (LSSF)
- Adaptive single-scale retinex (ASR)

Yukarıda sıralı olarak verilen yöntemlerin uygulandığı yüz imgeleri kullanılarak klasik k -ortalamalar kümeleme yöntemi sonucu elde edilen başarımlar oranları Çizelge 4.2 'de verilmiştir.

Çizelge 4.2 Verilen yöntemler ile normalize edilmiş veri üzerinde k -ortalamalar poz sınıflandırma başarımlar oranları

Yöntem	Başarımlar Oranı(%)	Test Süresi(sn)
Histogram Eşitleme	64.31	0.70
Uyarlamalı Histogram Eşitleme	45.14	0.74
SSR	47.37	0.72
MSR	47.04	0.71
HOMO	48.67	0.69
SSQ	62.20	0.70
MSQ	62.17	0.72
DCT	31.85	0.74
WA	57.32	0.71
WD	38.41	0.74
NLM	64.97	0.74
ANL	63.02	0.80
SF	51.86	0.73
GF	51.29	0.74
DoG	61.52	0.75
TT	62.51	0.74
WEB	60.70	0.71
MSW	61.72	0.71
LSSF	43.04	0.69
ASR	67.51	0.72

Herhangi bir normalizasyon uygulanmamış olan veri ile normalizasyon yöntemleri uygulanmış veri için k -ortalamlar yöntemi ile poz sınıflandırması karşılaştırıldığında başarımlarında ciddi bir iyileştirme olduğu gözlenmektedir. Özellikle NLM ve ANL yöntemleri sınıflandırma üzerinde büyük geliştirmeler sağlamıştır. Ancak k -ortalamlar yönteminin başarımının Çizelge 4.1'deki deneyler göz önüne alınarak geliştirilebileceği bilinmektedir. Bu sebeple ilk olarak PCA yöntemi ile temsil edilen verinin k -ortalamlar yöntemi ile poz ayrıştırması yapılmıştır. Kullanılan aydınlatma normalizasyon yöntemleri ile elde edilen başarımlar ve her bir imgenin işlenmesi için geçen süreler Çizelge 4.3'de verilmiştir.

Çizelge 4.3 Verilen yöntemler ile normalize edilmiş veri üzerinde PCA+ k -ortalamlar poz sınıflandırma başarımları

Yöntem	Başarımları Oranı(%)	Test Süresi(sn)
Histogram Eşitleme	64.22	0.88
Uyarlamalı Histogram Eşitleme	44.23	0.70
SSR	46.64	0.73
MSR	46.89	0.71
HOMO	50.23	0.66
SSQ	61.88	0.75
MSQ	63.84	0.72
DCT	30.54	0.72
WA	57.24	0.72
WD	41.20	0.73
NLM	64.43	0.72
ANL	62.74	0.72
SF	52.44	0.70
GF	50.92	0.75
DoG	61.55	0.72
TT	62.37	0.72
WEB	61.33	0.74
MSW	61.45	0.71
LSSF	42.57	0.69
ASR	67.57	0.72

Çizelge 4.4 Verilen yöntemler ile normalize edilmiş veri üzerinde NMF+
 k -ortalamalar poz sınıflandırma başarımları

Yöntem	Başarımları Oranı(%)	Test Süresi(sn)
Histogram Eşitleme	59.39	0.85
Uyarlamalı Histogram Eşitleme	48.63	0.84
SSR	46.72	0.76
MSR	47.70	0.68
HOMO	52.54	0.71
SSQ	62.43	0.83
MSQ	60.69	0.82
DCT	29.36	0.69
WA	55.30	0.70
WD	37.47	0.78
NLM	61.20	0.80
ANL	59.24	0.69
SF	45.44	0.71
GF	43.59	0.70
DoG	55.75	0.74
TT	56.71	0.70
WEB	59.34	0.77
MSW	57.86	0.69
LSSF	44.61	0.77
ASR	68.04	0.72

Bu yöntem sayesinde yapılan geliştirmelerin artırılabilceği düşünüldüğünden GNMF yöntemi de uygulanmıştır.

Çizelge 4.5 Verilen yöntemler ile normalize edilmiş veri üzerinde GNMF+
k-ortalamalar poz sınıflandırma başarımları

Yöntem	Başarımları(%)	Test Süresi(sn)
Histogram Eşitleme	61.63	0.70
Uyarlamalı Histogram Eşitleme	50.29	0.74
SSR	62.55	0.72
MSR	63.18	0.71
HOMO	55.75	0.78
SSQ	69.41	0.69
MSQ	73.40	0.70
DCT	37.27	0.72
WA	54.93	0.74
WD	45.08	0.71
NLM	68.31	0.74
ANL	64.87	0.74
SF	59.91	0.80
GF	55.39	0.73
DoG	61.35	0.74
TT	63.97	0.75
WEB	65.11	0.74
MSW	60.66	0.71
LSSF	47.00	0.71
ASR	73.90	0.69

4.3 Yüz Hizalama Deneyleri

Poz ayrıştırmasının ardından yüz imgelerinin hizalanması gerekmektedir. Tez çalışması dahilinde kullanılan hizalama yöntemleri RASL, SIFT, SIFT-akış ve SIFT-akış+afin dönüşüm yöntemleridir. Hizalama yöntemlerinin görsel başarımları ölçütü olarak hizalama öncesi ve sonrası ortalama yüzler kullanılabilir.

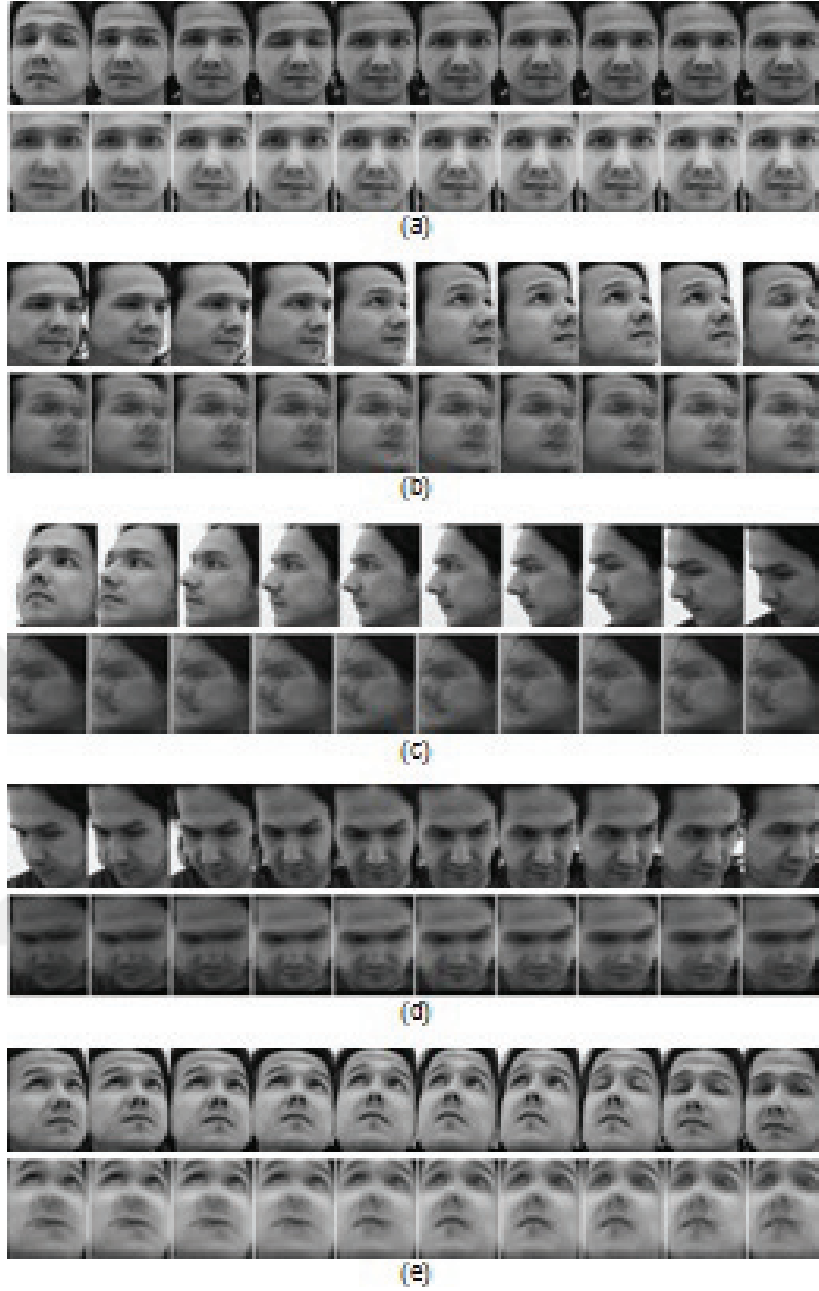
RASL yöntemi, hazırladığımız video veri tabanı üzerinde video karelerindeki yüz imgelerinin hizalanması amacıyla denenmiştir. Deneyler sırasında RASL yöntemi ile imgeler 80×60 piksel boyutunda şablonlar şeklinde hizalanarak dış göz kenarları arası 50 piksel olacak şekilde normalleştirilmiştir.

Alınan sonuçlarda en dikkat çeken ayrıntı gözlerin kapalı olduğu yüz imgelerinde hizalama sonrası gözlerin açık duruma gelmiş olmasıdır. Bu durum, yöntemin kısmi kapanmalar üzerindeki başarısını gözler önüne sermektedir. Şekil 4.2 incelediğinde cepheden görünümlü pozlar için yöntem oldukça başarılı hizalama tahminleri, aşağı ve yukarı bakış durumlarında kısmen başarılı sayılabilecek hizalama tahminleri döndürmesine rağmen, sağ ve sol profil imgeleri için başarılı tahminler döndüremediği görülmektedir. Şekil 4.3'de verilen hizalama öncesi ve sonrasında kullanılan imgelerden oluşturulan ortalama imgeler, bu çıkarsamayı doğrulamaktadır.

Yüz imgelerinin hizalanması için denenen diğer yöntem SIFT öznitelikleri kullanılarak hizalama yöntemidir. Bu yöntemde imge kümesinde bulunan her bir imgeye ait seyrek SIFT öznitelikleri çıkarılıp, bu öznitelikler kullanılarak yüzlerin hizalanması amaçlanmıştır. Oluşturulan veritabanı üzerine bu yöntem uygulanırken, veritabanındaki poz çeşitliliği yüzünden yine kategorileştirme yapılmıştır. Ardışıl video çerçevelerinde birbirini takip eden imgeler benzeşmektedir. Bu yüzden referans olarak ilk imge alınarak başlanmış, takip eden $N-1$ çerçeve ilk çerçevedeki SIFT öznitelik noktalarının konumlarına eşleştirilecek şekilde bir haritalama yapılmıştır. Böylece her bir kategoride N imge değerlendirilmiştir. Hizalama sonuçları $N = 3, 5, 10, 15$ ve 20 değerleri için elde edilmiştir.

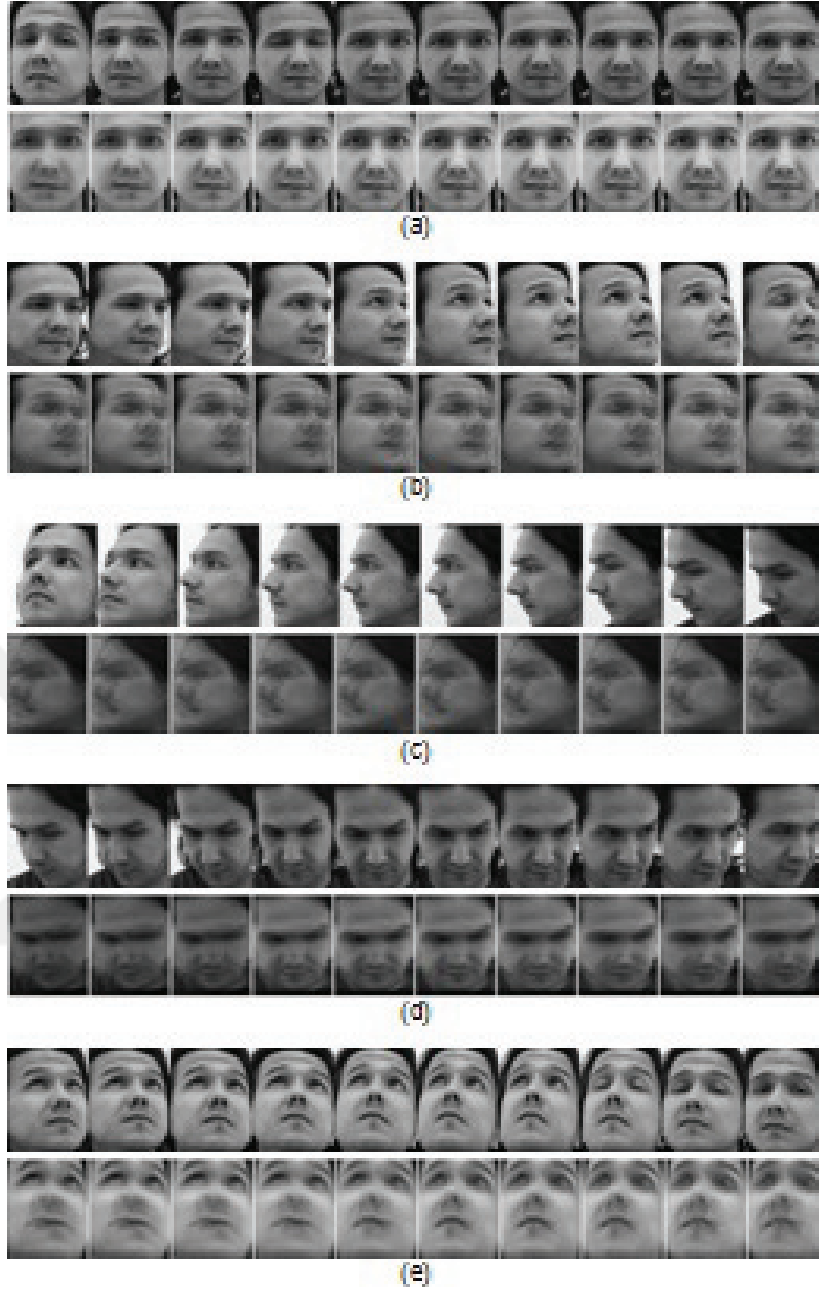
Seyrek SIFT öznitelikleri kullanılarak yapılan deneylerde Şekil 4.4'de gözlenebileceği gibi hizalama adına tutarlı sonuçlar gözlenememiştir. Bunun en önemli sebebi SIFT özniteliklerinin aydınlatma farklılıklarına yüksek oranda bağımlılığıdır. Özellikle senaryo-3 ve 4 için art arda gelen imgelerde dahi eşleşmesi gereken noktaların farklı yerlerde bulunabildiği tespit edilmiştir. Başarı kistası olarak Şekil 4.5'de verilen hizalama öncesi ve sonrası ortalama yüz imgelerine bakıldığında hizalama doğruluğu açısından fark edilir bir gelişme olmadığı gözlemlenmektedir.

Seyrek SIFT öznitelikleri kullanılarak hizalama yapılırken en önemli sorun nirengi noktalarının SIFT özniteliklerinde kendine karşılık bulamamasıdır. Özellikle büyük boyutlu gölgelenmelerin olduğu senaryo 3 ve 4'te, bazı imge kümelerinde SIFT öznitelikleri yüzdeki aydınlanma farkını bir kenar gibi görerek tüm noktaları belli bir çizgi üzerinde tespit etmiştir. Bu hatanın giderilmesi için daha sık aralıklarla referans seçiminin yapılması ($N=2, 3$) durumunda ise özniteliklerinin yaklaşık olarak aynı yerlerde bulunmasından dolayı herhangi bir dönüşüme gerek olmadığı sonucuna ulaşılması sebebiyle hizalama yapılamadığı gözlemlenmiştir. Bu yöntem kullanıldığında bir diğer önemli sorun, bazı noktalar için algoritmanın doğru eşlenmiş olarak kabul ettiği noktaların aslında yanlış eşleşmelere denk gelmesi yüzünden yanlış dönüşümler yapılmasıdır.



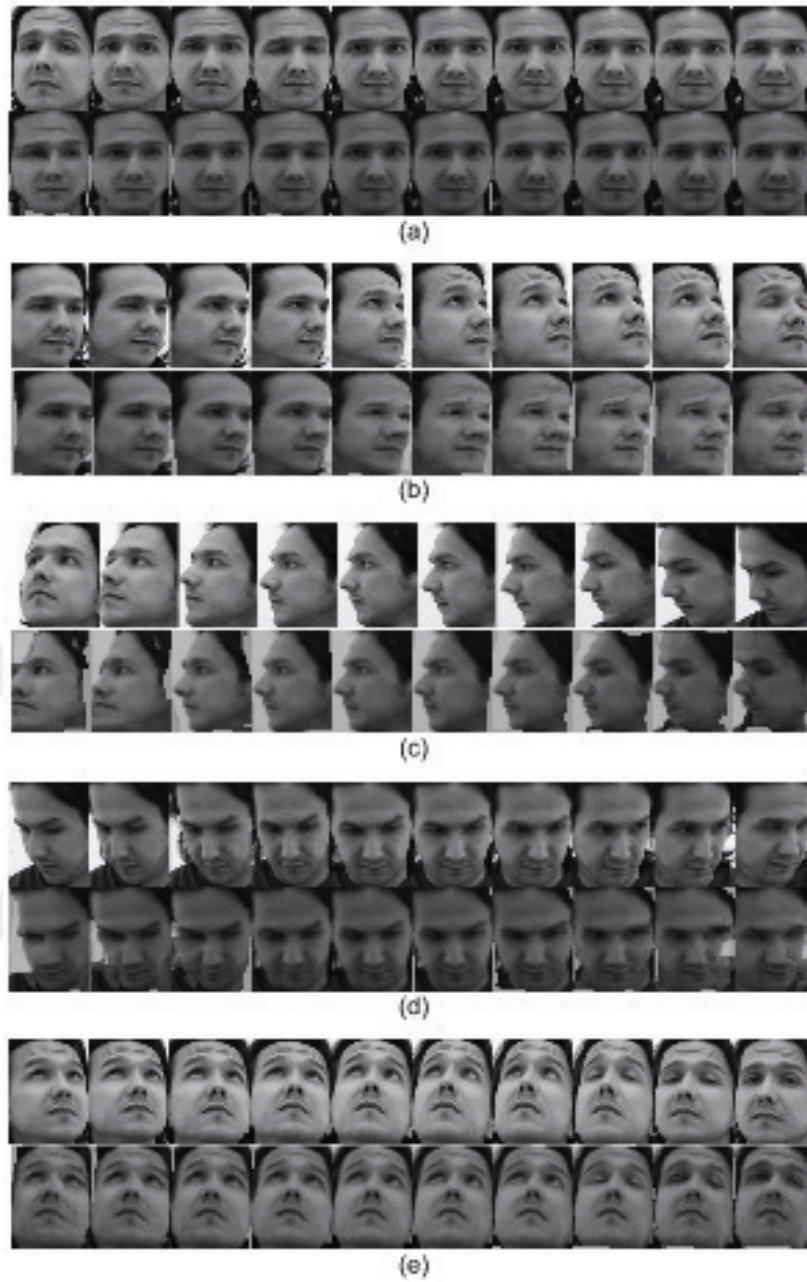
Şekil 4.2 RASL uygulama ile (pozlar ayrıştırılarak) hizalanmış imge kümeleri örnekleri, Poz kategorisi: a) cephe (b) sağ profil (c) sol profil (d) aşağı doğru (e) yukarı doğru, Birinci sıra: giriş imgeleri, İkinci sıra: hizalanmış imgelerden örnekler

Seyrek SIFT öz nitelikleri yerine tez çalışmasında kullanılabilecek diğer bir SIFT tabanlı yöntem SIFT-flow yöntemidir. Bu yöntemde optik akış algoritması temel alınarak piksel piksele bağımlılık yerine SIFT betimleyicilerinin bağımlılığı üzerine çalışılmıştır. SIFT betimleyicilerinin eşlenmesi mantığına dayalı bu yöntemde iki imge seçilir. İmgelerdeki her bir piksel değeri için referans pikselin komşulukları 4×4 'lük hücre matrislerine bölünür. Ardından her bir hücre 8 farklı yönelme için incelenerek, bir piksel



Şekil 4.3 RASL: Hizalama öncesi (birinci imge) ve sonrası (ikinci imge) ortalama yüz imgeleri Poz kategorisi: a)cephe (b)sağ profil (c)sol profil (d)aşağı bakış (e)yukarı bakış

değeri için $4 \times 4 \times 8 = 128$ boyutlu bir vektör elde edilir. Ardından optik akış algoritmasına benzer olarak, SIFT betimleyicilerinin akış vektörleri doğrultusunda eşlenmesi için bir amaç fonksiyonu tanımlanır. Bu amaç fonksiyonu optik akış algoritmasından farklı olarak çift katmanlı belief propagation yöntemi ile optimize edilir. SIFT-flow yönteminde giriş olarak bir referans imgesi seçilir ve eşleşmeler bu referans imgesine ait SIFT betimleyicileri kullanılarak yapılır. Bu yöntem kullanılarak yapılan hizalama sonuçlarına örnekler Şekil 4.6’da verilmiştir.



Şekil 4.4 SIFT öznelilikleriyle hizalanmış imge kümeleri örnekleri. Satırlardaki imge kümeleri sırasıyla Senaryo 1, 2, 3 ve 4'e aittir.

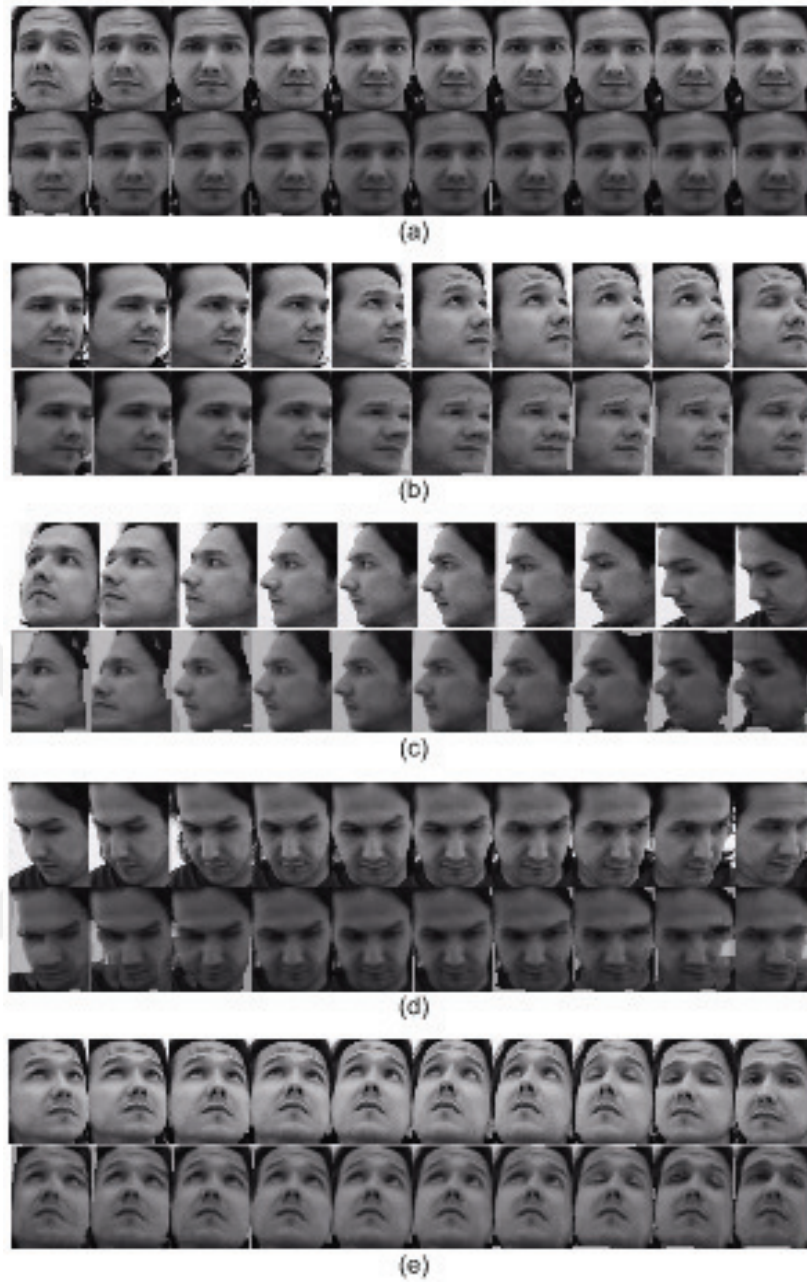
Aynı uygulama için hizalama öncesi ve sonrası ortalama yüz imgeleri de Şekil 4.7'de gösterilmiştir. RASL ile karşılaştırıldığında, hizalanmış ortalama imge görüntüleri açısından SIFT-flow yönteminin daha iyi hizalama yaptığı söylenebilir. Ancak pozları ayrıştırılmış imgeler tekil olarak incelendiğinde SIFT-flow ile hizalanan imgelerin bazılarında RASL ile hizalanma durumuna göre daha fazla deformasyon olduğu görülmektedir. SIFT-flow yönteminin bazı imgelerde oluşturduğu bölgesel deformasyonlar yüz tanıma performansının düşmesine sebep olabilir.



Şekil 4.5 SIFT öz nitelikleriyle hizalama öncesi (birinci imge) ve hizalama sonrası (ikinci imge) ortalama yüz imgeleri (a)Senaryo-1, (b) Senaryo-2, (c)Senaryo-3, (d)Senaryo-4

RASL, seyrek SIFT öz nitelikleri ve SIFT-flow yöntemleri ile yapılan hizalamalar karşılaştırıldığında seyrek SIFT öz niteliklerinin diğer iki yönteme göre daha kötü sonuçlar verdiği gözlenmiştir. RASL ve SIFT-flow yöntemlerinin çıktıları tek taraftan aydınlatma kullanılan bazı profil pozları dışında başarılıdır. Hizalama yöntemlerinin şablona eşleştirmesi dışında, başarımları da önemlidir. Veritabanından seçilen 50 kişinin imge kümeleri üzerinde toplam hizalama süreleri aynı bilgisayar ile aynı koşullarda ölçülmüş ve toplam imge sayısına bölünerek ortalama hizalama hesap süresi elde edilmiştir. Buna göre RASL yönteminin hesap süresi imge başına 10,6 sn., SIFT-flow yönteminin hesap süresi imge başına 1,3 saniyedir. Matris ayrıştırmasına dayalı hesaplamalar yapması yüzünden RASL yönteminin hizalama için harcadığı zaman daha fazladır ve gerçek zamanlı tanıma uygulamaları için uygun değildir. Şunu da belirtmek gerekir ki, RASL yönteminin hızı oluşturulan yığın matrisine bağlıdır. İmge kümeleri çok sayıda çerçeve içerdiği zaman, hesaplama hızı buna göre büyük olacaktır. Oluşturduğumuz veritabanı imge kümesi başına ortalama yaklaşık 335 video karesi içermektedir, bu RASL yönteminin gerçek zamanlı kullanılması için oldukça büyük bir sayıdır. Az sayıda imge içeren kümelere RASL daha hızlı çalışabilir.

RASL ve SIFT-flow yöntemlerinden iyi sayılabilecek hizalama sonuçları elde edilmesine rağmen, yüzde daha az deformasyon yapabilen bir hizalama biçimi belirleyebilmek amacıyla, yukarıda verilen hizalama yöntemlerine ilave olarak, kendi tasarladığımız şablonlara eşleştirme yapan bir başka hizalama biçimi de kullanılmıştır. Kaynak imge üzerinde döndürülen bu noktalar, hedef olarak belirlenen bir şablona haritalanabilir. İmge hizalaması, test imgesinin bakış açısına en yakın olan yüz şablonuna eşleştirilmesiyle tamamlanır. Şablon sayısını arttırmak mümkündür ancak belirlenen şablon sayısı hizalama için yeterli görülmüştür. Her bir poz kategorisinde en iyi yüz şablonunu oluşturabilmek için öncelikle örneklem imgeleri göz kontrolü yapılarak seçilmiş, bu imgeler üzerinde yüz sezicinin döndürdüğü nirengi noktalarının ortalamaları hesaplatılarak poz kategorilerinin şablonları oluşturulmuştur. Oluşturulan şablonlar, ortalama imgeler üzerinde kontrol edilmiş ve görsel olarak güvenilirlikleri test edilmiştir. Yedi poz kategorisi için deneysel olarak oluşturduğumuz şablonlar Şekil 4.8’de gösterilmiştir.

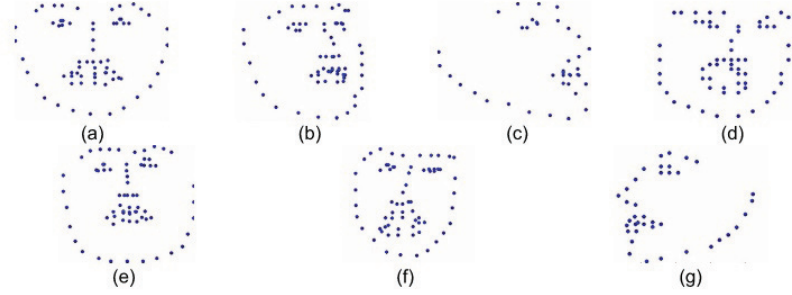


Şekil 4.6 SIFT-flow yöntemiyle (pozlar ayrıştırılarak) hizalanmış imge kümeleri örnekleri. Poz kategorisi: a)cephe (b)sağ profil (c)sol profil (d)aşağı bakış (e)yukarı bakış. Birinci sıra: giriş imgeleri, İkinci sıra: hizalanmış imgelerden örnekler

Otomatik yüz hizalama süreci, otomatik yüz sezimi sonucunda döndürülen nirengi noktalarının yukarıda verilen şablonlara haritalanmasıyla tamamlanır. Hedef şablonları belli olduğu için haritalama süreci RASL ve SIFT-flow yöntemlerine göre çok daha kısa sürmektedir. Hesap süresi diğerleriyle aynı koşullarda test edildiğinde imge başına 0,9 saniye olarak ölçülmüştür. Bu yöntem kullanılarak elde edilen hizalanmış imgeler Şekil 4.9'da gösterilmiştir.

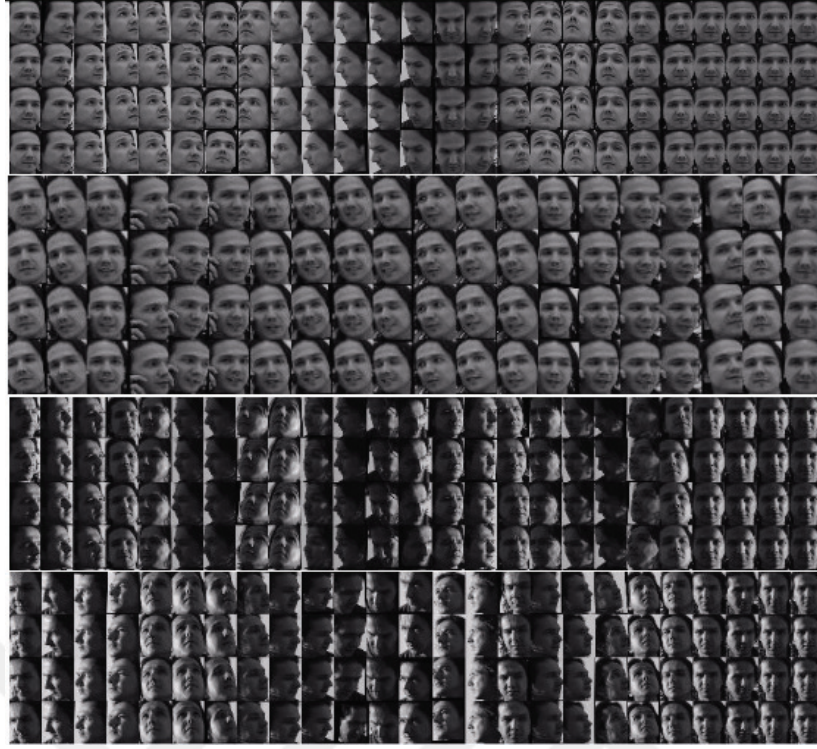


Şekil 4.7 SIFT-flow (uygulama-2): Hizalama öncesi (birinci imge) ve sonrası (ikinci imge) ortalama yüz imgeleri. Poz kategorisi: a)cephe (b)sağ profil (c)sol profil (d)aşağı bakış (e)yukarı bakış



Şekil 4.8 Nirengi noktaları için tasarlanan hedef şablonları (a) cephe, (b) sağa doğru, (c) sağ profil (d) aşağı doğru (e) yukarı doğru (f) sola doğru (g) sol profil.

Şablon hizalama yöntemiyle hizalanan imgelerde, RASL ve SIFT-flow yöntemlerinde görülen deformasyon ortadan kalkmıştır. Kullanılan yüz hizalama yöntemlerinin hizalama başarımları göz önüne alındığında, aynı poz kategorisindeki imgelerin birbirlerine benzeşen biçimde hizalandıkları tespit edilmiştir. Otomatik hizalama için test edilen yöntemler birbirleriyle kıyaslandığında, hız ve doğruluk açısından en verimli yöntemin en son kısımda verilen tasarlanmış şablona hizalama yöntemi olduğu tespit edilmiştir.



Şekil 4.9 Şablon hizalama yöntemiyle hizalanmış imge kümeleri örnekleri. Satırlardaki imge kümeleri sırasıyla Senaryo 1, 2, 3 ve 4'e aittir.

4.4 Hizalanmış İmge Kümelerinin Yüz Tanıma Deneylerinde Kullanılması

Hizalanmış yüz imgelerinin başarısını ölçmek için literatürde yer alan çeşitli yöntemler kullanılmıştır. Bu yöntemler şu şekilde sıralanabilir: AZM, KZM, İndirgenmiş Konveks Zarf Modelleme (İKZM) ve bunun hızlandırılmış uyarlaması olan İndirgenmiş Konveks Zarf Modelleme+İkili Hiyerarşik Karar Ağaçları (İHKA) (Yalcin vd., 2015), Mutual Subspace Method (MSM) (Yamaguchi K., Maeda, K.-i. ve Kansai Res. Labs. Kobe, 1998), Sparse Approximation Nearest Points (SANP) (Hu vd., 2012), Regularized Nearest Points (RNP) (Yang vd., 2013), Collaboratively Regularized Nearest Points (CRNP) (Wu vd., 2013), Manifold-to-Manifold distance (MMD) (Wang vd., 2012) ve Self-Regularized Non-Negative Coding and Adaptive Distance Metric Learning (SRN-ADML) (Mian vd., 2013).

Deneylerde birinci oturumda çekilen videolar eğitimde kullanılırken ikinci oturumda çekilen videolar test aşamasında kullanılmıştır. Manuel olarak kesilen yüzler üzerinde doğru tanıma oranları tespit edilmiştir. Ardından literatürden seçilen yöntemler kullanılarak

hizalanan yüz imgelerinden oluşan imge kümeleri kullanılarak aynı deneyler tekrarlanmıştır. Yüz imgelerinin yukarıda bahsedilen tüm yöntemler kullanılarak hizalanmasının ardından yüz tanıma yöntemlerinin hepsi tekrar uygulanmış ve tüm doğru tanıma oranları Çizelge 4.6'da verilmiştir.

Çizelge 4.6 ESOGU-285 Yüz Video Veri Tabanı Üzerindeki Doğru Tanıma Oranları(%)

Yöntem	Doğru Tanıma Oranları (Hizalama Öncesi)(%)	Doğru Tanıma Oranları (RASL) (%)	Doğru Tanıma Oranları (SIFT-akış)(%)	Doğru Tanıma Oranları (Yöntem)(%)
AUM	44.3	46.4	42.5	50.2
KZM	55.1	57.2	53.8	58.9
MSM	50.1	51.8	48.3	54.8
SANP	51.9	53.3	49.6	55.8
RNP	46.7	48.6	44.6	50.1
CRNP	OOM*	OOM*	OOM*	OOM*
SRN-ADML	45.4	47.1	42.9	49.2
MMD	52.0	53.9	50.1	56.0
Kernel AUM	62.1	64.5	60.3	66.9
Kernel KZM	62.1	64.3	60.2	66.8
Kernel İKZM	61.2	63.2	59.7	65.1
Kernel İKZM-İHKA	60.8	62.9	58.6	64.8

*OOM: Out Of Memory

Tablonun birinci satırındaki sonuçlar, yüz tanıma literatüründe birçok araştırmacının giriş olarak kullandığı, yüz sezicisinin döndürdüğü yüz bölgeleri kullanılarak (hizalama yapılmadan) elde edilen tanıma sonuçları, ikinci satırdaki sonuçlar RASL yöntemi ile hizalama yapılan imgeler kullanıldığında, üçüncü satırdaki sonuçlar SIFT-akış yöntemi ile hizalama yapılmış olan imgeler kullanıldığında, son satırdaki sonuçlar ise önerilen yöntem ile hizalanmış imgeler kullanıldığında elde edilen tanıma oranlarını göstermektedir. SIFT-akış yöntemi, yüz imgelerinin hizalama yapılmadan kullanılması durumundan daha düşük tanıma oranları vermiştir. Bunun en büyük sebebi, daha öncede bahsedildiği gibi bu yöntemde yüz imgelerinde yüksek deformasyonların oluşmasıdır. RASL yöntemiyle hizalama yapılması yüz tanıma başarımını küçük bir miktar artırıcı yönde etki yapmıştır.

Tez çalışmasında yapılan değerlendirmeler sonucunda SIFT-akış + afin dönüşüm yönteminin en yüksek başarımları döndüğü tespit edilmiştir. Başarılı bir yöntem olan SIFT-akış yöntemindeki deformasyonların etkisi minimuma indirilerek yüz tanıma başarımları arttırılmıştır.

5. SONUÇ VE ÖNERİLER

Yüz hizalama süreci, yüz tanıma performansını önemli derecede geliştirebilecek bir süreç olsa da birçok araştırmacı bu aşamayı zor bir süreç olması sebebiyle göz ardı etmektedir. Otomatik hizalama yüz tanıma yöntemleri için önemli bir önışleme adımıdır.

Bu çalışmada literatürde yer alan imge hizalama için denenmiş bazı yöntemler kullanılarak yüz tanıma performansları ölçülmüştür. Yapılan deneyler sonucunda, yöntemlerin yüz tanıma başarımlarını nasıl etkilediği ortaya konmuştur. Ayrıca poz ayırıştırmasının otomatik bir karar verme süreci ile gerçekleştirilerek, yüz imgelerinin belirli şablonlara otomatik olarak hizalanması ile poz çeşitliliği yüksek veri tabanlarında başarımın daha fazla arttırılması sağlanmıştır. Buna ek olarak, aydınlatma farklılıklarından yüksek oranda etkilenen yüz imgelerinin bu etkilerden arındırılması ile başarım oranının arttırılması hedeflenmiştir.

Hizalama işlemleri poz çeşitliliği ve aydınlatma farklılıkları içeren ve laboratuvarımızda oluşturulan ESOGU-285 Video Veri tabanı üzerinde denenmiştir. Bu veritabanı üzerine uygulanan iyileştirmeler literatüde yer alan imge kümesi tabanlı yüz tanıma yöntemleri ile denenerek başarımları ölçülmüştür. Bu deneyler, poz sınıfları için oluşturulan şablonlara yapılan hizalamanın diğer hizalama yöntemlerine göre başarımı bir miktar daha fazla arttırdığı kanıtlamıştır.

Seyrek SIFT özneteliklerinin kullanılması yüz imgelerinde farklı dönüşümlerin uygulanmasına sebep olduğundan bu yöntem yüz hizalama deneylerinde kullanılmamıştır. SIFT-akış yönteminde ise yüz imgelerinde oluşan belirgin deformasyonlar tanıma performansını düşürücü yönde etki yapmıştır. RASL yöntemi kısmen başarılı sonuçlar döndürmekle birlikte şablon tabanlı hizalamaya göre daha düşük tanıma oranları elde edilmesine sebep olmuştur. Nirengi noktalarının bulunması konusunda geliştirilmeler yapıldığı takdirde bu deneylerin başarımlarının yükselmesi beklenmektedir. Ayrıca nirengi noktalarının bulunması, yüz imgelerinin aydınlatma değişimlerinden arındırılması, poz kategorilerine ayrılması ve yüz tanıma işlemleri birleştirilerek tam otomatik bir sistem oluşturulması deneylerdeki başarımın gerçek zamanlı uygulamalarda kullanılabilmesi için ne kadar verimli olduğunu ortaya koymakta daha sağlıklı sonuçlar verecektir.

KAYNAKLAR DİZİNİ

- Arandjelović, O., Zisserman A., 2005, Automatic face recognition for film character retrieval in feature-length films, Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, *CVPR 2005.*, s. 860–867.
- Arandjelovic, O., Shakhnarovich, G., Fisher, J., Cipolla, R., Darrell, T., 2005, Face recognition with image sets using manifold density divergence, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, s. 581–588.
- Ashraf, A. B., Lucey S., Chen, T., 2008, Learning patch correspondences for improved viewpoint invariant face recognition, 26th IEEE Conference on Computer Vision and Pattern Recognition.
- Asthana, A., Marks, T. K., Jones M. J., Tieu, K. H., Rohith, M., Fully, 2011 automatic pose-invariant face recognition via 3D pose normalization, International Conference on Computer Vision, s. 937–944.
- Blanz, V., Vetter, T., “Face recognition based on fitting a 3D morphable model, 2003, IEEE Transactions on Pattern Analysis and Machine Intelligence 25.9, s. 1063–1074.
- Cai, D., He, X., Wu, X., Han, J., 2008, Non-negative Matrix Factorization on Manifold, Eighth IEEE International Conference on Data Mining, s. 63–72.
- Cao, X., Wei, Y., Wen, F., Sun J., 2013, Face alignment by explicit shape regression, International Journal of Computer Vision 107.2, s.177–190.
- Cevikalp, H., Triggs, W., 2010, Face recognition based on image sets, Computer Vision and Pattern Recognition CVPR 2010 IEEE Conferenceon 365.2, s. 2567–2573.
- Cevikalp, H., Yavuz H. S., Edizkan, R., Gündüz, H., Kandemir, C. M., 2011, Comparisons of features for automatic eye and mouth localization, INISTA International Symposium on INnovations in Intelligent SysTems and Applications, s.576–580.
- Chang, J. M., Kirby, M., Peterson, C., 2007, Set-to-set face recognition under variations in pose and illumination, Biometrics Symposium.

KAYNAKLAR DİZİNİ (devam)

- Chen, S., Sanderson, C., Harandi, m. T., Lovell B., C., 2013 Improved image set classification image set classification via joint sparse approximated nearest subspaces, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, s.452–459.
- Chen, W., Meng-Joo, E., Shiqian, W., 2005, Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain, Systems, Man, and Cybernetics, s. 458–466.
- Chen, Y., Patel, V. M., Shekhar, S., Chellappa, R., Phillips, P. J., 2013, Video-based face face recognition via joint sparse representation, 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition.
- Chow, C., Liu C., 1968, Discrete Probability Distributions with Dependence Trees, *IEEE Transactions on Information Theory* 14, s. 462–467.
- Cui, Z., Shan, S., Zhang, H., Lao, S., Chen, X., 2012, Image sets alignment for video Based Face Recognition, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, s. 2626–2633.
- Cui, Z., Chang, H., Shan, S., Ma, B., Chen, X., 2013, Joint sparse representation for video-based face recognition, *Neurocomputing* 135, s.306–312.
- Ding, L., Martinez A. M., ,Features versus context: An approach for precise and detailed detection and delineation of faces and facial features, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32.11, s. 2022–2038.
- Du, S., Ward, R., 2005, Wavelet-based illumination normalization for face recognition, *IEEE International Conference on Image Processing*, s. 2–5.
- Fitzgibbon, A.W., Zisserman, A., 2003, Joint manifold distance: a new approach to appearance based clustering, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2003. Proceedings. s. I–26–I–33.
- Hamm, J., Lee D., Grassmann Discriminant Analysis : a Unifying View on Subspace Based Learning, Proceedings of the 25th international conference on Machine learning, s. 376–383.

KAYNAKLAR DİZİNİ (devam)

- Hamm, J., Lee, D., 2009, Extended Grassmann kernels for subspace-based learning, *Advances in Neural Information* , s. 1–8.
- Harandi, M. T., Sanderson, C., Shirazi S., Ovell, B C., 2011, Graph embedding discriminant analysis on Grassmannian manifolds for improved image set matching, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, s. 2705–2712.
- Harandi, M. T., Sanderson, C., Shirazi S., Ovell, B C., 2013, Kernel analysis on Grassmann manifolds for action recognition, *Pattern Recognition Letters*, s. 1906–1915.
- Hayat, M., Bennamoun, M., An, S., 2014, Learning Non-Linear Reconstruction Models for Image Set Classification, s. 1915–1922.
- Hayat, M., Bennamoun, M., Ani S., 2014, Deep reconstruction models for image set classification, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, s. 713–727.
- Hines, G., Rahman, Z., Woodell, G., Single-Scale Retinex Using Digital Signal Processors, *Global Signal Processing Conference*, s. 1–6.
- Hu, Y., Mian, A. S., Owens, R., 2011, Face recognition using sparse approximated nearest points between image sets, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34.10, s. 1992–2004.
- Jobson, D.J., Rahman, Z. U., Woodell, G. A.. 1997, A multiscale retinex for bridging the gap between color images and the human observation of scenes, *IEEE Transactions on Image Processing* 6.7, s. 965–976.
- Jolliffe, I T, *Principal Component Analysis*, Second Edition, *Encyclopedia of Statistics in Behavioral Science* 30.3, s. 487.
- Kim, T. K., Arandjelovic, O., Cipolla, R., Boosted manifold principal angles for image set-based recognition, *Pattern Recognition* 40.9, s. 2475–2484.
- Liu, C., Yuen, J., Torralba, A., 2010, SIFT flow: Dense correspondence across scenes and its applications, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.5, s. 978–994.

KAYNAKLAR DİZİNİ (devam)

- MacQueen, J. B., ,Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, s. 281–297.
- Mian, A., Hu, Y., Hartley, R., Owens, R.,2013, Image Set Based Face Recognition Using Self-Regularized Non-Negative Coding and Adaptive Distance Metric Learning, IEEE Transactions on Image Processing 22.12, s. 5252–5262.
- Naseer, T., Ruhnke,M., Stachniss, C., Spinello, L., Burgard, W., 2015, Robust visual SLAM across seasons, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), s. 2529–2535.
- Nilufar, S., Ray, N., Zhang, H., 2012, Object Detection with DoG Scale-Space: A Multiple Kernel Learning Approach, IEEE transactions on image processing : a publication of the IEEE Signal Processing Society 21.8, s. 3744–3756.
- Ohkawa, Y., Suryanto. C. D., Fukui, K., ,Image set-based hand shape recognition using camera selection driven by multi-class AdaBoosting, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) s. 555–566.
- Peng, Y., Ganesh, A., Wright, J., Xu, W., Ma, Y., 2011, RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images, IEEE Transactions on Pattern Analysis and Machine Intelligence 34.11, s. 2233–2246.
- Rahimzadeh A., Kittler, S., 2010, Energy normalization for pose-invariant face recognition based on MRF model image matching, IEEE Transactions on Pattern Analysis and Machine Intelligence 33.6, s. 1274–1280. I
- Shen, X., Lin, Z., Brandt, J., Wu, Y., 2013, Detecting and aligning faces by image retrieval Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, s. 3460–3467.
- Shi, J., Malik, J.,Normalized Cuts and Image Segmentation, IEEE Trans.Pattern Anal. Mach. Intell. 22.8, s. 888–905.
- Štruc, V., Pavešić, N., 2009,I llumination Invariant Face Recognition by Non-Local Smoothing, Proceeding of the BIOID Multicomm 2009, s. 1–8.

KAYNAKLAR DİZİNİ (devam)

- Viola, P., Jones, M., Robust real-time object detection, *International Journal of Computer Vision* 57, s. 137–154.
- Wang, B., Li, W., Yang, W., Liao, Q., 2011, Illumination normalization based on weber's law with application to face recognition, *IEEE Signal Processing Letters* 18.8, s. 462–465.
- Wang, H, Wang, Y., Cao, Y., Video-based face recognition: A survey, *World Academy of Science, Engineering*, s.136–139.
- Wang, H., Li, S. Z., Wang, Y., Zhang, J., 2004, Self quotient image for face recognition, *International Conference on Image Processing*, s. 1397–1400.
- Wang, O., Schroers, C., Zimmer, H., Gross, M. H., Sorkine-Hornung, A., VideoSnapping: interactive synchronization of multiple videos., *ACM Trans. Graph.*
- Wang, R., Shan, S., Chen, X., Dai, Q., Gao, W., Manifold-manifold distance and its application to face recognition with image sets, *IEEE Transactions on Image Processing*, s. 4466–4479.
- Wibowo, M. E., Tjondronegoro, D., Chandran, V., 2012, Probabilistic matching of image sets for video-based face recognition”, *International Conference on Digital Image Computing Techniques and Applications*.
- Wibowo, M. E., Tjondronegoro, D., Zhang, L., Himawan, I., Heteroscedastic probabilistic linear discriminant analysis for manifold learning in video-based face recognition, *Proc. of IEEE Works. on Applications of Computer Vision*, s. 46–52.
- Williams, D. B., Madisetti, V., *Digital Signal Processing Handbook*, 1st. Boca Raton, FL, USA: CRC Press, Inc.
- Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., Ma, Y., 2008, Robust face recognition via sparse representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, s. 210–227.
- Wu, K. K., Wang, L., Soong, F. K., Yam, Y., 2011, A sparse and low-rank approach to efficient face alignment for photo-real talking head synthesis, *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing Proceedings*, s. 1397–1400.

KAYNAKLAR DİZİNİ (devam)

- Wu, Y., Minoh, M., Mukunoki, M., 2013, Collaboratively regularized nearest points for set based recognition, BMVC2013, s. 1–10.
- Wu, Y., Shen, B., Ling, H., 2012, Online robust image alignment via iterative convex optimization, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, s. 1808–1814.
- Xie, X., Zheng, W. S., Lai, J., Yuen, P. C., Suen, C. Y., 2010, Normalization of face illumination based on large-and small-scale features, IEEE Transactions on Image Processing, s. 1807–1821.
- Xiong, X., De La Torre, F., 2013, Supervised descent method and its applications to face alignment, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, s. 532–539.
- Yalcin, M., Cevikalp, H., Yavuz, H. S., 2015, Towards Large-Scale Face Recognition Based on Videos, 2015 IEEE International Conference on Computer Vision Workshop (ICCVW), s. 1078–1085.
- Yamaguchi K., Maeda, K., 1998, Face recognition using temporal image sequence Third IEEE International Conference on Automatic Face and Gesture Recognition, s. 318–323.
- Yang, M., Zhu, P., Van Gool, L., Zhang, L., 2013, Face recognition based on regularized nearest points between image sets, 2013 10th IEEE Int. Conference and Workshops on Automatic Face and Gesture Recognition.
- Yang, M. H., Kriegman, D. J., Ahuja, N., Detecting faces in images: A survey, IEEE Transactions on Pattern Analysis and Machine Intelligence 24.1, s. 34–58.
- Yavuz, H. S., Effects of different face template designs for appearance based face, AWER Procedia Information Technology & Computer Science, s. 899–904.
- Yavuz, H. S., Çevikalp, H., Edizkan, R., 2013, Frontal imgelerden otomatik yüz tanıma, 2013 21st Signal Processing and Communications Applications Conference, SIU 2013.

KAYNAKLAR DİZİNİ (devam)

- Zhang, T., Tang, Y. Y., Fang, B., Shang, Z., Liu, X., 2009, Face recognition under varying illumination using gradientfaces, *IEEE Transactions on Image Processing* 18.11, s. 2599–2606.
- Zhang, T., Fang, B., Yuan, Y., Tang, Y. Y., Shang, Z., Li, D., Lang, F., 2008, Multiscale facial structure representation for face recognition under varying illumination, *Pattern Recognition* 42.2, s. 251–258.
- Zhao, W., Chellappa, R., Phillips, P. J., Rosenfeld, A., Face recognition: A literature survey, *Acm Computing Surveys* 35.4, s. 399–458.
- Zhou, E., Fan, H., Cao, Z., Jiang, Y., Yin, O., 2013, Extensive facial landmark localization with coarse-to-fine convolutional network cascade, *Proceedings of the IEEE Int. Conference on Computer Vision*, s. 386–391.
- Zhu, P., Zuo, W., Zhang, L., Shiu, S. C. K., Zhang, D., Image Set based Collaborative Representation for Face Recognition, *IEEE Transactions on Information Forensics and Security* 9.7, s. 1120–1132.
- Zhu, X., Ramanan, D., Face detection, pose estimation and landmark localization in the wild *Computer Vision and Pattern Recognition (CVPR)* Providence, Rhode Island, June 2012.

EK AÇIKLAMALAR

Ek Açıklamalar - A: Viola-Jones yüz sezimi uygulaması

Ek Açıklamalar - B : Zhu & Ramanan yüz sezimi uygulaması

Ek Açıklamalar - C : Face++ yüz sezimi uygulaması

Ek Açıklamalar - D: Eğitimli İniş Metodu yüz sezimi uygulaması

Ek Açıklamalar - E: Video çekimi için tasarlanan senaryolar



EK AÇIKLAMALAR - A: Viola-Jones yüz sezimi uygulaması



Şekil 6.1 Viola-Jones yüz sezimi uygulaması

EK AÇIKLAMALAR - B : Zhu & Ramanan yüz sezimi uygulaması



Şekil 6.2 Zhu & Ramanan tarafından önerilen yüz sezici ile yüz sezimi sonuçları

EK AÇIKLAMALAR - C : Face++ yüz sezimi uygulaması



Şekil 6.3 Face++ algoritması ile alınan sonuçlar

EK AÇIKLAMALAR - D: Eğitmenli İniş Metodu yüz sezimi uygulaması



Şekil 6.4 Eğitmenli İniş Metodu kullanılarak alınan sonuçlar

EK AÇIKLAMALAR - E: Video çekimi için tasarlanan senaryolar



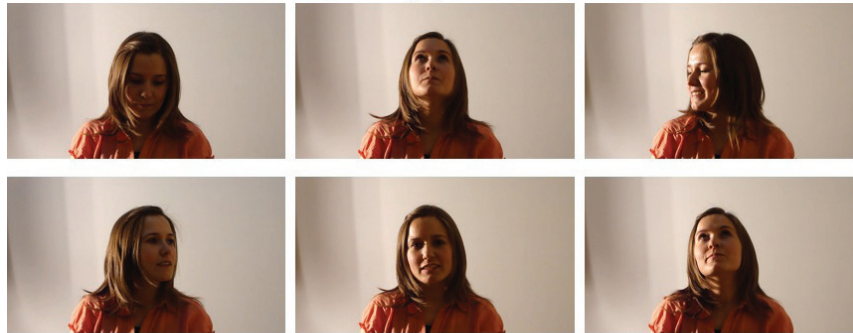
(a)



(b)



(c)



(d)

Şekil 6.5 Video çekimi için tasarlanan senaryolar (a) Senaryo-1 (b) Senaryo-2 (c) Senaryo-3 (d) Senaryo-4