

KANONİK KORELASYON ANALİZİ İLE YÜZ İFADESİNDEN DUYGU TANIMA İÇİN ÖZİNİTELİK ÇIKARIMI

FEATURE EXTRACTION FOR FACIAL EXPRESSION RECOGNITION BY CANONICAL CORRELATION ANALYSIS

C. Okan Şakar¹, Olcay Kurşun², Ali Karaali³, Çiğdem Eroğlu Erdem^{3*}

¹Bilgisayar Mühendisliği Bölümü

³Elektrik-Elektronik Mühendisliği Bölümü

Bahçeşehir Üniversitesi, İstanbul

{okan.sakar, ali.karaali, cigdem.eroglu}@bahcesehir.edu.tr

²Bilgisayar Mühendisliği Bölümü

İstanbul Üniversitesi, İstanbul

okursun@istanbul.edu.tr

ÖZETÇE

Verilen bir imgedeki yüz ifadesinden kişinin duygu durumunu kestirebilmek için, imgenin farklı yöntemlerle işlenmesi sonucu elde edilen farklı gösterimler literatürdeki çalışmalarda değişik şekillerde kaynaştırılmıştır; ancak aralarındaki bağımlılık ve bağıntıların analizi yeterince irdelenmemiştir. Bu çalışmada, farklı gösterimler arasındaki ilintileri ortaya çıkaran Kanonik Korelasyon Analizi (KKA) ile elde edilen eşdeğişkenlerin, duygu tanıma için yüksek başarıma sahip oldukları gösterilmektedir. Çıkarttığı az sayıda özneliklerin kullanılmasıyla yüksek duygu tanıma başarımu elde edilebildiğinden dolayı, KKA iyi bir boyut indirgeme yöntemi olarak öne çıkmaktadır. Benzetim çalışmamızda kullandığımız CK+ veri kümesinde, fark imgeleri ve geometrik öznelikler içeren iki farklı gösterim üzerinden elde edilen eşdeğişkenlerin yüksek kestirim doğruluğu verdiği gösterilmiştir.

ABSTRACT

Although several methods have been proposed for fusing different image representations obtained by different preprocessing methods for emotion recognition from the facial expression in a given image, the dependencies and relations among them have not been much investigated. In this study, it has been shown that covariates obtained by Canonical Correlation Analysis (CCA) that extracts relations between different representations have high predictive power for emotion recognition. As high prediction accuracy can be achieved using a small number of features extracted by it, CCA is considered to be a good dimensionality reduction method. For our simulations, we used the CK+ database and showed that covariates obtained from difference-images and geometric-features representations have high prediction accuracy.

1. GİRİŞ

Verilen bir imgedeki yüz ifadesinden kişinin duygusal durumunu kestirme, insan-bilgisayar etkileşimi (İBE) alanındaki temel problemlerden birisidir ve yoğun bir araştırma konusudur [1]. Birçok İBE problemindeki gibi

duygu tanıma probleminde de, farklı işaret kipleri ve bir kipteki bilginin farklı yöntemlerle işlenmesi sonucu elde edilen farklı gösterimler kullanılmaktadır. Her ne kadar literatürde bu çok gösterimli/kipli işaretlerin kaynaştırılması çalışılmış bir konu olsa da [2, 3], gösterimler arasındaki bağımlılık ve bağıntıların analizi fazla irdelenmemiştir. Bu çalışmada, duygu tanıma için kullanılan görüntü kipi üzerinde tanımlı iki farklı gösterim arasındaki doğrusal bağıntılar, Kanonik Korelasyon Analizi (KKA) [4-6] ile ortaya çıkarılmıştır. Bu gösterimlerden ilki, görünüm tabanlı (appearance-based) olup, duygunun tepe noktasındaki yüz ifadesi ile nötr yüz ifadesi arasındaki fark imgesinden oluşur. Gösterimlerden ikincisi ise, geometrik tabanlı olup, yüz üzerinde takip edilen 68 nirengi noktasının koordinatlarından, nötr imgedeki koordinatların çıkarılmasıyla elde edilmiştir. Her iki gösterim de duygu tanıma çalışmalarında yaygın olarak kullanılan CK+ veri kümesi [7, 8] üzerinde tekil olarak kullanıldıklarında yüksek başarıma sahip olmakla beraber, gösterimlerdeki özneliklerin boyutları bu tip çalışmalarda tipik olarak kullanılan Temel Bileşen Analizi (TBA) ile azaltıldığında, bu başarımlar düşmektedir. Bunun nedeni Temel Bileşen Analizi (TBA) yönteminin, tek bir gösterim üzerinden öznelik çıkarımı (boyut azaltma) için kullanılmasıdır. Eğer imgelerden birden fazla gösterim elde edildiyse, öznelik çıkarımı için bu gösterimlerin TBA yöntemiyle birbirinden bağımsız olarak kullanılması, tahmin edilmek istenen duygu bilgisinin kaybolmasına neden olabilir, çünkü TBA gözetimsiz bir yöntemdir ve duygu tanımaya yönelik sınıf bilgisi gibi bir ek bilgi kullanamaz. Bunun yerine doğrudan sınıf bilgisini kullanmaması nedeniyle gözetimsiz olduğu söylenebilecek Kanonik Korelasyon Analizi (KKA) kullanılarak çıkarılan öznelikler [9, 10] gösterimler arasındaki ortak bilginin en çok olduğu vektörler olarak belirlenmektedir. Bu nedenle, KKA ile elde edilen eşdeğişkenler (covariates) ile boyut azaltıldığında kestirim doğruluğunda fazla bir düşüş olmamaktadır.

Bu çalışmada, imge kipi üzerinde tanımlı iki farklı gösterimden, TBA ve KKA yöntemleriyle çıkartılan özneliklerin Destek Vektör Makineleri (DVM) ile yüz ifadelerini tanımadaki başarımları ölçülmüş, bu yöntemlerle iki boyuta indirgenerek görselleştirmeleri karşılaştırılmış ve

* Bu bildiriye ilişkin çalışmalar TÜBİTAK – EEAG tarafından 110E056 numaralı proje kapsamında desteklenmiştir.
978-1-4673-0056-8/12/\$26.00 ©2012 IEEE

KKA yöntemi ile gösterimler arasındaki bağımlılık ve bağıntılar analiz edilmiştir.

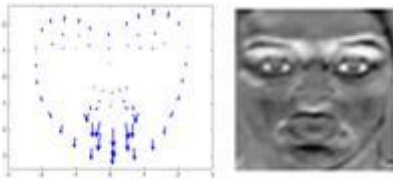
2. VERİ

Bu çalışmada kullanılan CK+ veritabanında [7, 8], 118 farklı kişiden kayıt edilmiş ve duygu etiketi taşıyan toplam 327 video klipi bulunmaktadır. CK+ veritabanında toplam 7 farklı duyguya ait videolar yer alır, bunlar (1) kızgınlık, (2) küçümseme, (3) iğrenme, (4) korku, (5) mutluluk, (6) üzüntü, (7) şaşırma duygularıdır. Bu videolardaki rol yapılarak oluşturulan yüz ifadeleri, ilk çerçevede nötr ifadeden başlar ve ifadenin en belirgin olduğu tepe noktasında son bulur (Şekil 1).



(a) 1. çerçeve (b) 2. çerçeve (c) 3. çerçeve
Şekil 1: CK+ veritabanında yer alan mutluluk duygusuna ait örnek video ve izlenen yüz nirengi noktaları

Yüz ifadesinden duygu tanıma çalışmalarında kullanılan öznitelikler, temel olarak, görünüme ait öznitelikler ve geometrik öznitelikler olarak ikiye ayrılabilirler [1]. Görünüme ait öznitelikler yüzün dokusu hakkında bilgi verirler: doğal kırışıklıklar, gülmeden, kaş çatmadan doğan göz, alın kırışıklıkları ve şişkinlikler gibi. Bu çalışmadaki görünüme ait öznitelikler, duygunun tepe noktasına ait yüz ifadesini içeren en son çerçeve ile ve nötr yüz ifadeli ilk çerçeve arasındaki fark imgesinden oluşur (Şekil 2(b)). Geometrik öznitelikler ise yüze ait bileşenlerin şekli (gözler, dudaklar vb.) ve yüzdeki belirgin nirengi noktalarının pozisyonları olarak tanımlanabilir (Şekil 1). Bu çalışmada kullandığımız geometrik öznitelikler, Şekil 1'de görülen yüzdeki 68 nirengi noktasının en son çerçevedeki koordinatlarından, ilk çerçevedeki koordinatların çıkarılmasıyla elde edilir (Şekil 2(a)). Bu çıkarma işlemlerinden önce imgeler arasındaki dönme, öteleme ve büyüklük farklarını en aza indirmek amacıyla bir hizalama işlemi uygulanmaktadır.



Şekil 2: Şaşkınlık duygusu için (a) geometrik özniteliklere bir örnek (b) görünüm tabanlı özniteliklere bir örnek

3. KULLANILAN KANONİK KORELASYON ANALİZ YÖNTEMİ

Kanonik Korelasyon Analizi (KKA) [4-6], iki değişken kümesi arasındaki doğrusal ilişkiyi maksimum yapan doğrusal dönüşümleri bulmayı amaçlayan bir yöntemdir. Bir başka deyişle, KKA veri kümelerinin değişkenleri üzerilerine

indürlendiğinde, aradaki korelasyonu maksimum yapan baz vektörlerinin bulunmasını amaçlar: X ve Y sırasıyla u ve v boyutlu çok değişkenli veri kümeleri ($X \in \mathbb{R}^u$ ve $Y \in \mathbb{R}^v$) ise

$$\rho = \max_{w_x, w_y} \text{corr}(Xw_x, Yw_y) \quad (1)$$

denklemini sağlayan w_x, w_y baz vektörleri kanonik vektörler, bunların X ve Y değişkenleriyle çarpımları ise kanonik eşdeğişkenler olarak adlandırılır. Değişken sayısı daha az olan veri kümesinin değişken sayısı kadar kanonik vektör bulunabilir.

Çalışmamızda, duygu tanıma için kullanılan imge kipi üzerinde tanımlı görünüm tabanlı olan fark imge gösterimi ve geometrik tabanlı olan nirengi noktalarının koordinat farklarına dayalı gösterim arasında KKA yöntemi kullanılarak öznitelikler çıkarılmıştır. Bunun için, tipik olarak yapıldığı gibi, kanonik vektörler bulunurken kullanılan ortak değişim matrislerini tekillikten kurtarmak için, gösterimlerin boyutları TBA yöntemi kullanılarak düşürülmüştür. Bu çalışmadaki deneylerde, bu ilk boyut düşürme her bir veri kümesinin değişimlerinin %95'i korunacak şekilde uygulanmıştır. Bu durumda, 4096 boyutlu birinci gösterim $u = 176$ boyuta, 134 boyutlu olan ikinci gösterim ise $v = 22$ boyuta indirgenmiştir. Daha sonra, bu boyutu düşürülmüş veri kümelerine KKA uygulanarak aralarındaki ortak bilgileri içeren farklı sayıda öznitelikler çıkarılmıştır. KKA ile çıkartılan öznitelikler, yalnızca TBA uygulanarak elde edilen K sayıda öznitelik ile duygu tanıma başarımları bakımından karşılaştırılmıştır. Bu özniteliklerin sınıflandırıcı başarımlarının karşılaştırılması için DVM yönteminin LIBSVM [11] gerçekleştirmesi, doğrusal ve çokterimli (3. derece) işlev çekirdeklerinin varsayılan parametre değerleriyle kullanılmıştır. Sınıflandırıcının sınanması için, veri kümesindeki tüm kişiler için sırasıyla, bir kişiye ait tüm imgeleri dışarıda bırakarak kalan imgelerle sınıflandırıcının eğitilip, dışarıda bırakılan kişinin imgeleri üzerinde modelin sınıandığı *Bir-Bireyi-Dışarıda-Bırakma* (BDB) yöntemi kullanılmıştır.

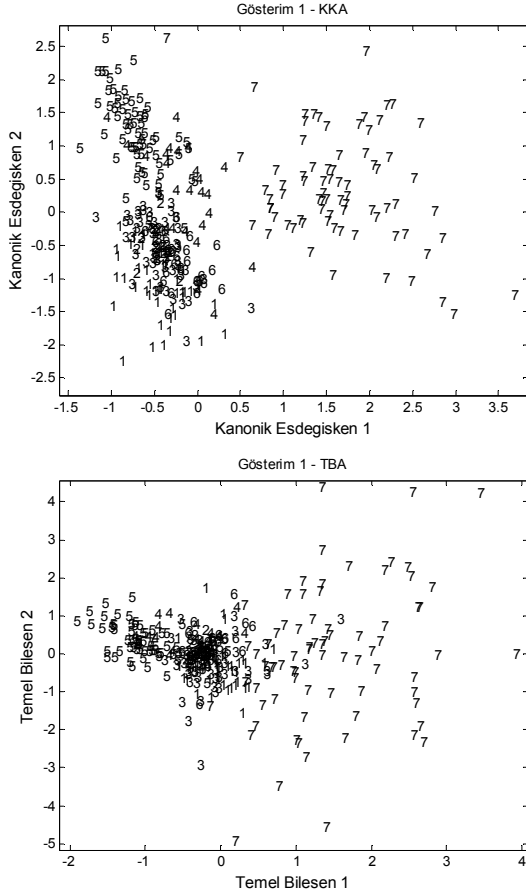
4. DENEYSEL BULGULAR ve TARTIŞMA

CK+ veri kümesinden TBA ve KKA yöntemleriyle çıkartılan özniteliklerin doğrusal ve çokterimli çekirdekli DVM sınıflandırıcılar kullanımıyla elde edilen kestirim doğruluk yüzdeleri Tablo 1'de gösterilmiştir. Tablo 1'de ilk 10 ve ilk 15 öznitelik için KKA yöntemi ile bulunan özniteliklerin DVM sınıflandırma doğruluk oranları TBA özniteliklerinden daha yüksek olduğu görülmektedir. Burada, öznitelik sayısı olarak en az 10 kullanılmasının nedeni, veri boyutunun daha az sayıda özneliğe indirgendiğinde elde edilen kestirim başarısının, boyut indirgeme yapılmadan asıl öznitelikler kullanıldığında elde edilen başarıya göre belirgin bir şekilde düşük olmasıdır. KKA'nın geometrik öznitelikler içeren G_2 gösteriminden çıkarttığı özniteliklerin, görünüm tabanlı G_1 gösteriminden çıkarttığı özniteliklere göre daha doğru sonuçlar verdiği gözlemlenmiştir.

Tablo 1. Özniteliklerin doğrusal çekirdekli DVM (D. DVM) ve çokterimli çekirdekli DVM (Çt. DVM) doğruluk yüzdeleri

	10 öznitelik		15 öznitelik	
	D. DVM	Çt. DVM	D. DVM	Çt. DVM
G_1 - KKA	86.88	82.81	88.44	80.00
G_1 - TBA	81.87	78.75	80.00	77.19
G_2 - KKA	87.81	87.19	90.31	86.25
G_2 - TBA	85.00	77.50	87.81	79.69

Şekil 3’de, G_1 gösteriminden elde edilen ilk iki kanonik eşdeğişkene ve, karşılaştırma amacıyla, TBA ile elde edilen ilk iki temel bileşene olan izdüşümler gösterilmiştir. Hem KKA, hem de TBA ilk 2 boyutlarında ‘şaşıрма’ duygusuna ait sınıfı (sınıf 7) diğer sınıflardan büyük ölçüde ayırmışlardır. Ancak, KKA’nın bu sınıfı diğerlerinden çok net bir biçimde ayırdığı görülürken, TBA’nın izdüşümünde ‘iğrenme’ (sınıf 3) sınıfının bazı örneklerinin ‘şaşıрма’ sınıfının örnekleriyle karıştığı görülmektedir.



Şekil 3: Görünüm tabanlı fark imgesi gösteriminin farklı yöntemlerle çıkartılan ilk iki öz niteliğe olan izdüşümleri. (a) KKA (b) TBA.

Tablo 2 ve Tablo 3’te sırasıyla KKA ile elde edilen 1. ve 2. kanonik eşdeğişkenlerinin artı ve eksi değerlerinin duygulara göre dağılımı gösterilmiştir. Birinci kanonik eşdeğişken ‘şaşıрма’ sınıfına ait örnekleri (artı değerler) diğer sınıfların örneklerinden (eksi değerler) ayırmıştır. Tablo 2’de görüldüğü üzere bu tek boyut ele alındığında, ‘şaşıрма’ sınıfıyla en çok karıştırılan örnekler ‘korku’ sınıfına ait örneklerdir. Toplam 25 ‘korku’ sınıfına ait verinin %28’ine ‘şaşıрма’ sınıfında olduğu gibi artı değerler verilirken, bu oran diğer sınıflar için çok daha azdır (öfke %4, aşağılama %0, iğrenme %3, mutluluk %0, üzüntü %17). Yalnızca, ikinci kanonik eşdeğişken ele alındığında ise (Tablo 3), bu değişkenin ‘mutluluk’ duygusunu diğer sınıflardan ayırdığını, bununla birlikte ‘şaşıрма’ sınıfı örneklerini birçok seferinde ‘mutluluk’ duygusuyla aynı kategoriye koyduğu gözlenmektedir. Aynı şekilde, bu değişken, ‘korku’ sınıfı örneklerinin %64’üne de artı değerler

vermiştir. Birinci kanonik eşdeğişkende birçok örneği eksi, ikincisinde ise artı değerler alan ‘korku’ (sınıf 4) ve ‘mutluluk’ (sınıf 5) örnekleri, Şekil 3(a)’da görüleceği üzere, yalnızca bu iki boyutun kullanılması durumunda, birbirlerine oldukça yakın görünmektedirler.

Elde edilen deneysel bulgular, duygu tanıma için kullanılan imge kipinden çıkartılan iki farklı gösterimin, beraberce ve gözetimsiz olarak KKA yöntemi ile kullanılmasıyla elde edilen eşdeğişkenlerin kestirim başarımının, temel bileşenler (TBA) ile elde edilenden daha yüksek olduğunu göstermektedir. Benzer şekilde, KKA eşdeğişkenlerinin TBA bileşenlerine göre daha iyi veri görselleştirme olanağı sunduğu da görülmüştür.

Tablo 2. KKA ile elde edilen 1. kanonik eşdeğişkenin artı ve eksi değerlerinin duygulara göre dağılımı.

	Ö	A	İ	K	M	Ü	Ş
artı	2	0	2	7	0	5	75
eksi	43	18	56	18	69	23	2

Tablo 3. KKA ile elde edilen 2. kanonik eşdeğişkenin artı ve eksi değerlerinin duygulara göre dağılımı.

	Ö	A	İ	K	M	Ü	Ş
artı	0	1	4	16	67	0	51
eksi	45	17	54	9	2	28	26

5. KAYNAKÇA

- Zeng Z. H., Pantic M., Roisman G. I. and Huang T.S., “A survey of affect recognition methods: audio, visual and spontaneous expressions”, *IEEE Tran. Pattern Analysis and Machine Intelligence*, cilt 31(1), s.39-58, 2009.
- Ruping S., Scheffer T., “Learning with multiple views”, *ICML 2005 Workshop*, 2005.
- Alpaydın E., Yapay öğrenme, Boğaziçi Üniversitesi Yayınevi, 2011.
- Hotelling H., “Relations between two sets of variates”, *Biometrika*, cilt 28, s.321-377, 1936.
- Hardoon D., Szedmak S. and Shawe-Taylor J., “Canonical correlation analysis: an overview with application to learning methods”, *Neural Computation*, cilt 16, s.2639-2664, 2004.
- Kurşun O., Alpaydın E., Fovorov O., “Canonical correlation analysis using within-class coupling”, *Pattern Recognition Letters*, cilt 32, s.134-144, 2011.
- Lucey P., Cohn J. F., Kanade T., Saragih J., Ambadar Z. and Matthews I., “The extended Cohn-Kanade dataset (CK+): a complete dataset for action unit and emotion-specified expression”, *Proc. of IEEE Workshop on CVPR for Human Communicative Behaviour Analysis*, 2010.
- Kanade T., Cohn J., Tian Y., “Comprehensive database for facial expression analysis”, *Proc. of the Fourth IEEE Inter. Conf. On Automatic Face and Gesture Recognition (FG’00)*, Grenoble, France, 2000.
- Chaudhuri K., Kakade S. M., Livescu K. and Sridharan K., “Multi-view clustering via canonical correlation analysis”, *ICML*, 2009.
- Sargin M. E., Yetmez Y., Erzin E., Tekalp M. A., “Audiovisual synchronization and fusion using canonical correlation analysis”, *IEEE Transactions on Multimedia*, cilt 9(7), s.1396-1403, 2007.
- Lin C. J. and Hsu C. W., “A comparison of methods for multi-class support vector machines”, *IEEE Trans. Neural Networks*, cilt 13, s. 415-425, 2002.