

Doğrusal-Olmayan Ayırt Edici Ortak Vektörler Nonlinear Discriminant Common Vectors

Hakan Çevikalp¹, Marian Neamtu²

¹Vanderbilt Üniversitesi, Elektrik ve Bilgisayar Mühendisliği Bölümü, 37235 Nashville, A.B.D

²Vanderbilt Üniversitesi, Matematik Bölümü, 37235 Nashville, A.B.D

hakan.cevikalp@vanderbilt.edu, neamtu@math.vanderbilt.edu

Özetçe

Bu bildiriye Kernel Ayırt Edici Ortak Vektör (Kernel AEOV) yöntemi olarak adlandırdığımız yeni bir yöntem sunulmuştur. İlk olarak veri uzayı kernel fonksiyonları kullanılarak doğrusal-olmayan daha yüksek boyutlu bir uzaya eşlenmiş, ve bu yeni uzayda doğrusal Ayırt Edici Ortak Vektör (AEOV) yöntemi uygulanmıştır. Önerilen yöntem doğrusal-olmayan uzaya eşlenen verilerin sınıflar içi saçılımının sıfır altuzayından seçilen izdüşüm vektörlerini, verilerin öznitelik vektörlerinin elde edilmesi için kullanmaktadır. Bu işlem sonucunda her bir sınıftaki örnek, o sınıfı temsil eden ayırt edici ortak vektörü üretmektedir. Bu da eğitim setindeki tüm örneklerin doğru olarak sınıflandırılmasını sağlar. Deney sonuçları önerilen yöntemin test setindeki örnekleri de diğer doğrusal-olmayan kernel yöntemlere oranla daha doğru olarak sınıflandırdığını göstermektedir.

Abstract

In this paper we propose a novel method called the Kernel Discriminative Common Vector (Kernel DCV) method. Firstly, the original input space is mapped nonlinearly to a higher-dimensional feature space through a kernel mapping. Then, the linear Discriminative Common Vector (DCV) method is applied in the transformed space. The proposed method employs the projection vectors from the null space of the within-class scatter matrix of the transformed samples for feature extraction. The same discriminative common vector for all samples in each class is obtained after feature extraction. Therefore, a 100% recognition rate is always guaranteed for the training set samples. The experiments on the test sets show that the generalization ability of the proposed method is also superior to other kernel approaches.

1. Giriş

Öznitelik özütleme veri tanıma alanındaki en önemli konulardan biri olmuştur. Öznitelik özütleme yöntemleri daha çok doğrusal yöntemler üzerinde yoğunlaşmıştır. PCA [1], FLDA [2], Direct-LDA [3] ve Ayırt Edici Ortak Vektör (AEOV) yöntemleri [4] doğrusal öznitelik özütleme yöntemlerinden bazılarıdır. Bu yöntemlerde amaç, verilerin sınıflandırılması için önemli bilgileri taşıyan öznitelikleri seçmektir. Genellikle, veri uzayı bu yöntemler kullanılarak daha küçük boyutlu bir uzaya aktarılır. Bu işlemin sonucunda bir çok durumda “curse of dimensionality” olarak adlandırılan olgunun zararlı etkileri

azaltılmaktadır [1]. Bazen doğrusal özütleme yöntemleri daha karmaşık doğrusal-olmayan yöntemler için ön-işlemci olarak kullanılmaktadır. Bununla birlikte, doğrusal öznitelik özütleme yöntemleri doğrusal-olmayan dağılıma sahip verilerin sınıflandırılmasında yetersiz kalmaktadır (örnek olarak “exclusive-or” problemi). Bu sebeple Kernel yöntemler olarak adlandırılan doğrusal-olmayan yöntemler önerilmiştir. Kernel PCA [5], Kernel FDA [6] ve Kernel GDA [7] bu yöntemlerden bazılarıdır.

Bu bildiriye Kernel Ayırt Edici Ortak Vektör (Kernel Discriminative Common Vector) olarak adlandırdığımız ve doğrusal Ayırt Edici Ortak Vektör yöntemini doğrusal-olmayan uzaya eşlenen verilere uygulayan yöntem sunulmuştur. Değiştirilmiş Fisher ölçütü,

$$J_{MFLDA}(W_{opt}) = \arg \max_W \frac{|W^T S_B W|}{|W^T S_T W|}, \text{ önerilen yöntem}$$

kullanılarak elde edilen öznitelik vektörlerinden oluşan matris sayesinde en büyük değeri olan 1 sayısına ulaşır. Bu eşitlikte S_B sınıflar arası saçılım matrisini, S_T tüm örneklere ait toplam saçılım matrisini ifade etmektedir. Bu nedenle önerilen yöntem öznitelik özütleme için optimal izdüşüm vektörlerini bulur. Bu da eğitim setindeki tüm örneklerin doğru olarak sınıflandırılmasına olanak sağlar.

2. Optimal Ayırt Edici Altuzayı Kavramı

Öznitelik özütlemeye kullanılan izdüşüm vektörleri, sınıflar içi saçılım matrisi S_w ’nin sıfır altuzayından seçilirse değiştirilmiş

$$\text{Fisher ölçütü, } J(W_{opt}) = \arg \max \frac{|W^T S_B W|}{|W^T S_T W|}, \text{ en büyük değeri}$$

olan 1 sayısına ulaşır. Bununla beraber, sınıflar içi saçılım matrisi S_w tekil bir matrisse (örnek uzayı boyutu eğitim setindeki örnek sayısından fazla ise S_w matrisi teki bir matris olacaktır) $W^T S_w W = 0$ ve $W^T S_B W \neq 0$ şartlarını sağlayan bir çok matris bulunabilir. Fakat bu matrislerin bir çoğu test setindeki örneklerin doğru olarak sınıflandırılmasında yetersiz kalır [8]. Bu sebeple eğitim setindeki örnek sayısının örnek uzayı boyutundan büyük olduğu durumlarda aşağıda verilen ölçüt kullanılmalıdır [4], [9].

$$J(W_{opt}) = \arg \max_{|W^T S_w W|=0} |W^T S_B W| = \arg \max |W^T S_{com} W| \quad (1)$$

Bu ölçüt kullanılarak sadece bir tane izdüşüm vektör matrisi bulunur ve bu matris aynı zamanda değiştirilmiş Fisher ölçütünü en büyük değeri olan 1 sayısına ulaştırır. Bu sebeple, bu ölçüt kullanılarak elde edilen izdüşüm vektörleri optimal olacaktır. Optimal izdüşüm vektörlerini bulmak için, eğitim setindeki örneklerin sınıflar içi saçılım matrisinin sıfır altuzayındaki izdüşümleri bulunmalıdır. Daha sonra, bu uzaya aktarılan örnekler PCA yöntemi uygulanarak optimal izdüşüm vektörleri bulunur. Bu işlem sonucunda elde edilen izdüşüm vektörleri sınıflar içi saçılım matrisinin sıfır altuzayı $N(S_w)$ ile toplam saçılım matrisinin erim uzayının $R(S_T)$ kesişim uzayını doğuran vektör setini oluştururlar [8]. Bu kesişim uzayının $R(S_T) \cap N(S_w)$ erim kümesini doğuran optimal izdüşüm vektörleri, öznelik özütleme için gerekli optimal öznelikleri ortaya çıkardığından, bu kesişim uzayını *optimal ayırt edici altuzayı* olarak isimlendirdik.

Optimal ayırt edici altuzayını doğuran optimal izdüşüm vektörlerini bulmak için başlıca iki temel yol izlenebilir. İlk olarak eğitim setindeki örneklerin $N(S_w)$ uzayındaki izdüşümleri bulunur ve ardından PCA uygulanırsa optimal izdüşüm vektörleri elde edilebilir. Ayırt Edici Ortak Vektör yöntemi bu yolu kullanmaktadır. Diğer yöntemde ise, ilk olarak eğitim setindeki örnekler PCA uygulanır ve bu işlemin ardından, yeni uzaydaki verilerin sınıflar içi saçılım matrisinin sıfır altuzayını doğuran vektör seti bulunur. PCA+Null Space yöntemi [10] bu yolu kullanmaktadır.

3. Kernel Ayırt Edici Ortak Vektör Yöntemi

Eğitim setinde her biri N_i örnek içeren C tane sınıf olduğunu varsayalım. Her bir sınıftaki m 'nci örneği d boyutlu örnek uzayında x_m^i ile gösterelim. Bu durumda eğitim setinde toplam $M = \sum_{i=1}^C N_i$ örnek olacaktır. Kernel fonksiyonları kullanan yöntemlerde ilk olarak örnek uzayı çok daha büyük boyutlu bir uzaya eşlemlenir. $\Phi(x_1^1), \Phi(x_2^1), \dots, \Phi(x_{N_1}^1), \Phi(x_1^2), \dots, \Phi(x_{N_C}^C)$ vektörlerinin bu uzaya eşlemlenen örnekleri, \mathfrak{S} uzayında daha büyük boyutlu bu yeni uzayı ifade ettiğini varsayalım. Bu durumda \mathfrak{S} uzayındaki örneklerin sınıflar içi saçılım matrisi S_w^Φ , sınıflar arası saçılım matrisi S_B^Φ ve toplam saçılım matrisi S_T^Φ aşağıdaki eşitlikler kullanılarak bulunabilir:

$$\begin{aligned} S_w^\Phi &= \sum_{i=1}^C \sum_{m=1}^{N_i} (\Phi(x_m^i) - \mu_i^\Phi)(\Phi(x_m^i) - \mu_i^\Phi)^T \\ &= (\Phi - \Phi G)(\Phi - \Phi G)^T, \end{aligned} \quad (2)$$

$$\begin{aligned} S_B^\Phi &= \sum_{i=1}^C N_i (\mu_i^\Phi - \mu^\Phi)(\mu_i^\Phi - \mu^\Phi)^T \\ &= (\Phi U - \Phi L)(\Phi U - \Phi L)^T, \end{aligned} \quad (3)$$

$$\begin{aligned} S_T^\Phi &= \sum_{i=1}^C \sum_{m=1}^{N_i} (\Phi(x_m^i) - \mu^\Phi)(\Phi(x_m^i) - \mu^\Phi)^T \\ &= (\Phi - \Phi \mathbf{1}_M)(\Phi - \Phi \mathbf{1}_M)^T = S_w^\Phi + S_B^\Phi. \end{aligned} \quad (4)$$

Bu eşitliklerde μ^Φ , \mathfrak{S} uzayındaki örneklerin ortalama değer vektörünü, μ_i^Φ , i 'nci sınıfa ait ortalama değer vektörünü ve Φ ise her bir sütunu \mathfrak{S} uzayına eşlemlenmiş eğitim seti örneklerinden oluşan matrisi ifade etmektedir. Eşitlik (2)'de verilen $G = \text{diag}[G_1, \dots, G_C] \in R^{M \times M}$ matrisi blok köşegen matrisi olup her bir matris $G_i \in R^{N_i \times N_i}$ 'nin tüm elemanları $1/N_i$ sayısına eşittir. Eşitlik (3)'de verilen $U = \text{diag}[u_1, \dots, u_C] \in R^{M \times C}$ matrisi blok köşegen matrisi olup her bir vektör $u_i \in R^{N_i \times 1}$ 'nin tüm elemanları $1/N_i$ sayısına eşittir. Eşitlik (3)'de verilen $L = [l_1, \dots, l_C] \in R^{M \times C}$ matrisinin sütunlarını oluşturan her bir vektör $l_i \in R^{M \times 1}$ 'nin tüm elemanları $\sqrt{N_i}/M$ sayısına eşittir. $\mathbf{1}_M \in R^{M \times M}$ matrisi ise tüm elemanları $1/M$ sayısına eşit olan matristir.

Eşlemlenmiş \mathfrak{S} uzayındaki örneklerin sınıflar içi saçılım matrisi S_w^Φ genellikle tekil bir matristir. Bu nedenle S_w^Φ matrisinin sıfır altuzayı $N(S_w^\Phi)$ ile S_T^Φ matrisinin erim uzayının $R(S_T^\Phi)$ kesişim uzayını doğuran birim vektör seti, öznelik özütleme için gerekli optimal izdüşüm vektörlerinden oluşur. \mathfrak{S} uzayındaki optimal izdüşüm vektörlerini bulmak için daha önceden de belirttiğimiz gibi başlıca iki temel yol vardır. Birinci yol olarak, \mathfrak{S} uzayındaki örneklerin $N(S_w^\Phi)$ uzayındaki izdüşümleri bulunur ve ardından PCA uygulanırsa optimal izdüşüm vektörleri elde edilebilir. Diğer bir alternatif olarak da, eğitim setindeki örneklere PCA uygulanarak bu örneklerin $R(S_T^\Phi)$ uzayındaki izdüşümleri bulunur ve bu işlemin ardından yeni uzaydaki verilerin sınıflar içi saçılım matrisinin sıfır altuzayını doğuran vektör seti bulunur. Fakat ilk anlatılan yöntemdeki algoritmalar çok büyük boyutlu \mathfrak{S} uzayını belirttik olarak kullandığından, bu yolun izlenmesi mümkün değildir. Bu yüzden \mathfrak{S} uzayındaki optimal izdüşüm vektörlerini bulmak için ikinci yol izlenmelidir. Eğitim setindeki örneklerin $R(S_T^\Phi)$ uzayındaki izdüşümleri, Kernel PCA yöntemi uygulanarak kolayca bulunabilir. Daha sonra, bu yeni uzaydaki örneklerin sınıflar içi saçılım matrisinin sıfır altuzayını doğuran vektör seti kullanılarak optimal izdüşüm vektörleri bulunur. Bu vektörler kullanılarak elde edilen öznelik vektörleri her sınıftaki örnekler için aynıdır. Bu vektörleri doğrusal Ayırt Edici Vektör yönteminde olduğu gibi *ayırt edici ortak vektörler* olarak adlandırdık. Burada anlattığımız yöntem aşağıdaki gibi özetlenebilir:

Adım 1: Kernel PCA yöntemi kullanarak eğitim setindeki örneklerin $R(S_T^\Phi)$ uzayındaki izdüşümlerini bul.

$$\tilde{K} = K - \mathbf{1}_M K - K \mathbf{1}_M + \mathbf{1}_M K \mathbf{1}_M \in R^{M \times M} = P \Lambda P^T \quad (5)$$

Bu eşitlikte Λ matrisi köşegen matris olup bu matrisin her bir köşegen elemanı sıfırdan farklıdır. $K \in R^{M \times M}$ matrisi ise

$K = \Phi^T \Phi = (K^{ij})_{\substack{i=1,\dots,C \\ j=1,\dots,C}}$ eşitliği ile bulunabilir. Bu eşitlikteki

her bir $K^{ij} \in R^{N_i \times N_j}$ matrisi aşağıdaki eşitlikle bulunabilir:

$$K^{ij} = (k_{mn}^{ij})_{\substack{m=1,\dots,N_i \\ n=1,\dots,N_j}} = \langle \Phi(x_m^i), \Phi(x_n^j) \rangle = k(x_m^i, x_n^j)_{\substack{m=1,\dots,N_i \\ n=1,\dots,N_j}}$$

Bu durumda eğitim setindeki örnekleri $R(S_T^\Phi)$ uzayına eşleyen matris $(\Phi - \Phi 1_M) P \Lambda^{-1/2}$ olacaktır. Bu yeni uzaya eşlenen örneklerin toplam saçılım matrisi ve sınıflar arası saçılım matrisi aşağıda verildiği gibidir:

$$\begin{aligned} \tilde{S}_T^\Phi &= ((\Phi - \Phi 1_M) P \Lambda^{-1/2})^T S_T^\Phi (\Phi - \Phi 1_M) P \Lambda^{-1/2} \\ &= \Lambda^{-1/2} P^T P \Lambda P^T P \Lambda P^T P \Lambda^{-1/2} = \Lambda, \end{aligned} \quad (6)$$

$$\begin{aligned} \tilde{S}_W^\Phi &= ((\Phi - \Phi 1_M) P \Lambda^{-1/2})^T S_W^\Phi (\Phi - \Phi 1_M) P \Lambda^{-1/2} \\ &= \Lambda^{-1/2} P^T \tilde{K}_W \tilde{K}_W^T P \Lambda^{-1/2}. \end{aligned} \quad (7)$$

Eşitlik (7)'de verilen matris \tilde{K}_W aşağıdaki eşitlikle bulunabilir:

$$\tilde{K}_W = K - K G - 1_M K + 1_M K G = (K - 1_M K)(I - G). \quad (8)$$

Adım 2: $\tilde{S}_W^\Phi \in R^{M \times M}$ matrisinin sıfır altuzayını doğuran birimlik vektör setini bul. Bu işlem özvektör ayrıştırma işlemi ile gerçekleştirilebilir. \tilde{S}_W^Φ matrisinin sıfır özdeğerlerine karşılık gelen özvektörler bu matrisin sıfır altuzayını doğuran birimlik vektör setini verir. V matrisinin sütunlarının \tilde{S}_W^Φ matrisinin sıfır altuzayını doğuran vektörlerden oluştuğunu varsayalım.

$$V^T \tilde{S}_W^\Phi V = 0 \quad (9)$$

Adım 3 (isteğe bağlı): Eğer $V^T \tilde{S}_B^\Phi V$ matrisinin sıfır altuzayı varsa bu altuzayı çıkar.

$$V^T \tilde{S}_B^\Phi V = V^T \tilde{S}_T^\Phi V = V^T \Lambda V = L \tilde{\Lambda} L^T \quad (10)$$

Bu durumda öznitelik vektörlerinin elde edilmesi için kullanılacak optimal izdüşüm vektörlerinden oluşan matris

$$W = (\Phi - \Phi 1_M) P \Lambda^{-1/2} V L \quad (11)$$

olacaktır. Optimal izdüşüm vektörlerinin sayısı en çok $C-1$ olabilir. Bu matris kullanılarak elde edilen öznitelik vektörleri her bir sınıf için ayırdır ve bu da eğitim setindeki tüm örneklerin doğru olarak sınıflandırılmasını sağlar.

5. Deneysel Sonuçları

Kernel Ayırt Edici Ortak Vektör yönteminin performansını denemek için, ORL yüz veri tabanı [11] kullanılmıştır. Bu veri tabanı 40 deneğe ait toplam 400 imge içermektedir. Şekil 1'de görüldüğü gibi her deneğe ait yüz ifadesi ve kafa yönelmesi

farklılıkları içeren 10 imge bulunmaktadır. Veri tabanındaki imgelere herhangi bir ön-işleme uygulanmamıştır. Her bir imge 92×112 boyutludur.

Belli amaçlar için kernel fonksiyonların nasıl seçileceği hala süregelen bir problem olmuştur. Bu çalışmamızda 2. dereceden kernel polinom fonksiyonu $k(x, y) = (\langle x, y \rangle)^2$ ile Gaussian kernel fonksiyonu $k(x, y) = \exp(-\|x - y\|^2 / \gamma)$ kullanılmıştır. Gaussian kernel fonksiyonunun parametresi γ , 1.06e8 olarak seçilmiştir. Kernel Ayırt Edici Ortak Vektör yönteminin haricindeki tüm yöntemler için, tanıma sırasında 1-en yakın komşu algoritması kullanılmıştır. Önerilen yöntem içinse, tanıma sırasında test örneğinin öznitelik vektörü ile ayırt edici ortak vektörler Öklid uzaklığı kullanılarak karşılaştırılmış ve test örneği en yakın uzaklığı veren sınıfa atanmıştır. Kernel PCA yönteminde seçilen en büyük özdeğer vektörlerine karşılık gelen özdeğerlerin toplamı toplam enerjinin % 95'ini oluşturmaktadır [12].



Şekil 1: ORL yüz veri tabanındaki bazı deneklere ait imgeler.

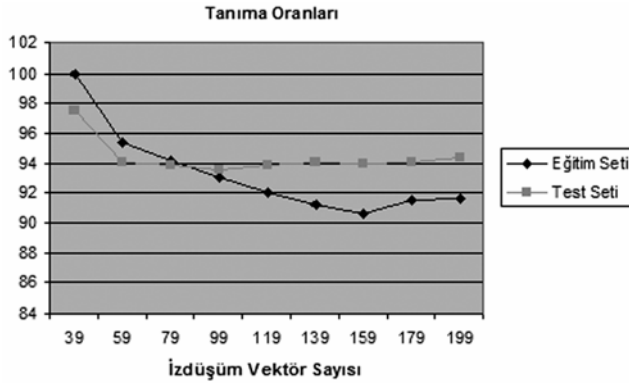
Önerilen yöntemin denenmesi için her sınıftan rastgele $N = 3, 5, 7$ örnek seçilerek eğitim seti oluşturulmuş ve geriye kalan $(10 - N)$ örnek test seti için kullanılmıştır. Daha sonra tanıma oranları bulunmuştur. Bu işlem 6 defa tekrarlanmış ve tanıma oranlarının ortalaması alınarak en son tanıma oranları elde edilmiştir. Sonuçlar Tablo 1'de verilmiştir. Tablo 1'de de görüldüğü gibi tüm durumlarda en iyi tanıma oranları önerilen yöntem ile elde edilmiştir. Bu da yöntemin oldukça başarılı olduğunu göstermektedir.

Daha öncedinde belirttiğimiz gibi Kernel Ayırt Edici Ortak Vektör yöntemi ile bulunan optimal izdüşüm vektörleri $W^T S_W W = 0$ ve $W^T S_B W \neq 0$ şartlarını sağlar ve bu vektörler aynı zamanda optimal ayırt edici altuzayını doğuran birimlik vektör setini oluştururlar. Bu çalışmada aynı zamanda, optimal izdüşüm vektörlerine optimal ayırt edici altuzayının dışından yeni izdüşüm vektörleri ekleyerek önerilen yöntemin performansının artırılıp artırılmayacağı sorusunu cevaplamaya çalıştık. Bu amaçla kaynak [13]'de verilen yol izlenmiştir. Bu deneyler için eğitim ve test setlerinde 5 örnek bulunan veri setleri kullanılmıştır. Kernel PCA adımının ardından \tilde{S}_W^Φ matrisi kullanılarak, bu matrisin sıfır altuzayı ile erim uzayını doğuran birimlik vektör seti bulunmuştur. Sıfır altuzayını doğuran birimlik vektör seti optimal izdüşüm vektörleridir ve bu vektörlerin sayısı $C - 1 = 39$ olarak bulunmuştur. Bu vektörlere daha sonra, \tilde{S}_W^Φ matrisinin erim uzayını doğuran birimlik vektör setinden aşamalı olarak yeni vektörler eklenmiş ve bu

Tablo I
Kernel Yöntemleri Kullanılarak Elde Edilen Tanıma Oranları

Kernel Yöntemleri	Eğitim Setindeki Sınıfların Örnek Sayısı					
	N = 3		N = 5		N = 7	
	Polinom kernel fonksiyonu	Gaussian kernel fonksiyonu	Polinom kernel fonksiyonu	Gaussian kernel fonksiyonu	Polinom kernel fonksiyonu	Gaussian kernel fonksiyonu
Kernel PCA	86.20	86.25	93.33	93.75	96.94	97.08
Kernel FDA	90.05	89.99	96.33	96.50	97.49	98.47
Kernel GDA	87.60	91.31	94.16	96.66	96.66	98.19
Kernel AEOV	91.38	91.60	97	97.50	98.61	98.75

vektörler için tanıma oranları bulunmuştur. Sonuçlar Şekil 2’de gösterilmiştir. Şekilde de görüldüğü üzere optimal ayırt edici altuzayı dışından yeni vektörlerin eklenmesi önerilen yöntemin performansını düşürmüştür. Optimal izdüşüm vektörlerine optimal ayırt edici altuzayının dışından vektörlerin eklenmesi aynı zamanda sistemin yeni test örneklerinin sınıflandırma hızını da olumsuz etkileyecektir. Çünkü bu yeni vektörler her sınıftaki örnekler için aynı olan ayırt edici ortak vektörleri üretmezler. Bu durumda yeni bir test örneğini sınıflandırmak için, test örneğinin öznelik vektörü eğitim setindeki tüm örneklerin öznelik vektörü ile karşılaştırılmalıdır. Bu da yöntemin hızını oldukça düşürecektir. Özetlenecek olursa optimal ayırt edici altuzayından seçilen izdüşüm vektörleri doğruluk ve hız bakımından en uygun izdüşüm vektörleridir.



Şekil 2. Farklı sayılarda izdüşüm vektörü kullanılarak elde edilen tanıma oranları.

6. Sonuçlar

Bu çalışmada kernel fonksiyonları kullananan doğrusal-olmayan yeni bir yöntem sunulmuştur. İlk aşamada, kuramsal olarak optimal izdüşüm vektörlerinin verilerin sınıflar içi saçılım matrisinin sıfır altuzayı ile toplam saçılım matrisinin erim uzayının kesişim altuzayından geldiği ispatlanmıştır. Daha sonra, verilerin doğrusal-olmayan eşleşmiş uzayındaki optimal izdüşüm vektörlerini bulmak için Kernel Ayırt Edici Ortak Vektör yöntemi sunulmuştur. Yapılan deneysel çalışmalar önerilen yöntemin tüm durumlarda diğer kernel yöntemlerden daha iyi sonuçlar ürettiğini göstermiştir. Bu da kuramsal ispatların doğruluğunu desteklemektedir. Önerilen yöntem kullanılarak elde edilen öznelik vektörleri her sınıftaki örnekler için aynıdır. Bu nedenle test imgesinin tanınması esnasında, test imgesi her sınıf için o sınıfı temsil eden tek bir ayırt edici vektörle karşılaştırılır. Bu sebeple Kernel Ayırt Edici Vektör yöntemi gerçek zamanlı yüz tanıma sistemleri için ideal bir yöntemdir.

7. Kaynakça

- [1] C.M. Bishop, *Neural Networks for Pattern Recognition*. Oxford University Press, 1995, pp. 11-17, 295-297, 318-319.
- [2] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *In Annals of Eugenics*, vol. 7, pp. 179-188, 1936.
- [3] H. Yu and J. Yang, "A direct LDA algorithm for high-dimensional data with application to face recognition," *Pattern Recognition*, vol. 34, pp. 2067-2070, 2001.
- [4] H. Cevikalp, M. Neamtu, M. Wilkes, and A. Barkana, "Discriminative common vectors for face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 4-13, January 2005.
- [5] B. Schölkopf, *Support Vector Learning*, Ph. D. thesis, Informatik der Technischen Universität, Berlin, 1997.
- [6] S. Mika, G. Rätsch, J. Weston, B. Schölkopf, and K.-R. Müller, "Fisher discriminant analysis with kernels," in *Neural Networks for Signal Processing IX*, Y.-H. Hu, J. Larsen, E. Wilson, and S. Douglas, Eds. Piscataway, NJ: IEEE, pp. 41-48, 1999.
- [7] G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach," *Neural Computation*, vol. 12, pp. 2385-2404, 2000.
- [8] H. Cevikalp, M. Neamtu, and M. Wilkes, "Discriminative common vectors with kernels," *IEEE Transaction on Neural Networks*, in review.
- [9] P.N Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, 1997.
- [10] R. Huang, Q. Liu, H. Lu, and S. Ma, "Solving the small size problem of LDA," in *Proceedings of 16th International Conference on Pattern Recognition*, August 2002, vol. 3, pp. 29-32.
- [11] The ORL Database of Faces, AT&T Laboratories Cambridge. Available: <http://www.uk.research.att.com/facedatabase.html>
- [12] D. L. Swets and J. Weng, "Using discriminant eigenfeatures for image retrieval," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 831-836, August 1996.
- [13] J. Yang, D. Zhang and J-Y Yang, "A generalised K-L expansion method which can deal with small sample size and high-dimensional problems," *Pattern Analysis & Applications*, vol. 6, pp. 47-54, April 2003.
- [14] V. N. Vapnik, *The Nature of Statistical Learning Theory*. 2nd edition, Springer-Verlag, New York, 1995, pp. 30-31.