



İmge Kümeleriyle Yüz Tanıma için Otomatik Önişleme

Autonomous Preprocessing for Image Set Based Face Recognition

Hasan Serhan Yavuz¹ , Meltem Seyirt² 

¹Eskişehir Osmangazi Üniversitesi, Elektrik-Elektronik Mühendisliği Bölümü, 26480, Eskişehir, TÜRKİYE

²TÜBİTAK BİLGEM, Barış Mah. Dr. Zeki Acar Cad. No:1 PK: 74, 41470, Gebze / Kocaeli, TÜRKİYE

Başvuru/Received: 09/01/2019

Kabul/Accepted: 26/03/2019

Son Versiyon/Final Version: 30/06/2019

Öz
Otomatik yüz tanıma süreci son yıllarda popülerliğini arttırmış bir konudur. İmge tabanlı yaklaşımların hâkim olduğu yüz tanıma süreci, kamera ve hesaplama 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 çalışmada, özellikle aydınlatma ve poz çeşitliliklerinin var olduğu durumlarda, küme tabanlı yüz tanıma yöntemlerinin başarımlarının nasıl iyileştirilebileceği araştırılmıştır. Ön işleme basamağında, aydınlatma farklılıklarını giderildikten sonra imgeler öncelikle yüz pozuna göre sınıflandırılmıştır. Pozlara göre ayrıştırılan yüzler, sınıf içi değişimlerinin azaltılması için ilgili pozun şablonuna hizalanmıştır. Yapılan deneyler sonucunda, önişleme basamağında önerilen otomatik poz hizalama yöntemi kullanıldığında, video tabanlı yüz tanıma deneylerinin başarımlarında gelişmeler sağlandığı tespit edilmiştir.

Anahtar Kelimeler

“Yüz sezimi, yüz hizalama, yüz tanıma, imge kümeleri”

Abstract

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 started to leave its place in video-based approaches by the development of camera and computing technologies. 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 study, we investigate how to improve the accuracies of set based face recognition methods in case of lighting and face pose variations. At the pre-processing stage, after the illumination differences are refined, the images are firstly classified according to face exposure. The faces that are separated according to the poses are aligned to the corresponding canonical pose patterns to reduce intra class variations. Experimental results demonstrate that set based recognition methods give higher correct recognition rates when the proposed methodology schemes have been applied as a preprocessing stage.

Key Words

“Face detection, face alignment, face recognition, image sets”

1. Giriş

Bilgisayar ortamında bir kişinin kimliğinin tespit edilebilmesi için birçok biyometrik özellik kullanılabilir. Yüz biyometrisi, bireyin yüz imgelerinin bir kamera ile kolay bir biçimde yakalanabilmesiyle gerçekleştirildiği için son zamanlarda tercih edilen bir biyometrik kimlik tespit sistemidir (Tistarelli & Champod, 2017). Yüz tanıma, askeri, güvenlik ve hukuki amaçların yanı sıra personel katılım/devam takibi gibi ticari amaçlarla kullanılabilen geniş bir uygulama sahasına sahiptir. Popüler bir araştırma konusu olmasına rağmen 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 (Hassaballah & Aly, 2015).

Görünüş tabanlı bir otomatik yüz tanıma sistemi genellikle dört ana bileşenden oluşmaktadır: yüz algılama, hizalama, öznelik çıkarma ve sınıflandırma (Li & Jain, 2011). Geleneksel yüz tanıma yöntemlerinde her bir yüz imgesi bir veri örneği olarak ele alınmaktadır. Veri sınıflarının ayrıştırılması da ayırt edici alt uzaylar belirlenerek yapılmaktadır. Buradaki en önemli problem veri boyutunun oldukça büyük olması (örnek olarak 100×90 boyutlarındaki gri seviye bir yüz imgesi, \mathbb{R}^{9000} uzayında bir vektör olarak temsil edilmektedir) ve örnek sayısının oldukça küçük olmasıdır (genellikle veritabanında bir bireye ait olan sadece birkaç adet imge mevcuttur). Bu problemlerin üstesinden gelen ve görünüş tabanlı yüz tanıma sistemlerinin en klasik yöntemleri olarak bilinen yöntemler temel bileşen analizi (Principal Component Analysis-PCA) (Turk & Pentland, 1991) ve doğrusal diskriminant analizi (Linear Discriminant Analysis-LDA) (Chelali et al., 2009) tabanlı yöntemlerdir. Bu tür yaklaşımlarda, eğitim verilerinden elde edilen optimal alt uzaylar kullanılarak yüz imgeleri sınıflandırılır. İlerleyen yıllarda bu yöntemlerin doğrusal olmayan versiyonları (Kernel PCA (Scholkopf et al., 1997) ve Kernel LDA (Mika et al., 1999)) ve bunların birçok varyantları (Li et al., 2009; Liu, 2004; Lu et al., 2003; Nhat & Lee, 2007) önerilmiştir. Bu yöntemler genellikle fazla sayıda sınıf içermeyen durumlarda ve özellikle gri seviye öznelikleri yerine, veri uzayı üzerine izdüşürüldüğünde daha iyi ayrıştırılabilen Gabor dalgacık öznelikleri (Vinay et al., 2015), EGMB (elastic graph bunch matching) (Wiskott et al., 1997) veya LBP (local binary patterns) (Ahonen et al., 2006) öznelikleri ile kullanıldığında oldukça başarılı sonuçlar vermektedir.

Yüz tanımda tekil imgeleri kullanan klasik yöntemlerin etkinlikleri sınıf sayısı çoğaldığında veya sınıf içi imge varyasyonları fazlaştığında (örneğin veritabanında bir bireye ait az sayıda yüz imgesi yerine aynı kişiye ait aydınlatma şiddeti farklılıkları, poz görünüş farklılıkları içeren yüzlerce imge bulunması durumunda) azalmaktadır. Bu durumlarda, yakın geçmişte önerilmiş olan ve tekil imgelerden alt uzaylar oluşturmak yerine, aynı sınıfa ait olan imgelerin birleşimini bir imge kümesi olarak ele alıp, imge kümelerini modelleyen küme tabanlı yaklaşımlar kullanılmaktadır (Lu et al., 2017; Yang et al., 2017; Zhao et al., 2018). Küme tabanlı yaklaşımlarda, küme modellemeleri genellikle ayrıık imgelerin bir öznelik uzayında noktalar biçiminde, imge kümelerinin ise küme içindeki noktaları içine alan bölgeler biçiminde temsil edilmesiyle gerçekleştirilir. Böylece, her bir kimlik çok boyutlu öznelik 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 ulaşılabilir özelliğidir. Standart ve kontrollü ortamlar yerine video kamerasının yerleştirildiği herhangi bir noktadan kayıt yapılabilen ve yeterli büyüklüklerde veri elde edilebilmektedir.

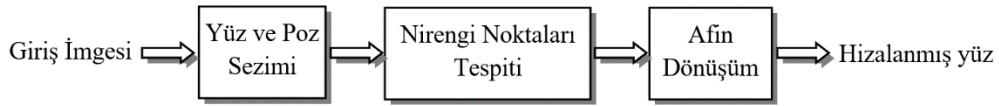
Video görüntüleri tekil imgelere göre çok daha fazla bilgi içermektedir. Bu da yüz tanıma sistemlerinde başarımları arttırıcı yönde etki göstermektedir. Ayrıca büyük veri işleme kapasitesi günümüzdeki teknolojiyle çok daha kolay hale gelmiştir. Cevikalp ve Triggs (2010) tarafından önerilen yöntemde, video görüntülerinden elde edilen imge dizilerine hesapsal yaklaşımlar, imge kümeleri ile yüz tanımanın temellerini oluşturmaktadır. İmge kümeleri kullanılan sistemlerde kişi sınıfları imge kümelerinden öğrenilen alt uzaylarla temsil edilirler (Wang et al., 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 et al., 2013). Kişilere ait imge kümelerinin sınıflandırılması için bazı benzerlik ölçütleri kullanılır ve karşılaştırma gerçekleştirilir. Benzerlik ölçüsü olarak doğrusal uzaylar kullanılmışsa açısal uzaklıklar, afin uzaylar kullanılmışsa geometrik uzaklıklar (bölgeler arası en yakın mesafe uzunluğu) kullanılmaktadır. İmge kümeleri tabanlı yöntemlerin farklılıkları nokta bu kümelerin modellemesi ve kümeler arası uzaklık ölçüsüdür.

İmge kümeleri ile yüz tanıma çalışmalarında karşılaşıldığı kadarıyla imge kümeleri oluşturulurken yüz sezme algoritmalarının döndürdüğü çıktılar ya hiç hizalanmadan ya da kısmen insan eliyle hizalanarak kullanılmıştır. Yüz tanıma başarımında yüz hizalamanın etkisi önemli olabilmektedir (Shi et al., 2018). Kameradan yakalanan görüntülerde yüzün kameraya uzaklığına göre değişim gösteren ölçekleme farklılıkları, yüzün dönme açısı veya eğiklikleri anıma 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, et al., 2013).

Bu çalışmada, video görüntülerinden yakalanan yüz imgeleri kullanılarak yüz tanıma uygulaması yapılmıştır. Farklı aydınlatma ve yüksek poz çeşitlilikleri altında kaydedilen video görüntüleri üzerine bir önışleme yöntemi tanımlanmış; bu sayede aydınlatma ve poz farklılıklarının doğru tanıma üzerindeki olumsuz etkisi azaltılmaya çalışılmıştır. Önerilen yöntem ikinci bölümde tanımlanmıştır. Deneysel çalışmalar, bulgular ve tartışma üçüncü bölümde verilmiştir. Dördüncü bölümde çalışmanın ana sonuçları özetlenmiştir.

2. Yöntem

Otomatik yüz hizalama süreci geleneksel olarak yüz sezimi, sezilen yüz bölgesi içinde göz, ağız gibi bazı önemli noktaların bulunması ve yüz bölgesinin önceden belirlenen bir yüz şablonuna haritalanması biçiminde gerçekleştirilir. Yüz sezicilerin başarımları, kameranın yüze bakış açısı, aydınlatma şiddeti farklılıkları, yüz görünümünü kapatan diğer nesnelere varlığı gibi durumlardan olumsuz yönde etkilenmektedir. Genellikle ortalama ışık altında ve ön cephe görünümü pozlarda tüm seziciler çok başarılı sonuçlar döndürmesine rağmen poz açısı değiştiğinde yöntemlerin başarımları düşmektedir. Poz değişikliklerine ilave olarak yüksek ışık veya karanlık ortamlarda alınan çekimlerde yukarıda bahsedilen sezici yöntemlerin başarımları oldukça kötüleşmektedir. Bu çalışmada, özellikle aydınlatma farklılıkları ve poz çeşitlilikleri söz konusu olduğunda, otomatik yüz hizalama gerçekleştirmek üzere, Şekil-1’de blok diyagramı verilen dört aşamalı ön işleme yöntemi önerilmiştir. Giriş imgesinde ilk olarak Haar benzeri öznelikler (Lienhart & Maydt, 2002) kullanarak eğitimi gerçekleştirdiğimiz bir yüz sezici algoritma ile imgedeki yüz bölgesi belirlenir. Tespit edilen yüz bölgesi kesildikten sonra yüz bölgesindeki gölgelenmeler gibi tanıma üzerinde olumsuz etki oluşturabilecek aydınlatma farklılıklarının etkisini ortadan kaldırmak için histogram eşitleme yapılır. Bir sonraki aşamada, yüzdeki nirengi noktaları tespit edilir. Bu noktaların konumları kullanılarak yüzün önceden belirlenmiş olan yedi poz kategorisindeki hangi poz sınıfına ait olduğu tespit edilir. Son aşamada ise yüzdeki nirengi noktaları, ilgili poz sınıfındaki kanonik şablon konumlarına eşleşecek şekilde bir konformal dönüşüm üzerinden geçirilerek yüz hizalama tamamlanır.

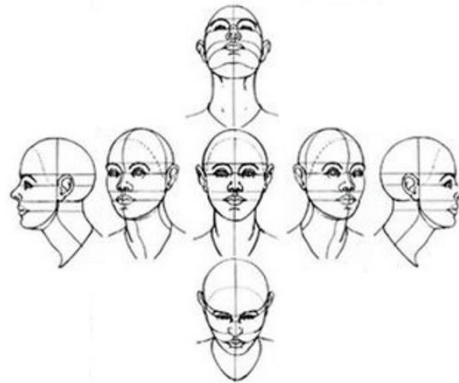


Şekil 1. Ön işleme yöntemi basamakları

2.1. Yüz ve Poz Sezimi

Yüz sezimi, imge içindeki yüz bölgelerinin tespit edilmesi olarak tanımlanır ve yüz sezme amacıyla kullanılan çok farklı yöntemler mevcuttur (Yang et al., 2002). Bu yöntemler içinde başarımları ve hız performansı oldukça iyi olan Viola-Jones yüz sezicisi (Viola & Jones, 2001) bilinen en etkili yüz sezicilerindendir. Etkili diğer bazı yüz sezici algoritmalarında, yüz bölgesi tespit edilirken yüzü oluşturan önemli nirengi noktaları da tespit edilebilmektedir (Xiong & De La Torre, 2013; Zhu & Ramanan, 2012). Genellikle göz, kaş, ağız, burun ve yüz sınırından oluşan bu noktalar, yüzle ilgili detaylı bilgiler de sunmaktadır. Son zamanlarda konvolüsyonel sinir ağları mimarisine dayanan derin öğrenme tekniklerinin nesne tespiti uygulamalarında kullanımıyla oldukça başarılı yüz sezicileri tasarlanmıştır. Açık kaynak kodlu bir yazılım olarak da kullanıcılarla paylaşılan Face++ yöntemi (Zhou et al., 2013) yüz sezimi ve nirengi noktaları tespit edilmesinde oldukça başarılıdır.

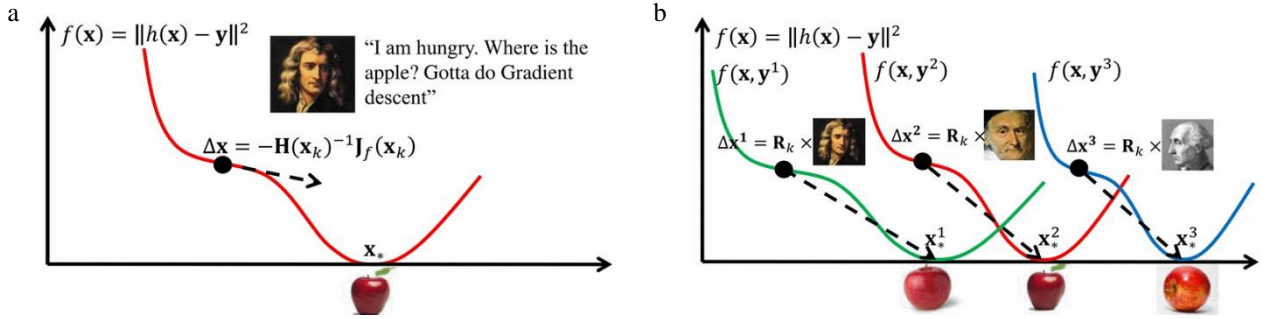
Yüksek aydınlatma farklılıkları ve pozlama çeşitlilikleri altında bahsedilen yöntemlerin yüz sezimi probleminde bazı durumlarda yetersiz kaldığı gözlenmiştir. Bu yüzden elde edilen veriler de kullanılarak, zor koşullar altında da başarılı sezimler yapabilen bir yüz sezicisi eğitilmiştir. Öncelikli olarak yüz sezimi problemi, 7 sınıflı bir nesne sezimi problemi olarak ele alınmıştır. Her bir sınıf, bir poz kategorisi sınıfı olarak tasarlanmış ve poz sınıfı temsili görselleri Şekil 2’de verilmiştir. Sınıflar: 1)Ön cephe, 2)yarım sol profil, 3)tam sol profil, 4)yarım sağ profil, 5)tam sağ profil, 6)yukarı bakış 7)aşağı bakış biçimindedir. Böylece bu nesne sınıflarından birisi sezildiğinde hem yüz sezimi hem de ilgili yüzün poz tespiti yapılmaktadır.



Şekil 2. Temsili poz sınıfları, toplam 7 sınıf (Sakarya, 2013).

2.2. Şablona Haritalama

Şablona haritalama süreci, sezilen yüz bölgesindeki göz, ağız, burun ve benzeri nirengi noktalarının tespit edilmesi ve tespit edilen noktaların önceden belirlenmiş olan bir kanonik şablona haritalanması işlemlerinden oluşmaktadır. Yüz sezimi başarıyla yapıldıktan sonra, yüzdeki nirengi noktalarının tespit edilmesi daha kolay bir problemidir. Bu konuda yapılan çeşitli çalışmalar mevcuttur (Wu & Ji, 2018). Değişken veya sabit sayılarda tespit edilen nirengi noktalarının konumları, yüz tanıma problemlerinde doğrudan öznel olarak veya bu konumlarla belirlenen bölgelerden öznel olarak çıkarılarak kullanılmaktadır. Biz bu çalışmada nirengi noktalarının konumlarını yüz hizalama amacıyla kullanmaktayız. Bu yüzden, nirengi noktalarını tespit etmek için bu konuda oldukça başarılı olan; sabit sayıda nirengi noktası konumu döndüren ve açık kaynak kodlu olarak akademik kullanıma sunulan INTRAFACE (Intraface, 2018) yöntemini kullandık. Bu yöntem eğitmenli iniş metodu (Supervised Descent Method-SDM) kullanılmaktadır (Xiong & De La Torre, 2013). Denetimli bir şekilde genel iniş haritaları öğrenerek, nirengi noktalarını başarılı bir şekilde tespit edebilmeyi amaçlar. Yöntemin ana fikri Şekil 3’de özetlenmiştir.



Şekil 3. Eğitmenli iniş yönteminin özeti (Xiong & De La Torre, 2013)

Şekil 3.a’da doğrusal olmayan en küçük kareler probleminin Newton yöntemiyle çözümü gösterilmektedir. Burada $f(\mathbf{x})$ imge öznel özelliklerini içeren doğrusal olmayan fonksiyon, \mathbf{y} bilinen bir vektör (şablon) ve \mathbf{x} hareket parametreleri (döndürme, ölçek, katı olmayan hareket) olmak üzere geleneksel Newton yöntemiyle Hessian ve Jacobian matrisleri hesaplanmaktadır. Şekil 3.b’de ise Eğitmenli İniş Metodunun genel yapısı betimlenmiştir. Eğitim verisinde gerçek nirengi konumları bilinen farklı örnekler için bir dizi fonksiyon $\{f(\mathbf{x}, \mathbf{y}^i)\}$ oluşmakta ve minimum değerler (\mathbf{x}_*^i) bilinmektedir. SDM bu eğitim verisini kullanarak doğrusal 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. Test sırasında ise aranan optimum \mathbf{x} , verilen \mathbf{y} vektörü için eğitimde bulunan genel iniş yönünde \mathbf{y} ’ye özgü bileşene haritalanmasıyla hesaplanır. Yöntemle ilgili detaylı bilgilendirme için bkz. (Xiong & De La Torre, 2013).

Yüz bölgesinde N adet nirengi noktasının konumları tespit edildikten sonra, şablona haritalama süreci 6 parametrelili afin dönüşüm kullanılarak yapılmıştır. Basit olarak ifade etmek gerekirse, birinci imgede bulunan nirengi noktalarının konumları (x_i, y_i) , $i=1,2,\dots,N$ ve aynı noktaların hedef şablondaki yeni konumları $(\tilde{x}_i, \tilde{y}_i)$, $i=1,2,\dots,N$ olsun. Uygulanan afin dönüşüm, $\mathbf{p}=(x_i, y_i, 1)$ konumunu, $\tilde{\mathbf{p}}=(\tilde{x}_i, \tilde{y}_i, 1)$ konumuna haritalamak üzere tasarlanmakta ve Denklem 1’de gösterildiği şekilde sadeleştirilmektedir.

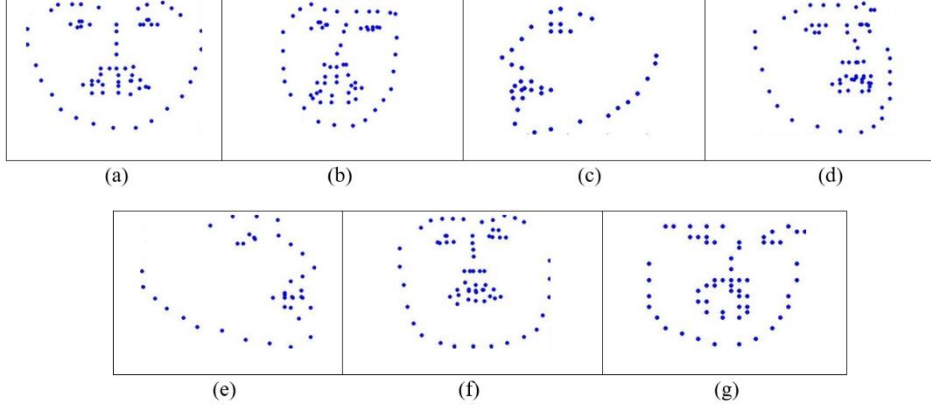
$$\begin{bmatrix} x_1 & y_1 & 1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ x_N & y_N & 1 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \end{bmatrix} = \begin{bmatrix} \tilde{x}_1 \\ \cdot \\ \cdot \\ \tilde{x}_N \end{bmatrix} \quad \text{ve} \quad \begin{bmatrix} x_1 & y_1 & 1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ x_N & y_N & 1 \end{bmatrix} \begin{bmatrix} a_{21} \\ a_{22} \\ a_{23} \end{bmatrix} = \begin{bmatrix} \tilde{y}_1 \\ \cdot \\ \cdot \\ \tilde{y}_N \end{bmatrix} \quad (1)$$

Bu eşitlik sistemi $N \geq 3$ için, en küçük hata kareleri toplamı çözümü kullanılarak, dönüşüm parametrelerini $(a_{11}, a_{12}, a_{13}; a_{21}, a_{22}, a_{23})$ hesaplamak üzere çözdürülmektedir. Elde edilen dönüşüm matrisi \mathbf{T} ile gösterildiğinde, Denklem 2’de verildiği formda olacaktır. Nirengi noktalarının konumları kullanılarak hesaplanan bu dönüşüm, yüz imgesindeki diğer tüm konumlara uygulandığında yüz imgesi, belirlenen şablona göre hizalanmış olur.

$$\mathbf{T} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Bu çalışmada 7 poz kategorisi belirlendiği için her bir poz kategorisinin şablonu farklıdır. Yüz imgeleri, önceki bölümde verilen yüz ve poz sezimi süreci çıktısı olarak hangi poz kategorisi olarak sınıflandırılmışlarsa, ilgili kategorideki şablona haritalanmaktadır. Poz kategorilerinin kanonik şablonları ise tarafımızca manuel olarak yapılan deneylerde elde edilen konumların ortalamaları alınarak belirlenmiştir. Bunun için öncelikle veritabanından yüz biçimleri farklı olan 30 kişi seçilmiş ve

bu bireylerin 7 farklı poz kategorisindeki 10'ar imgeleri kullanılarak nirengi noktaları INTRAFACE algoritmasıyla tespit edilmiştir. Nirengi noktaları belirlenen yüz imgelerinde ilave olarak yüz bölgesini belirleyen dikdörtgen çerçeve kenarları elle işaretlenmiştir. Yüz bölgesi dikdörtgenin uzun kenar ve kısa kenar uzunlukları oranı sabit olacak şekilde standart bir değere ölçeklenmiş ve nirengi noktalarının merkezi referans noktası olacak şekilde bir çerçeveye oturtulmuştur. Tüm imgeler için nirengi noktaları konumlarının ortlaması, kanonik şablon nirengi noktası olarak kabul edilmiştir. Her bir poz sınıfı için ortaya çıkan kanonik şablonlar Şekil 4'de verilmiştir. Bu şablonlardaki her bir nokta, Denklem-1'de konumu (\hat{x}_i, \hat{y}_i) biçiminde tanımlanan değişkenlerdir.



Şekil 4. Poz sınıfları için kanonik şablonlar: (a) ön cephe, (b) yarım sol, (c) tam sol, (d) yarım sağ, (e) tam sağ, (f) yukarı, (g) aşağı.

3. Yüz Tanıma Deneyleri

İmge kümeleri tabanlı yüz tanıma uygulamaları, video görüntülerinden yüz kareleri yakalayıp her bir videoyu bir imge kümesi olarak modelleme amacı güder. Bu uygulamalar yakın zamanlarda literatürde denenmeye başlamıştır ve tanıma performansları genellikle Honda/UCSD (Lee et al., 2005), MoBo (Gross & Shi, 2001) ve Youtube Celebrities (Kim et al., 2008) gibi küçük hacimli veri kümeleri üzerinde test edilmiştir. İmge kümeleri oluşturulurken, çalışmaların çoğunda yüz sezicinin döndürdüğü çıktılar, herhangi bir ilave işleme yapılmadan kullanılmaktadır. Fazla sayıda katılımcı içeren, büyük hacimli ve aydınlatma şiddeti ve poz farklılığı gibi zorlayıcı durumlar içeren videolar söz konusu olduğunda küme tabanlı tanıma yöntemlerinin başarımları düşmektedir. Bu çalışmada, küme tabanlı yüz tanıma yöntemlerinde imge kümeleri oluşturulurken, önerilen yöntemle hizalanmış yüz imgeleri kullanıldığında tanıma yöntemlerinin başarımları üzerinde olumlu etkileri olup olmadığı araştırılmıştır. Deneyler, geniş hacimli ve zorlayıcı durumları içinde barındıran ESOGÜ-285 video veritabanı üzerinde gerçekleştirilmiştir.

3.1. Veritabanı

ESOGU-285 video veritabanı (Yalcin et al., 2015), yüz ve mimik çeşitliliklerini içeren kısa videolardan (yaklaşık 1 dakikalık süreler) oluşmaktadır. Bireylerin, iki farklı oturumda, 4'er farklı senaryo altında video çekimleri mevcuttur. Birinci senaryo, Honda/UCSD veritabanının tasarımına benzetilerek doğal ışık ortamında kişinin yüzünü yavaşça sağa, sola ve aşağı yukarı hareket ettirecek yüzün çeşitli açılardan görüntülediği senaryodur. İkinci senaryo, yine doğal ışık ortamında ancak bu kez konuşma ve mimiklerin yoğun olarak canlandırıldığı bir senaryodur. 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. Veritabanında, 285 kişi için iki oturumda 4'er senaryo (toplam 8 senaryo) çekimlerinden oluşturulan veritabanı 2280 videodan oluşmaktadır. Veritabanından bir oturum için senaryo örnekleri Şekil 5'de gösterilmiştir.



Şekil 5. ESOGU-285 video veritabanındaki 4 temel senaryo

3.2. İmge Kümeleri Tabanlı Tanıma Yöntemleri

İmge kümeleri tabanlı yüz tanıma kavramı, ilk olarak Cevikalp and Triggs (2010) tarafından önerilmiştir. Genellikle bir videodan yakalanan aynı bireye ait yüz imgeleri, basit geometrik yapılar (afin veya konveks zarflar) kullanılarak modellenir. Eğitim aşamasında, eğitim verileri ile öğrenilen modeller, test aşamasında test verisi ile karşılaştırılırken geometrik yapılar (manifoldlar) arasındaki uzaklıklara bakılır. Test verisi, eğitim aşamasında öğrenilen modellerden hangisine en yakınsa, ilgili sınıfa atanır. Doğrusal konveks modellerin uygun olmadığı durumlarda ise kernel hilesi kullanılmakta ve bu sayede doğrusal olmayan imge manifoldlarının doğru bir şekilde modellenmesi sağlanmaktadır.

Bir bireye ait yüz imgesinin öznitelik vektör gösterimi $\mathbf{x}_{ci} \in \mathcal{R}^d$ biçiminde olsun. Burada $c = 1, \dots, C$ sınıf numarasını; $i = 1, \dots, n_c$ ise c 'inci sınıfta bulunduğu varsayılan toplam n_c adet imge örneğini belirten örnek numarasını temsil eden indis, d ise imge vektörü uzayının boyutunu ifade eden değişkendir. Yöntem, eğitim verisi olarak ayrılan örnekleri kullanarak bir geometrik belirlemeyi amaçlar. Geometrik modeller basit afin veya konveks zarf modelleri olarak seçilmektedir. Herhangi bir sınıfı temsil edecek olan afin (affine) model, Eşitlik 3'de verildiği üzere, veri örneklerini içerecek olan en küçük alt uzay biçiminde tanımlanmaktadır.

$$H_c^{afin} = \left\{ \mathbf{x} = \sum_{i=1}^{n_c} \alpha_{ci} \mathbf{x}_{ci}; \left| \sum_{i=1}^{n_c} \alpha_{ci} = 1 \right. \right\}, \forall c \quad (3)$$

Bu tanımın parametrik karşılığı Eşitlik 4'de verilmiştir. Burada $\boldsymbol{\mu}_c \in \mathcal{R}^d$: ilgili sınıf örneklerinin ortalaması ($\boldsymbol{\mu}_c = \frac{1}{n_c} \sum_{i=1}^{n_c} \mathbf{x}_{ci}$),

$\mathbf{U}_c \in \mathcal{R}^{d \times l}$: afin uzayı geren birimlik vektörlerden seçilen ilk l ($l \ll d$) adet taban vektörlerinin oluşturduğu birimlik taban, $\mathbf{v}_c \in \mathcal{R}^l$: örneklerin \mathbf{U}_c dönüşümü üzerinden indirgenmiş koordinatlardaki karşılıklarını veren parametre vektörüdür. \mathbf{U}_c dönüşüm matrisi, sıfır ortalamalı veri örnekleri ($\{\mathbf{x}_{c1} - \boldsymbol{\mu}_c; \mathbf{x}_{c2} - \boldsymbol{\mu}_c; \dots; \mathbf{x}_{cn_c} - \boldsymbol{\mu}_c\}$) üzerinde tekil değer ayrıştırması yapılarak bulunur. Tekil değer ayrıştırması sonucunda bulunan tüm özvektörler yerine, indirgenmiş boyutlar elde edebilmek için özdeğerleri büyükten küçüğe doğru sıralandığında ilk l adet etkili özdeğere karşılık gelen özvektörler kullanılır. Burada, sıfıra karşılık gelen özvektörler, sıfır uzayını girdiği için kullanılmaz. Özdeğerleri sıfıra yakın olan birçok vektör de parazit, gürültü gibi gereksiz bilgileri temsil etmekte olduğu için kullanılmaz. Kullanılan özvektör sayısı l , veri uzay boyutunun yanında oldukça küçük bir değerdir.

$$H_c^{afin} = \{ \mathbf{x} = \boldsymbol{\mu}_c + \mathbf{U}_c \mathbf{v}_c \}, \forall c \quad (4)$$

Konveks zarf modelleri ise, veri örneklerinin afin uzaylar kadar geniş olmayan, veri örneklerini içine alan en küçük konveks bölge modellerini belirleme amacı güder. Herhangi bir sınıfı temsil edecek olan konveks model, Eşitlik 5'de tanımlanmıştır. Afin zarf modelleriyle karşılaştırıldığında en önemli farklılığın, veri örneklerinin önünde verilen parametrelerdeki kısıtlamadır. Afin modelde, bu çarpanlar, toplamları 1 olacak şekilde herhangi bir değerde seçilebilir ($-\infty < \alpha_{ci} < \infty$). Konveks modelde ise çarpanların pozitif olma zorunluluğu vardır ($0 \leq \alpha_{ci}$).

$$H_c^{konveks} = \left\{ \mathbf{x} = \sum_{i=1}^{n_c} \alpha_{ci} \mathbf{x}_{ci}; \sum_{i=1}^{n_c} \alpha_{ci} = 1, 0 \leq \alpha_{ci} \right\}, \forall c \quad (5)$$

Konveks zarf modellerinin parametrik ifadesi basitçe, eşitlik 6'da verildiği biçimde gösterilebilir. Burada $\mathbf{X}_c \in \mathfrak{R}^{d \times n_c}$ matrisi, sütunları c 'nci sınıfın veri örneklerini içeren veri matrisi, $\boldsymbol{\alpha}_c \in \mathfrak{R}^{n_c}$ ise Eşitlik 5'de verilen kısıt altında hesaplanan katsayılardan oluşturulan katsayı vektörüdür. Farklı iki sınıfa ait olan konveks modeller H_i ve H_j olarak verilirse, bu modellere karşılık gelen katsayılar Eşitlik 7'de verilen optimizasyon problemi çözdürülerek hesaplanır.

$$H_c^{konveks} = \{ \mathbf{x} = \mathbf{X}_c \boldsymbol{\alpha}_c \} \quad (6)$$

$$(\boldsymbol{\alpha}_i^*, \boldsymbol{\alpha}_j^*) = \arg \min_{\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j} \left\{ \left\| \mathbf{X}_i \boldsymbol{\alpha}_i - \mathbf{X}_j \boldsymbol{\alpha}_j \right\|^2; \sum_{k=1}^{n_i} \alpha_{ik} = 1, 0 \leq \alpha_{ik} < \infty; \sum_{k=1}^{n_j} \alpha_{jk} = 1, 0 \leq \alpha_{jk} < \infty \right\} \quad (7)$$

Bu problem, klasik destek vektör makineleri optimizasyonuna oldukça benzemektedir. Bu yüzden, ayrıntıları (Cevikalp & Triggs, 2010) kaynakçasında belirtildiği üzere, bir destek vektör makineleri eğiticisi kullanılarak problem çözdürülebilir.

Afin veya konveks yapılar, eğitim verisi kullanılarak, her bir sınıf için yukarıda belirtilen yöntemlerle belirlendikten sonra bu modeller, ilgili sınıfın imge kümelerini temsil eden matematiksel modeller olarak kullanılır. Test aşamasında, hangi sınıfa ait olduğu bilinmeyen bir imge kümesinin modeli, eğitim aşamasında belirlenen modellere olan uzaklığı cinsinden karşılaştırılır ve uzaklığı en küçük olan eğitim modelinin sınıfına atanır.

3.3. Uygulama

Bu çalışmada, tasarlanan yüz hizalama sürecinin, küme tabanlı yüz tanıma yöntemlerinin verimliliği üzerine etkileri test edilmiştir. Küme tabanlı tanıma yöntemi olarak, güncel birçok yöntem kullanılmıştır. Bu yöntemler, Mutual Supspace Metodu (MSM) (Yamaguchi et al., 1998), Afin Zarf Modelleme (AZM) yönteminin doğrusal ve kernel modelleri (Cevikalp & Triggs, 2010), Konveks Zarf Modelleme (KZM) yönteminin doğrusal ve kernel modelleri (Cevikalp & Triggs, 2010), Regularized Nearest Points (RNP) (Yang et al., 2013), Sparse Approximation Nearest Points (SANP) (Hu et al., 2012), Manifold-to-Manifold distance (MMD) (Wang et al., 2012) yöntemleridir. Literatürde, yöntemlerin başarımları test edilirken imge kümeleri çoğunlukla elle kesilmiş olan yüz imgeleri kullanılarak veya doğrudan yüz sezicinin döndürdüğü çıktılar kullanılarak gerçekleştirilmiştir (yüz hizalama yapılmamıştır). Bu çalışmada, önerilen yöntem kullanılarak yüz hizalama yapıldıktan sonra imge kümeleri oluşturulmuştur. Önerilen yöntemin olumlu etkileri olup olmadığı, elle kesilen yüz imgeleri ile oluşturulan imge kümelerinin tanıma sonuçlarıyla karşılaştırılmıştır.

İmge kümeleri oluşturma süreci her iki durum için de benzer süreçlerle işletilmiştir. Birinci durumda, Şekil 5'de örnekleri gösterilen veritabanı imgelerinde yüz bölgeleri bir uzman tarafından elle işaretlenerek kesilmiş, kesilen bölge 8-bit gri seviye imgeye dönüştürülmüş, aydınlatma şiddeti farklılıklarının etkisi ortadan kaldırılmak üzere histogram eşitleme yapılmış ve son olarak, tanıma yöntemleri çalıştırılırken herhangi bir hafıza sorunu yaşanmaması için 40×30 gri seviye imgeleri olarak ölçeklenmiştir. İkinci durumda, veritabanı üzerinde önerilen yöntem Şekil 1'de verilen basamaklar üzerinden uygulanarak pozlama durumuna göre uygun şablona haritalanarak hizalanmış yüz imgeleri oluşturulmuştur. Daha sonra hizalanmış yüz imgesi, birinci durumda kullanılan aynı strateji üzerinden 8-bit gri seviye imgeye dönüştürülmüş, aydınlatma şiddeti farklılıklarının etkisi ortadan kaldırılmak üzere histogram eşitleme yapılmış ve son olarak 40×30 boyutlarına ölçeklenmiştir. Veritabanındaki her bir video için bu süreç işletilmiş ve ilgili video çerçevelerinden elde edilen 40×30 gri seviye imgeleri topluluğu, o videonun imge kümesini oluşturmuştur.

3.4. Bulgular ve tartışma

Küme tabanlı yüz tanıma deneylerinde, ESOGU-285 veritabanı için (Yalcin et al., 2015) kaynakçasında önerilen test yapısı kullanılmıştır. Buna göre, veritabanının birinci oturumunda çekilen 4 video eğitim kümeleri olarak, ikinci oturumda çekilen 4 video ise test kümeleri olarak kullanılmıştır. Eğitim aşamasında, eğitim için ayrılan kümelere sınıf modelleri öğrenilmekte, test aşamasında ise test modeli ile eğitimde öğrenilmiş olan küme modelleri arasındaki uzaklıklara bakılmaktadır. Test kümesi ataması ise minimum uzaklık prensibine göre yapılmaktadır. Bu çalışmada kullanılan küme tabanlı yöntemler, farklı teoriler kullanarak modelleme yapmakta ve bu yüzden modeller arası uzaklık hesaplamasını farklı biçimlerde gerçekleştirmektedir. İmge kümeleri oluşturulurken hizalama yapılmayan birinci durumdaki tanıma oranları ve önerilen yöntemle hizalama yapılarak oluşturulduğu ikinci durum yüz tanıma başarımları Tablo 1'de verilmiştir.

Tablo 1. ESOGU-285 video yüz veritabanı üzerindeki küme tabanlı yüz tanıma oranları

Yöntem	Hizalanasız durum	Hizalamalı durum
	Doğru tanıma oranı (%)	Doğru tanıma oranı (%)
1. MSM	50,1	54,8
2. AZM	44,3	50,2
3. KZM	55,1	58,9
4. RNP	46,7	50,1
5. SANP	51,9	55,8
6. MMD	52,0	56,0
7. Kernel AZM	62,1	66,9
8. Kernel KZM	62,1	66,8

Tablonun birinci satırındaki sonuçlar, yüz tanıma literatüründe birçok araştırmacının imge kümelerini oluştururken, yüz sezicisinin döndürdüğü veya elle kesilen yüz bölgeleri kullanılarak (hizalama yapılmadan) elde edilen tanıma oranlarını, ikinci satırdaki sonuçlar ise önerilen yöntem ile hizalanmış imgeler kullanıldığında elde edilen tanıma oranlarını yüzde cinsinden göstermektedir. Elde edilen sonuçlar incelendiğinde, önerilen hizalama yöntemi uygulandığında tüm tanıma yöntemlerinin tanıma oranlarında iyileşmeler gözlenmiştir. Bu iyileşmeler, en az % 3,4 oranla RNP yönteminde; en çok % 5,9 oranla AZM yönteminde olmuştur. Bu çalışmada kullanılan sekiz yöntemin tanıma oranları, önerilen hizalama yöntemi ile hizalanan imge kümeleri ile karşılaştırıldığında, ortalama % 4,4 oranında artmıştır. Sonuç olarak, hizalanmış yüz imgeleri ile hizalanmamış yüz imgeleri, çeşitli imge kümeleri tabanlı yüz tanıma yöntemleriyle test edildiğinde, uygulanan görsel normalizasyon sürecinin tanıma yöntemlerinin başarımını iyileştirdiği gözlenmiştir.

4. Sonuç

Bu çalışmada, yüz tanıma uygulamalarında kullanılmak üzere bir otomatik yüz hizalama yöntemi önerilmiştir. Yüz hizalama aşaması kolay bir süreç olmaması sebebiyle araştırmacılar tarafından şablon eşleştirmeye dayanan yüz tanıma yöntemleri dışında çoğunlukla uygulanmamaktadır çünkü birçok tanıma yöntemi için hizalama gerekli değildir. Özellikle video karelerinden yakalanan yüz imgeleri ile yüz tanıma uygulamalarında, her bir video karesi için yüz sezicinin döndürdüğü yüz bölgeleri herhangi bir hizalama yapılmadan kullanılmaktadır (Hu et al., 2012; Wang et al., 2012; Yang et al., 2013). Önerilen hizalama yönteminin kullanılması, veri üzerinde bir çeşit görsel normalizasyon işlemi uygulanması anlamına gelmektedir. Hizalama sayesinde, video imgelerinde benzer yüz pozlarının aynı şablona haritalanması, veri örneklerinin sınıf içi benzerliklerini arttırmaktadır. Sınıf içi benzerliklerin artmasının tanıma oranlarında iyileştirmeler vermesi beklenmektedir. Yapılan yüz tanıma deneylerinde, hizalanmış ve hizalanmamış yüz imgeleri, literatürde yaygın olarak bilinen çeşitli küme tabanlı tanıma yöntemlerinde kullanılmıştır. Deney sonuçlarında, hizalanmış yüz imgeleri kullanıldığında daha başarılı tanıma oranları elde edildiği gözlenmiştir. Önerilen yöntemle hizalama gerçekleştirildiğinde, hizalanmamış imgelere göre tanıma oranlarında ortalama %4,4 artış elde edilmiştir. Bu durum hizalama sürecinin tanıma oranları üzerinde olumlu bir etkisinin bulunduğunu göstermektedir.

Teşekkür

Bu çalışma, TÜBİTAK tarafından 114E014 numaralı proje kapsamında desteklenmiştir.

Referanslar

- Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28 (12), 2037-2041. doi: 10.1109/tpami.2006.244
- Cevikalp, H., & Triggs, B. (2010). Face recognition based on image sets. In *2010 IEEE computer society conference on computer vision and pattern recognition*. IEEE. doi: 10.1109/cvpr.2010.5539965
- Chelali, F., Djeradi, A., & Djeradi, R. (2009). Linear discriminant analysis for face recognition. In *2009 international conference on multimedia computing and systems*. IEEE. doi: 10.1109/mmcs.2009.5256630
- Gross, R., & Shi, J. (2001). *The cmu motion of body (mobo) database* (Tech. Rep. No. CMU-RI-TR-01-18). Pittsburgh, PA: Carnegie Mellon University.
- Harandi, M. T., Sanderson, C., Shirazi, S., & Lovell, B. C. (2013). Kernel analysis on Grassmann manifolds for action recognition. *Pattern Recognition Letters*, 34 (15), 1906-1915. doi: 10.1016/j.patrec.2013.01.008

- Hassaballah, M., & Aly, S. (2015). Face recognition: challenges, achievements and future directions. *IET Computer Vision*, 9 (4), 614-626. doi: 10.1049/iet-cvi.2014.0084
- Hu, Y., Mian, A. S., & Owens, R. (2012). Face recognition using sparse approximated nearest points between image sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34 (10), 1992-2004. doi: 10.1109/TPAMI.2011.283
- Intraface. (2018, oct). *Intraface software, the human sensing laboratory of carnegie mellon university*. Retrieved from <http://www.humansensing.cs.cmu.edu/software>
- Kim, M., Kumar, S., Pavlovic, V., & Rowley, H. (2008). Face tracking and recognition with visual constraints in real-world videos. In *2008 IEEE conference on computer vision and pattern recognition*. IEEE. doi: 10.1109/cvpr.2008.4587572
- Lee, K.-C., Ho, J., Yang, M.-H., & Kriegman, D. (2005). Visual tracking and recognition using probabilistic appearance manifolds. *Computer Vision and Image Understanding*, 99 (3), 303-331. doi: 10.1016/j.cviu.2005.02.002
- Li, S. Z., & Jain, A. K. (Eds.). (2011). *Handbook of face recognition*. Springer London.
- Li, Z., Lin, D., & Tang, X. (2009). Nonparametric discriminant analysis for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31 (4), 755-761. doi: 10.1109/tpami.2008.174
- Lienhart, R., & Maydt, J. (2002). An extended set of haar-like features for rapid object detection. In *proceedings of international conference on image processing*. IEEE. doi: 10.1109/icip.2002.1038171
- Liu, C. (2004). Gabor-based kernel pca with fractional power polynomial models for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26 (5), 572-581. doi: 10.1109/tpami.2004.1273927
- Lu, J., Plataniotis, K., & Venetsanopoulos, A. (2003). Face recognition using kernel direct discriminant analysis algorithms. *IEEE Transactions on Neural Networks*, 14 (1), 117-126. doi: 10.1109/tnn.2002.806629
- Lu, J., Wang, G., & Zhou, J. (2017). Simultaneous feature and dictionary learning for image set based face recognition. *IEEE Transactions on Image Processing*, 26 (8), 4042-4054. doi: 10.1109/tip.2017.2713940
- Mika, S., Ratsch, G., Weston, J., Scholkopf, B., & Mullers, K. (1999). Fisher discriminant analysis with kernels. In *Neural networks for signal processing IX: Proceedings of the 1999 IEEE signal processing society workshop (cat. no.98th8468)*. IEEE. doi: 10.1109/nnsf.1999.788121
- Nhat, V. D. M., & Lee, S. (2007). Kernel-based 2dpc for face recognition. In *2007 IEEE international symposium on signal processing and information technology*. IEEE. doi: 10.1109/isspit.2007.4458104
- Sakarya, Z. (2013). Karakalem Yüz Çizme - Charcoal Face Drawing [blog post]. Retrieved from <http://karakalem-portre-photoshop.blogspot.com/2013/10/karakalem-yuz-cizme-charcoal-face.html>
- Scholkopf, B., Smola, A., & Muller, K.-R. (1997). Kernel principal component analysis. In *Lecture notes in computer science* (pp. 583-588). Springer Berlin Heidelberg. doi: 10.1007/bfb0020217
- Shi, B., Bai, X., Liu, W., & Wang, J. (2018). Face alignment with deep regression. *IEEE Transactions on Neural Networks and Learning Systems*, 29 (1), 183-194. doi: 10.1109/tnnls.2016.2618340
- Tistarelli, M., & Champod, C. (Eds.). (2017). *Handbook of biometrics for forensic science*. Springer International Publishing. doi: 10.1007/978-3-319-50673-9
- Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3 (1), 71-86. doi: 10.1162/jocn.1991.3.1.71
- Vinay, A., Shekhar, V., Murthy, K. B., & Natarajan, S. (2015). Face recognition using gabor wavelet features with PCA and KPCA - a comparative study. *Procedia Computer Science*, 57 , 650-659. doi: 10.1016/j.procs.2015.07.434
- Viola, P., & Jones, M. (2001). Robust real-time object detection. *International Journal of Computer Vision*, 57, 137-154. doi: 10.1023/B:VISI.0000013087.49260.fb
- Wang, H., Wang, Y., & Cao, Y. (2009). Video-based face recognition: A survey. *World Academy of Science, Engineering and Technology*, 2 , 136-139. doi: 10.1007/978-1-84882-385-3_8
- Wang, R., Shan, S., Chen, X., Dai, Q., & Gao, W. (2012). Manifold-manifold distance and its application to face recognition with image sets. *IEEE Transactions on Image Processing*, 21 (10), 4466-4479. doi: 10.1109/tip.2012.2206039

- Wiskott, L., Fellous, J.-M., Kruger, N., & von der Malsburg, C. (1997). Face recognition by elastic bunch graph matching. In *Computer analysis of images and patterns* (pp. 456-463). Springer Berlin Heidelberg. doi: 10.1007/3-540-63460-6_150
- Wu, Y., & Ji, Q. (2018). Facial landmark detection: A literature survey. *International Journal of Computer Vision*. doi: 10.1007/s11263-018-1097-z
- Xiong, X., & De La Torre, F. (2013). Supervised descent method and its applications to face alignment. In *Proceedings of the IEEE computer society conference on computer vision and pattern recognition* (pp. 532-539). doi: 10.1109/CVPR.2013.75
- Yalcin, M., Cevikalp, H., & Yavuz, H. S. (2015). Towards large-scale face recognition based on videos. In *2015 IEEE international conference on computer vision workshop (ICCVW)*. IEEE. doi: 10.1109/iccvw.2015.141
- Yamaguchi, O., Fukui, E., & Maeda, K. (1998). Face recognition using temporal image sequence. In *proceedings third IEEE international conference on automatic face and gesture recognition*. IEEE. doi: 10.1109/afgr.1998.670968
- Yang, M., Wang, X., Liu, W., & Shen, L. (2017). Joint regularized nearest points for image set based face recognition. *Image and Vision Computing*, 58, 47-60. doi: 10.1016/j.imavis.2016.07.008
- Yang, M., Zhu, P., Gool, L. V., & Zhang, L. (2013). Face recognition based on regularized nearest points between image sets. In *2013 10th IEEE international conference and workshops on automatic face and gesture recognition (FG)*. IEEE. doi: 10.1109/fg.2013.6553727
- Yang, M. H., Kriegman, D. J., & Ahuja, N. (2002). Detecting faces in images: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 (1), 34-58. doi: 10.1109/34.982883
- Yavuz, H. S., Cevikalp, H., & Edizkan, R. (2013). Automatic face recognition from frontal images. In *2013 21st signal processing and communications applications conference (SIU)*. IEEE. doi: 10.1109/SIU.2013.6531215
- Zhao, J., Han, J., & Shao, L. (2018). Unconstrained face recognition using a set-to-set distance measure on deep learned features. *IEEE Transactions on Circuits and Systems for Video Technology*, 28 (10), 2679-2689. doi: 10.1109/tcsvt.2017.2710120
- Zhou, E., Fan, H., Cao, Z., Jiang, Y., & Yin, Q. (2013). Extensive facial landmark localization with coarse-to-fine convolutional network cascade. In *Proceedings of the IEEE international conference on computer vision* (pp. 386-391). doi: 10.1109/ICCVW.2013.58
- Zhu, X., & Ramanan, D. (2012). Face detection, pose estimation, and landmark localization in the wild. In *Proceedings of the IEEE computer society conference on computer vision and pattern recognition* (pp. 2879-2886). doi: 10.1109/CVPR.2012.6248014