

İmge Setleriyle Yüz Tanıma

Face Recognition from Sets of Images

Hakan Çevikalp

Elektrik - Elektronik Mühendisliği Bölümü Eskişehir Osmangazi Üniversitesi
hakan.cevikalp@gmail.com

Özetçe

Bu bildiride çoğul imgeleri kullanan yeni bir yüz tanıma yöntemi önerilmiştir. Çoğul imgeler söz konusu olduğunda, yüz tanıma problemi bilinmeyen bir kişiye ait çoğul imgelerden oluşan bir imge setini alıp, eğitim setindeki etiketli imge setleri arasında bu imge setine en çok benzeyeni bulma olarak tanımlanır. Önerilen yöntemde her bir imge seti, bu seti oluşturan imgeler kullanılarak geometrik ve konveks bir modelle (affine ve konveks zarflar) yakınsanmıştır. Bu yaklaşım kullanılarak oluşturulan iki konveks model arasındaki uzaklık, bu modeller üzerindeki en yakın noktaları bularak belirlenmiştir. Çekirdek hilesi kullanılarak, yöntem doğrusal olmayan durumlar için genelleştirilerek, karmaşık ve doğrusal olmayan topolojiye sahip imge setlerinin doğru olarak modellenmesi sağlanmıştır. Farklı veri tabanları üzerinde yapılan deneylerde yöntemin literatürde bu amaçla kullanılan en iyi yöntemlerden daha iyi sonuçlar verdiği gözlenmiştir.

Abstract

This paper introduces a novel method for face recognition based on multiple images. When multiple images are considered, the face recognition problem is defined as taking a set of face images from an unknown person and finding the most similar set among the database of labeled image sets. Our proposed method approximates each image set with a geometric convex model (affine/convex hulls) by using the images in these sets. For any pair of models of this form, the distance between them is determined based on the distance between the closest points in these models. By using the kernel trick, the method is extended to the nonlinear case, which allows us to approximate and match complex and nonlinear face image manifolds. The experiments on different databases show that our proposed method outperforms the current state-of-the-art methods in many cases.

1. Giriş

Geleneksel yüz tanıma yöntemlerinde kişiler tanınırken tek bir test imgesi kullanılır ve bu yöntemler genellikle kontrollü ortamlarda çekilen imgelerle eğitilirler. Fakat, bu şekilde eğitilen yüz tanıma sistemleri, günlük hayatta farklı poz ve aydınlatma koşulları altında çekilen imgelerle (örnek olarak güvenlik kameraları yada film videolarından elde edilen imgeler gibi) pek iyi sonuçlar vermemektedirler.

Son zamanlarda tek bir imge yerine imge setlerini kullanan yüz tanıma sistemleri giderek önem kazanmaya başlamıştır [1-4]. Bu yöntemlerde bilinmeyen kişiye ait tek bir imge yerine birden fazla (çoğul) imge sağlanır; aynı şekilde eğitim setindeki kişilere ait birden fazla imge bulunur.

Yüz tanıma sistemi, eğitim setindeki imge setleri arasında bilinmeyen kişiye ait imge setine en çok benzeyen seti bulur ve bilinmeyen kişi bu imge setine karşılık gelen kişi olarak sınıflandırılır. Gerek eğitim gerekse test setindeki imge setlerini oluşturan yüz imgeleri farklı poz, aydınlatma ve ölçek farklılıkları içerebilir. Örnek olarak, güvenlik amaçlı çekilen videolarda bir kaç dakika boyunca alınan imgelerde arka plan değişmesi bile, imgelerde büyük ölçüde hem poz hem de ölçek farklılıkları görülür.

Yüz tanıma imge setlerinin kullanılmasının çeşitli faydaları vardır. Öncelikle imge setlerini kullanan yüz tanıma yöntemlerinin tek bir imge kullanan yöntemlerden daha başarılı olması beklenir. Çünkü, imge setleri 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 setleri kolaylıkla elde edilir. Bu yüzden imge setlerini kullanan yöntemlerle çalışmak daha doğal ve kolaydır.

İmge setlerini kullanarak yüz tanıma, geleneksel tek bir test imgesi kullanan yüz tanıma göre daha yeni bir gelişme olmakla birlikte, yüksek çözünürlüklü video kameralar ve imge setlerini saklamak için gerekli büyük ölçekli depolama araçlarının gelişmesiyle popüleritesi artmıştır. Literatürde imge setlerini kullanan yüz tanıma yöntemleri, imge setlerini modelleme şekillerine göre ve bu setler arasındaki uzaklığı (yada benzerliği) ölçmek için seçilen ölçütlere göre farklılıklar göstermektedir. Fitzgibbon ve Zisserman [1] ardışıl video karelerinden elde edilen yüz imgelerini kişilere ait imge setlerini oluşturmada, imge setlerini de filmlerdeki ana karakterlerin otomatik olarak bulunmasında kullanmışlardır. Bu imge setleri affine zarflar kullanılarak yakınsanmış ve affine zarflar arasındaki uzaklıkları bulmak için *Joint Manifold Distance* adı verilen uzaklık ölçütü kullanılmıştır. Yamaguchi ve arkadaşları [2] ise Mutual Subspace yöntemi adını verdikleri yöntemde, imge setlerini doğrusal altuzaylar kullanarak modellemişler ve benzerlik ölçütü olarak bu altuzaylar arasındaki *kanonik (canonical) açılı* kullanılmışlardır. Daha yeni tarihli çalışmalarda ise, imge setlerini oluşturan yüz imgelerinin doğrusal olmayan düzgün (smooth) bir manifold oluşturduğu varsayılmış, ve bu doğrusal olmayan manifold, doğrusal lokal yamaların birleştirilmesi suretiyle yakınsanmıştır. Bu imge manifoldları arasındaki uzaklıklar lokal doğrusal yamalar arasındaki uzaklıklara bakılarak belirlenmiştir. Örnek olarak, Hadid ve Pietikainen [3] öncelikle Locally Linear Embedding

yöntemini kullanarak giriş uzay boyutunu indirgemişler ve daha sonra da k -en yakın merkez toplama algoritması kullanılarak lokal yamaları oluşturan yüz imgelerini belirlemişlerdir. Bu lokal yamalar arasındaki uzaklıkları ölçmek için toplama algoritmasıyla bulunan topak merkezleri arasındaki uzaklıklar kullanılmıştır. Wang ve arkadaşları [4] ise en yakın komşu benzerliklerini kullanan bir toplama algoritmasıyla lokal yamaları oluşturan örnekleri belirlemişler ve daha sonra bu lokal yamalar doğrusal altuzaylarla yakınsanmıştır. İmge manifoldları arasındaki uzaklıklar bu lokal altuzaylar arasındaki benzerlikler kullanılarak belirlenmiştir.

Bu çalışmada biz, yukarıda bahsedilen yöntemlerden farklı olarak, her bir imge setini bu sete ait örneklerin oluşturduğu basit geometrik konveks yapıları (affine ve konveks zarflar – affine/convex hulls) kullanarak modelledik. Bu modellerin arasındaki benzerlikleri ölçmek için de bu geometrik modeller arasındaki uzaklıkları kullandık. Doğrusal konveks modellerin uygun olmadığı durumlarda ise kernel hilesi kullanılmış ve bu sayede doğrusal olmayan karmaşık imge manifoldlarının doğru bir şekilde modellenmesi sağlanmıştır.

2. Yöntem

Yüz imge setlerini oluşturan imgeleri $\mathbf{x}_{ci} \in \mathfrak{R}^d$ ile gösterelim. Burada $c = 1, \dots, C$ imge setlerini (dolayısıyla kişileri) $i = 1, \dots, n_c$ ise c . imge setindeki örnek sayısını belirtmektedir. Önerilen yöntemde her bir imge seti konveks bir modelle (affine ve konveks zarflar) yakınsanmıştır. Tanıma sırasında bilinmeyen seti sınıflandırmak için bu konveks modeller arasındaki uzaklıklara bakılır ve test imge seti en yakın uzaklığı veren imge setine ait kişiye atanır. Konveks setler arasındaki uzaklığı bulmak için kullandığımız strateji aşağıdaki gibi özetlenebilir: Elimizde doğrusal olarak ayrıştırılabilen (linearly separable) iki konveks model H_i ve H_j olduğunu varsayalım. Bu iki model arasındaki uzaklık, bu modeller üzerindeki herhangi iki nokta arasındaki uzaklıkların en küçüğü ile ifade edilir, yani

$$D(H_i, H_j) = \min_{\mathbf{x} \in H_i, \mathbf{y} \in H_j} \|\mathbf{x} - \mathbf{y}\|. \quad (1)$$

2.1. Affine Zarf Modelleme

İlk olarak imge setlerinin bu setleri oluşturan örneklerin oluşturduğu affine zarflarla modellendiğini varsayalım. Affine zarflar imge örneklerinin oluşturduğu en küçük affine altuzay olup, aşağıdaki formülle bulunurlar

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

Normal altuzayların aksine affine zarfların orijinden geçme şartı olmadığından, affine zarflar yüz örneklerini içine alacak şekilde kaydırılmış altuzaylar olarak görülebilirler. Yukarıdaki formülden de görüldüğü üzere affine zarflar örneklerin affine kombinasyonlarını kullanarak sonsuz sayıda yeni örnekler oluşturmakta ve orijinal örneklerin altuzaydaki konumlarını kısıtlamamaktadırlar. Bu yüzden imge setlerini yakınsamak için oldukça gevşek sayılabilecek modellerdir.

Doğrusal olarak ayrıştırılabilen durumlarda affine zarflar üzerinden herhangi bir referans noktası $\boldsymbol{\mu}_c$ seçip, mesela affine zarfı oluşturan örneklerden herhangi biri yada onların

ortalaması gibi, affine zarfları aşağıdaki gibi yazmak mümkündür

$$H_c^{aff} = \left\{ \mathbf{x} = \boldsymbol{\mu}_c + \mathbf{U}_c \mathbf{v}_c \mid \mathbf{v}_c \in \mathfrak{R}^l \right\}, \quad c = 1, \dots, C. \quad (3)$$

Burada \mathbf{U}_c affine altuzayı geren birimlik vektörlerinin oluşturduğu matrisi, \mathbf{v}_c ise affine zarf içindeki örneklerin birimlik vektörlere göre koordinatlarını yazmak için gerekli katsayıları içeren vektörü ifade etmektedir. Numerik olarak \mathbf{U}_c matrisi $\{\mathbf{x}_{c1} - \boldsymbol{\mu}_c, \dots, \mathbf{x}_{cn_c} - \boldsymbol{\mu}_c\}$ fark vektörlerine tekil değer ayrıştırması (Singular Value Decomposition) uygulanarak bulunur. Genellikle sıfıra yakın çok küçük özdeğerlere karşılık gelen özvektörler verilerdeki gürültüye karşılık geldiğinden birimlik vektörlere dahil edilmez.

Elimizde doğrusal olarak ayrıştırılabilen iki affine zarf olduğunu ($\{\mathbf{U}_i \mathbf{v}_i + \boldsymbol{\mu}_i\}$ ve $\{\mathbf{U}_j \mathbf{v}_j + \boldsymbol{\mu}_j\}$) varsayalım (Affine zarfların doğrusal olarak ayrıştırılabilmesi bu affine zarfların kesişmediğini gösterir). Bu durumda bu affine zarfların üzerlerindeki en yakın iki nokta aşağıdaki optimizasyon problemi çözümlenerek bulunur

$$\min_{\mathbf{v}_i, \mathbf{v}_j} \|(\mathbf{U}_i \mathbf{v}_i + \boldsymbol{\mu}_i) - (\mathbf{U}_j \mathbf{v}_j + \boldsymbol{\mu}_j)\|^2. \quad (4)$$

Yukarıdaki optimizasyon probleminde $\mathbf{U} = (\mathbf{U}_i - \mathbf{U}_j)$ ve

$\mathbf{v} = \begin{pmatrix} \mathbf{v}_i \\ \mathbf{v}_j \end{pmatrix}$ olarak tanımlanırsa, aşağıdaki optimizasyon problemi elde edilir.

$$\min_{\mathbf{v}_j} \|\mathbf{U} \mathbf{v} + (\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)\|^2. \quad (5)$$

Bu optimizasyon problemini en iyileyen \mathbf{v} vektörü $\mathbf{v} = (\mathbf{U}^T \mathbf{U})^{-1} \mathbf{U}^T (\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)$ olacaktır. Bu durumda iki affine zarf arasındaki uzaklık aşağıdaki formülle bulunur

$$D(H_i^{aff}, H_j^{aff}) = \|(\mathbf{I} - \mathbf{P})(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)\|. \quad (6)$$

Burada \mathbf{P} matrisi her iki affine zarfı geren uzayın birimlik izdüşüm matrisini ifade eder ve nümerik olarak $\mathbf{P} = \mathbf{U}(\mathbf{U}^T \mathbf{U})^{-1} \mathbf{U}^T$ eşitliği ile bulunur. $(\mathbf{I} - \mathbf{P})$ matrisi ise bu uzayın sıfır altuzayının birimlik izdüşüm matrisidir.

2.2. İndirgenmiş (Reduced) Affine Zarf Modelleme

Affine zarfların doğrusal olarak ayrıştırılmadığı durumlarda (affine zarfların birbirleriyle kesiştiği durumlar), iki affine zarf arasındaki uzaklık her zaman sıfır olacağından bir önceki bölümde verilen formüller işe yaramayacaktır. Affine zarflar birkaç sebepten dolayı kesişebilirler: İlk olarak imge setlerini oluşturan imgeler arasında aykırı örnekler olabilir (yanlış yada kötü çekilmiş imgeler aykırı örnek olarak görülebilirler). Bu durumda aykırı değerlerin olduğu imge setlerine karşılık gelen affine zarflar çok büyük olacaktır. Bu problemi çözmek için daha gürbüz affine zarf modelleme yöntemleri kullanılmalı veya aykırı değerler tespit edilip imge setlerinden atılmalıdır. İkinci olarak yüz imgelerinden elde edilen öznitelikler kötü seçilmişse, bu öznitelikler imgeleri ayırtmak 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 affine zarflar imge setlerinin giriş uzayında gömülü olduğu

bölgelerin kestirimi için oldukça gevşek modellerdir ve imge setlerinin kesiştiği yada öznelik uzayında birbirine çok yakın olduğu durumlarda bunlara karşılık gelen affine zarflarda kesişebilir. Bütün bu problemleri tek bir hamlede çözmek için aşağıdaki anlatılan yaklaşımı kullandık.

Affine zarfları daha sıkılaştırmak amacıyla eşitlik (2)'de verilen α_{ck} katsayılarını sınırlandırırız, bu durumda indirgenmiş affine zarfları elde ederiz

$$H_c = \left\{ \mathbf{x} = \sum_{k=1}^{n_c} \alpha_{ck} \mathbf{x}_{ck} \mid \sum_{k=1}^{n_c} \alpha_{ck} = 1, L \leq \alpha_{ck} \leq U \right\}. \quad (7)$$

Bu eşitlikte (L, U) en alt ve en üst sınırları belirleyen parametrelerdir. Gevşek affine zarflarda bu sınırlar aktif değildir, yani $(L, U) = (-\infty, \infty)$. Eğer $L = 0$ olarak seçilirse indirgenmiş affine zarflar konveks zarflara karşılık gelir. Benzer şekilde altsınır $-\infty < L < 0$ ve üst sınır $U = \infty$ olarak seçilirse, buna karşılık gelen indirgenmiş affine zarf konveks zarflardan daha geniş olmasına rağmen normal affine zarflara oranla daha küçüktür.

İndirgenmiş affine zarflar içindeki noktalar $\mathbf{x} = \mathbf{X}_c \mathbf{a}_c$ olarak gösterilebilirler. Burada \mathbf{X}_c , sütun vektörleri imge setindeki yüz imgelerine karşılık gelen özellik vektörlerini içeren matrisi, \mathbf{a}_c ise α_{ck} katsayılarını içeren vektörü ifade etmektedir. İndirgenmiş affine zarfların üstündeki en yakın iki noktayı aşağıdaki karesel (quadratic) optimizasyon problemini çözerek buluruz

$$\begin{aligned} (\mathbf{a}_i^*, \mathbf{a}_j^*) &= \arg \min_{\mathbf{a}_i, \mathbf{a}_j} \|\mathbf{X}_i \mathbf{a}_i - \mathbf{X}_j \mathbf{a}_j\|^2 \\ \sum_{k=1}^{n_c} \alpha_{ik} &= \sum_{k'=1}^{n_c} \alpha_{jk'} = 1, L \leq \alpha_{ik}, \alpha_{jk'} \leq U. \end{aligned} \quad (8)$$

Yukarıdaki problem konveks bir problem olup, evrensel (global) optimal çözümü vardır. Optimal $\mathbf{a}_i^*, \mathbf{a}_j^*$ vektörlerinin bulunmasının ardından, indirgenmiş affine zarflar arasındaki uzaklık $D(H_i, H_j) = \|\mathbf{X}_i \mathbf{a}_i^* - \mathbf{X}_j \mathbf{a}_j^*\|$ formülü ile bulunur.

Eşitlik (8)'de verilen optimizasyon problemi imge setlerini oluşturan yüz örneklerinin iç çarpımları şeklinde yazılabilir. Bu durum bize, affine zarfların doğrusal olarak ayrıştırılmadığı durumlarda kernel hilesinin kullanılmasına olanak verir. Kernel hilesini kullanarak, yani örnekler arası iç çarpımları $\langle \mathbf{x}_{ck}, \mathbf{x}_{c'k'} \rangle$ kernel fonksiyonları $k(\mathbf{x}_{ck}, \mathbf{x}_{c'k'}) = \langle \phi(\mathbf{x}_{ck}), \phi(\mathbf{x}_{c'k'}) \rangle$ ile değiştirerek, örnekler giriş uzayı \mathfrak{R}^d 'den $\phi: \mathfrak{R}^d \rightarrow \mathfrak{Z}$ fonksiyonu kullanılarak çok daha büyük boyutlu bir uzaya aktarılmış olurlar. Bu şekilde doğrusal olmayan karmaşık topolojilere sahip imge setleri arasındaki uzaklıklar daha doğru bir şekilde yakınsanabilir.

2.3. Konveks Zarf Modelleme

Affine zarflara alternatif olarak imge setlerini modellemek amacıyla konveks zarflar kullanılabilir. Konveks zarflar imge örneklerinin konveks bileşimlerinden meydana gelir ve affine zarflardan tek farkları bileşen katsayılarının negatif olmamasıdır. Bir başka deyişle indirgenmiş affine zarflarda $L = 0$ alınır (bu durumda otomatikmen $U = 1$ olacaktır), konveks zarflar elde edilmiş olur

$$H_c^{con} = \left\{ \mathbf{x} = \sum_{k=1}^{n_c} \alpha_{ck} \mathbf{x}_{ck} \mid \sum_{k=1}^{n_c} \alpha_{ck} = 1, \alpha_{ck} \geq 0 \right\}, \quad c = 1, \dots, C. \quad (9)$$

Konveks zarflar imge setlerinin gömülü olduğu en küçük konveks yapılar olduklarından oldukça sıkı sayılabilecek yapılardır.

Eşitlik (8)'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 Destek Vektör Makineleri (DVM) -Support Vector Machines- 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 setini en iyi şekilde ayıran hiper-düzlemin normal vektörüdür.

Aykırı değerlerin etkisini azaltmak için indirgenmiş affine zarflarda olduğu gibi, üst sınır birden küçük bir sayıya atanır, $U < 1$. Kernel hilesi indirgenmiş affine zarflarda olduğu gibi konveks zarflarda da kullanılabilir ve bu şekilde doğrusal olmayan topolojiye sahip imge setleri arasındaki uzaklıklar daha doğru bir biçimde yakınsanabilir.

3. Deneysel Çalışmalar

Önerdiğimiz yöntemler olan Affine Zarf Modellemeye Dayalı İmge Set Tanıma (AZMDİST) ve Konveks Zarf Modellemeye Dayalı İmge Set Tanıma (KZMDİST) yöntemlerini Honda/UCSD [5] ve CMU MoBo [6] veri tabanları üzerinde test ettik. Bu veri tabanlarındaki yüz görüntülerini içeren videolardan, Viola ve Jones'un önerdiği yüz sezme yöntemi [7] kullanılarak yüzler bulunmuş ve imge setleri oluşturulmuştur. Doğrusal AZMDİST yönteminde affine altuzayları veren birimlik vektörler seçilirken özdeğerlerin toplam enerjisinin %98'lik bölümünü oluşturan en büyük özdeğerlere karşılık gelen özvektörler kullanılmıştır. Doğrusal olmayan kernel yöntemlerde ise kernel fonksiyonu olarak Gaussian fonksiyonu kullanılmıştır. Ayrıca doğrusal olmayan AZMDİST yönteminde α_{ck} katsayılarını sınırlandıran alt ve üst sınırlar eşit seçilmiştir, $-L=U$. Önerdiğimiz yöntemler Mutual Subspace (MS) yöntemi [2], Hadid ve Pietikainen [3] tarafından önerilen Manifold Öğrenme (MÖ) yöntemi ve Wang ve arkadaşları [4] tarafından önerilen MÖ yöntemi ile karşılaştırılmıştır.

3.1. Honda/UCSD Veri Tabanı Üzerindeki Deneyler

Honda/UCSD veri tabanı video tabanlı yüz tanıma için toplanmıştır. Bu veri tabanı 20 kişiye ait 59 videodan oluşmaktadır ve her bir videoda yaklaşık olarak 300-500 video karesi bulunmaktadır. Veri tabanındaki 20 video görüntüsü eğitim için ayrılmış olup, geriye kalan 39 video testler için saklanmıştır. Video karelerinden elde edilen yüz imgelerini 40×40 boyutlu olacak şekilde örnekledikten sonra histogramlar eşitlenmiştir. Bu imgeler daha sonra yüz imge setlerini oluşturmak için kullanılmışlardır. Şekil 1'de video karelerinden elde edilen bazı yüz görüntüleri gösterilmiştir.

Deneyler sonucunda elde edilen tanıma oranları Tablo 1'de özetlenmiştir. Bu veri tabanı için imge setlerini tanıma oldukça kolay olup, test edilen tüm yöntemlerle çok iyi tanıma oranları elde edilmiştir. Bu durumu daha zorlaştırmak amacıyla, sistematik olarak temiz imge setlerine diğer kişilere

ait imge setlerinden imgeler eklemek suretiyle imge setleri kirlenmiştir. Önerilen yöntemlerle tüm durumlar için en başarılı tanıma oranları elde edilmiştir. Özellikle hem eğitim hem de test setinin kirlenmiş olduğu durumda Kernel AZMDİST yöntemi diğer yöntemlere oranla çok daha başarılıdır. Doğrusal olmayan kernel yöntemlerle doğrusal yöntemlere oranla daha iyi tanıma oranları elde edilmiştir.



Şekil 1: Honda/UCSD veri tabanındaki iki kişiye ait videolardan elde edilen bazı imgeler.

Tablo I: Honda/UCSD veri tabanına ait tanıma oranları (%). ES: Eğitim Seti, TS:Test Seti, MÖ: Manifold Öğrenme, MS: Mutual Subspace.

Doğrusal Yöntemler	Temiz ES+TS	Kirli ES	Kirli TS	Kirli ES+TS
AZMDİST	97.4	97.4	92.3	87.2
CZMDİST	94.9	92.3	92.3	82.1
MS	97.4	97.4	87.2	76.9
Kernel Yöntemler				
AZMDİST	97.4	97.4	92.3	92.3
CZMDİST	100	97.4	92.3	82.1
MÖ [4]	97.4	97.4	89.7	79.5
MÖ [3]	94.9	89.7	84.6	79.5

3.2. CMU MoBo Veri Tabanı Üzerindeki Deneyler

CMU Mobo veri tabanı 24 kişiye ait 96 tane video görüntüsünden oluşmaktadır. Videolar farklı kameralarla çekildiğinden yüz görüntüleri arasında büyük poz ve ölçek farklılıkları vardır. Bir önceki deneyde olduğu gibi video karelerinden elde edilen yüz imgeleri 40×40 boyutlu olacak şekilde örneklendikten sonra histogramlar eşitlenmiş ve imge setleri oluşturulmuştur. Bu veri tabanında eğitim ve test setleri için videolar ayrılmadığından, deneylerimizde her bir sınıfa ait imge setlerinden rastgele bir set eğitim için seçilmiş ve geriye kalan 3 imge seti test amacıyla kullanılmıştır. Bu işlem 10 defa tekrarlandıktan sonra elde edilen tanıma oranlarının ortalaması alınarak en son tanıma oranları elde edilmiştir. CMU MoBo veri tabanından elde edilen tanıma oranları Tablo II'de verilmiştir. Tanıma oranları üzerindeki * sembolü en iyi tanıma oranlarını elde eden ve koyu renkte gösterilmiş yöntemden istatistiksel olarak daha kötü sonuçları veren (%95 güvenirlilik düzeyinde) yöntemleri işaret etmektedir.

Tabloda görüldüğü üzere önerdiğimiz kernel CZMDİST yöntemiyle en iyi tanıma oranları elde edilmiştir. Bunu doğrusal CZMDİST yöntemi takip etmekte ve yine önerdiğimiz kernel AZMDİST yöntemi ise üçüncü olmuştur. Bir önceki deneyde olduğu gibi kernel yöntemler doğrusal yöntemlere oranla daha iyi sonuçlar vermiştir. En kötü tanıma oranları Hadid ve Pietikainen [3] tarafından önerilen (MÖ) yöntemi ile elde edilmiştir. Genel olarak önerdiğimiz

yöntemler literatürde imge setlerinden kişi tanıma amacıyla kullanılan en iyi yöntemlerden çok daha iyi sonuçlar vermişlerdir.

Tablo II: CMU MoBo veri tabanına ait tanıma oranları (%).

Doğrusal Yöntemler	Tanıma Oranları (%)
AZMDİST	92.7*+3.3
CZMDİST	94.2+2.7
MS	92.0*+3.0
Kernel Yöntemler	
AZMDİST	93.8+2.8
CZMDİST	95.3+2.2
MÖ [4]	88.2*+5.0
MÖ [3]	85.5*+4.4

4. Sonuçlar

Bu bildiriye, yüz tanımda imge setlerini kullanan yeni yöntemler önerilmiştir. Bu yöntemlerde her bir imge seti geometrik konveks bir model (affine ve konveks zarflar) kullanılarak yakınsanmış ve bu modeller arasındaki uzaklıklar bunların üzerindeki en yakın noktaları bularak belirlenmiştir. Kernel hilesinin kullanılması, daha karmaşık ve doğrusal olmayan yapılarla sahip imge setlerinin daha doğru bir şekilde yakınsanmasına olanak vermiştir. Honda/UCSD ve CMU MoBo veri tabanları üzerinde yapılan deneylerde önerdiğimiz yöntemler literatürde bu amaçla kullanılan en iyi yöntemlerden çok daha iyi sonuçlar vermişlerdir.

5. Teşekkür

Bu çalışma Türkiye Bilimler Akademisi tarafından TÜBA-GEBİP/2010-11 programı çerçevesinde desteklenmiştir.

6. Kaynakça

- [1] A. W. Fitzgibbon and A. Zisserman, "Joint manifold distance: a new approach to appearance based clustering," IEEE Comp. Soc. Conf. On Computer Vision and Pattern Recognition, 2003.
- [2] K. Fukui and O. Yamaguchi "Face recognition using multi-viewpoints patterns for robot vision," Int. Symposium of Robotics Research, 2003.
- [3] A. Hadid, M. Pietikainen, "From still image to video based face recognition: an experimental analysis," Int. Conf. On Automatic Face and Gesture Recognition, 2004.
- [4] R. Wang, S. Shan, X. Chen, W. Gao, "Manifold-manifold distance with application to face recognition based on image sets," IEEE Comp. Soc. Conf. On Computer Vision and Pattern Recognition, 2008.
- [5] K. C. Lee, J. Mo, M. H. Yang, D. Kriegman, "Video based face recognition using probabilistic appearance manifolds," IEEE Comp. Soc. Conf. On Computer Vision and Pattern Recognition, 2003.
- [6] R. Gross, J. Shi, "The cmu motion of body (mobo) database," Technical Report, Carnegie Mellon Univ. 2001.
- [7] P. Viola, M. Jones, "Robust real-time face detection," International Journal of Computer Vision, 57:137-154, 2004.