

2004-591



**TÜRKİYE BİLİMSEL VE
TEKNOLOJİK ARAŞTIRMA KURUMU**

**THE SCIENTIFIC AND TECHNOLOGICAL
RESEARCH COUNCIL OF TURKEY**

**BAĞIMSIZ BİLEŞEN ANALİZİNİN İLETİŞİM,
İMGE İŞLEME VE JEOFİZİKTESKİ UYGULAMALARI
(APPLICATIONS OF INDEPENDENT COMPONENT
ANALYSIS in COMMUNICATIONS, IMAGE
PROCESSING AND GEOPHYSICS**

**72747
PROJE NO: 102E027**

Elektrik, Elektronik ve Enformatik Araştırma Grubu
Electric, Electronics and Informatics Research Grant Committee

**BAĞIMSIZ BİLEŞEN ANALİZİNİN İLETİŞİM,
İMGE İŞLEME VE JEOFİZİKTESKİ UYGULAMALARI
(APPLICATIONS OF INDEPENDENT COMPONENT
ANALYSIS in COMMUNICATIONS, IMAGE
PROCESSING AND GEOPHYSICS**

72747
PROJE NO: 102E027

PROF. DR. BÜLENT SANKUR

**OCAK 2006
İSTANBUL**

ÖNSÖZ

Bu projede, birbirleriyle karışmış, farklı ağırlıklarla harmanlanmış işaretlerin ayrıştırılması ve işaret kaynaklarının kökenlerine inilmesi problemi ele alınmıştır. Ele alınan işaret kaynağı problemi imgeler, deprem işaretleri, beyin sinyalleri, astrofizik imgeleri gibi çok ve çeşitli olabilir. Kimileyin işaret karışımlarının fiziksel bir gerçekliği vardır; örneğin, astrofizik işaretleri gerçekten big-bang'den kalan ışınım, kozmik ışınım, alet gürültüsü gibi farklı kaynaklara karşılık düşer. Kimileyin ise, ikili siyah-beyaz imgelerde olduğu gibi fiziksel gerçekliği olmayan "algoritmik" ürünlerdir.

Bu çalışmada söz konusu kaynakların birbirlerinden istatistikselce bağımsız oldukları varsayılmıştır. Bunun dışında kalan önemli bir problem sınıfında ise kaynakların bağımlı oldukları gerçeği yatar. Ancak bağımsız kaynak problemleri de yeterince ilginç ve büyük bir problem sınıfını oluştururlar.

Özetle bu çalışmada iki hedefe ulaşılmıştır:

- 1) Kaynak ayrıştırma tekniklerinde algoritmik ve kuramsal ilerlemeler: Burada gürültülü ve durağan olmayan durumlara genelleştirmeler yapılmıştır.
- 2) Kaynak ayrıştırma tekniklerini yeni uygulamalarını gerçekleştirmek: Burada biometri amaçlı el şekli silüetlerine, astrofizik ve deprem işaretlerine, ve nihayet işlevsel yakın kızılaltı bölgesinde ölçülen beyin işaretlerine uygulanmıştır.

Bu proje TÜBİTAK EEEAG tarafından desteklenmiştir.

İÇİNDEKİLER

1. Şekil listesi	4
2. Öz	5
3. Proje özet bilgi formu	7
4. Anahtar Kelimeler	9
5. Projeden Kaynaklanan Yayınlar	9
a. Makaleler	9
b. Uluslararası konferans bildirileri	9
c. Ulusal konferans bildirileri	11
d. Tezler	12
6. Yapılan Çalışmaların Özeti	13
a. Yöntemsel Katkılar	13
i. Durağan Olmayan Ortamlarda Ayrıştırma	13
ii. Çevrimiçi Bağımsız Bileşen Ayrıştırılması için ROTAPEX Algoritması	15
iii. Kaynak Ayrıştırılmasında Farklı Parçacık Süzgeci Yapılarının Karşılaştırılması	16
b. Kaynak Ayrıştırma Uygulamalarına Katkılar	17
i. Jeofizik İşaret İşleme	17
ii. Beyinde Bilişsel İşaretler	18
iii. Kod Bölüşümlü Çoklu Erişim (CDMA) İletişiminde Kaynak Ayrıştırma	22
iv. Biyometride Kaynak Ayrıştırma	23
v. El biyometrisi	23
vi. Yüz Biyometrisi:	26
7. Elde Edilen Sonuçlar	28
a. Yazılımlar	28
b. Eğitime Katkısı	28
c. Birikimlerin Başka Projeye Aktarımı	28
d. Genç Araştırmacıların Eğitimi	28
e. Geleceğe Yönelik Çalışmalar	29

ŞEKİL LİSTESİ

Şekil 1. Değişik alpha değerleri için özbağımlı süreç parametresinin kestirimi (a) $\alpha = 0.5$, (b) $\alpha = 1$, (c) $\alpha = 1.5$, (d) $\alpha = 2$. Parametre sinüsoidal olarak zamanla değişmektedir.

Şekil 2. Değişik alpha değerleri için özbağımlı süreç parametresinin kestirimi (a) $\alpha = 0.5$, (b) $\alpha = 1$

Fig. 3. (a) Özgün kaynakların; (b) karışımların; (c) ROTAPEX yöntemi ile elde edilen kaynakların saçılım grafikleri Satır 1: Karışım 1; Satır 2: Karışım 2; Satır 3: Karışım 3.

Şekil 4 Prototip iYKÖS cihazının şeması

Şekil 5 NIROXCOPE 201'in görüntüsü

Şekil 6: İnsan yüzlerinin PCA analizine göre bileşenleri (birinci sıra) ve ICA analizine göre bileşenleri (ikinci sıra)

Şekil 7: İnsan yüzünün bağımsız bileşenlerden oluşumu

Şekil 8: Bağımsız Bileşenlerin yüz tanımadaki başarımı

TABLO LİSTESİ

Tablo 1. Farklı öznitelik türlerine ve artan nüfusa göre tanıma başarımları (%)

Tablo 2. Farklı öznitelik türlerine ve artan nüfusa göre doğrulama başarımları (%)

ÖZ

Bu projede literatürde kaynak ayrıştırma olarak bilinen istatistiksel işaret işleme yöntemleri hem kuramsal açıdan, hem de uygulamaları açısından ele alınmıştır.

Kuramsal çalışmalarda, kaynakların uzamda ve zamanda birbirlerinden bağımsız olmadıkları düşüncesiyle parçacık süzgeçleri yöntemi ele alınmıştır. Bu yaklaşımda süreç hakkında elimizde varolan önsel bilgilerin algoritmaya ağırdırılması mümkün olmuştur. Ayrıca durağan olmayan süreçlerin de ayrıştırılma problemi ele alınmış, farklı yeniden örnekleme ve önem fonksiyonları seçerek literatürdeki teknikleri aşan sonuçlar bulunmuştur.

Uygulamaya dönük çalışmalarda ise, jeofizik işaret işlemede klimatolojik verilere bakılmış ve gerek Kuzey Atlantik Salınımı diye adlandırılan olgu irdelenmiştir. Ancak topografik etkiler istenen sonuca ulaşılmasına engel olmuştur. Öte yandan sıcaklık verileri faktör analizi ile incelenmiş ve volkanik hareketlerin etkisi ortaya çıkarılmaya çalışılmıştır.

Beyin işaretleri konusunda özellikle prefrontal korteks bölgesinden bilişsel süreçlerle ilgili araştırmalar yapılmıştır. Deneklerin biliş esnasında beyindeki kan dengesinin, oksijen ve deoksi hemoglobinin değişimleri yakalanmaya çalışılmıştır. Bağımsız bileşen analizinin bu türlü dalga biçimlerini ortaya çıkarmakta çok etkili bir araç olduğu gösterilmiştir.

Biyometri alanında, bağımsız bileşenler analizi hem yüz imgelerine hem de el imgelerine uygulanmıştır. Bağımsız bileşenlerin deneklerin kimlik bilgilerini taşıyan ve bozucu etkilere karşı en dayanıklı öznitelikler olduğu görülmüştür.

Nihayet bağımsız bileşen analizi CDMA: kod bölüşümlü çoklu erişim işaretlerine uygulanmış ve sönümlü kanallardaki alıcının performansını iyileştirici tasarımlar irdelenmiştir.

ABSTRACT

In this project, we have investigated the theoretical aspects of the statistical tool called ICA: Independent Component Analysis, and we have applied it to a diverse set of areas.

In the theoretical study, we first addressed the difficult problem where the sources are not independent temporally and spatially. Particle filters can bring a solution to this problem and it enables one to incorporate a priori information into the solution. Second, nonstationary processes are considered. With judicious choice of re-sampling scheme and of the importance function we improved over the results in the literature.

In application-oriented studies, climatological data was investigated. The first problem was the so-called North Atlantic Oscillation: however, it was found that topological effects perturb and in fact preclude any ICA-based solution. The world temperature data was analyzed with a view to determine the effect of volcanic activity on the temperature time series.

Brain signals provide a fertile ground for the application of source separation techniques. Cognitive signals were sensed from several subjects in task-oriented experimental protocols. ICA was instrumental in extracting the brain hemodynamic response waveform from the mixture of brain background activity and non-cortex signals.

Biometry can have different modalities and two of them, hand and face were considered in this project. Decomposition of human faces and hand silhouettes into their independent components proved a very robust method of feature extraction and of combating adverse effects of disturbances.

Finally, application of ICA to the code division multiple access schemes in mobile communication was considered. Designs for improved performance in fading channels were analyzed.

1. cilt

PROJE ÖZET BİLGİ FORMU

1. Proje Kodu : TÜBİTAK-CNR 102E027

2. Proje Başlığı : BAĞIMSIZ BİLEŞEN ANALİZİNİN İLETİŞİM, İMGE İŞLEME VE JEOFİZİKTEKİ UYGULAMALARI

(APPLICATIONS of INDEPENDENT COMPONENT ANALYSIS in COMMUNICATIONS, IMAGE PROCESSING and GEOPHYSICS)

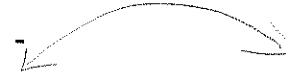
3. Proje Yürütücüsü ve Yardımcı Araştırmacılar :

Ar. Gör. Deniz Genççağa (BÜ)
Ar. Gör. Ender Konukoğlu (BÜ)
Ar. Gör. Erdem Yörük (BÜ)
Ar. Gör. Hazım K. Ekenel (BÜ)
Ar. Gör. Olcay Kalkan (İYTE)
Ar. Gör. Ömer Şaylı (BÜ)
Ar. Gör. Uzay Emir (BÜ)
Doç. Dr. Aysın Ertüzün (BÜ)
Doç. Dr. Sinan Özeren (İTÜ)
Dr. Ercan Kuruoğlu (ISTI-CNR)
Prof. Bülent Sankur (BÜ)
Yrd. Doç. Dr. Ata Akın (BÜ)
Yrd. Doç. Dr. Mustafa Altinkaya (İYTE)
Yük. Müh. Ceyhun Burak Akgül (BÜ)

4. Projenin Yürütüldüğü Kuruluş ve Adresi :

Boğaziçi Üniversitesi
Elektrik-Elektronik Mühendisliği Bölümü
Bebek, İstanbul

5. Destekleyen Kuruluş(ların) Adı ve Adresi :



6. Projenin Başlangıç ve Bitiş Tarihleri : 1 Ocak 2002 – 31 Aralık 2004

7. Öz : (en çok 70 kelime)

İstatistiksel işaret işlemede sıkça farklı işaretlerin doğrusalca harmanlanmış halleriyle karşılaşılır. Bağımsız bileşen analizi (BBA) yöntemleri, giriş işaretlerini ya da karıştıran sistemi de tanımadan bu bileşenleri ayırıştırma teknikleridir. Bu çalışmada gerek BBA'nın jeofizik, imge işleme, biometri, beyin işaretlerindeki uygulamalarına bakılmış, gerekse kuramsal irdelemeler sonucunda BBA tekniği ilerletilmiş, bağımlı veya durağan olmayan işaretlere genelleştirilmiştir.

8. Proje sorumluları

Türk tarafı sorumlusu Prof Bülent Sankur	İtalyan tarafı sorumlusu Dr Ercan E. Kuruoglu
Bağlı olduğu kuruluş Boğaziçi Üniversitesi Elektrik-Elektronik Mühendisliği Bölümü	Bağlı olduğu kuruluş Istituto di Scienza e Tecnologie dell'Informazione – CNR, Pisa
Adres Boğaziçi Üniversitesi, Elektrik-Elektronik Mühendisliği Bölümü 34342 Bebek/İstanbul telefon (+90) 212-3581540-1455 faks (+90) 212-2872465 e-posta sankur@boun.edu.tr	Adres Area della Ricerca CNR di Pisa, via G. Moruzzi 1, 56124, Pisa telefono 050 315 3128 fax 050 315 2810 e-mail ercan.kuruoglu@isti.cnr.it
Başka Türk kuruluşları 1) İstanbul Teknik Üniversitesi, Yerbilimleri Enstitüsü 2) İzmir Yüksek Teknoloji Enstitüsü, Elektrik- Elektronik Mühendisliği Bölümü 3) Boğaziçi Üniversitesi, Biomedikal Enstitüsü	Başka İtalyan kuruluşları

PROJE ADI : TÜBİTAK-NSF 102E027 "APPLICATIONS of INDEPENDENT COMPONENT ANALYSIS in COMMUNICATIONS, IMAGE PROCESSING and GEOPHYSICS"

PROJE NO.:	Başlama Tarihi	Bitiş Tarihi							
102E027	...01/.01./2003....	...31/.12./2004....							
Gelişme Raporu No: 4	Raporun Kapsadığı Dönemin Tarihleri ..01/..08 /2004 - ..31/.01./2005								
TÜBİTAK Tarafından Sağlanan Destek Miktarı	Gelişme Raporu Döneminde Harcanan	Şimdiye Kadar Harcanan	KALAN						
15.232.000.000	4.232.165.675	9.148.757.077	6.083.242.923						
Destek Sağlayan Diğer Kuruluşların Katkısı	Gelişme Raporu Döneminde Harcanan	Şimdiye Kadar Harcanan	KALAN						
-	-	-	-						
TÜBİTAK'ca SAĞLANAN DESTEKTEN ŞİMDİYE KADAR YAPILAN HARCAMALARIN FASILLARA GÖRE DAĞILIMI (*) X 1.000 TL									
Telif	Personel	Tecihaz ve Sabit Yatırım	Sarf Malzeme si	Seyahat ve Nakliye	Hizmet Alımı	Bilgi İşlem Harcama- ları	Kırtasiye ve Basım	Diğer	TOPLAM
		4.601.800. 000	564.970.. 000	3.982.167.0 00					9.148.757.07 7

Proje Yürütücüsünün Adı-Soyadı	İMZASI	TARİH
Bülent Sankur		31 Ocak 2005

9. Anahtar Kelimeler:

Bağımsız bileşenler, Ana bileşenler, Bayeşçi yöntemler, Biometri, Beyin hemodinamik tepkisi

10. Projeden Kaynaklanan Yayınlar :

MAKALELER

- 1) Ceyhun Burak Akgül, Bülent Sankur, Ata Akın, "Selection of relevant frequency bands in functional near infrared spectroscopy", *Journal of Computational Neuroscience* (18) 67-83, 2005.
- 2) Ceyhun Burak Akgül, Ata Akın, Bülent Sankur, "Extraction of cognitive-activity related waveforms form functional near infrared spectroscopy signals," *Annals of Biomedical Engineering*'e yollandı, Ocak 2005
- 3) Sinem Tiveci, A. Akın, T. Çakır, H. Saybaşılı, K. Ülgen, "Modeling of Calcium dynamics in brain energy metabolism and Alzheimer's disease," *kabul edildi, Computational Biology and Chemistry*, Mart 2005
- 4) H.K. Ekenel, B. Sankur, Feature selection in the independent component subspace for face recognition, *Pattern Recognition Letters*, 25, 1377-1388, 2004.
- 5) H.K. Ekenel, B. Sankur, Multiresolution Face Recognition, *Image and Vision Computing*, 23(5), 469-477, 2005.
- 6) E. Konukoğlu, E. Yoruk, J. Darbon, B. Sankur, Shape-Based Hand Recognition, *IEEE Image Processing*, (kabul edildi), 2004.
- 7) E. Yörük, H. Dutagaci, B. Sankur, Hand Biometrics, *Image and Vision Computing*, (under review), 2004.

ULUSLARARASI KONFERANS BİLDİRİLERİ

- 8) Genççağa, D., Kuruoğlu E. E., Ertüzün A., "Estimation of Time-Varying Autoregressive Symmetric Alpha Stable Processes by Particle Filters", *EUSIPCO'2005, The 13th European Signal Processing Conference, Antalya*.
- 9) Ata Akın, Uzay E. Emir, Serhan Kalsın, Ömer Şayli, "Frequency components in breathholding experiments," *Proc. SPIE, Photonics West 22-26 Ocak 2005'te San Jose, USA*
- 10) Akın, S. Bunce, "Behavioral reaction times are correlated with frequency shifts in optical imaging signals," *Proc. SPIE, Photonics West 22-26 Ocak 2005'te San Jose, USA*
- 11) Ata Akın, Uzay E. Emir, Didem Bilensoy, Gulin Erdogan, Selcuk Candansyar, Hayrunnisa Bolay, "fNIRS measurements in migraine," *Proc. SPIE, Photonics West 22-26 Ocak 2005'te San Jose, USA*
- 12) Emir E. U., Şayli Ö., Aydın K., Demiralp T., Ademoğlu A., Öztürk C., Akın A., "Effect of repetitive breathholding on BOLD and fNIRS signals," *BMES 2004 Fall Meeting, 14-16 Ekim*

2004, Philadelphia, USA

- 13) Uzey Emir, Ömer Şayli, Ata Akın, "Investigation of breatholding effect on cerebral hemodynamics by functional near-infrared spectroscopy" AGORA 2004, Kuşadası, Turkey
- 14) Ceyhun B. Akgül, Bülent Sankur, Ata Akın, "Evidence of cognitive activity in fNIRS data," OSA Biomedical Optics Topical Meeting, 2004
- 15) Uzey E. Emir, Ceyhun B. Akgül, Ata Akın, Aysin Ertüzün, Bülent Sankur, Kerem Harmancı, "Wavelet Denoising Vs Ica Denoising For Functional Optical Imaging", Proceedings of the 1st International IEEE EMBS Conference on Neural Engineering Capri Island, Italy, March 20-22, 2003, 384-387
- 16) Ata Akın, S. Bunce, "Quantification of Brain Hemodynamic Response Measured by functional Optical Imaging", Proceedings of the 1st International IEEE EMBS Conference on Neural Engineering Capri Island, Italy, March 20-22, 2003, 503-506
- 17) M. S. Özeren, E. Kuruoğlu, B. Uzunoğlu, An IFA-based machine learning algorithm for the separation of climate signals, EUSIPCO'2005, The 13th European Signal Processing Conference, Antalya
- 18) M.A. Altinkaya, O. Kalkan ve E.E. Kuruoğlu, "Independent Factor Analysis for Estimating Fading Channels in CDMA Communication", IEEE Workshop on Statistical Signal Processing (SSP'05)'ya yollandı, Bordeaux, Fransa, Temmuz 17-20, 2005.
- 19) Erdem Yörük, Ender Konukoğlu, Bülent Sankur, and Jerôme Darbon, Person Authentication Based On Hand Shape, EUSIPCO'2004: European Conf. On Signal Processing, 1233-1236, Vienna, September 2004.
- 20) E. Yoruk, H. Dutagaci, B. Sankur, Hand-Based Biometry, SPIE Electronic Imaging Conference: Image and Video Communications and Processing, 16-20 January 2005, San Jose, USA.

ULUSAL KONFERANS BİLDİRİLERİ

- 21) Gençğa, D., Kuruoğlu E. E., Ertüzün A., "Zamanla Değişen Özbağlanımlı Cauchy Süreçlerinin Parçacık Süzgeçleri ile Kestirimi", *SIU 2005: 13. Sinyal İşleme ve İletişim Uygulamaları Kurultayı'na sunulmak üzere gönderildi.*
- 22) Gençğa D. ve A. Ertüzün, "Çevrimiçi Bağımsız Bileşen Ayrıştırılması", *11. İşaret İşleme ve Uygulamaları Kurultayı (SIU2003)*, Haziran, 2003, Türkiye, sayfa. 7-10.
- 23) M. S. Özeren, H. Tatlı, H.N. Dalfes, Kuzey Atlantik salınımının klimatolojilere etkisinin BBA (ICA) ile çözümlenmesi, SIU 2003: 11. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, Koç Üniversitesi.
- 24) Ceyhun Burak Akgül, Ata Akın, Bülent Sankur, "İncelevel Optik İşaretlerde Bilisel Etkinliğin

Varlığı", SIU 2004: 12. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, Kuşadası.

- 25) Ceyhun Burak Akgül, Ata Akın, Bülent Sankur, "İşlevsel Optik İşaretlerden Bilişsel Etkinlikle İlişkili Dalga biçimlerinin Çıkarılması", SIU 2004: 12. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, Kuşadası.
- 26) Uzay Emrah Emir, Ömer Şayli, Kubilay Aydın, Tamer Demiralp, Ahmet Ademoğlu, Cengizhan Öztürk, Ata Akın, "Nefes tutmanın KOSB fMRI ve iYKÖS işaretlerine etkisi," BİYOMUT 2004
- 27) Ceyhun B. Akgül, Bülent Sankur, Ata Akın, "İşlevsel Optik Görüntüleme Zamansal Olay Sezimi" SIU 2003: 11. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, Koç Üniversitesi.
- 28) Ata Akın, "İşlevsel Optik Görüntüleme Sisteminde İşaret Modelleri" SIU 2003: 11. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, Koç Üniversitesi.
- 29) O. Kalkan ve M.A. Altınkaya, "Kod bölüşümlü çoklu erişim (CDMA) iletişimde Gauss olmayan sönümlü kanal kestirimi için Pearson sistemine dayalı gözü kapalı kaynak ayrıştırma yöntemi", IEEE 12. Sinyal İşleme ve İletişim Uygulamaları Kurultayı SIU 2004, 534-537, Kuşadası, 2004.
- 30) Hazım K. Ekenel, Bülent Sankur, Bağımsız Bileşen Analizi ile Yüz Tanımada Öznitelik Seçimi, 11. SIU Kurultayı, s. 15-18, Koç Üniversitesi, İstanbul, 18-20 Haziran, 2003.
- 31) Hazım K. Ekenel, Bülent Sankur, Çoklu Çözünürlük Analizine Dayalı Yüz Tanıma Sistemi, 11. SIU Kurultayı, s. 640-643, Koç Üniversitesi, İstanbul, 18-20 Haziran, 2003.
- 32) E. Konukoglu, E. Yoruk, B. Sankur, J. Darbon, El Şekillerine Dayalı Kişi Tanıma, 12. SIU Kurultayı, s. 152-155, Kuşadası, 28-30 Nisan, 2004

TEZLER

- 33) H. Kemal Ekenel, Expression and Illumination Insensitive Independent Components and Wavelet Subbands for Face Recognition, MS Tezi, B.Ü., 2003.
- 34) Ceyhun Burak Akgül, Analysis of functional Near Infrared Spectroscopy Signals, MS Tezi, B.Ü., 2004
- 35) Erdem Yörük, Hand Based Biometry, MS Tezi, B.Ü., 2004.
- 36) Senem Tiveci, Role of Calcium Dynamics in brain energy metabolism, MS Tezi, Boğaziçi Üniversitesi, Eylül 2004
- 37) Meryem Çakıroğlu, Analysis of brain hemodynamic response by functional near infrared spectroscopy, MS Tezi, Eylül 2003
- 38) Uzay E. Emir, System Characterization for a fast optical imager, MS Tezi, Boğaziçi Üniversitesi, MS Tezi, Haz. 2003
- 39) O. Kalkan, "Independent Component Analysis Applications in CDMA Systems", MS Tezi, İzmir Yüksek Teknoloji Enstitüsü, Temmuz, 2004.

11. Bilim Dalı :

Doçentlik B. Dalı Kodu :

12. Yapılan Çalışmaların Özeti

İşaret ya da kaynak ayrıştırma birçok gerçek problemde karşımıza çıkan bir sorundur. Kaynak ayrıştırma uygulamalarına konuşma işlemeden astrofizik imgelerin sınıflandırılması, biometrik veri analizinden işlevsel biyolojik işaretlere, finansal zaman serilerinden Internet trafik analizine, mobil iletişim alıcılarından otomatik belge okumaya kadar geniş bir spektrumda rastlanılır.

Kaynak ayrıştırma araştırmaları son 10 yıldan beri hız kazanmış olup, kaynakların gerçekten bağımsız olup gürültüye bulaşmamış olduğu durumlarda iyi sonuç veren birçok algoritma vardır. Ne var ki gerçek yaşamda karşılaşılan problemlerin çoğunda ya kaynaklar tümüyle bağımsız değildir, ya doğrusal-olmayan işlemlere maruz kalarak birleşmişlerdir ya da gürültülüdür. Üstelik kaynak sayısı gözlemci sayısı ile aynı olmayabilir ya da kaynaklar durağan olmayabilirler.

Bu çalışmanın iki ana hedefi şunlardır:

- 3) Kaynak ayrıştırma tekniklerinin yeni uygulamalarını gerçekleştirmek
- 4) Kaynak ayrıştırma tekniklerini gürültülü, bağımlı, durağan-olmayan durumlara genelleştirmek.

12A. Yöntemsel Katkılar

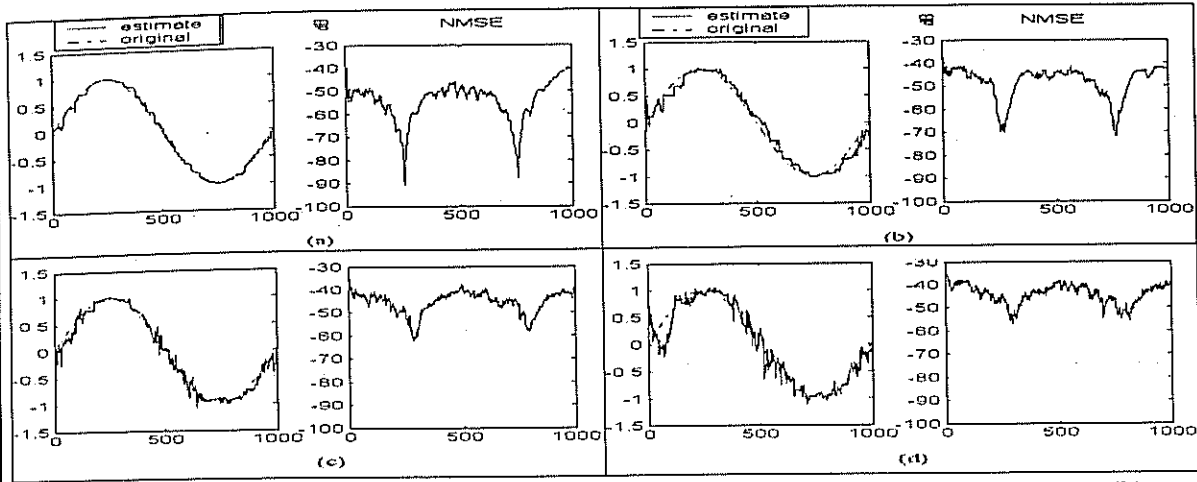
Kaynak ayrıştırmada yaptığımız yöntemsel yenilikler şunlardır:

12A.1. Durağan Olmayan Ortamlarda Ayrıştırma *Proje grubu: Aysin Ertüzün, Deniz Genççağa, Ercan Kuruoğlu*

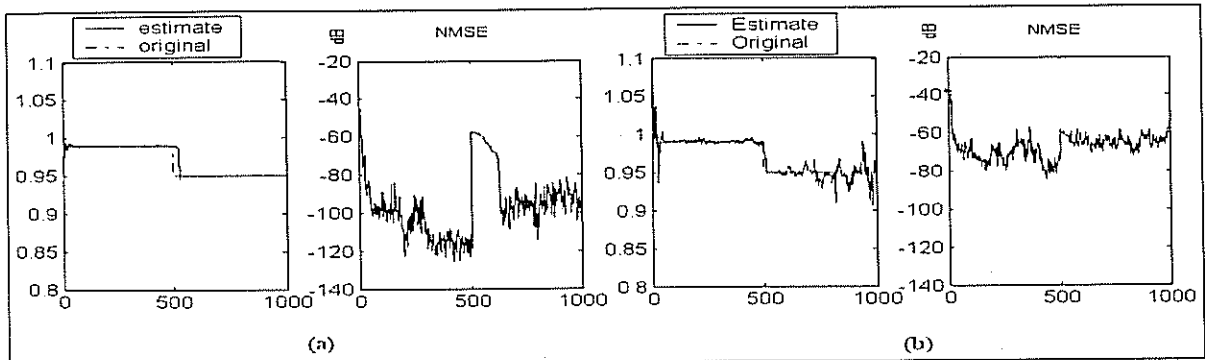
Literatürdeki yöntemlerin tümünde, karışım sistemindeki gürültü durağan olarak varsayılmaktadır. Literatürdeki bazı çalışmalarda, kaynakların durağan olmama durumuna değinilmiş, fakat hiçbirinde kaynak ayrıştırmada kullanılmak üzere bu durağansızlık parametrik olarak modellenmemiştir. Bu çalışmada biz, literatürde ilk olarak, durağan olmayan kaynakların, durağan olmayan gürültülü durumlarda ayrıştırılmasına ilişkin bir çözüm getirdik. Bu yöntem, ayrıca, karışım matrisinin de zamanla değiştiği durumlara da çözüm olanağı sağlamaktadır.

Gauss dağılımının söz konusu olduğu durumlarda, çözüm Kalman süzgeci ile verilebilir. Fakat, Gauss olmayan durumlar için, nispeten yeni bir teknik olan parçacık süzgeçlerinden yararlanılmıştır. Benzer bir yaklaşım Ahmed v.d. [1] tarafından kullanılmıştır. Biz, onun yaklaşımını farklı önsel dağılımlar, yeni önem fonksiyonları (importance function) ve yeniden örnekleme (resampling) yapıları kullanarak genişlettik. Ayrıca, onun çalışmasını imge ayrıştırılmasına uyguladık [2]. Buna ilaveten, aşağıda belirtildiği gibi, tekniğin astrofizik imge ayrıştırılmasına uygulanması üzerinde de çalıştık [3]. Burada, durağansızlığın, [4]'teki gibi karışımında değil de, işaretlerde ve gürültüde olduğuna dikkat edilmelidir.

Bunlara ilaveten, bu proje kapsamında, zamanla değişen, özbağlanımlı (autoregressive) süreç kestirimi problemi de incelenmiş ve literatürde ilk olarak zamanla değişen özbağlanımlı Cauchy süreçlerinin [5] ve daha genel bir durum olan zamanla değişen özbağlanımlı alpha-karalı süreçlerin [6] kestirimi gerçekleştirilmiştir. Bu deneylerden bazılarının sonuçları Şekil 1 ve 2'de verilmiştir.



Şekil 1 Değişik alpha değerleri için özbağlanımlı süreç parametresinin kestirimi (a) $\alpha = 0.5$, (b) $\alpha = 1$, (c) $\alpha = 1.5$, (d) $\alpha = 2$. Parametre sinüsoidal olarak zamanla değişmektedir.



Şekil 2. Değişik alpha değerleri için özbağlanımlı süreç parametresinin kestirimi (a) $\alpha = 0.5$, (b) $\alpha = 1$ Parametre $t=500$ anında aniden 0.99'dan 0.95'e değişmektedir

- [1] A. Ahmed, C. Andrieu, A. Doucet, P. J. W. Rayner, On-line Non-stationary ICA Using Mixture Models. Proc. IEEE ICASSP, (2000) Vol. 5, 3148–3151.
- [2] M. Costagli, E. E. Kuruoglu, "Image Separation Using Particle Filters", Teknik Rapor, ISTI-2004-TR-46, ISTI-CNR, Pisa, Italya, 2004.
- [3] M. Costagli, E. E. Kuruoglu, A. Ahmed, "Astrophysical separation images using particle filters," International Conference on Independent Component Analysis and Blind Signal Separation (ICA), n.5 (Granada, Spagna, Eylül 22-24, 2004), Lecture Notes in Computer science (LNCS), v. 3195 pp. 930-937, García Puntonet, Carlos and Prieto, Alberto, (ed), Springer Verlag 2004.
- [4] Everson R., Roberts S., "Particle Filters for Non-Stationary ICA" in *Advances in Independent Component Analysis*, Ed. Mark Girolami, Springer, 2000.
- [5] D. Genççağa, E.E. Kuruoğlu, Ertüzün A., "Zamanla Değişen Özbağlanımlı Cauchy Süreçlerinin Parçacık Süzgeçleri ile Kestirimi", Submitted to the 13th IEEE Signal Processing and Communication Applications Conference (SIU2005), Kayseri, Turkey..

[6] Genççağa, D., Kuruoğlu E. E., Ertüzün A., "Estimation of Time-Varying Autoregressive Symmetric Alpha Stable Processes by Particle Filters", *Submitted to EUSIPCO 2005 Conference*.

12A.2. Çevrimiçi Bağımsız Bileşen Ayrıştırılması için ROTAPEX Algoritması

Proje grubu: Ayşın Ertüzün, Deniz Genççağa, Ercan Kuruoğlu

Bilindiği gibi, Bağımsız Bileşen Ayrıştırılması (BBA) sonucu elde edilen kaynaklar ile bulunması gereken orijinal halleri arasında bir permütasyon ve ölçek farkı vardır[1]. Bu çalışmada, bahsi geçen ölçek (değişim) farkının önlenildiği yeni bir metod geliştirilmiştir [2]. Önerilen algoritma, temel olarak iki-aşamalı bir BBA metodu olup, ardışıl olarak önce beyazlatma, ardından da döndürme işlemlerinin çevrimiçi olarak gerçekleştirilmesi esasına dayanır [2]. Buradaki amaç, BBA'nın çevrimiçi bir şekilde gerçekleştirilmesi olduğu için, ilintisizleştirme işlemi için Uyarlamalı Temel Bileşen Çıkarımı (APEX) yöntemi kullanılmış, ardından da beyazlatma için çevrimiçi bir güç normalizasyonu uygulanmıştır [2]. Daha sonra, kaynakları elde edebilmek için Ardışıl Givens Döndürmelerinden çevrimiçi bir şekilde yararlanılmıştır. Bilgisayar simülasyon sonuçları, geliştirilen metodun, BBA'nın doğasında bulunan ölçek farkının önlenilmesinde başarılı olduğunu ve bu perspektifte, çevrimiçi BBA uygulamalarında oldukça ümit verici olduğunu göstermiştir. Bu yöntemle a yırıştırılan kaynaklar 3'te verilmiştir.

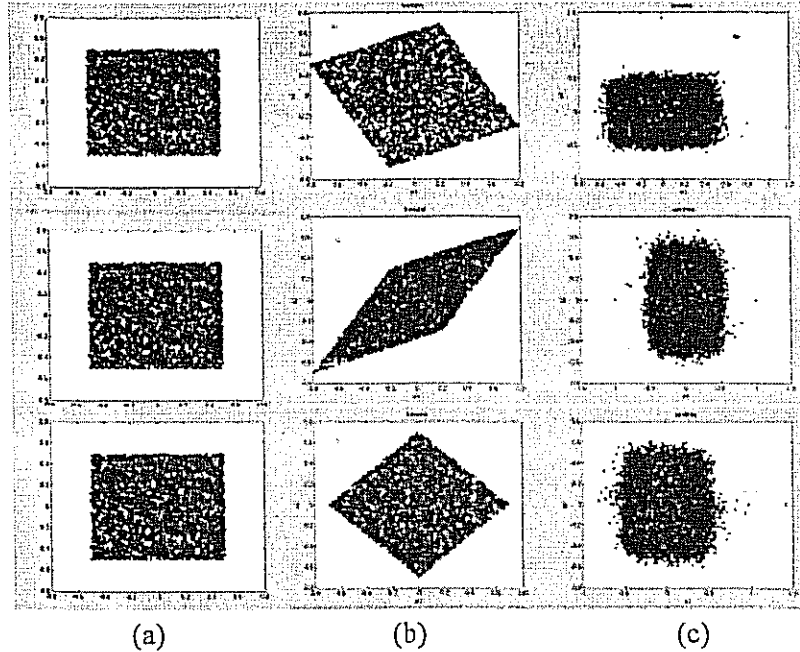


Fig. 3. (a) Özgün kaynakların; (b) karışımların; (c) ROTAPEX yöntemi ile elde edilen kaynakların saçılım grafikleri Satır 1: Karışım 1; Satır 2: Karışım 2; Satır 3: Karışım 3.

[1] Hyvarinen A., Karhunen J. and E. Oja, *Independent Component Analysis*, John Wiley, 2001.

[2] Genççağa D. ve A. Ertüzün, "Çevrimiçi Bağımsız Bileşen Ayrıştırılması", 11. İşaret İşleme ve Uygulamaları Kurultayı (SIU2003), Haziran, 2003, Türkiye, sayfa. 7-10.

12A.3. Kaynak Ayırıştırılmasında Farklı Parçacık Süzgeci Yapılarının Karşılaştırılması

Proje grubu: Ayşın Ertüzün, Deniz Genççağa, Ercan Kuruoğlu

Projenin bu bölümünde, Bağımsız Bileşen Ayırıştırılması (BBA) teorik açıdan ele alınmış ve hem zamansal hem de uzamsal önsel bağımlılık bilgilerinin geliştirilen yönteme nasıl entegre edilebileceği üzerinde durulmuştur. Bu amaçla, araştırma kapsamında ardışıl Monte Carlo yöntemleri olan “Parçacık Süzgeçleri” kullanılmış ve elimizde kaynaklar hakkında bulunan önsel bilgilerin sisteme bu sayede Bayesçi bir modelleme ile entegrasyonu sağlanabilmiştir. Bu amaçla geliştirilen “Parçacık Süzgeçleri” yazılımı, BBA problemine iki farklı şekilde uygulanmış (Algoritma 1 ve 2) [1] ve bağımsız kaynakların bulunmasına çalışılmıştır. Bu inceleme sonucunda, geliştirilen yöntemin çok tatmin edici bir başarıyı sağlayamamasının muhtemel sebepleri içerisinde bulunan “en iyi önem fonksiyonu” seçimi konusu üzerinde durulmuş ve şu sonuçlar çıkarılmıştır: Modellenen sistemde, durum değişkenlerinin zamansal evrimini belirten “durum geçiş denklemi” (process equation) hakkında önsel bir bilgimizin olmaması, Bootstrap algoritmasının başarımını olumsuz yönde çok etkilemektedir. Bahsi geçen zamansal evrim bilgisi olmadan da “en iyi önem fonksiyonu”nun bulunması imkansız olduğu için, kullanılan parçacıkların geçmiş anlardaki veriye bağlı olarak gerçekleştirilen güncellemeleri, modelin bilinmediği durumlarda kullanılabilir. Bu amaçla, “durum geçiş denklemi” bir “rasgele yürüyüş” (random walk) modeli kullanılarak modellenebilir ve evrimi sağlayan durum güdültüsü de sıfır ortalamalı Normal bir dağılımla modellenebilir [2]. Bu Normal dağılımın değışintisi de, geçmiş zamanlardaki parçacıkların modellediği olasılık yoğunluk fonksiyonunun değışintisinden özyineli (iterative) olarak bulunabilir. Bu şekilde bir “önem fonksiyonu” seçilmesinin başarımdaki olumlu etkisi deneylerle gösterilmiştir [2].

Önem fonksiyonunun seçilmesi problemi bu şekilde aşıldıktan sonra, kullanılan parçacık süzgeç modelinin BBA için elverişli olup olmadığı üzerinde daha detaylı bir inceleme yapılmıştır. Bu inceleme sonucunda, iki gözlem modeli [2]’de karşılaştırılmıştır. Bu modeller “Form 1” ve “Form 2” olarak adlandırılmış ve karşılaştırılmıştır. Kısaca, birinci yapılarda, gözlem verisi direkt olarak kendisinin zamanda gecikmiş değerlerine bağlıdır. Bu model altında geliştirdiğimiz yöntemler oldukça başarılı bir şekilde çalışmaktadır [3,4]. Araştırmanın diğer bir kolu da, parçacık süzgeçlerinde kullandığımız modelin, Form 2 tipi uygulamalardaki başarımının neden yeterli düzeyde olmadığına olası sebeplerinin incelenmesi üzerinde yoğunlaşmıştır. BBA, bu grup içerisinde yer almaktadır. Bu amaçla, daha önceki yaklaşımlarımız olan ve Rao-Blackwellization işleminden yararlanan algoritmalar yerine, [5]’de önerilen ve integrasyon yaklaşımı kullanan diğer bir yöntem denenmiş ve tek-değişkenli (univariate) durum için başarılı sonuçlar elde edilmiştir [2]. Dolayısıyla da şu ana kadar problem çıkan Form 2 tipi modellerin çözülebilmesinde önemli bir gelişim sağlanmıştır [2]. Bu yöntemin şu sıralarda çok-değişkenli (multivariate) durumlara, kaynaklar arasında olasılık yoğunluk fonksiyonları üzerinden bağımlılık bilgisini sokmak suretiyle, genelleştirilmesi üzerinde çalışılmaktadır. Bu çalışmanın bağımlı kaynakların ayırıştırılabilmesindeki tıkanıklığın önünü açacağı düşünülmektedir.

- [1] Genççağa, D., Probabilistic Filter Theory and Source Separation, *Doktora Tezi İkinci Altı Aylık Gelişim Raporu*, Boğaziçi Üniversitesi, Ağustos 2004.
- [2] Genççağa, D., Particle Filters and their Applications, *Doktora Tezi Üçüncü Altı Aylık Gelişim Raporu*, Boğaziçi Üniversitesi, Ocak 2005.
- [3] Genççağa, D., Kuruoğlu E. E., Ertüzün A., "Estimation of Time-Varying Autoregressive Symmetric Alpha Stable Processes by Particle Filters", *EUSIPCO 2005 Konferansına sunulmak üzere gönderildi*.
- [4] Genççağa, D., Kuruoğlu E. E., Ertüzün A., "Zamanla Değişen Özbağlanımlı Cauchy Süreçlerinin Parçacık Süzgeçleri ile Kestirimi", *SİU 2005 Konferansına sunulmak üzere gönderildi*.
- [5] Everson R., Roberts S., "Particle Filters for Non-Stationary ICA" in *Advances in Independent Component Analysis*, Ed. Mark Girolami, Springer, 2000.

12B. Kaynak Ayırıştırma Uygulamalarına Katkılar

12B.1. Jeofizik İşaret İşleme

Proje grubu: Sinan Özeren, Ercan Kuruoğlu

Jeofiziksel sinyal çözümlemesi çalışmalarımız şu ana kadar daha çok klimatolojik sinyallerin ayrıştırılması üzerinde yoğunlaştı. İlk adım olarak Bağımsız Bileşen Analizi (BBA) ile Kuzey Atlantik Salınımı'nın (KAS) Türkiye'deki yağış ölçümlerine olan etkisini ayırştırmak için kullandık. KAS Kuzey Atlantik'de etkili olan önemli bir atmosfer salınımı olayını betimlemeye yarayan bir basınç indisi. Bu salınımın geniş ölçekte sadece yağışları değil başka olayları da etkilediği klimatologlar ve atmosfer bilimciler tarafından iyi biliniyor (örneğin KAS, Kuzey Denizi-Hazar salınımıyla kuplaja girebiliyor). Biz, BBA için gereken, karışmış sinyallerin sayısı için bir varsayım yaparak uzay-zaman ortamında (yağış kayıtları tek bir yerden değil birkaç bölgeden alınıyor). Analiz sonuçları 2003 SIU toplantısındaki bildirimizde de belirttiğimiz gibi ICA'nın bu tür çözümlemelerde pek işe yaramadığını ortaya çıkarmıştır (sonuçları anlamlandırmak mümkün ise de ICA, atmosfer bilimlerinde sıkça kullanılan tekil spectrum analizi gibi boyut indirgeme metodlarının ötesinde bir işlevsellik göstermemiştir). Bu durumun bizce temel olarak üç nedeni vardır

- 1- Kullandığımız FASTICA algoritması gürültü faktörü için bir stratejiye sahip değildir
- 2- KAS sinyalinin Avrupa'nın Orta ve Doğu Avrupa'daki yağış rejimlerine etki ettiği basit istatistiksel yöntemlerle bile tespit edilebilir olsa da sinyal ayıklama bağlamında problem çok daha zordur. KAS sinyali aslen lineer olmayan bir atmosferik filtreden geçerek yağış verilerine yansımaktadır. Söz konusu non-lineerlik atmosferin hareketini betimleyen denklemlerin konvektif terimler içermesi ve bu nedenle de tek bir harmonikten çok sayıda harmonic oluşturabilmesinden kaynaklanmaktadır. Bunun pratik anlamı original sinyalden birbirinden bağımsız (Fourier serisi cinsinden ifade edildiklerinde her biri farklı bir moda tekabül eden) bir çok ölçeğin ortaya çıkmasının mümkün olmasıdır.
- 3- Başlangıçta veriye Temel Bileşen Analizi (TBA) uygulanarak problemin asal olarak kaç boyut içerdiği konusunda bir fikir edinilip BBA'da bu sayı kullanılsa bile topografik etkiler bunu yapmayı anlamsızlaştırabilmektedir, zira lokal olarak yağış topografik değişimlerden etkilenebilir. Bu raporun

yazarı, yukarıda bahsedilen çalışmadan edindiği tecrübeye göre topografik etkilerin bulunduğu bir ortamda BBA ile sinyal ayrıştırması yapmaya kalkmanın anlamsız bir iş olacağına inanmaktadır.

Çalışmanın ikinci aşaması olarak küresel sıcaklık verileri üzerinde çalışmaya başladık (herkese açık olan CDC verileri; <http://www.cdc.noaa.gov/PublicData/>). Bu ikinci çalışmada Bağımsız Faktör Analizi (BFA) kullanmaya karar verdik. Bu seçimin birinci nedeni, şartlı olasılık argümanları ile bir minimizasyon algoritması kurmaya elverişli olması ve BBA gibi karışma matrisinin bir kare matris olması kısıtlamasının olmaması. Global sıcaklık verileri (bir çok yükseklik için) standart faktörlerin yanında (güneş radyasyonu vs) KAS, EL_NINO ve volkanik patlamalardan da etkileniyor. Bunların arasında spekülasyona en uygun olanı volkanik patlama etkileri, sayısal iklim simülasyonları bu etkiyi bir dereceye kadar ortaya çıkarabiliyor. Volkanik etki çok kısa bir süre önce bir grup Amerikalı araştırmacı tarafından BBA ile çözümlenmeye çalışıldı (http://www.llnl.gov/CASC/sapphire/sep_climate/) ancak olayın doğası, bu raporun yazarına göre, kare karışma matrisi kullanmaya elverişli değil. Biz, çalışmanın ilk aşamasında CDC verilerinin enlem ortalamalarını hazırladık (bu oldukça standart bir iş, zira volkanik etkinin enlemler boyunca değil kuzey-güney ekseninde bir çeşitlilik yaratmasını bekliyoruz). Daha sonra bu işlemi aylık, mevsimsel, yıllık ve dört yıllık zaman ortalamaları içeren imgeler üzerinde tekrarladık. Her durumda bir imge bir eksende boylamları, öbür ekseninde de basınç seviyelerini (bir anlamda yükseklikleri) içeriyor. İmgeler bu anlamda düşey düzlemde (x-z) tanımlanıyorlar. Söz konusu işlemler 1979-2000 arasındaki tarihler için yapıldı, bu dönem içinde iki büyük volkanik patlama var (1982 El Chicon and 1991 Pinatubo patlamaları). Şu anda KAS, EL_NINO ve volkanik patlamalara tekabül eden orijinal sinyallere ulaşmak için BFA bazlı bir öğrenme algoritması üzerinde çalışıyoruz (bu çalışmanın birincil sonuçlarını EUSIPCO 2005 kongresinde sunmak üzere gönderdik, makale şu sırada değerlendiriliyor). Öğrenme algoritması, kaynaklar için a-priori kabul edilen olasılık dağılımı fonksiyonunun parametrelerini ayarlayarak Kullback-Leibler integral ifadesini minimize ediyor, bu minimizasyon, bilindiği gibi, kaynaklar için bütün mertebelerden istatistiksel bağımsızlığı garantiliyor.

1. M. S. Özeren, H. Tatlı, H.N. Dalfes, Kuzey Atlantik salınımının klimatolojilere etkisinin BBA (ICA) ile çözümlenmesi, 2003 11. Sinyal İşleme ve İletişim Uygulamaları Kurultayı (SIU).
2. An IFA-based machine learning algorithm for the separation of climate signals, M. S. Özeren, E. Kuruoğlu, B. Uzunoglu, (under review for EUSIPCO-2005, Antalya)

12B.2. Beyinde Bilişsel İşaretler

Proje grubu: Ata Akın, Büilent Sankur, Koray Çiftçi, Uzay Emir, Ceyhan Burak Akgül

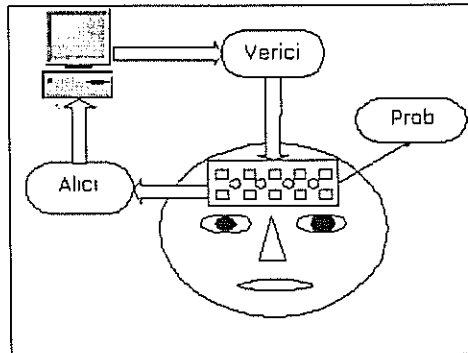
Beyin dinamiklerinin incelenmesinde sinirgörüntüleme teknikleri hızla gelişmektedir. Özellikle bilişsel etkinliklerin, normal ve sağlıklı beyinlerde yol açtığı elektriksel, biyokimyasal ve damarsal değişimleri gözlemlemek yararlı olmaktadır. Ne var ki beyin gibi karmaşık bir organ fizyolojik ölçümlerle gözlemlendiğinde tek bir işlevsel tepkiyi ölçmek mümkün olmamakta, aynı anda damar-sinir sisteminden gelen birden çok olgu kaydedilmektedir. Yeni bir nöral görüntüleme (neuro-imaging)

tekniki, işlevsel yakın kızılötesi spektroskopisi sistemi (iYKÖS) (fNIRS: functional Near Infrared Spectroscopy) sayesinde alnın üstüne yerleştirilen problar ile prefrontal kortekste fizyolojik değişimleri izlemek mümkün olmaktadır. Probon yaydığı belirli ışık dalgaboyları kandaki oksijen ve deoksijen hemoglobin moleküllerine (HbO ve Hb) duyarlıdır ve bu moleküllerin belirli bir hacim içerisindeki derişim değişiklikleri yansıtır. Özellikle yoğun bilişsel etkinlik içerisinde giren beyinde çalışan sinir hücrelerinin gereksinim duyduğu oksijeni taşıyan bu moleküllerin derişimindeki değişimler, beynin hangi bölgesinin ne kadar etkin bir şekilde çalıştığı hakkında bilgiler sunmaktadır.

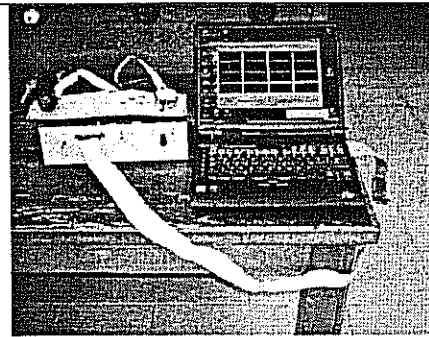
Bu doğrultuda çalışmalarımız: 1) iYKÖS donanımını gerçekleştirmek, 2) Taşınabilir iYKÖS sistemini tasarlamak; 3) MR uyumlu iYKÖS sistemi tasarlamak; 4) Çeşitli bilişsel etkinlikler altında deneklerden ölçümler almak; 5) Bağımsız bileşenler analizi, kanonik korelasyon analizi, entropik topaklandırma gibi yöntemlerle kaynakları ayırıştırma; 6) Bilişsel iYKÖS işaretlerinin spektral analizini yapma. Gerçekleştirilmiş işler şöyle sıralanır:

Donanım tasarımı ve gerçekleştirme:

1. NIROSCOPE 101, 201 ve 301 modelleriyle üç sürüm halinde iYKÖS sistemi gerçekledik. İlk sistem (101) laboratuvarında çalışabilecek bir sistem iken 201 kutulanmış ve hasta başına götürülebilecek bir sistem halini almıştır.



Şekil 4 Prototip iYKÖS cihazının şeması



Şekil 5 NIROSCOPE 201'in görüntüsü

Bu sistem ile çeşitli öncül ve klinik ölçümler alınmıştır. Özellikle nefes tutma gibi temel fizyolojik uyarılar esnasında yaptığımız ölçümlerde net olarak beyin dinamiklerini görme ve nicelendirme şansını yakaladık. Bunun dışında aynı sistemle migren hastalarından, depresif hastalardan klinik ölçümler alırken lise öğrencilerinden ise bilişsel süreçler esnasında ölçümler aldık. Çalışmalarımızda aynı zamanda kaslar üzerinden de ölçümler aldık. Sporcular ve sedanterler üzerinden aldığımız ölçümlerde özellikle damar dinamiklerinin ve kasların oksijen tüketim hızlarının bu iki grup arasındaki farklarını gözlemlemeye çalıştık.

2. NIROSCOPE 301 sistemi 201'in aksine bilgisayara ek olarak takılan veri toplama ünitesini sistem içerisinde tutmaktadır. Bu haliyle sistem USB tabanlı bir iYKÖS sistemi olmuştur. Bu sistem şu an veri toplamaya hazır hale gelmiştir. Bundan sonra tüm ölçümlerimizi bu sistem ile

gerçekleştireceğiz. 201'in benzerleri dünyada bir gurup arařtırmacı tarafında yapılmıřken, USB tabanlı bu sistem henüz hiřbir yerde yoktur.

3. Sistemlerimizin ölçümlerini doęrulamak için ayrıca MR uyumlu fiberoptik kablo tabanlı bir sistem daha gerçekleřtirdik. Bu sistem, MR aygıtı içerisinde MR ile eř zamanlı olarak Hb ve HbO ölçümleri alabilecek duyarlılıktadır. Bu haliyle sistem tasarımı bitmiř ve denemeye hazır hale gelmiřtir. Önümüzdeki aylarda ilk klinik ölçümlerimizi almayı planlamaktayız.

İřaret İřleme Adımları:

- (i) *Biliřsel etkinlięin nicelendirilmesi amacıyla yapılan insanlı deneylerde, iYKÖS yoluyla toplanan oksihemoglobin (HbO₂) deęiřimi zaman-dizisi verilerinin istatistiksel analizi.*

İstatistiksel analiz baęlamında, iřaretler üzerinde istatistiksel duraęanlık ve Gaussluk testleri yapıldı. Buna göre, HbO₂ iřaretlerinin genel anlamda duraęan bir süreçten gelmedięi, buna karřılık kısa-dönemli duraęanlıęın korunması gereken bir varsayım olduęu anlařıldı. Uygulanan üç ayrı Gaussluk testi (Kolmogorov-Smirnov, Jarque-Bera ve Hinich) sonucunda, HbO₂ iřaretlerinin yok denecek kadar az bir yanılma payıyla Gauss bir süreçten kaynaklanmadıęı doęrulandı.

- (ii) *HbO₂ iřaretlerinin zaman-frekans analizi, bu iřaretlerden farklı nörofizyolojik bileřenlere dair farklı frekans altbantlarının uyarlamalı bir biçimde sečilmesi.*

İlk ařamada kısa-dönemli Fourier analizinin görsel olarak incelenmesi yoluyla yürütölen analizin ilk sonuçlarının çok bilgilendirici olmaması üzerine, uyarlamalı bir spektrum bölütleme algoritması geliřtirildi. Sözkonusu algoritma, 10 mHz'lik darbantlarda hesaplanan zaman-frekans öznitelik zaman-dizilerinin benzerlięinden ya da farklılıęından hareketle, sıradüzensel bir aęaç yapısı ve topaklandırma kullanarak, darbantları birleřtirmektedir. Her bir iřaret için bu yolla ayrı bir spektrum bölütlemesi elde edilmektedir. Daha sonra farklı iřaretler üzerinden bulunan bölütlemeler, çoęunluk oylama yoluyla elenmekte ya da korunmaktadır. HbO₂ iřaretleri üzerinde yapılan deneyler sonucunda, spektrumu bütünüyle kaplayan ve birbirleriyle örtüřmeyen dört frekans bandı keřfedilmiřtir: *A*-bandı (0-30 mHz), *B*-bandı (30-40 mHz), *C*-bandı (40-250 mHz), *D*-bandı (250-850 mHz). *A*-bandının asıl aradıęımız biliřsel etkinlik dalga biçimlerinin içine gömülmüř olduęu, arkaplan etkinlięinin bir bölümünü içerdigi düşünölmektedir. Buna ek olarak, *B*-bandının biliřsel etkinlikle ilgili cınlama frekansıyla ilgili olduęu, *C*-bandının da biliřsel etkinlik bilgisini dolaylı olarak tařıdıęı doęrulanmıřtır. *D*-bandındaysa, birtakım yüksek frekans salınımları ve kalp atıřı iřaretinin kalıntısının bulunduęu varsayılmaktadır.

- (iii) *HbO₂ iřaretlerinin biliřsel etkinlięe iliřkin bilgi içerip içermedięinin deneysel olarak kanıtlanması, dolayısıyla FOI yönteminin biliřsel etkinlik ölçümü açısından geđerlenmesi.*

Biliřsel etkinlięin ölçülmesi amacıyla yapılan insanlı deneylerde kullanılan protokoller, neredeyse-dönemli görsel ya da iřitsel uyarın dizileri içerir. Beklenen deneęin her uyarana yanıt vermesi ve elde edilen HbO₂ iřaretinde de bu yanıtın belirmesidir. Buna göre, protokoldeki dönemlilięin, doęrusal sistem varsayımı altında, ölçüle iřaretlere de yansımaları savlanabilir. Gerçekten de yaptığımız dönemlilik kestirimi deneyleri sırasında, bazı denek ve sezicilerden elde edilen iřaretlerin, yüksek bir güvenilirlik skoruyla protokoldekiyle neredeyse aynı deęere sahip bir dönemlilik tařıdıęını gördük. Bu sayede hem FOI'nin biliřsel etkinlięi ölçtüęünü geđerledik, hem de elimizdeki verilerin biliřsel

etkinlik açısından kalitesini nicelendiren bir yöntem elde etmiş olduk. Bunlarla birlikte (ii) maddesinde sözü edilen *B* ve *C* bantlarında önsüzgeçleme yapıldığında güvenilirlik skorlarının gözle görülür bir biçimde arttığını gözlemledik.

- (iv) *BBA ve dalga biçimi topaklandırması yoluyla kısa-dönemli HbO₂ işaretlerinden bilişsel etkinlikle ilişkili dalga biçimlerinin kestirimi.*

Bilişsel etkinliğin ölçülmesinde asıl amaç, bu etkinlikle ilişkili dalga biçimlerinin kestirilmesi. Bu dalga biçimleri üzerinden hesaplanan birtakım parametrelerin, denegin yorgunluğu ve zihinsel durumunu gözlemeye olanak vereceği düşünülmektedir. Bu konuda geniş bir bilgi birikimine sahip olan fMRI yönteminin aksine FOI için geliştirilmiş yöntemler bu çalışmaya kadar bulunmuyordu. Uyarın anlarının hemen ardından gelen zaman-dizisi örneklerini kullanarak, iki ayrı yöntemle dalga biçimleri kestirdik. İlk yöntem olan BBA, girişteki dalga biçimlerini birbirlerinden bağımsız katsayılarla sahip birden çok taban vektörlerine ayrıştırmaktadır. Bu taban vektörlerinden bir ya da en çok ikisinin bilişsel etkinlik ilişkili bir dalga biçimi olduğu öngörülebilir. Diğer yandan, dalga biçimi topaklandırmasında girişteki dalga biçimlerinin her biri için B-spline katsayı seti hesaplanmakta, daha sonra bu katsayı setleri topaklandırılmaktadır. Sonuçta elde edilen topak merkezlerinden biri ya da ikisi yine bilişsel etkinlikle ilişkili olabilir. Kestirilen BBA taban vektörlerinin ya da topak merkezlerinin bilişsel etkinlikle ilişkili olup olmadığı, fMRI'da sıkça kullanılan Gamma dalga biçimine benzerlikleri yoluyla belirlenmiştir. Buna göre, kestirilen dalga biçimi üzerine zaman sabiti, gecikmesi ve şiddeti eniyileme değişkeni olan bir Gamma dalga biçimi oturtulmuş, daha sonra kestirilen dalga biçimiyle bu yeni Gamma modeli arasındaki ilinti katsayısı hesaplanmıştır. Yüksek ilinti katsayılı dalga biçimleri bilişsel etkinlik ilişkili olarak atanmıştır. Bir bütün olarak değerlendirildiğinde algoritma yarı-parametrik. Elimizdeki veriler üzerinde yaptığımız yoğun deneyler sonucunda, Gamma modeline benzerlik nesnel bir standart ölçüt olarak düşünüldüğünde, kestirilen dalga biçimlerinin bilişsel etkinlik ilişkili olduğu saptanmıştır. Bu dalga biçimlerinin zaman-sabitleri, denekten deneye ve aynı denek için verili bir beyin bölgesinden diğerine, daha tutarlı olmakla birlikte, değişmektedir.

1. Ceyhun Burak Akgül, Bülent Sankur, Ata Akın, "Selection of relevant frequency bands in functional near infrared spectroscopy", *Journal of Computational Neuroscience* (18) 67-83, 2005.
2. Ceyhun Burak Akgül, Ata Akın, Bülent Sankur, "Extraction of cognitive-activity related waveforms from functional near infrared spectroscopy signals," *Annals of Biomedical Engineering*'e yollandı, Ocak 2005
3. Sinem Tiveci, A. Akın, T. Çakır, H. Saybaşı, K. Ülgen, "Modeling of Calcium dynamics in brain energy metabolism and Alzheimer's disease," *kabul edildi, Computational Biology and Chemistry*, Mart 2005
4. U. Emir, A. Ademoglu, C. Öztürk, K. Aydın, T. Demiralp, A. Kurt, A. Dincer, A. Akin, Design of an MR compatible fNIRS System" *Proc. SPIE, Photonics West 2005*, basımda
5. Ata Akın, Uzay E. Emir, Serhan Kalsın, Ömer Şayli, "Frequency components in breathholding experiments," *Proc. SPIE, Photonics West 2005*, basımda
6. A. Akın, S. Bunce, "Behavioral reaction times are correlated with frequency shifts in optical imaging signals," *Proc. SPIE, Photonics West 2005*, basımda
7. Ata Akın, Uzay E. Emir, Didem Bilensoy, Gulin Erdogan, Selcuk Candansyar, Hayrunnisa Bolay, "fNIRS measurements in migraine," *Proc. SPIE, Photonics West 2005*, basımda
8. Emir E. U., Şayli Ö., Aydın K., Demiralp T., Ademoğlu A., Öztürk C., Akın A., "Effect of repetitive

- breathholding on BOLD and fNIRS signals,” BMES 2004 Fall Meeting, 14-16 Ekim 2004, Philadelphia, USA
9. Şayli, Ö. Emir E. U., Akın A. “Oxidative metabolism differences in muscles measured by near infrared spectroscopy,” BMES 2004 Fall Meeting, 14-16 Ekim 2004, Philadelphia, USA
 10. Sefer B. Kacar, Ömer Şaylı, Kutlu Ülgen, Ata Akın, “Use of NVC model in hyperammonemia,” AGORA 2004, Kuşadası Turkey
 11. Uzay Emir, Ömer Şayli, Ata Akın, “Investigation of breathholding effect on cerebral hemodynamics by functional near-infrared spectroscopy ” AGORA 2004, Kuşadası, Turkey
 12. Ceyhun B. Akgül, Bülent Sankur, Ata Akın, “Evidence of cognitive activity in fNIRS data,” OSA Biomedical Optics Topical Meeting, 2004
 13. Sefer Burak KACAR, Ömer ŞAYLI and Ata AKIN, “Implementation of a Neurovascular Coupling Model,” 13th Balkan Biochemical Biophysical Days, Kuşadası, Kasım 2003.
 14. Uzay E. Emir, Ceyhun B. Akgül, Ata Akın, Aysin Ertüzün, Bülent Sankur, Kerem Harmancı, “Wavelet Denoising Vs Ica Denoising For Functional Optical Imaging”, Proceedings of the 1st International IEEE EMBS Conference on Neural Engineering Capri Island, Italy, March 20-22, 2003, 384-387
 15. Ata Akın, S. Bunce, “Quantification of Brain Hemodynamic Response Measured by functional Optical Imaging”, Proceedings of the 1st International IEEE EMBS Conference on Neural Engineering Capri Island, Italy, March 20-22, 2003, 503-506
 16. Ceyhun Burak Akgül, Ata Akın, Bülent Sankur, “İşlevsel Optik İşaretlerde Bilişsel Etkinliğin Varlığı”, SIU 2004
 17. Ceyhun Burak Akgül, Ata Akın, Bülent Sankur, “İşlevsel Optik İşaretlerden Bilişsel Etkinlikle İlişkili Dalga biçimlerinin Çıkarılması”, SIU 2004
 18. Uzay Emrah Emir, Ömer Şayli, Kubilay Aydın, Tamer Demiralp, Ahmet Ademoğlu, Cengizhan Öztürk, Ata Akın, “Nefes tutmanın KOSB fMRI ve iYKÖS işaretlerine etkisi,” BİYOMUT 2004
 19. Burak S. Kacar, Ömer Şayli, Ata Akın, “Damar sinir birleşim modeli ile hiperamonyağın modellenmesi” BİYOMUT 2004
 20. Ömer Şayli, Ahu Akdemir, Yağız Ataklı, Uzay E. Emir, Pınar Ö. Çıtlak, L. Sevil Cengiz, Ata Akın, “Kaslardaki oksidatif metabolizma farkının işlevsel yakın kızıl ötesi spektroskopisi ile incelenmesi,” BİYOMUT 2004
 21. Ceyhun B. Akgül, Bülent Sankur, Ata Akın, “İşlevsel Optik Görüntüleme Zamansal Olay Sezimi” 11. SİU Kurultayı, 18-20 Haz. 2003, İstanbul, 11-14
 22. Ata Akın, “İşlevsel Optik Görüntüleme Sisteminde İşaret Modelleri” 11. SİU Kurultayı, 18-20 Haz. 2003, İstanbul, 597-600
 23. Uzay Emrah EMİR, Ömer Şayli, Ata AKIN, "Kasta Oksijenlenme Değişiminin Ölçülmesine Yönelik Sürekli Dalga Yakın Kızılötesi Spektroskopi Yönteminin İGO Karakterizasyonu", 5. Ulusal Optik, Elektro-Optik ve Fotonik Çalışma Toplantısı, ODTÜ, 12 Aralık 2003.
 24. Uzay Emrah EMİR, Ata AKIN, “Hızlı Optik Görüntüleme Yönteminde Gürültü Modellenmesi” BİYOMUT 2003
 25. Sefer Burak Kacar, Ömer Şayli, Tunahan Çakır, Kutlu Ülgen, Ata Akın, “Beynin Enerji Metabolizmasının Damar-Sinir Birleşimi İle Modellenmesi” BİYOMUT 2003

12B.4. Kod Bölüşümlü Çoklu Erişim (CDMA) İletişiminde Kaynak Ayırıştırma

Proje grubu: Mustafa Aziz Altınkaya, Olcay Kalkan

Bir kod bölüşümlü çoklu erişim (code division multiple access, CDMA) sisteminde istasyondan kullanıcıya olan iletimde her bir kullanıcıya gelen sinyal kendisine ait sinyal ile diğer kullanıcıların sinyallerinin toplamından oluşan bir karışımdır. Kullanıcıların genellikle kendileri dışındaki kullanıcılara ait yayma

kodlarını bilemeyecekleri kabul edildiğinden dolayı bu bir gözü kapalı kaynak ayrıştırma (blind source separation, BSS) problemidir. Bu alt projede bağımsız bileşen ayrıştırma (independent component analysis, ICA) yöntemlerinin, özellikle kanalın sönümlenme katsayılarının dürtün (impulsive) bir niteliği olduğu durumlarda istasyondan kullanıcıya CDMA iletişimindeki BSS problemine uygulaması araştırılmıştır [2].

BSS için yeni bir yaklaşım olan Pearson Sistemi'ne dayalı ICA (PS-ICA) yöntemi, Gauss olmayan çokyollu sönümlenmeli kanalın katsayılarının kestirilmesinde kullanılmıştır [1]. PS-ICA yönteminin diğer klasik ICA yöntemlerine kıyasla Gauss'tan biraz daha dürtün ve alfa-kararlı dağılımlarla modellediğimiz kanal sönümlenmesi katsayılarını kestirmekte daha başarılı olduğu görülmüştür. CDMA sistemleri üzerindeki araştırmalarımızda, bağımsız etmen analizi (independent factor analysis, IFA) yöntemi yüksek gürültü seviyesi toleransı ve Gauss karışımı modeli kullanarak kaynak dağılımlarını modellemedeki esnekliği nedeniyle iki kullanıcı ve tek yoldan iletimi içeren basitleştirilmiş bir CDMA sistemi uygulamasında en öne çıkan BSS yaklaşımı olmuştur. Pratik bir CDMA sisteminde kullanıcı sayısı 10'lar seviyesinde ve sönümlenmeli kanaldaki yol sayısı en az 3 olarak kabul edilebileceği için, işlem karmaşıklığı bu sayıların çarpımının üstel bir fonksiyonuyla artan IFA'nın orijinal algoritmasının bu BSS probleminde kullanılması mümkün değildir. Bu nedenle Attias tarafından önerilen [3], çarpanlara ayrılan değişimsel yaklaşıklık (factorized variational approximation, FVA) kullanan IFA çözümü geliştirilmeye başlanmıştır. Önerilen çözüm beklentinin enbüyüklenmesi (expectation maximization, EM) algoritmasının E-adımında probleme neden olan gerçek sonsalın yerine gerçek sonsalın mantıklı bir yaklaşımı olan ve bağımsız etmen parametrelerinden ayrı olarak FVA yöntemiyle kendi parametreleri kestirilen farklı bir yoğunluk fonksiyonunun kullanılmasıdır [4].

[1] O. Kalkan ve M.A. Altinkaya, "Kod bölüşümlü çoklu erişim (CDMA) iletişiminde Gauss olmayan sönüm kestirimi için Pearson sistemine dayalı gözü kapalı kaynak ayrıştırma yöntemi", IEEE 12. Sinyal İşleme ve Uygulamaları Kurultayı SIU 2004, 534-537, Kuşadası, 2004.

[2] O. Kalkan, "Independent Component Analysis Applications in CDMA Systems", Y.L. tezi, İzmit Teknoloji Enstitüsü, Temmuz, 2004.

[3] H. Attias, "Independent Factor Analysis", Neural Computation, 11:803-855, 1999.

[4] M.A. Altinkaya, O. Kalkan ve E.E. Kuruoğlu, "Independent Factor Analysis for Estimating Fading Ch CDMA Communication", IEEE Workshop on Statistical Signal Processing (SSP'05)'ya yollandı, Bordeaux Temmuz 17-20, 2005.

12B.4. Biyometride Kaynak Ayrıştırma

Proje grubu: Bülent Sankur, Hazım Kemal Ekenel, Erdem Yörük

Biyometri, insanların bazı fizyolojik özelliklerine dayanan kimlik tanıma algoritmalarıdır. Bu özellikler yüz fizyonomisi, el şekli, parmak izi, aya izi, ses tınısı, iris örüntüsü, yürüyüş dinamiği vb. olabilir. Çoğunlukla güvenlik uygulamalarında karşımıza çıkan biyometrinin isterleri arasında kullanıcıya dost bir arayüz bulunması, çevre koşulları ve zaman faktöründen etkilenmemesi, hızlı ve gerçekleştirilebilir bir algoritmaya dayanması, alandan kolay olmaması, kayıt sisteminin zahmetsiz olması, büyük nüfuslara hizmet verebilmesi sayılabilir. Çeşitli biyometri seçenekleri arasında biz el ve yüz biyometrisine eğildik.

Kaynak ayrıştırmanın biyometrideki rolü, biyometrik ipuçları arasındaki bazen çok küçük ama ayırt edici

olabilen öznitelikleri ortaya çıkarmaktır. Örneğin insane elinin silüetini kullanarak tanıma yapmaya kalkıştığımızda, sisteme kaydedilmiş denek kişi sayısı yüzlerce olduğunda, ayırt edici öznitelikleri bulmak çok güçleşmektedir. Ancak kaynak ayrıştırma yöntemleri ile deneklerin yüzlerce arasında, bunlar birbirinden ayırt edilmesine elveren şekilsel özellikleri bulup çıkarabilmektedir. Benzer bir tarzda, insane yüzlerinin çoğu ortalama bir prototipe benzemekte ve bundan ufak nüanslarla ayrılmaktadır. BBA algoritmaları bu görünüşe dayalı farklılıkları ortaya çıkarmakta çok etkili olmuştur.

El biyometrisi

El imgelerine dayalı bir kişi tanıma sistemi önerilmiştir. İmgeler, ellerin duruşuna ve takılan aksesuarlara herhangi bir kısıt getirilmeksizin sıradan bir tarayıcı aracılığıyla alınmış olup, tüm el şekline ve teker teker parmaklara uygulanan öteleme ve döndürme işlemleriyle standart duruşa getirilmiştir. Bağımsız Bileşenler Analizi (ICA), Asal Bileşenler Analizi (PCA), Açısız Radyal Dönüşüm (ART) ve Uzaklık Dönüşümü (DT) gibi yöntemlerle salt şekle ve hem şekle hem de dokuya dayalı öznitelikler kullanılmıştır. Tanıma ve doğrulama başarımları memnun edici bulunmuş ve el'in kişi tanıma için, uygun bir seçenek olduğunu göstermiştir.

$$x_i = A(i,1) \times \hat{s}_1 + A(i,2) \times \hat{s}_2 \dots + A(i,N)$$

ICA1 representation for i^{th} hand: $a_i = [A(i,1), A(i,2), \dots, A(i,N)]$

(a)

$$x_i = a_1 \times \hat{p}(1,i) + a_2 \times \hat{p}(2,i) \dots + a_N \times \hat{p}(N,i)$$

ICA2 representation for i^{th} hand: $\hat{p}_i = [\hat{p}(1,i), \hat{p}(2,i), \dots, \hat{p}(N,i)]$

(b)

Şekil 3: Ellerin ICA1 ve ICA2 modeline göre bağımsız bileşenleri

Tablo 1. Farklı öznitelik türlerine ve artan nüfusa göre tanıma başarımları (%)

Öznitelik Türü	Nüfus			
	40	100	200	458
ICA (şekil)	99.19	99.09	98.55	98.40
ICA (şekil+doku)	99.68	99.65	99.58	99.49
PCA (şekil)	98.67	98.69	98.56	97.19
PCA (şekil+doku)	99.14	98.89	98.72	97.99
ART (şekil)	98.72	97.78	97.00	95.78
ART (şekil+doku)	99.28	98.72	98.06	97.67
DT (şekil)	99.17	98.22	96.22	95.99

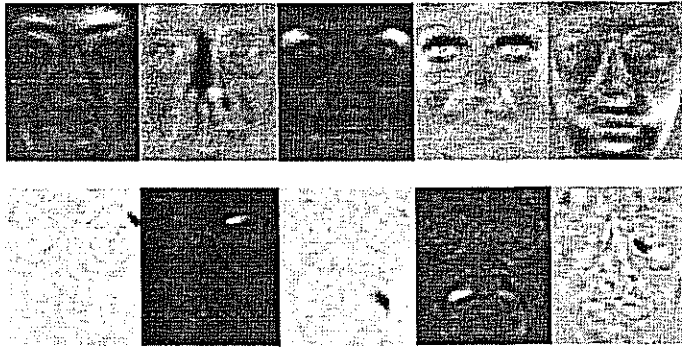
Tablo 2. Farklı öznitelik türlerine ve artan nüfusa göre doğrulama başarımları (%)

Öznitelik Türü	Nüfus			
	40	100	200	458
ICA (şekil)	97.49	98.97	99.41	99.45
ICA (şekil+doku)	97.94	98.93	99.49	99.74
PCA (şekil)	98.27	97.80	97.83	97.78
PCA (şekil+doku)	98.61	98.50	98.73	98.49
ART (şekil)	98.29	97.89	97.95	97.91
ART (şekil+doku)	97.50	97.28	97.36	97.51
DT (şekil)	98.31	98.03	98.08	98.34

Kullanılan veri tabanında, 458 farklı kişiden alınmış, kişi başına 3 örnek olmak üzere toplam 1374 el imgesi bulunmaktadır. Tarayıcıda parmakları ve elin tamamını yönlendiren herhangi bir kalıp kullanılmamış ve kişilerin kullandıkları yüzük gibi aksesuarlara herhangi bir kısıt getirilmemiştir. Yukarıda belirtilen poz çakıştırma aşamasından sonra uygulanan şekle ve dokuya dayalı sınıflandırıcılarla Tablo 1 ve 2'de görüleceği gibi oldukça yüksek tanıma ve doğrulama başarımlarına ulaşılmıştır.

Yüz Biyometrisi:

Bu çalışmada bağımsız bileşen analizi ile yüz tanıma öznelik seçimi problemi ele alınmıştır. Bağımsız bileşen analizinde, temel bileşen analizinde olduğu gibi öznelik seçimini özdeğerler büyüklüklerine dayanarak, ya da işaretin geriçatılmasındaki enerji korunumu kriteri uygulanarak sağlanamaz. Bu nedenle öznelik seçiminde, sınıflandırma başarımını arttıran, dolayısıyla en ayırt edici öznelik alt kümelerini arama yöntemleri kullanılmalıdır. Çalışmada bu yöntemlerden değişinti oranı, en iyi bireysel öznelik, ardışıl eklemeli seçim (SFS: Sequential Forward Selection), ve ardışıl eklemeli çıkarmalı seçim (SFFS: Sequential Floating Forward Selection) kullanılmıştır. Bu yöntemlerin sınıflandırma başarımları, karşılaştırmalı olarak, CMU PIE ve FERET veritabanlarından seçilen yüz imgeleri ile kurulan bir veri kümesi üzerinde sınanmıştır.



Şekil 1. Örnek BBA yüzleri

Şekil 6: İnsan yüzlerinin PCA analizine göre bileşenleri (birinci sıra) ve ICA analizine göre bileşenleri (ikinci sıra)

Diğer bir çalışmada da çoklu çözünürlük analizine dayalı bir yüz tanıma sistemi sunulmuştur. Önerilen yüz tanıma sistemi üç temel birimden oluşur. İlk birim –çoklu çözünürlük birimi- yüz imgesini farklı frekans altbantlarına ayırır. İkinci birim –alt-uzay birimi- elde edilen altbant imgelerini temel bileşen analizi ya da bağımsız bileşen analizi kullanarak daha düşük boyutlu öznelik vektörleriyle betimler. Son birim –sınıflandırma birimi- çıkartılan öznelik vektörlerine göre gelen yüz imgesinin kime ait olduğunu belirler. Geliştirilen yüz tanıma sisteminin başarımı CMU PIE (poz, aydınlatma ve ifade), FERET ve Yale veri tabanları ile kurulan veri kümeleri üzerinde, ifade ve aydınlatma değişimi parametreleri ayrı ayrı ele alınarak sınanmıştır.

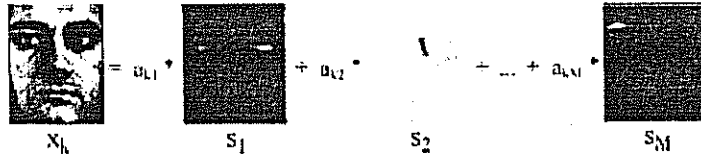


Fig. 5. Face representation using ICA1.

Şekil 7: İnsan yüzünün bağımsız bileşenlerden oluşumu

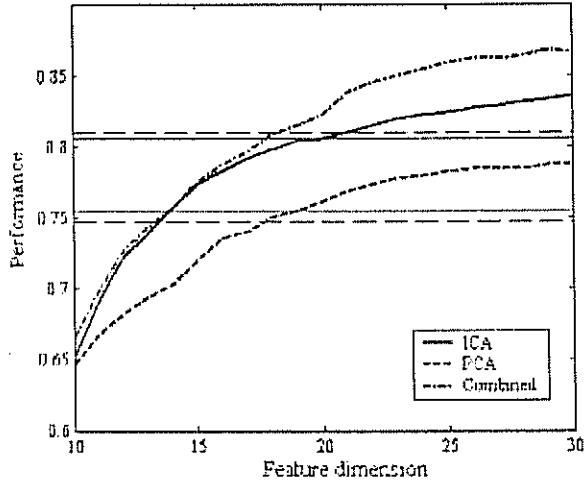


Fig. 5. Cumulative matching score curves of SFPS.

Şekil 8: Bağımsız Bileşenlerin yüz tanımadaki başarımı

13. Elde Edilen Sonular

1. Zamanla deęiŐen zbaęlanımlı alpha-kararlı srelerin parametrelerinin kestirimi iin yeni bir yntem geliŐtirilmesi
2. İki aŐamalı, lekleme belirsizlięiniz ortadan kaldıran evrimii alıŐan bir BBA algoritmasının geliŐtirilmesi
3. CDMA iŐaretlerinin kolay kod zlmesini saęlayan bir teknik.
4. Biyometride, el ve yz imgelerinden hareketle baęımsız bileŐenlerin kullanımı. Halen BBA el biyometrisi literatrdeki en baŐarılı sonuları iermektedir.
5. Beyindeki biliŐsel olgular karŐısında grlen hemodinamik tepkinin ayıklanabilmesi
6. Kuzey Atlantikte grlen iklimsel salınım olgusunun Trkiye'ye yansımaları incelenmiŐtir.

14. Yazılımlar

- 1) El Őekillerinin imge iŐlenmesi ve baęımsız bileŐenlerinin elde edilmesi yazılımı. Ayrıca 3000 civarında el imgesinden oluŐan bir veritabanı da hazırlanmıŐtır.
- 2) Kaynak ayrıŐtırılması, imge iyileŐtirilmesi ve zamanla deęiŐen srelerin parametrelerinin kestirimi iin paracık szgeleri kullanan yazılımlar

15. Eęitime Katkısı

Ercan Kuruoęlu Boęazici niversitesi'nde Kaynak AyrıŐtırma ve Paracık Szgeleri zerindeeęitim semineri vermiŐtir. Bu seminer notları geniŐletilerek Pisa niversitesi ve İzmir YksekTeknoloji Enstitsnde derse dnŐmektedir.

16. Birikimlerin BaŐka Projeye Aktarımı

Projenin birikimlerinin:

- 1) BIOSECURE adlı 6. ereve projesine katkısı vardır. Ayrıca el tanıma algoritmasının ticari bir rne dnŐme olasılıęı bulunmaktadır.
- 2) Paracık szgeleme yntemi Dr. Cemalettin Dnmez'in (İYTE) "Modal Deney Yntemi ile Dinamik Ykler Altında TaŐıyıcı ereve/Dolgu Duvar EtkileŐiminin İncelenmesi" adlı TBİTAK projesinde grlt gidermek iin kullanılacaktır.
- 3) Boęazici niversitesi Bilimsel AraŐtırma Projeleri tarafından desteklenen "İstatistiksel olarak Baęımlı Kaynakların Bayes Yntemi ile AyrıŐtırılması" projesine katkısı vardır.

17. Gen AraŐtırmacıların Eęitimi

- 1) Doktora ęrencisi, Deniz Gencaga 6 ay sresince ISTI'de kalıp paracık szgeleri ve baęımlı bileŐen

analizi üzerinde çalışmıştır. Bu sürenin bir ayı TÜBİTAK tarafından, 5 ayı da NATO bursu ile karşılanmıştır. Bu vesile ile Ercan Kuruoğlu Deniz Genççağ'a'nın doktora ortak danışmanı olmuştur.

- 2) BÜ Elektrik-Elektronik, BÜ Biomedikal, İYTE Elektrik-Elektronik bölümlerinde 7 öğrenci master tezlerini tamamlamıştır.
- 3) BÜ Elektrik-Elektronik Mühendisliği Bölümünde bir doktora öğrencisi tez çalışmalarına devam etmektedir.

18. Geleceğe Yönelik İşbirliğinin Sürdürülme Olasılığı

- a. Ercan Kuruoğlu'nun Boğaziçi Üniversitesi ve İzmir Yüksek Teknoloji Enstitülerindeki temasları İYTE'de bir ders verme ve BÜ'de doktora öğrenci tez danışmanlığı şeklinde sürecektir.
- b. Özellikle biometri alanında yapılan çalışmalar sayesinde Avrupa 6. Çerçeve programlarında (BIOSECURE projesi) katkıda bulunulmuştur.
- c. Daha genişletilmiş bir ekiple (örneğin, Anadolu Üniversitesinin dahil ile) ikinci bir TÜBİTAK-CNR proje başvurusu için çalışmalar yapılmaktadır.
- d. Beyin işaretlerinin işlenmesi projesi doğrultusunda bir TÜBİTAK projesi başvurusu yapılmış, ayrıca SIMILAR adlı bir Avrupa 6. Çerçeve Projesine de katkıda bulunulmuştur.

tarih: 08/02/06



(proje sorumlusunun imzası)

date:

(signature of the foreign project l



Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Pattern Recognition Letters 25 (2004) 1377–1388

Pattern Recognition
Letters

www.elsevier.com/locate/patrec

Feature selection in the independent component subspace for face recognition

H.K. Ekenel *, B. Sankur

Department of Electrical and Electronic Engineering, Bogazici University, Bebek 34342, Istanbul, Turkey

Received 2 October 2003; received in revised form 19 March 2004
Available online 11 June 2004

Abstract

This paper addresses the feature selection problem for face recognition in the independent component subspace. While there exists, at least, energy principle to guide the selection of the principle components, the independent components (ICs) are devoid of any energy ranking, and must therefore selected based on their discriminatory power. In addition the independent component features can be selected starting from a much larger pool, or from a combination pool of ICA and PCA features. Four feature selection schemes have been comparatively assessed, and feature subsets are tested on a face database constructed from CMU PIE and FERET databases. The discriminatory features from larger pools are observed to be concentrated around fiduciary spatial details of the nose, the eyes and the facial contour. Overall, face recognition benefits from the feature selection of ICA or PCA components and from the combination of ICA and PCA feature pools.

© 2004 Published by Elsevier B.V.

1. Introduction

Face recognition has become one of the most active research areas of pattern recognition since the early 1990s. The interest on face recognition is mainly fueled by the identification requirements for access control and for surveillance tasks against terrorism. This interest is still increasing, since face recognition is also seen as an important part of next-generation smart environments (Pentland and Choudhury, 2000).

Among the plethora of face recognition methods, the paradigm based on face appearance data, template-based algorithms and their concomitant subspace versions, such as PCA and LDA methods are the most popular (Turk and Pentland, 1991; Belhumeur et al., 1997). Recently a blind source separation technique, called independent component analysis (ICA) has been adopted for face recognition as an alternative subspace method. Face recognition algorithms, however, encounter several difficulties due to changes of the face appearances caused by such factors as occlusion, illumination, expression, pose, make-ups and aging. In fact the subsequent intra-individual variability of face images can be larger than the inter-individual variability (Gong et al., 2000).

* Corresponding author. Tel.: +90-212-358-15-00; fax: +90-212-287-24-65.

E-mail address: ekenelha@boun.edu.tr (H.K. Ekenel).

One expects to surmount these difficulties, or at least to mitigate their effect by judicious choice of features that are insensitive to the variations in the facial appearance. The purpose of this study is to explore subsets of ICA and PCA features for face recognition that increase recognition performance, and that are purportedly robust against expression variations and differences in illumination. The feature selection criterion used in PCA-based face recognition is the eigenvalue variance. On the other hand the independent components (ICs) are devoid of any energy/importance ranking, therefore relatively high-dimensional feature vectors are used in face recognition using ICA, i.e., 200 dimensions (Baek et al., 2001; Bartlett et al., 1998; Draper et al., 2003) with respect to PCA. To handle this problem, proportion of variance (PoV) has so far been the only feature selection technique used in the literature on ICA face recognition (Bartlett et al., 1998; Deniz et al., 2001; Yuen and Lai, 2002; Havran et al., 2002). Besides using the feature selection methods based on individual properties of the features, like proportion of variance and best individual feature, we also utilize more sophisticated feature selection schemes that take into account the combinatorial properties of the features such as sequential forward selection and sequential floating forward selection. We assess comparatively these four different feature selection schemes on the basis of the improvement they bring in the recognition performance. Furthermore, since the first face recognition architecture of ICA provides local features, we tried to discover which one of these local features has more discriminatory power, and thus contributes most to the person identification.

The paper is organized as follows. In Section 2 we review briefly PCA and ICA techniques. The feature selection methods are presented in Section 3. Experimental results are discussed in Section 4 and conclusions are drawn in Section 5.

2. Subspace analysis methods

Face data, as obtained from the raster scanning of the face images, constitutes a very high-dimensional space. However the intrinsic dimensionality

of the face space is known to be much smaller (Gong et al., 2000), this despite the variations in expression, pose and lighting. In fact the faces are believed to be clustered on some low-dimensional manifolds. The subspace techniques aim to reduce the inherent excessive dimensionality of scanned data to make the face recognition algorithms viable and to capture or approximate the underlying face manifolds. The most widely used subspace analysis tools are the principal component analysis (PCA), independent component analysis (ICA), linear discriminant analysis (LDA) and their non-linear varieties via kernel tools (Turk and Pentland, 1991; Belhumeur et al., 1997; Kim et al., 2002; Yang, 2002; Moghaddam, 2002).

In the sequel, lowercase and uppercase letters denote scalar values, bold uppercase letters denote matrices, and bold lowercase letters denote row vectors.

2.1. Principal component analysis (PCA)

PCA aims to determine a new orthogonal basis vector set that best reconstructs the face images, in other words with the smallest mean-square error for any given subspace dimensionality. These orthogonal basis vectors, also called eigenfaces, are the eigenvectors of the covariance matrix of the face images. The most parsimonious eigenvector set, say of dimension M , for the face reconstruction problem is chosen as the subset of the M most energetic eigenvectors, that is the eigenvectors corresponding to the first M rank ordered eigenvalues.

Consider the $K \times D$ -dimensional face data matrix X , where each D -dimensional row corresponds to the lexicographically ordered pixels of one of the faces, and where there are K face images. The PCA method tries to approximate this face space using an M -dimensional feature vector, that is using M eigenfaces, where typically $M \ll \min(D, K)$. These M eigenvectors span a face subspace, such that $\|X\|^2 - \|XV\|^2$ is minimum, where V is the $D \times M$ -dimensional matrix that contains orthogonal basis vectors of the face space in its columns. Once the projection bases V are formed, when a test image x , arrives, it is projected onto the face subspace to yield the feature vector,

$r_i = x, V$. The classifier decides for the identity of the individual, according to a similarity score between r_i and the feature vectors of the individuals in the database $\{r_1, r_2, \dots, r_K\}$.

2.2. Independent component analysis (ICA)

Independent component analysis is an unsupervised learning method based on high order statistics. Briefly, ICA is the separation of independent sources from their observed linear mixtures (Hyvarinen and Oja, 2000). The system model of ICA is given as

$$X = AS \quad (1)$$

where A denotes the mixing matrix, S denotes the source matrix containing statistically independent source vectors in its rows and X denotes the data matrix. In the ICA method, the only information we possess is the observations, and neither the mixing matrix nor the distribution of the sources is known. Under the assumptions that the sources are statistically independent and non-Gaussian (at most one of them can have Gaussian distribution), we find the unmixing matrix W by maximizing some measure of independence. In other words, a separation matrix, W , is estimated, which, under ideal conditions, is the inverse of the mixing matrix A .

$$Y = WX \quad \text{and} \quad W = A^{-1} \quad \text{and} \quad Y \approx S \quad (2)$$

In the context of face recognition, the use of ICA features was first proposed in (Bartlett et al., 1998) and several other studies followed in (Liu and Wechsler, 1999; Baek et al., 2001; Kwak et al., 2002; Deniz et al., 2001; Yuen and Lai, 2002; Ding et al., 2001; Havran et al., 2002; Draper et al., 2003). In (Bartlett et al., 1998) two different approaches are presented for face recognition. In the first approach (called ICA1 architecture), the face images are assumed to be a linear mixture of an unknown set of statistically independent source images. The source images obtained in this architecture are spatially local and sparse in nature. In the second approach (called ICA2 architecture), the representation (weighting) coefficients are assumed to be statistically independent. In this second architecture, while mixing coefficient

vectors are independent, source images tend to have global face appearance, as in the case of PCA. In contrast to the PCA method, where feature subset selection is based on energy criterion, the selection of an ICA basis subset is not immediately obvious since the energies of the independent components cannot be determined. One typically first trims the $\min(D, K)$ -dimensional space to M using a PCA stage, and then proceeds to extract the ICA components from this M -dimensional space.

The architecture of ICA1 and that of PCA have analogous structures, as they are both based on the statistical properties of the basis images. In ICA1, one tries to find statistically independent basis images, whereas in PCA one extracts uncorrelated basis images. ICA2 is based on the statistical properties of the representation coefficients, thus it has a different structure with respect to ICA1 and PCA. In this work, both for a fair comparison with PCA-based method and to explore the local face regions that are appropriate for identification, the first architecture of the ICA is used.

In summary, the face recognition algorithm using ICA1 architecture is as follows:

- (i) Prior to ICA, PCA is performed on a training set and the M eigenvectors, associated with the largest eigenvalues are selected to form the matrix V .
- (ii) The training face images are projected onto the PCA-based face space and the $K \times M$ matrix of their representation coefficients, $R = [r_1 | r_2, \dots, | r_K]^T$, are obtained:

$$R = X * V \quad (3)$$

- (iii) An ICA analysis is performed on V^T , where face eigenvectors form the rows of this matrix; the unmixing matrix, W , reveals the "ICA faces"

$$S = W * V^T \quad (4)$$

- (iv) By using PCA representation coefficients, R , and independent basis images, S , ICA representation coefficients of the faces in the training set, $\{a_1, a_2, \dots, a_K\}$, are calculated as follows:

$$A = R * W^{-1} \quad (5)$$

since $V^T = W^{-1} * S$, $X = R * V^T$, and $X = R * W^{-1} * S$. When a test image x_t arrives, it is projected onto the face subspace to extract its PCA representation $r_t = x_t V$, which is then multiplied with the inverse of the separation matrix to yield the ICA feature vector $a_t = r_t * W^{-1}$. This vector is compared vis-à-vis the feature vectors $\{a_1, a_2, \dots, a_K\}$ of the individuals in the face database.

3. Feature selection techniques

3.1. Rationale for feature selection for PCA/ICA faces

The main goal of this study is the automatic selection of the best feature subset for classification purposes given a high-dimensional ICA feature vector. Recall that the feature selection criterion used in PCA-based face recognition is the eigenvalue variance. While this criterion yields the most compact set for the reconstruction, it does not necessarily follow that it is optimal from face recognition point of view. The ICA methods do not have such an opportunity of feature selection, in that the number of ICA features is pre-determined in the PCA stage of data processing. It is conjectured that some feature selection scheme focused on “recognition” rather than on “reconstruction” could augment face classification performance. With this goal in mind, we have not initially reduced the subspace dimensionality via PCA to the target feature size, M . Instead we have reduced the initial dimensionality $\min(D, K)$ down to an intermediate size M' , $M < M' < \min(D, K)$. The judicious choice of M' , on the one hand, should make the implementation of the ICA algorithm viable and avoid overlearning effects (Bartlett et al., 1998; Liu and Wechsler, 1999). On the other hand M' should allow sufficient freedom or richness of choice for the feature selection algorithms to be effective. For example, if the target feature vector will be 30-dimensional, we do not effect a $\min(D, K)$ -to-30 PCA algorithm, but use, for example, a $\min(D, K)$ -to-200 PCA reduc-

tion. Following this, the feature selection algorithm proceeds to find the most discriminating $M = 30$ ICA features from the intermediate set of $M' = 200$.

We have comparatively assessed the recognition performances resulting from different feature selection algorithms, namely, proportion of variance (PoV), best individual feature (BIF), sequential forward selection (SFS) and sequential floating forward selection (SFFS) techniques (Feri et al., 1994). PoV has so far been the only feature selection technique used in the literature on ICA face recognition (Bartlett et al., 1998; Deniz et al., 2001; Yuen and Lai, 2002; Havran et al., 2002). PoV would be the best approach if the features' contributions to the face recognition performance were independent from each other. BIF is also based on the same assumption as PoV, the difference being that, while in PoV the variances of features are used, in BIF the classification performances of the individual features are taken into consideration. SFS technique is a special case of the well known “plus / – take away /” methods, and SFFS is an enhanced version of SFS. These last two techniques search for the best performing feature set while taking into consideration the correlation between features.

3.2. Proportion of variance (PoV)

In this technique, for all feature points the ratio of between-class variance to within-class variance is calculated and the feature points that have high ratio are selected. In this context “class” denotes the ensemble of images belonging to one individual. Thus the m th feature of the q th face image ($q = 1, \dots, Q$) of the k th ($k = 1, \dots, K$) individual will be expressed as $f_m(k, q)$. The class mean will be obtained as

$$\bar{f}_m(k) = \frac{1}{Q} \sum_{q=1}^Q f_m(k, q) \quad (6)$$

while the average of this feature over the database will be denoted as

$$\bar{\bar{f}}_m = \frac{1}{KQ} \sum_{k=1}^K \sum_{q=1}^Q f_m(k, q) \quad (7)$$

The ratio of between-class variance to within-class variance for the m th feature point in the feature vector is calculated as

$$r_m = \frac{\sigma_{\text{between},m}}{\sigma_{\text{within},m}} = \frac{\sum_{k=1}^K (\bar{f}_m(k) - \bar{\bar{f}}_m)^2}{\sum_{k=1}^K \sum_{q=1}^Q (f_m(k,q) - \bar{f}_m(k))^2} \quad (8)$$

This measure reveals out whether a change in facial appearance corresponds to a change in the m th feature point or not. Thus we rank the features according to their r_m ratio where $f_{r(m)}$ is the m th ranking feature, and select the highest ranking M features, $\{f_{r(1)}, f_{r(2)}, \dots, f_{r(M-1)}, f_{r(M)}\}$.

3.3. Best individual feature (BIF)

In this technique, classification performance of each feature point is calculated separately, that is, on individual basis, and the features giving rise to highest correct recognition rate are selected. Let P_m be the classification performance of the m th feature point, that is, P_m is the probability of correct detection using solely feature f_m . Then the feature points are ordered according to their individual classification performances P_m and the first M ones, having highest P_m values, are selected.

3.4. Sequential forward selection (SFS)

In the SFS method, features are selected successively by adding the locally best feature point, the feature point that provides the highest incremental discriminatory information, to the existing feature subset. The SFS technique starts as the BIF by identifying the first feature that has the highest discrimination power. It proceeds, however, by adding sequentially to the selected subset, those features that contribute most to the classification performance on top of the already selected ones. Thus, from a single initial BIF feature, the SFS subset grows to a pair, to a triple...till an M -fold subset is found.

3.5. Sequential floating forward selection (SFFS)

SFFS is similar to the SFS in adding features to the subset; however, in addition, it goes through

cleansing periods, in that features are removed systematically so long as the performance improves after pruning. Note that the performance of the pruned subset is compared with that of a previous one with the same population. This stratagem helps to avoid the nesting effect, which results when one is stuck with a suboptimal subset.

4. Results and discussion

4.1. Experimental setup

We have used a face database constructed with images chosen from CMU PIE and FERET databases (Phillips et al., 2000; Sim et al., 2002). In our experiment set, there are 214 individuals each having four different frontal face images, making a total of 856 images. 584 out of the 856 images are chosen randomly from the FERET database, fafb image set. These FERET-based images mainly differ in expression. The remaining 272 images derive from the CMU PIE database, which contain variations in illumination. All the face images are aligned with respect to the manually detected eye coordinates, scaled to 60×50 pixels resolution and histogram equalized. For each individual in the set, two of their images that contain normal facial expression and have frontal illumination are used for training, and the remaining two images that contain alternative facial expressions and illumination from left and right sides are used for testing purposes (Fig. 1).



Fig. 1. Sample face images: four images of each of two individuals with differing illumination effects, expression or accessories.

Feature vectors are extracted via independent component analysis and principal component analysis schemes. The FastICA algorithm (Hyvarinen and Oja, 2000) is used to perform independent component analysis. In either case we extracted 200-dimensional feature vectors, that is, the original 3000-dimensional image feature vectors consisting of raw pixel values were reduced to, respectively, 200 PCA features conserving 96.46% of the energy and 200 ICA features. Any further feature selection was carried on these intermediate sets of 200.

We used the nearest-neighborhood classifier in our feature selection study. We opted for the nearest-neighbor method, as it is a powerful non-parametric classifier without any costly training stage since otherwise one needs to re-design the classifier repetitively for each selected feature subsets. Finally we evaluated comparatively three different distance metrics, namely, the $L1$ norm, the $L2$ norm, and the normalized correlation coefficient, defined as follows:

$$L1 : d = \sum_{j=1}^M |f_{\text{training},j} - f_{\text{test},j}| \quad (9)$$

$$L2 : d = \left(\sum_{j=1}^M |f_{\text{training},j} - f_{\text{test},j}|^2 \right)^{1/2} \quad (10)$$

$$CC : d = \frac{f_{\text{training}} \cdot f_{\text{test}}}{\|f_{\text{training}}\| * \|f_{\text{test}}\|} \quad (11)$$

In Fig. 2, sample ICA and PCA basis images are shown. In the first row, samples of eigenfaces are seen. The basis images in the second row are instances of ICA faces obtained by applying source separation to 30 eigenfaces, whereas in the third row we display basis images that are obtained by separating 200 eigenfaces. It can be observed that, as the number of independent sources increases, the ICA images become more spatially localized and sparse. On the other hand, when a small number of sources are used, the ICA basis images resemble the eigenfaces.

4.2. Experimental results

In the sequel we present the classification results as well as the cumulative matching scores (CMS) of face images under various feature selection schemes and distance metrics.

In Table 1, the correct classification performances of ICA and PCA are given when, respectively, 30 and 200-dimensional feature vectors are used. As it has also been reported in the literature (Moghaddam, 2002), there are no significant performance differences between ICA and PCA. As

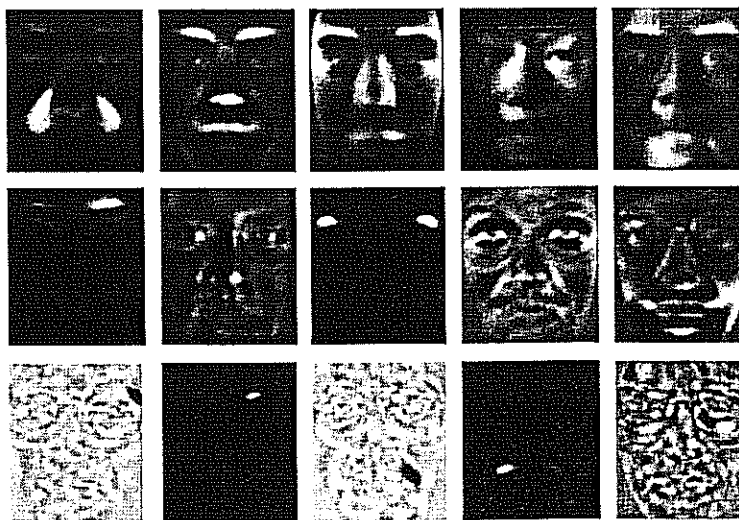


Fig. 2. Sample basis images: first row—eigenfaces; second row—typical instances of 30 ICA faces obtained from 30 eigenfaces; third row—typical instances of 30 ICA faces obtained from 200 eigenfaces.

Table 1
Correct classification performance of ICA and PCA

	L1	L2	CC
ICA-30	75.47	74.30	73.60
PCA-30	74.77	74.30	73.60
ICA-200	80.61	79.67	78.50
PCA-200	81.07	79.67	78.50

one should expect, the correct recognition rate gradually increases with the increase in the dimension of the feature vector. For both feature sets, the L1 norm gives the best results; hence the L1 distance metric is taken into consideration while performing feature selection. Recall that in this experiment the most energetic 30 (200) PCA components were selected as feature sets, and that the ICA features were simply obtained by further processing these PCA components via the Fast-ICA algorithm. The rationale for this choice is explicated in Fig. 3 on the basis of the relative

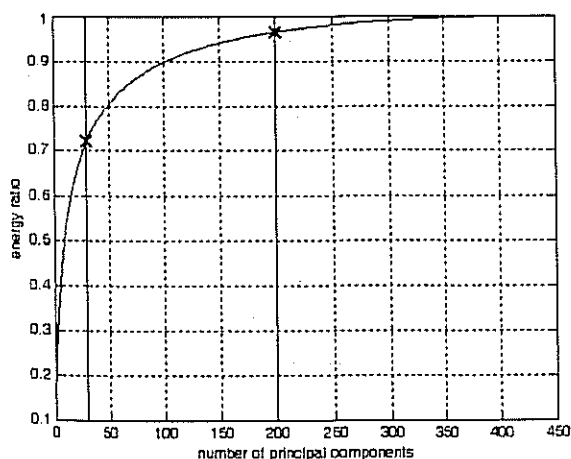


Fig. 3. Energy percentage accounted by the eigenfaces.

energy accounted for by the chosen number of eigenfaces.

To prove the conjecture that a judiciously selected subset from the individual feature sets or from the merged set of both ICA and PCA faces can yield better performance, we carried the following experiments. In the first experiment, we selected subsets of 30 features out of the initial set of 200, individually for the PCA and ICA cases using various selection algorithms. In the second experiment, we merged the two sets to constitute a superset of 400 dimensions (200 PCA + 200 ICA) and then proceeded to select the subset of 30 features from this augmented feature set. Note that before merging of the two feature sets, they were rendered commensurate by variance normalization of the ICA and PCA feature vectors. In Table 2, the performance values of the selected feature subsets are shown. Sequential selection methods perform better than the PoV and BIF methods. The reason can be attributed to the fact that, in classification, the class-conditional features are not independent. Therefore combinatorial feature selection proves to be more effective than selecting them based only upon their individual performances. In other words, selecting the feature points that have high individual classification performances does not imply the best subset for classification. For example, in Fig. 4, the ratio of between-class variance to within-class variance is shown. Although ICA coefficients have higher ratio as compared to those of PCA coefficients, this does not necessarily imply higher classification performance when considered jointly as can be observed from Table 2. These results show that the correct combination of the feature points in the selected subset is more important than any selection based on individual properties.

Table 2
Performance results of the feature selection techniques, when the initial feature set is reduced from 200 to 30. From 400 to 30 in the combined case

Feature selection method	PCA-30/200	ICA-30/200	Combined 30/400
PoV	74.30	71.73	72.90
BIF	74.07	75.93	71.73
SFS	77.34	80.84	79.67
SFFS	78.74	83.64	86.68

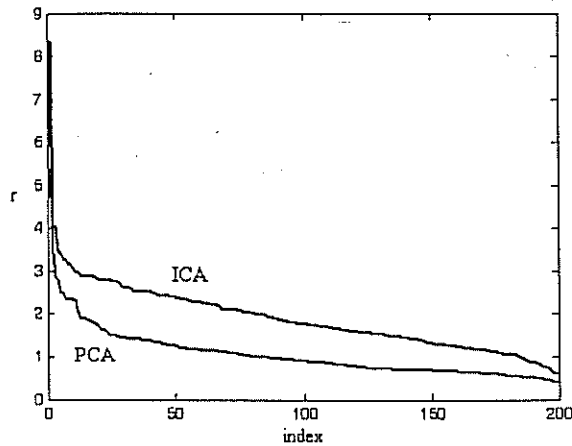


Fig. 4. The ratio of between-class to within-class variances.

One can observe from Table 2 that, in the sequential selection techniques, ICA achieves higher correct recognition rates as compared to PCA. While trying to bring down the dimensionality of the PCA feature set proves detrimental, the converse is true for the ICA set. As the ICA feature set is reduced from the original 200 components down to dimension 30, the performance improves by 3%, that is from the score of 80.61 for ICA-200 to the score of 83.64 for ICA-30 using the SFFS method. This increase in classification performance is a consequence of the removal of the ICA basis images localized on those face regions experiencing major changes in the appearance and having less discriminatory information.

It is interesting to observe that the subset selected from merged 400 dimensional ICA and PCA feature vectors perform even better than separate ICA or PCA feature vectors, despite the fact the ICA features were obtained from PCA features via a linear method. Finally we can observe that the SFFS feature selection method proves to be uniformly superior in all experiments. In fact the ranking in increasing performance of the feature selection methods in this experiment is as follows: PoV < BIF < SFS < SFFS. In Fig. 5, cumulative matching score curves of SFFS method are given. The horizontal lines in the figure represent the performance values of the 30 and 200-dimensional PCA and ICA feature vectors obtained without any selection. As can be seen

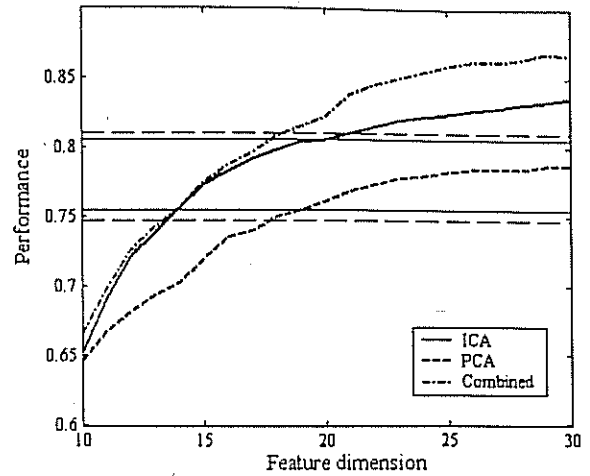


Fig. 5. Cumulative matching score curves of SFFS.

from the figure, in ICA, nearly a 20-dimensional selected feature vector can outperform the initial 200-dimensional feature vector, or similarly, a 14-dimensional selection of features is on a par with the 30-dimensional feature performance.

In Figs. 6–8, we show the indices of the selected 30 features from the original set of 200 PCA and ICA features and from the combined set using SFFS method. In Fig. 6, the features are ordered according to the energy index, that is the size of

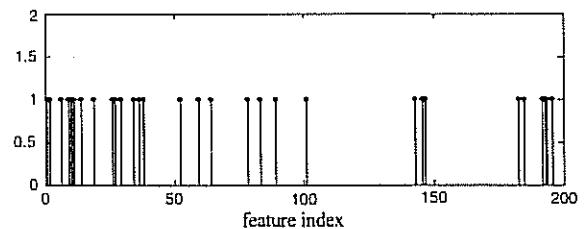


Fig. 6. Selected PCA feature points by SFFS.

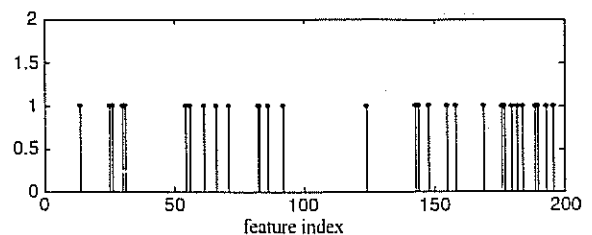


Fig. 7. Selected ICA feature points by SFFS.

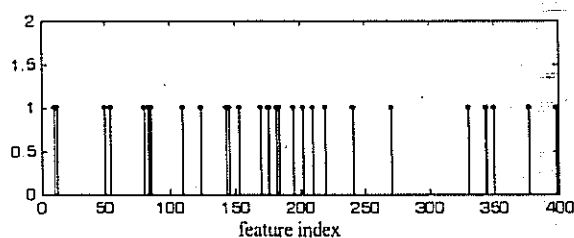


Fig. 8. Selected feature points from the combined set by SFES: note that the first 200 indices refer to the ICA features, and the second set of features, with indices from 201 to 400 refer to the PCA features.

their eigenvalue. In Fig. 7, the ordering is what results from the ICA algorithm, after the PCA reduction stage. Finally note that in Fig. 8, the ICA and PCA features are juxtaposed in a single graph, such that the first 200 indices correspond to the ICA features and the last 200 indices correspond to the PCA features. Contrary to the common practice of selecting the most energetic PCA features, as can be observed in Fig. 6, only two

thirds of the PCA feature are from the first 100 points. This shows that some of the less energetic PCA coefficients, which are normally discarded for image representation and compression, may still contain valuable discriminatory information. The selected ICA features are spread more evenly over the index range. Finally, it is interesting to observe that the number of selected feature indices from PCA and ICA feature pools are approximately equal.

In Fig. 9, the selected ICA features are presented. One notices that these features are characterized by sparse regions localized around the eyes, on the face contours and at the nose. One can interpret these regions in the selected ICA components as being less sensitive to the variations in the facial appearance and at the same time to contain more discriminatory information. These findings confirm the results of a recent study on optimal Gabor kernel location selection (Gokberk et al., 2003). Another observation from Fig. 9 is that, these basis images resemble the receptor fields

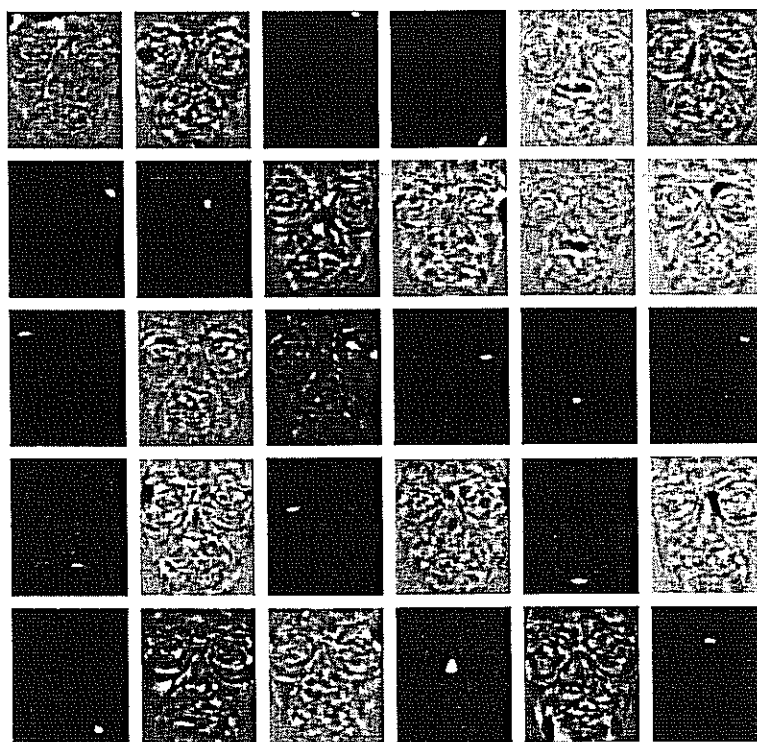


Fig. 9. Selected ICA faces with SFES method from a training set of normal face appearances.

obtained in local feature analysis (LFA) (Penev and Atick, 1996).

4.3. Discussion of the results

The rationale to resort to a feature selection scheme was to improve the face recognition performance and to mitigate the effects of intra-variation. We investigated the source of the attained improvement (compare for example, ICA-30 performances in Tables 1 and 2). Some of the face image, which were erroneously classified without feature selection in ICA, but were then correctly labeled after feature selection are shown in Fig. 10, in pairs side by side the test image and erroneously matched training image. As can be seen from these face pairs, their appearances are very similar and the individuals can be discriminated by only a careful consideration of the local face regions. Appropriate selection of the ICA basis images provides the utilization of this valuable local discriminatory information.

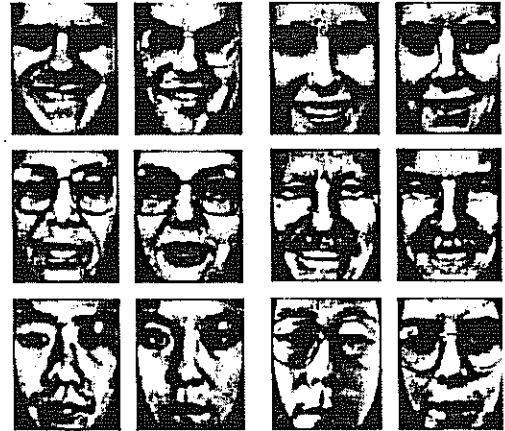


Fig. 10. Samples of misclassified face images. First image in the pair is the test image; second image in each pair is the matched training image.

A desirable aspect of the proposed scheme should be the robustness from one face database to another. In other words, one would like to encounter the same or very similar subset of fea-

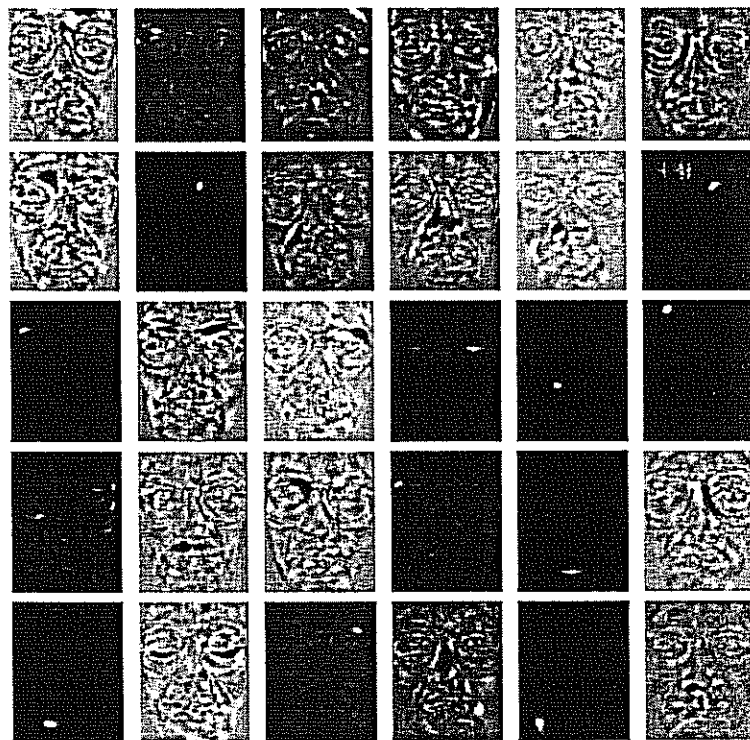


Fig. 11. Selected ICA faces with SFFS method from a training set of faces with expression and illumination variations.

tures, as one switches from one face database to another one. To validate the usefulness of the features we have selected, we perform ICA on a different training set. We swap the images in the test set with those in the training set. Thus, in this new arrangement, the face images in the training set contain differences in expression and variations in illumination, and the face images in the test set contain normal facial expression and have frontal illumination. The independent components that correspond to the selected ICA feature points with SFFS method are presented in Fig. 11. As can be observed from Fig. 11, again the ICA basis images that localize on the eye region, outlines of face and nose are selected.

5. Conclusions

In this study we have explored feature selection techniques on ICA and PCA bases for face recognition. Feature selection techniques are warranted especially for ICA features since these are devoid of any importance ranking based on energy content as the PCA components. The study was carried out on a face database that contains both facial expression and illumination variations. Four different feature selection techniques were used comparatively and the sequential floating forward selection method was observed to be uniformly superior in all cases, in that the maximum correct classification rate was obtained with its feature subset.

The major conclusion from this study was that the feature selection applied on ICA features definitely improves the recognition performance by 8.17%. Furthermore, if the features are selected from the augmented pool of both ICA and PCA features, the performance improvement becomes 11.21%. In other words, it pays to select ICA/PCA features subset from a larger set of them, rather than deciding a priori for the dimensionality of the final feature subset. Indeed, instead of selecting features on the basis of the first M most energetic PCA components or their ICA versions, it was more beneficial to search for the subset resulting in the best possible classification performance from a larger pool. It is interesting to note that even the

PCA features benefited from this approach, when 30 of them were selected from the set of 200 via the SFFS algorithm instead simply selecting the first most energetic ones, as is commonly done in the literature.

When the ICA features are selected from a larger initial set, we observe that the resulting features are more localized, as the sharp dark or bright spots in Figs. 9 and 11 indicate. In fact, not surprisingly the majority of these local accents are around the eyes and nose, as well as close to the facial contours.

References

- Baek, K., Draper, B.A., Beveridge, J.R., She, K., 2001. PCA vs. ICA: A comparison on the FERET data set. In: *Internat. Conf. on Computer Vision, Pattern Recognition and Image Processing*, North Carolina.
- Bartlett, M.S., Lades, H.M., Sejnowski, T.J., 1998. Independent component representations for face recognition. In: *Conf. on Human Vision and Electronic Imaging III*, San Jose, California.
- Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J., 1997. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Machine Intell.* 19 (7), 711–720.
- Deniz, O., Castrillon, M., Hernandez, M., 2001. Face recognition using independent component analysis and support vector machines. In: *Third Internat. Conf. on Audio and Video Based Biometric Person Authentication*, Sweden.
- Ding, P., Kang, X., Zhang, L., 2001. Personal recognition using ICA. In: *Eighth Internat. Conf. on Neural Information Processing*, China.
- Draper, B.A., Baek, K., Bartlett, M.S., Beveridge, J.R., 2003. Recognizing faces with PCA and ICA. *Computer Vision and Image Understanding* 91 (1–2), 115–137.
- Feri, F.J., Pudil, P., Hatef, M., Kittler, J., 1994. Comparative study of techniques for large-scale feature selection. In: *Pattern Recognition in Practice IV*. Elsevier, pp. 403–413.
- Gokberk, B., Irfanoglu, M.O., Akarun, L., Alpaydin, E., 2003. Optimal Gabor kernel location selection for face recognition. In: *Internat. Conf. on Image Processing*, Spain.
- Gong, S., McKenna, S.J., Psarrou, A., 2000. Dynamic vision: From images to face recognition. Imperial College Press.
- Havran, C., Hupet, L., Czyz, J., Lee, J., Vandendorpe, L., Verleysen, M., 2002. Independent component analysis for face authentication. In: *KES'2002 Proc.*, Italy, pp. 1207–1211.
- Hyvarinen, A., Oja, E., 2000. Independent component analysis: Algorithms and applications. *Neural Networks* 13 (4–5), 411–430.

- Kim, K.I., Jung, K., Kim, H.J., 2002. Face recognition using kernel principal component analysis. *IEEE Signal Process. Lett.* 9 (2), 40–42.
- Kwak, N., Choi, C., Ahuja, N., 2002. Face recognition using feature extraction based on independent component analysis. In: *Internat. Conf. on Image Processing*, Rochester.
- Liu, C., Wechsler, H., 1999. Comparative assessment of independent component analysis (ICA) for face recognition. In: *Second Internat. Conf. on Audio and Video Based Biometric Person Authentication*, Washington, DC.
- Moghaddam, B., 2002. Principal manifolds and probabilistic subspaces for visual recognition. *IEEE Trans. Pattern Anal. Machine Intell.* 24 (6), 780–788.
- Penev, P., Atick, J., 1996. Local feature analysis: A general statistical theory for object representation. *Network: Computation in Neural Systems* 7 (3), 477–500.
- Pentland, A., Choudhury, T., 2000. Face recognition for smart environments. *IEEE Computer* 33 (2), 50–55.
- Phillips, P.J., Moon, H., Rizvi, S.A., Rauss, P.J., 2000. The FERET evaluation methodology for face-recognition algorithms. *IEEE Trans. Pattern Anal. Machine Intell.* 22 (10), 1090–1104.
- Sim, T., Baker, S., Bsat, M., 2002. The CMU pose, illumination, and expression (PIE) database. In: *Proc. IEEE Internat. Conf. on Automatic Face and Gesture Recognition*, Washington, DC.
- Turk, M., Pentland, A., 1991. Eigenfaces for recognition. *J. Cognitive Sci.* 3 (1), 71–86.
- Yang, M.H., 2002. Kernel eigenfaces vs. kernel fisherfaces: Face recognition using kernel methods. In: *Fifth IEEE Internat. Conf. on Automatic Face and Gesture Recognition*, Washington, DC.
- Yuen, P.C., Lai, J.H., 2002. Face representation using independent component analysis. *Pattern Recognition* 35 (6), 1247–1257.



Multiresolution face recognition

Hazim Kemal Ekenel*, Bülent Sankur

Electrical and Electronic Engineering Department, Boğaziçi University, Bebek, İstanbul, Turkey

Received 30 January 2004; accepted 17 September 2004

Abstract

In this paper the contribution of multiresolution analysis to the face recognition performance is examined. We refer to the paradigm that in classification tasks, the use of multiple observations and their judicious fusion at the data, feature or decision level improves the correct decision performance. In our proposed method, prior to the subspace projection operation like principal or independent component analysis, we employ multiresolution analysis to decompose the image into its subbands. Our aim is to search for the subbands that are insensitive to the variations in expression and in illumination. The classification performance is improved by fusing the information coming from the subbands that attain individually high correct recognition rates. The proposed algorithm is tested on face images that differ in expression or illumination separately, obtained from CMU PIE, FERET and Yale databases. Significant performance gains are attained, especially against illumination perturbations.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Multiresolution Analysis; Discrete Wavelet Transform; Independent Component Analysis; Principal Component Analysis; Fusion

1. Introduction

Face recognition problem has become one of the most relevant research areas in pattern recognition. Face recognition debts its popularity to its potential application areas, ranging from human computer interaction to authentication and surveillance.

The holistic or appearance-based approach has been gaining popularity vis-à-vis anthropometrical feature-based approach in face recognition [1]. In the holistic approach, all the pixels in the entire face image are taken as a single signal, and processed to extract the relevant features for classification. Most of the appearance-based face recognition algorithms perform some kind of subspace analysis in the image space to extract the relevant feature vectors. The most widely used subspace analysis tools are the principal component analysis (PCA) [2], linear discriminant analysis (LDA) [3] and a blind source separation technique, called independent component analysis (ICA) [4]. All face recognition algorithms, however, witness a performance drop whenever face appearances are subject to variations by

factors such as occlusion, illumination, expression, pose, accessories and aging. In fact, often these factors lead to intra-individual variability of face images, to the extent that they can be larger than the inter-individual variability [5].

In this study, we apply multiresolution techniques in order to mitigate the loss of classification performance due to changes in facial appearance. We design experiments specifically to investigate the gain in robustness against illumination and facial expression changes. The underlying idea in the use of the multiresolution analysis is firstly, to obtain multiple evidences from the same face, and search for those components that are less sensitive to intrinsic deformations due to expression or due to extrinsic factors, like illumination. Secondly, our approach follows the paradigm of fusion that utilizes multiple evidences. Although at first sight, these evidences can appear somewhat redundant and may contain less information, their judicious combination can prove often to be superior for classification.

The most popular multiresolution analysis technique is the wavelet transform. Therefore in this study we use the 2D discrete wavelet transform in order to extract multiple subband face images. These subband images contain coarse approximations of the face as well as horizontal, vertical

* Corresponding author.

E-mail address: ekenelha@boun.edu.tr (H. Kemal Ekenel).

and diagonal details of faces at various scales. Subsequently, we extract PCA or ICA features from these subbands. We exploit these multiple channels by fusing their information for improved recognition. We have compared three fusion approaches, namely, fusion at the subband data level, fusion at the ICA/PCA feature level, and finally, fusion of the classifier decisions at the subband channel level. The main contribution of the paper is thus to search for most discriminative set of wavelet channels, and to construct face recognition schemes using fusion techniques at different levels of data processing.

Discrete wavelet transform has been used in various studies on face recognition [6–10]. In [6], three-level wavelet transform is performed to decompose the original image into its subbands, on which the PCA is applied. The experiments on Yale database show that third level diagonal details attain highest correct recognition rate. A wavelet transform-based speaker identification system in a teleconferencing environment is proposed in [7]. In this algorithm a three-level wavelet decomposition is performed. The scaling components at each level as well as the original image are used for classification. The classifier used in this study is a kind of neural network with one-class-in-one-network structure, that is, each subnet is trained separately and there is one subnet per individual. Wavelet packet analysis-based face recognition system is proposed in [8]. The original image is decomposed into its subbands by using two-level wavelet packet decomposition. Out of the 16 subbands, a 21-dimensional feature vector is obtained consisting of variances of 15 detail subbands and three mean values and three variances calculated from different parts of the approximation subband. From this 21 components, only the most meaningful components are selected resulting in a final feature vector size of 11. Bhattacharya distance between these statistical features is used to classify faces. In [9], three-level wavelet decomposition is performed and the resulting approximation subbands at each level are concatenated to produce a new data vector on which PCA is applied. Radial basis functions are used as the classifier of the system. Discriminant waveletfaces approach is proposed in [10]. In this study, third level approximation resulting from three level wavelet decomposition, called the waveletface, is used as the input of the LDA. For classification, presented nearest feature plane (NFP) and nearest feature space (NFS) classifiers are examined. Different from these previous studies, we put into evidence the contribution of wavelet subbands to combat, specifically, illumination and expression factors, and we investigate the interplay of subband information fusion styles, choice of metrics and of features. In other words, the thrust of the paper is to explore how the discriminatory ICA and PCA features can be desensitized or rendered more invariant to the effects of illumination and expression via the judicious selection of subbands and via fusion at various levels.

The paper is organized as follows. In Section 2, multiresolution analysis is briefly reviewed. Subspace

analysis tools (PCA, ICA) and fusion techniques used in the study are explained in Sections 3 and 4, respectively. In Section 5 experimental results against expression and illumination variations are presented separately. Finally, in Section 6 conclusions are given.

2. Multiresolution analysis

Multiresolution methods provide powerful signal analysis tools, which are widely used in feature extraction, image compression and denoising applications. Wavelet decomposition is the most widely used multiresolution technique in image processing. Images have typically locally varying statistics that result from different combinations of abrupt features like edges, of textured regions and of relatively low-contrast homogeneous regions. While such variability and spatial nonstationarity defies any single statistical characterization, the multiresolution components are more easily handled. Wavelet transform can be performed for every scale and translation, resulting in continuous wavelet transform (CWT), or only at multiples of scale and translation intervals, resulting in discrete wavelet transform (DWT). Since, CWT provides redundant information and requires a lot of computation, generally DWT is preferred.

The two-dimensional wavelet transform is performed by consecutively applying one-dimensional wavelet transform to the rows and columns of the two-dimensional data. In Fig. 1, the tree representation of one level, two-dimensional wavelet decomposition is shown. In this figure, G denotes high-pass filtering and H denotes low-pass filtering while $\downarrow 2$ represents downsampling by a factor of 2. In the final stage of the decomposition we have four $N/2 \times N/2$ resolution subband images: A_1 , the scaling component containing global low-pass information; and three wavelet components, H_1 , V_1 , D_1 , corresponding, respectively, to the horizontal, vertical and diagonal details. We can perpetuate this decomposition, either pursuing the same pattern along the scaling component, or obtaining the full-blown tree, or achieve some intermediate tree where 'interesting' branches

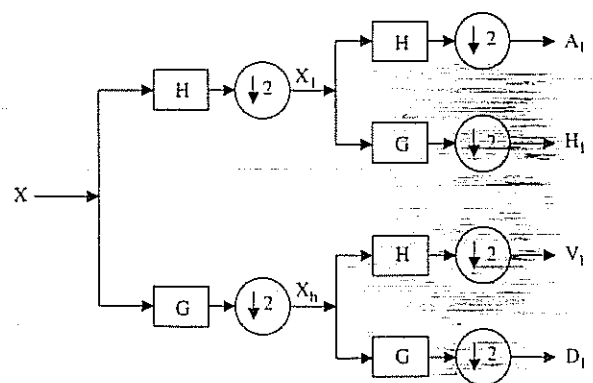


Fig. 1. Tree representation of one-level 2D wavelet decomposition.

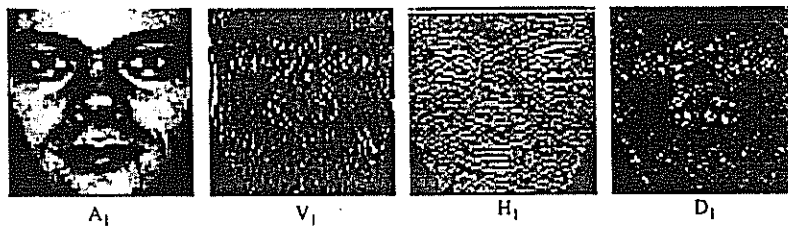


Fig. 2. Sample one-level wavelet decomposed image.

are grown using, for example, some projection pursuit scheme [11]. In Fig. 2, one-level wavelet decomposition of a face image is shown.

In Fig. 3 the schematics of the wavelet decomposition used in this study is shown. The letters in the figure serve to differentiate the scaling component or the orientations of the wavelet components, while the accompanying numbers denote the level of decomposition. If the subbands are obtained by decomposing the original image or any of the scaling components, then they are represented with single letter. If however, a subband is derived by decomposing one of the detail subbands, then these are denoted with two letters, where the first letter indicates the parent subband and the second letter denotes the orientation of the child.

In the first level, a 128×128 original face image is decomposed and four 64×64 pixels resolution subband images— A_1 , H_1 , V_1 and D_1 —are obtained. The H_1 , V_1 , and D_1 components are not further decomposed, because we found their classification performance figures to be very low. Consequently we proceed to decompose only A_1 , yielding four 32×32 subband images— A_2 , H_2 , V_2 and D_2 . In the third level, we decompose all components, A_2 , H_2 , V_2 and D_2 , producing 16×16 subband images. In summary, we obtain 24 different subband images from the original face image and input them into the classification scheme.

3. Subspace analysis

An $m \times n$ resolution face image can be considered as a point in an $N = m \times n$ dimensional image space. For example, a 128×128 face image corresponds to a point in

16,384-dimensional huge feature space. On the other hand, face images are very similar, and therefore highly correlated. It follows that they can be represented in a much lower dimensional feature subspace. PCA and ICA are the two popular methods to descend to such face subspaces.

3.1. Principal component analysis (PCA)

Principal component analysis (PCA) is based on the second-order statistics of the input image, which tries to attain an optimal representation that minimizes the reconstruction error in a least-squares sense. Eigenvectors of the covariance matrix of the face images constitute the eigenfaces. The dimensionality of the face feature space is reduced by selecting only the eigenvectors possessing largest eigenvalues. Once the new face space is constructed, when a test image arrives, it is projected onto this face space to yield the feature vector—the representation coefficients in the constructed face space. The classifier decides for the identity of the individual, according to a similarity score between the test image's feature vector and the PCA feature vectors of the individuals in the database.

3.2. Independent component analysis (ICA)

Independent component analysis (ICA) can be seen as a tool, based on higher order statistics, for extracting independent sources from an observed mixture, where neither the mixing matrix nor the distribution of the sources

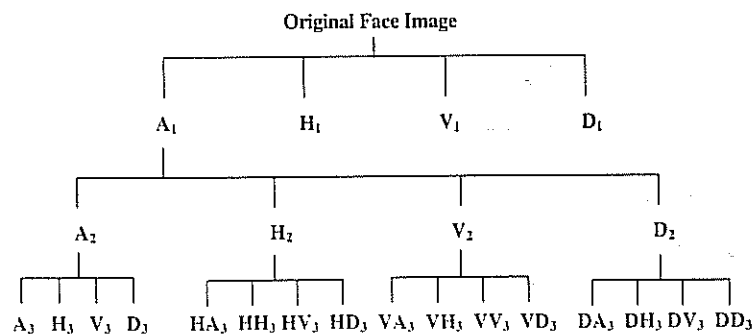


Fig. 3. Wavelet decomposition tree used in the study.

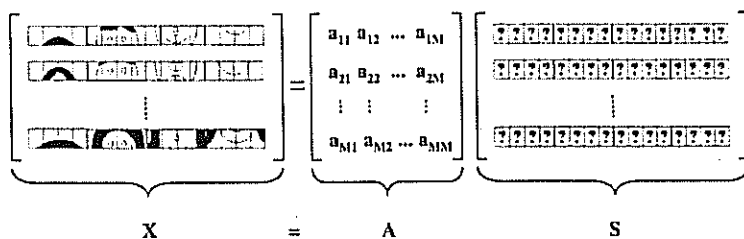


Fig. 4. First face recognition architecture of ICA (ICA1). X: observation, A: rows of mixing matrix, representation coefficients, S: statistically independent basis images.

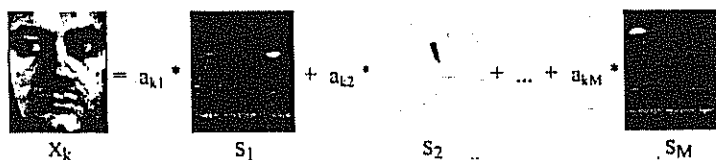


Fig. 5. Face representation using ICA1.

are known. The system model of ICA is given as

$$X = AS$$

where A denotes the mixing matrix, S denotes the source matrix containing statistically independent source vectors in its rows, and X denotes the observation matrix containing the 'linear mixtures' in its rows. The un-mixing matrix W is found by minimizing or maximizing some objective function, such as likelihood ratio, network entropy, mutual information or Kullback-Leibler divergence [12].

The separation matrix, W, under ideal conditions, is the inverse of the mixing matrix A

$$Y = WX \text{ and } W = A^{-1} \text{ and } Y \approx S$$

In the context of face recognition, the use of ICA features was first proposed in [4], where two different approaches were presented. In the first approach (called ICA1 architecture), the face images are assumed to be a linear mixture of an unknown set of statistically independent source images. Therefore, in this architecture,

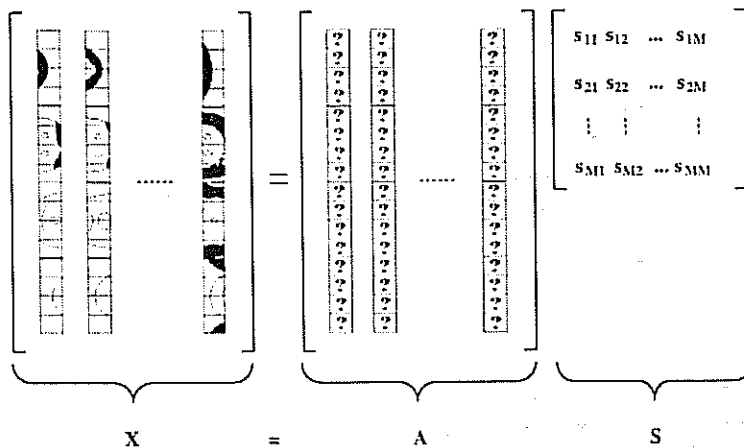


Fig. 6. Second face recognition architecture of ICA (ICA2). X, observations; A, columns of mixing matrix; S, statistically independent representation coefficients.

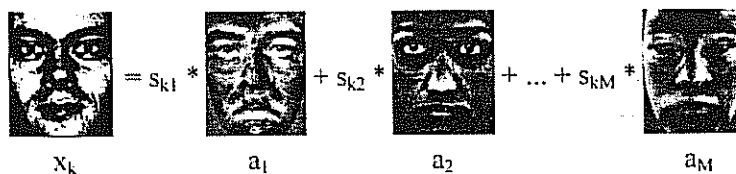


Fig. 7. Face representation using ICA2.

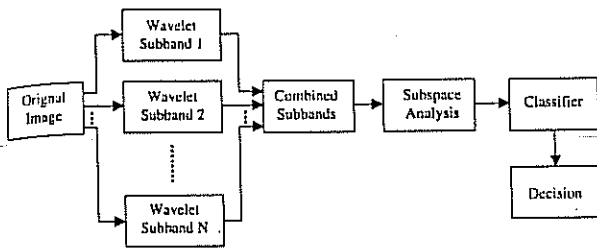


Fig. 8. Block diagram of the data fusion scheme.

the lexicographically ordered face images constitutes the rows of the observation matrix X , the statistically independent basis images constitutes the rows of the source matrix S , and the representation coefficients constitutes the rows of the mixing matrix A (Fig. 4). The source images obtained in this architecture are spatially local and sparse in nature (Fig. 5).

In the second approach (called ICA2 architecture), the representation (weighting) coefficients are assumed to be statistically independent. Therefore, in this architecture, the lexicographically ordered face images constitute the columns of the observation matrix X , the statistically independent representation or weighting coefficients constitute the columns of the source matrix S , and the basis images constitutes the columns of the mixing matrix A (Fig. 6). In this second architecture, while mixing coefficient vectors are independent, source images tend to have global face appearances, similar to the case of PCA (Fig. 7).

4. Fusion

The outcomes from the various wavelet channels are fused to achieve possibly higher correct recognition rates. We investigated three schemes, namely, fusing raw pixel values of the subbands, fusing ICA/PCA feature vectors extracted from the subbands, and fusing the classification decisions of the subbands.

4.1. Data fusion

In data fusion, lexicographically ordered pixels of the subband images are concatenated to construct a new data vector. Following this operation, the subspace projection

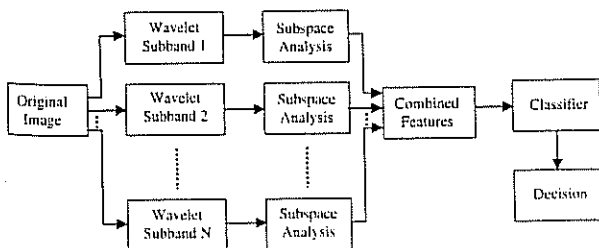


Fig. 9. Block diagram of the feature fusion scheme.

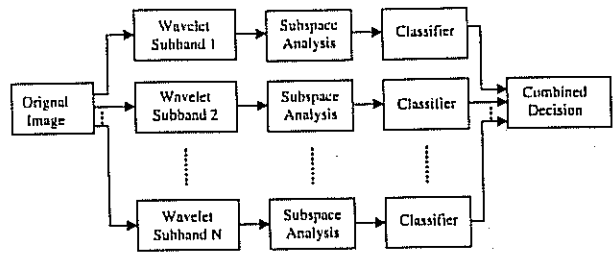


Fig. 10. Block diagram of decision fusion scheme.

and feature extraction are performed on the combined data vectors (Fig. 8).

4.2. Feature fusion

In feature fusion, subspace analysis tools are performed on each subband, and then the extracted feature vectors are concatenated to construct a new feature vector to be used for classification (Fig. 9).

4.3. Decision fusion

In decision fusion, face classification is run separately in each subband. According to the distance values between the test face feature vector and feature vectors of the individuals in the database, a confidence measure is calculated for each classifier. If we have K images in the database and if we define the distance between two feature vectors by the function $d(\cdot)$, then the confidence score, c_i , of a classifier's decision for i th class, is proportional to:

$$c_i = \frac{\sum_{k=1}^K d(x_{\text{test}}, x_{\text{database},k})}{d(x_{\text{test}}, x_{\text{database},i})}$$

The final decision is made through these confidence values by using sum rule, product rule or maximum rule [13]. Note that there are other fusion techniques based on the training with decision patterns of experts. For example, each subband can be considered to be a 'face recognizer expert' and their decisions could be fused via a neural



Fig. 11. Sample face images containing changes in expression—first row from CMU PIE, second row from FERET.

Table 1
Correct recognition rates of successful subband images against changes in expression

	PCA-120			ICA1-120			ICA2-120		
	L1	L2	CC	L1	L2	CC	L1	L2	CC
Original	92.33	90.33	92	93.33	90.33	92	87	87.67	96
A ₁	93	91	91.67	93.67	91	91.67	90	88.67	95.33
A ₂	94	93	93	93	93	93	90.33	92	96
A ₃	94	91.67	91.67	94.33	91.67	91.67	90.33	93.33	95

network. In our case, the limitations in the training data precluded this approach (Fig. 10).

5. Experiments

Two separate experiments are conducted to test the advantage of the wavelet-based face recognition scheme. In the first experiment, the subbands that are potentially insensitive to changes in expression are searched, whereas in the second experiment the subbands that are insensitive to variations in illumination are searched. In both the experiments, feature vectors are extracted from the subband images via PCA, ICA1 and ICA2. The FastICA algorithm [12] is used to perform ICA. Daubechies 4 wavelet is used in the study. The Daubechies wavelets, with their compact support and orthonormal nature, are one of the most widely used wavelet families [14]. Besides this, in [6] it is shown that Daubechies 4 wavelet performs best in terms of computation time and recognition performance with respect to other order Daubechies wavelets, and other well-known wavelets such as biorthogonal, Symlets and Lemarie.

We used the nearest neighborhood classifier in our study. We evaluated comparatively three different distance metrics, namely, the L1 norm, the L2 norm, and the normalized correlation coefficient, defined as follows

$$L1 : d = \sum_{m=1}^M |f_{\text{training},m} - f_{\text{test},m}|$$

$$L2 : d = \left(\sum_{m=1}^M |f_{\text{training},m} - f_{\text{test},m}|^2 \right)^{1/2}$$

$$CC : d = \frac{f_{\text{training}} f_{\text{test}}}{\|f_{\text{training}}\| \|f_{\text{test}}\|}$$

Table 2
Correct recognition percentages using fusion techniques for faces having changes in expression

	Best performing subband	Data fusion	Feature fusion	Decision fusion— sum rule	Decision fusion— product rule	Decision fusion— max. rule
PCA	94.00	94.00	93.67	94.00	93.67	94.67
ICA1	94.33	94.33	94.00	94.67	94.67	94.33
ICA2	96.00	95.67	96.33	96.33	96.33	96.67

where $f_{\text{training},m}$ is the m th ($m=1,\dots,M$) component of the training feature vector, and similarly for $f_{\text{test},m}$.

5.1. Experiments with changes in expression

The experimental data we used to test the performances of subbands against expression changes consists of 600 face images of 150 individuals (four images per individual) (Fig. 11). The images were chosen from CMU PIE [15] and FERET [16] databases. Two hundred and seventy-two of these images belong to CMU PIE database, and remaining 328 images belong to FERET database, fafb image set. Facial expression changes in the images occur due to smiling, blinking or talking in CMU PIE database and due to so-called alternative expressions in FERET database. It would be desirable to have a more extensive set of expressions that cover the whole gamut of human emotions. Nonetheless, these experiments allow us to show the proof of concept, that is that subbands and fusion bring in improvements in the recognition performance. All the face images are aligned with respect to the manually detected eye coordinates, scaled to 128×128 pixels resolution, and histogram equalized. For each individual in the set, two of their images are used for training, and the remaining two images are used for testing purposes. The images that contain neutral facial expression are put in the training set. For recognition, 120-dimensional feature vectors, conserving 91.55% of the energy, are used (Fig. 11).

In Table 1, the correct recognition rates from selected 'successful' subbands are given. Note that under expression change, only the scaling components A₁, A₂, A₃ are selected, and in fact, none of the detail bands qualifies in the recognition competition. For faces subject to expression changes only, we found out that performing PCA or ICA1 on scaling components slightly increases the correct recognition rate. On the other hand, in ICA2, no improvement is observed. With PCA and ICA1 features, the L1 norm gives the best results, whereas for ICA2, the best



Fig. 12. Sample face images containing variations in illumination—first row from CMU PIE, second row from Yale.

results are obtained with normalized correlation measure. ICA2 proves overall superior to ICA1 and PCA in all subbands, its best being realized on the A_2 component.

We next tried to fuse the information contained in the selected subbands, A_1 , A_2 and A_3 , at the data level, the feature level and the decision level as documented in Table 2. For each feature type we took into consideration the distance metric with which they performed best. Therefore $L1$ norm is used for PCA and ICA1, and CC is used for ICA2. In data and decision fusions 120-dimensional feature vectors are used. In feature fusion, 360-dimensional feature vectors are constructed by concatenating the individual feature vectors of the subbands. The best result is obtained with ICA2 by using the decision fusion, based on the maximum rule principle. Since the individual subbands already have relatively high recognition rates, only a small improvement in the performance is achieved. If we take the performance of the PCA on the original image as a reference, the correct recognition rates increases from 92.33 to 96.67, thus an overall 4.33% improvement is achieved. On the other hand, if we take as a reference the performance of ICA2 on the original sized image, the improvement is a meagre 0.67%. Since the performance of the original 128×128 image is on a par with those of the lower-resolution (64×64 , 32×32 , 16×16) versions,

a computational advantage would accrue by selecting, let us say, the much smaller 32×32 images (A_2 component).

5.2. Variations in illumination

The experimental data we used to test the performance of subbands against illumination variations consists of 332 face images of 83 individuals (four images per individual) (Fig. 12). The images were chosen from CMU PIE [15] and Yale databases. Two hundred and seventy-two of these images belong to CMU PIE database and remaining 60 images belong to Yale database. Illumination variations in the images occur due to the intensity and direction of the light. All the face images are aligned with respect to the manually detected eye coordinates, scaled to 128×128 pixels resolution and histogram equalized. For each individual in the set, two of their images that contain frontal illumination with different amounts of light are used for training, and the remaining two images that contain illumination from sides are used for testing purposes. Eighty-dimensional feature vectors, which conserve 92.74% of the energy, are used (Fig. 12).

The correct classification rates of subbands selected on the basis of their success are given in Table 3. It is interesting to observe that the horizontal detail subbands (H_2 , H_3 , HH_3) attain higher correct classification rates as compared to the scaling components. In this respect, PCA and ICA1 features extracted from the three horizontal detail images lead to better results than the ICA2 features. The normalized correlation measure, CC, has proved to be superior in all of the three feature extraction methods. In contrast, ICA2 features performed better when extracted from the scaling components (A_1 , A_2 , A_3) as well as the original face image, though the recognition rate remained overall inferior to those attained with the horizontal components H_2 , H_3 , HH_3 . As can be observed from Table 3, a significant performance improvement, of the order of 40%, is achieved by using horizontal details of wavelet subbands. If we again take as a reference the performance of the PCA on the original image (54.82%

Table 3

Correct recognition rates of sample subband images in the presence of illumination variations (first five rows correspond to the successful subband images)

	PCA-80			ICA1-80			ICA2-80		
	L1	L2	CC	L1	L2	CC	L1	L2	CC
Original	54.82	52.41	51.20	51.81	52.41	51.20	57.83	56.02	66.87
A_1	56.02	53.01	51.20	51.20	53.01	51.20	59.64	56.63	67.47
A_2	57.83	53.01	51.20	51.81	53.01	51.20	56.02	60.24	63.86
H_2	34.94	33.13	71.69	33.73	33.13	71.69	34.94	30.12	71.08
A_3	57.23	51.81	50.60	51.81	51.81	50.60	57.83	60.24	60.24
H_3	65.06	64.46	72.29	67.47	64.46	72.29	58.43	59.64	62.65
HH_3	45.18	45.78	68.67	45.78	44.58	69.28	40.36	42.17	60.24
V_2	21.08	21.08	37.95	23.49	21.08	37.95	18.67	18.67	48.80
D_2	6.02	8.43	36.75	8.43	8.43	36.75	6.63	6.02	39.16
V_3	43.37	43.98	45.78	43.37	43.98	45.78	42.17	44.58	50
D_3	33.73	34.94	56.02	36.14	34.94	56.63	34.34	32.53	49.40

Table 4
Correct recognition rates of fusion techniques against variations in illumination

	Best performing subband	Data fusion	Feature fusion	Decision fusion— <i>sum rule</i>	Decision fusion— <i>product rule</i>	Decision fusion— <i>max. rule</i>
PCA	72.29	75.90	77.11	77.11	77.11	75.30
ICA1	72.29	75.90	75.90	77.11	77.11	75.90
ICA2	71.08	77.11	72.89	75.90	75.30	77.71

score), the correct recognition rate increases to 72.29, by using only a 16×16 third-level horizontal subband component. It is quite interesting that the vertical and diagonal details bring in no improvement at any level of fusion. In Table 3, the last four rows are reserved for the vertical and diagonal subband components on two successive scales, where one can observe the poor performance with these components.

Next, we carried out fusion experiments, with fusions realized at the data level, feature level and decision level. In all fusion experiments, we used the correlation coefficient, CC, since it was overall the best performing distance metric. The subbands involved in the fusion were the three components, H_2 , H_3 and HH_3 using PCA and ICA1 features, and the six components A_1 , A_2 , H_2 , A_3 , H_3 and HH_3 using ICA2 features. These subbands were selected on the basis of their performance in single-band experiments. Note that since the energy levels of the scaling components and of the horizontal detail components are very different, for the data fusion experiments, they must be rendered commensurate. In other words, to prevent the low-frequency subbands containing higher energy from dominating the horizontal details, energy normalization is applied by scaling each component by its respective standard deviation.

In the data fusion, the concatenated subband coefficients form longer vectors. For example, when the A_1 , A_2 , H_2 , A_3 , H_3 and HH_3 data are fused, the resulting vector becomes $4096 + 1024 + 1024 + 256 + 256 + 64 = 6656$ dimensional. These longer concatenated vectors are, however, still reduced to dimension 80, as shown in Fig. 8, after PCA projection (recall that ICA scheme has a PCA preprocessing stage.) In the feature fusion case, on the other hand, 240-dimensional feature vectors are formed in PCA and ICA1, which are constructed by concatenating the feature vectors of H_2 , H_3 and HH_3 subbands. In ICA2, 480-dimensional feature vectors result from concatenation of the features of the A_1 , A_2 , H_2 , A_3 , H_3 and HH_3 . Finally, in the case of decision fusion, the confidence scores from the different channels (three channels in the case of PCA and ICA1 and six channels in the case of ICA2) are combined to a final decision by using sum rule, product rule and maximum rule [13].

Table 4 shows the improvements in correct recognition rate achievable as a result of fusion schemes. All features (PCA, ICA1, ICA2) seem to benefit from fusion, whether it is data, feature or decision fusion. The first column of Table 4 gives the best attainable 'pre-fusion' score for

a comparison. The highest classification performance increase is obtained with ICA2 using decision fusion based on the maximum rule principle. The recognition performance amounts to 77.71 in a database where faces are subject to lighting variations. The contribution of fusion techniques in this case is a 5.42% improvement vis-à-vis the single best performing subband and feature combination (PCA, ICA1, H_3).

6. Conclusions

In this study, we searched for the frequency subbands that qualify as being insensitive to expression differences and illumination variations on faces. Briefly, it was observed that the frequency subbands containing coarse approximation of the images are successful against expression differences, whereas the subbands containing horizontal details are successful against illumination variations.

Since the recognition performance is not in the first place very adversely affected by changes in facial expression, the performance improvement brought about by the multi-resolution analysis and/or fusion remains unimpressive. One interesting observation made is that the performance with the original image and scaling components at various levels, namely, the A_1 , A_2 and A_3 components, respectively, at the 128×128 , 64×64 , 32×32 and 16×16 resolutions remain almost the same, hence low-resolution versions should be preferred for computational simplicity.

The search for alternative wavelet channels for faces in the presence of illumination variations proves much more effective. The horizontal wavelet components, obtained after either one or two stage of low-pass filtering, were found to be the channels yielding the highest performance scores. One can conjecture that horizontal wavelet removes any horizontal illumination pattern, e.g. one cheek darker, the other lighter. Alternatively, one can conjecture that the horizontal details emphasize the left-right asymmetry on faces, and the facial asymmetry was shown to be a very good feature in face recognition [17]. Fusing a number of channels further improves the highest performance achieved by a single channel, by another 5.42%. The Tables 5–9 summarize the recognition performance improvements instrumented by the choice of metrics, choice of wavelet components and choice of fusion scheme. The comparison reference is the original face image with PCA features and

Table 5
Improvement due to choice of features and metrics vis-à-vis the PCA-L1 performance

Facial expression change (%)	Illumination change (%)
3.67	12.05

Both improvements are attained with the max-rule fusion of the ICA2 features of the A_1 , A_2 , A_3 , bands in lieu of the original images (compare first rows of Tables 1 and 3).

Table 6
Improvement due to wavelet decomposition vis-à-vis the full-scale performance

	Facial expression change (%)	Illumination change (%)
PCA	1.66	17.47
ICA1	1.00	19.88
ICA2	–	4.21

Compare corresponding columns in Tables 1 and 3: for example, PCA on original image performance is 54.82, while PCA on the H_3 band performance is 72.79 resulting in 17.47% improvement.

Table 7
Improvement due to fusion vis-à-vis the best subband component

	Facial expression change (%)	Illumination change (%)
PCA	0.67	4.82
ICA1	0.33	4.82
ICA2	0.67	6.63

Table 8
Overall improvement vis-à-vis the PCA-L1 performance

Facial expression change (%)	Illumination change (%)
4.33	22.89

Table 9
Decrease in the error rate vis-à-vis the PCA-L1 performance

Facial expression change (%)	Illumination change (%)
56.45	50.66

L1 metric (referred to, simply as PCA-L1), as this is the most commonly occurring technique in the literature.

In conclusion, it can be said that natural expression changes, i.e. smiling, blinking, talking, do not cause severe performance reduction in face recognition. The attained correct recognition rates are already relatively high; therefore, multiresolution analysis and fusion provide a small improvement. On the other hand, it is observed that recognition of faces, subject to illumination changes is a more sensitive task. Utilizing multiresolution analysis and fusion is quite effective in combating the detrimental effects of illumination variations. For example, it would be intriguing to consider the illumination variations coupled with expression variations. However, not only the effects of illumination variations are more dominant vis-à-vis

the expression variations, but also we have found that the subbands useful in combating illumination variations subsumes already subbands effective for expression variations.

Similar studies can be conducted for other facial factors, of aging, accessories or pose, provided adequate databases become available. In such relatively more complicated tasks, where more intra-class variability and/or more number of classes can be encountered, we believe the multiresolution scheme can also be beneficial.

References

- [1] R. Brunelli, T. Poggio, Face recognition: features versus templates, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15 (10) (1993) 1042–1052.
- [2] M. Turk, A. Pentland, Eigenfaces for recognition, *Journal of Cognitive Science* 1991; 71–86.
- [3] P.N. Bellhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (7) (1997) 711–720.
- [4] M.S. Bartlett, H.M. Lades, T.J. Sejnowski, Independent component representations for face recognition, *Proceedings of the SPIE Conference on Human Vision and Electronic Imaging III*, San Jose, CA, 1998 pp. 528–539.
- [5] Y. Moses, Y. Adini, S. Ullman, Face recognition: the problem of compensating for changes in illumination direction, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (7) (1997) 721–732.
- [6] G.C. Feng, P.C. Yuen, D.Q. Dai, Human face recognition using PCA on wavelet subband, *SPIE Journal of Electronic Imaging* 9 (2) (2000) 226–233.
- [7] J. Tang, R. Nakatsu, S. Kawato, J. Ohya, A wavelet-transform based asker identification system for smart multi-point tele-conferences, *Journal of the Visualization Society of Japan* 20 (1) (2000) 303–306.
- [8] C. Garcia, G. Zikos, G. Tziritas, Wavelet packet analysis for face recognition, *Image and Vision Computing* 18 (4) (2000) 289–297.
- [9] B. Li, Y. Liu, When eigenfaces are combined with wavelets, *International Journal of Knowledge-Based Systems* 15 (5/6) (2002) 343–347.
- [10] J.T. Chien, C.C. Wu, Discriminant waveletfaces and nearest feature classifiers for face recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (12) (2002) 1644–1649.
- [11] P.J. Phillips, Matching pursuit filters applied to face identification, *IEEE Transactions on Image Processing* 7 (8) (1998) 1150–1164.
- [12] A. Hyvärinen, E. Oja, Independent component analysis: algorithms and applications, *Neural Networks* 13 (2000) 411–430.
- [13] J. Kittler, M. Hatef, R.P. Duin, J.G. Matas, On combining classifiers, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (3) (1998) 226–239.
- [14] I. Daubechies, Ten lectures on wavelets, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 61, SIAM Press, Philadelphia, 1992.
- [15] T. Sim, S. Baker, M. Bsat, The CMU pose, illumination, and expression (PIE) database, *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, May 2002.
- [16] P.J. Phillips, H. Moon, S.A. Rizvi, P.J. Rauss, The FERET evaluation methodology for face-recognition algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (10) (2000) 1090–1104.
- [17] Y. Liu, J. Palmer, A quantified study of facial asymmetry in 3D faces, *Proceedings of the 2003 IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, in conjunction with the 2003 International Conference of Computer Vision (ICCV '03), October 2003.



Spectral Analysis of Event-Related Hemodynamic Responses in Functional Near Infrared Spectroscopy

CEYHUN BURAK AKGÜL AND BÜLENT SANKUR

Electrical and Electronics Engineering Department, Boğaziçi University, Bebek, Istanbul, Turkey

ATA AKIN*

Institute of Biomedical Engineering, Boğaziçi University, Bebek, Istanbul, Turkey

ataakin@boun.edu.tr

Received January 15, 2004; Revised June 29, 2004; Accepted August 24, 2004

Action Editor: Xiao-Jing Wang

Abstract. The goal of this paper is to design experiments that confirm the evidence of cognitive responses in functional near infrared spectroscopy and to establish relevant spectral subbands. Hemodynamic responses of brain during single-event trials in an odd-ball experiment are measured by functional near infrared spectroscopy method. The frequency axis is partitioned into subbands by clustering the time-frequency power spectrum profiles of the brain responses. The predominant subbands are observed to confine the 0–30 mHz, 30–60 mHz, and 60–330 mHz ranges. We identify the group of subbands that shows strong evidence of protocol-induced periodicity as well as the bands where good correlation with an assumed hemodynamic response models is found.

Keywords: near infrared spectroscopy, cerebrovascular dynamics, time-frequency distribution, oscillatory brain dynamics

1. Introduction

Functional near infrared spectroscopy (fNIRS) has been proposed as a non-invasive and rapid tool to monitor the cerebrovascular changes during cognitive tasks (for a review, see (Chance et al., 1998; Obrig et al., 2000b; Villringer and Chance, 1997). Their results have indeed confirmed that fNIRS shows explicitly the response pattern of a deoxyhemoglobin (Hb) decrease and an oxyhemoglobin (HbO₂) increase when monitored over the same area where the maximal blood oxygen level-dependent (BOLD) signal increase has been observed in functional magnetic resonance imaging (fMRI) studies. Similar to fMRI, fNIRS measures any cognitive activity indirectly via

the coupling of neuronal activation to blood flow and oxygen delivery. However in contrast, fNIRS is capable to monitor these activities on a millisecond basis at the expense of lost spatial details. The fNIRS systems use multi-wavelength illumination to extract both Hb and HbO₂ concentration changes (Boynton et al., 1996; Obrig et al., 2000a, b).

In this paper, we aim to characterize the spectrum of fNIRS signals during a cognitive task and identify the relevant frequency bands. We define as relevant frequency bands those intervals of frequency, which are most correlated with neuronal activity associated with cognitive tasks and where protocol-induced periodicity can be observed.

Several researchers in the field of neuroimaging have investigated oscillatory behavior of hemodynamic activity. Efforts for characterizing the components in the

*To whom all correspondence should be addressed.

spectra have concentrated on establishing a physiological correspondence with the peaks or the energy bands. Functional MRI and transcranial Doppler sonography studies have proposed several association mechanisms of vasomotor dynamics while emphasizing the fact that a strong signal, the brain hemodynamic response, exists and dominates the lower portion of the spectrum computed from temporal neuroimaging data (Obrig et al., 2000a; Franceschini et al., 2000; Prince et al., 2003; Toronov et al., 2000).

In all these studies, the main goal has been to identify the spectral peaks or bands related to a given task, hence elucidate the underlying physiological dynamics and their relation to performance. This provides a means to investigate the coupling between cerebral energy metabolism and cerebrovascular dynamics (namely, neurovascular coupling). The common view is that while some of the oscillatory dynamics occur independently of any task and give rise to distinct spectral bands uncorrelated with other physiological activities (e.g. breathing, heart beat etc.), other bands are affected by psychological or pathological conditions (or vice versa) through which degradation of task performance can be monitored (Kim and Uğurbil, 2002; Franceschini et al., 2000; Obrig et al., 2000a; Toronov et al., 2000).

Several researchers used spectral estimation techniques and transfer function analysis to relate the brain hemodynamic response to various stimuli and physiological events (Villringer and Chance, 1997; Svensen et al., 2000; Wobst et al., 2001; Hu et al., 1999; Kuo et al., 1998; Giller et al., 1999). Most studies that are performed by transcranial Doppler sonography suggest low frequency domination due to cerebral autoregulation of hemodynamical activity where the increase in acidity levels (increase on carbondioxide content) leads to this very slow vasomotor activity forming the Mayer waves or the V-signal (Obrig et al., 2000a).

Functional NIRS systems collect cerebrovascular information by spectroscopic methods tuned to the main chromophores in the near infrared range (Hb and HbO₂). The simplicity of measurement system allows the fast sampling of such changes, an advantage over fMRI systems. Moreover, the ease of application of the probe and the freedom to move around are factors that alleviate the discomfort felt by the subjects during these studies. The increased level of comfort and non-invasive nature of fNIRS studies can be considered as minimization of the nuisance factors by suppressing the disrupting effects of emotions, stress and physical limitations on hemodynamic signals. Major limitation

of fNIRS systems when compared against fMRI, is the low spatial resolution. Thus in any interpretation of the fNIRS signals, the fact that these signals derive from a much wider support must be taken into consideration.

To summarize, the goals of the paper are to (i) Parse the fNIRS signal spectrum in an unsupervised and serendipitous manner based on the time-frequency spectrum, (ii) Identify the frequency bands that are most informative with respect to stimulus, and (iii) Develop a framework for choosing the set of optodes that carry relevant information on the single-event trial cognitive task. Once these fNIRS bands are identified, we believe that behavioral parameters can be better correlated with these spectral bands.

2. Preprocessing of fNIRS Signals

2.1. Data Collection

For all tasks, participants (with written informed consents) were seated in a comfortable chair in front of a computer screen in a sound-attenuated, electrically shielded room and were fitted for EEG and fNIR monitoring. Functional NIRS measurements are taken by a custom-built system developed at Dr. Britton Chance's lab at University of Pennsylvania. The system houses four LED multi-wavelength light sources and twelve photodetectors that when time and source-multiplexed constitute four non-overlapping quadruples of photodetectors (see Fig. 1). Hence in one scan of the forehead a total of sixteen measurements at each wavelength can be acquired totaling to 48 optic signals. The source and detectors are equidistantly placed on the probe as seen in Fig. 1. The probe is positioned such that the base of it aligns with the eyebrows of the subject and the middle with the Fz location from EEG electrode placement and a sports bandage is used to secure it on its place and eliminate background light leakage. Taking into consideration Firbank and et al.'s study (Firbank et al., 1998), a pre-determined source detector separation of 2.5 cm was used to account for an average adult cortex depth around 1.5 cm that allowed us to probe the first couple of millimeters of the gray matter (Firbank et al., 1998; Boas et al., 2001). Notice, however, that only one fNIRS sequence results from the processing of the recordings at three wavelengths.

It was argued that a modified version of the Beer-Lambert law could be used to determine the concentrations of hemoglobin agents from raw fNIRS measurements (Villringer and Chance, 1997). We

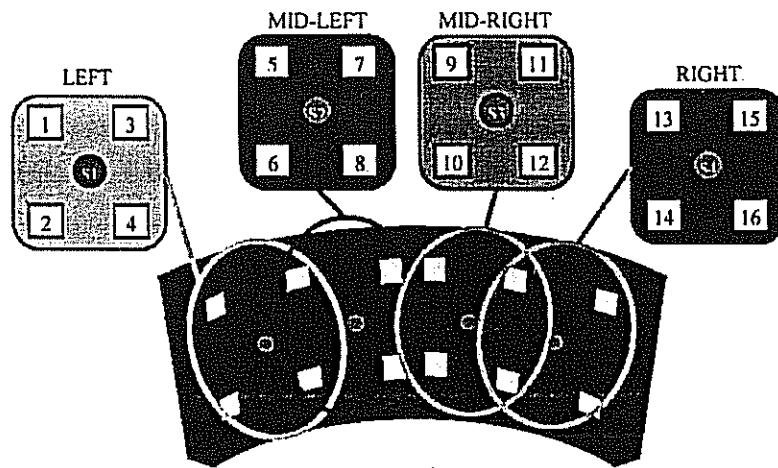


Figure 1. Source-detector configuration on the brain probe and nomenclature of photodetectors.

considered the oxyhemoglobin (HbO_2) agent, based on a recent study, where it was demonstrated that a strong correlation exists between blood-oxygenation-level-dependent fMRI data and diffuse optical HbO_2 data (Strangman et al., 2002). Thus, our measurement data consist of the time series of the HbO_2 signal samples.

Target categorization or "oddball task" is a simple discrimination task in which subjects are presented with two classes of stimuli in a Bernoulli sequence in the center of the screen. The probability of one stimulus is less than the other (e.g. 20% of trials for the "target" or "oddball" stimulus, versus 80% of trials for the "typical" or "context" stimulus); the participants have to press a button when they see the less frequent of the two events. Stimulus categories are varied, beginning with the letters "XXXXX" versus the letters "OOOOO". 1024 stimuli are presented 1500 ms apart (total time, 25 minutes); a target is presented on 64 trials, with a minimum of 12 context stimuli in between to allow for the hemodynamic response to settle. Therefore the rest of the stimuli (960) are context cases. Subjects are asked to press the left button on a mouse when they see "OOOOO" and right button when they see the target "XXXXX" (McCarthy et al., 1997). This timing parameter is used as the behavioral reaction parameter tracking the performance of the subjects. Five male subjects with an age range of 22–50 are recruited for the preliminary test. Protocol is approved by IRB of Drexel University and MCP Hahnemann University (now Drexel University College of Medicine). We have the following additional specifications for target stimuli:

- (i) In a given experiment, overall 64 target stimuli are presented. The stimuli follow a block periodic temporal pattern, where in every block there are 8 stimuli with randomly jittered locations. However, the same pattern is repeated in every one of the eight blocks during the course of the 25-minute experiment. In other words, the inter-arrival patterns between the 1st and 8th target stimuli repeat themselves successively between the 9th to 16th, 17th to 24th and so on up to 57th and 64th.
- (ii) Inter-target interval is a random variable uniformly distributed on the (30 to 50) sample interval, or alternately on the (18 to 29) second interval.

Duration of stimuli of both context and target types is 500 ms, hence there are blank intervals of 1 second. Prior to digitization, analog optical density signals are filtered by an RC filter with 330 mHz cut-off frequency. Hence although recording is done at a sampling rate of 1.7 Hz, (the Nyquist bandwidth is 850 mHz), we consider that the 3-dB bandwidth is 330 mHz.

We had five human subjects for which the target categorization experiments were carried out. We observed that some of the detector outputs were not usable, due to either severe motion artifacts or occasional defects of the sensor. Our rejection criteria were based on visual investigation: one sort of defect was clipping in the signal amplitudes due to saturation of sensors; another was the outliers due to head movements of subjects. We also observed that, in some cases, sensors did not give any measurement at all due to hair occlusions. After eliminating the improper measurements, we ended

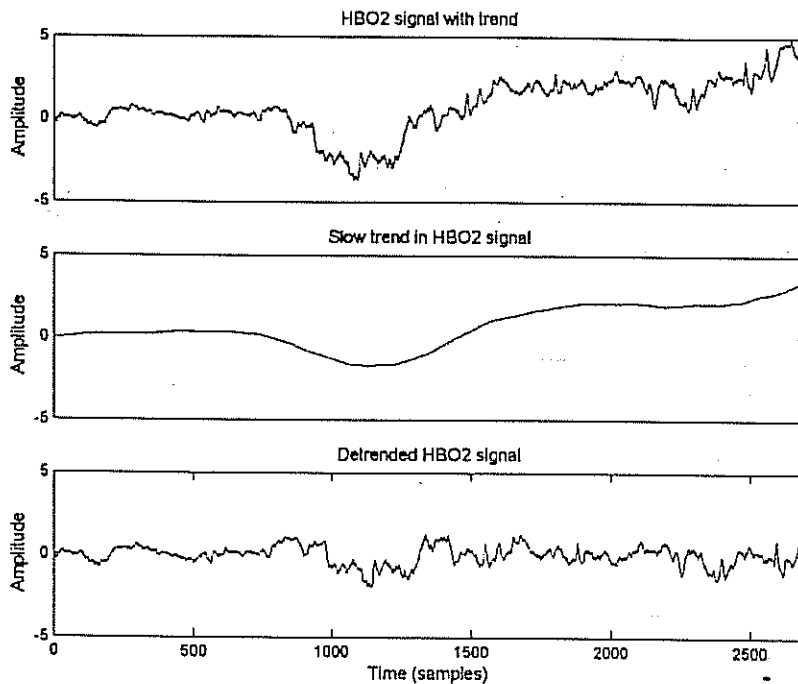


Figure 2. The effect of detrending on a typical HbO₂ signal.

up in a collection of 72 fNIRS-HbO₂ time series out of the planned $5 \times 16 = 80$ recordings. These signals were detrended, which effectively removed the very low frequency components below 3 mHz. Simple moving average filtering and subtraction of the local average perform the trend removal: a frame of support 500 samples (corresponding to 4.9 minutes of data) is slid continuously over the time-series and the mean value of the samples inside the frame is subtracted from the actual value at the frame position. The effect of such a detrending scheme is illustrated in Fig. 2. In summary, for signal characterization of the fNIRS signals we have used the 72 detrended series, each consisting of approximately 2700 samples, corresponding to duration of the cognitive task experiment of about 25 minutes.

2.2. Signal Characterization

In this section, we discuss features that characterize the fNIRS signals $\{s(t)\}$. Specifically, we search for informative bands and then explore periodicities in these bands corresponding to the target quasi-periodicity.

We conjecture that the fNIRS time-frequency spectrum can be partitioned into characteristic subbands.

The feature that we use in characterizing these subbands is the relative power time series for each band. Similar time profiles of subband energies were used in the analysis of epileptic seizure-induced EEG series (Blanco et al., 1995).

We start by estimating the time-frequency representation (TFR) in terms of the short-time Fourier transform (STFT), defined as

$$S(\tau, f) = \int_{-\infty}^{\infty} s(t)w_D(t - \tau)e^{-j2\pi ft} dt \quad (1)$$

where $s(t)$ denotes the fNIRS signal of interest and $w_D(t)$ is a window of finite support D . The short-time Fourier transform (STFT) in (1) is actually computed using the *discrete Fourier transform* (DFT), so that the TFR is discrete in both time and frequency, respectively, with time resolution Δt and frequency sampling interval Δf . A short-time transform is warranted since the signals are nonstationary and also because we want to capture and characterize local events, like brain hemodynamic response in the course of the fNIRS process. The windowing $w_D(t)$ guarantees the local nature of the spectral analysis and its support is chosen so that within that D interval the process can be considered

Table 1. Parameters of TFR, sampling rate $F_s = 1700$ mHz.

Parameter	Value	Comment
Window type	Rectangular	Rectangular window has the best frequency resolution
Window length D	60 samples ≈ 35 s	This window size provides for 28.3 mHz frequency resolution
Time resolution Δt	5 samples ≈ 3 s	The analysis is slid by steps of 5 samples, an adequate time resolution.
Frequency sampling Δf	1 mHz	Thus each 30 mHz subband has 30 representative samples.

to be at least wide-sense stationary. Table 1 gives the parameters used in the TFR analysis.

We remark that the frequency resolution is given by the effective window length, hence it is of the order of $F_s/D \approx 28$ mHz, while the 1 mHz frequency sampling rate Δf is obtained by interpolation, that is padding the windowed time series with zeroes.

We consider the evolutionary power spectral density within the n th frequency band at the instant t as

$$B_n(t, f) = S^*(t, f)S(t, f) \quad \text{in } f \in (f_{n,l}, f_{n,h}) \quad (2)$$

where $f_{n,l}$ and $f_{n,h}$ denote, respectively, the lower and upper limits of the band and the superscript * stands for complex conjugation. In our case we took $f_{n,h} - f_{n,l} = 30; \forall n$. Thus this initial partitioning of the frequency spectrum has 11 bands of width 30 mHz, which collectively cover the 330 mHz bandwidth. The 30 mHz initial width of the bands is dictated by the achievable resolution after windowing the time-series data. The window size of 60 samples provides frequency resolution of 28 mHz.

The total power in a band as a function of time can be calculated by

$$I_n(t) = \int_{f_{n,l}}^{f_{n,h}} B_n(t, f) df \quad (3)$$

Similarly the total instantaneous power in the whole signal bandwidth $I(t)$ is defined as

$$I(t) = \int_{f_l}^{f_h} B(t, f) df \quad (4)$$

where, in our case, the integration goes from 3 to 330 mHz. The lower limit is dictated by detrending, while the upper limit by the cut-off frequency of the RC filter which operates prior to digitization. Finally,

the relative power profile in the n th band as a function of time becomes:

$$R_n(t) = \frac{I_n(t)}{I(t)} \quad (5)$$

The relative power profile per band reflects the temporal evolution of the relative power in each band. Two bands are considered to be distinct if the evolutions of their relative power profiles are dissimilar. Conversely, two bands are merged into one if their $R_n(t)$ responses are close to each other. Dissimilar bands that have different time evolution of the profiles, $R_n(t)$ are considered to provide different information.

3. Selection of Relevant Frequency Bands by Clustering

3.1. Band Selection Methodology

We adopt an agglomerative approach, by starting with a fine partitioning of the frequency spectrum, and then grouping bands similar in their evolutionary power profiles, $R_n(t)$, into wider bands that hopefully capture significant signal information. To express the relative power time-series from different bands and detectors/subjects, let's adopt the following notation:

$R_n^m(t)$: time-series of relative power profile time-series for the n th band of the m th fNIRS signal.

Thus the subscript n denotes the frequency band of interest, where $n = 1, \dots, 11$, which covers the frequency range $[(n-1) \times 30, n \times 30]$ mHz. The superscript m points to one of the $m = 1, \dots, M (M = 72)$ time-series in the database. Recall that these time series were obtained from the 16 detectors of the 5 subjects, after some pruning. We will refer to the superscript m as simply the m th measurement. The time index t runs with the lags of $\Delta t = 5$ samples, $t = 1, \dots, T$. It will be convenient to express the whole time-series $R_n^m(t)$, $t = 1, \dots, T$ in vector notation as \mathbf{R}_n^m . Thus the T -dimensional \mathbf{R}_n^m vector denotes the time series of the m th detector/subject in the n frequency band. Notice that we have a total of $N \times M = 11 \times 72$ such \mathbf{R}_n^m vectors, each detector being expanded onto 11 bands, and conversely, there are 72 representative time-series for each band.

We search for the formation of the bands by a clustering procedure. In fact, we use a scheme based on agglomerative clustering and majority voting as described below. We apply clustering to the N bands of each measurement m so that the set of N subbands

$R^m = \{R_n^m\}_{n=1, \dots, N}$ are clustered into C subbands $Q^m = \{Q_c^m\}_{c=1, \dots, C}$. Specifically the 11 initially chosen subbands from any detector/subject are clustered into $C = 3$ subbands. We have decided for this target number of three clusters to allow for a possibly very low frequency band, a high frequency band and potentially a single "interesting" mid-band.

The agglomerative clustering on data set R^m has the following steps:

- (i) Initialize: Assign each vector R_n^m to its own cluster.
- (ii) Compute all pairwise distances between clusters.
- (iii) Merge the two clusters that are closest to each other.
- (iv) Return to step (ii) until there are only three clusters left.

There are two important aspects in such a clustering algorithm: the *metric* used to compute distances and the *closeness criterion* between vectors. In this study we adopted the *One-Minus-the-Normalized correlation coefficient* as the distance metric

$$d(R_p^m, R_q^m) = 1 - \frac{\langle R_p^m, R_q^m \rangle}{\|R_p^m\| \cdot \|R_q^m\|} \quad (6)$$

where $\langle \cdot, \cdot \rangle$ stands for the inner product of two vectors and $\|\cdot\|$ for the Euclidean norm. The vectors involved in the computation are made zero-mean by subtracting their mean value. Furthermore, we adopt the *single linkage criterion* as the closeness criterion. Accordingly, the pair of adjacent bands (p, q) for which $d(R_p^m, R_q^m)$ is minimum are merged.

The end product of clustering the R^m set is a dendrogram D^m , a hierarchical tree that helps us to visualize cluster relationships. An example is shown in Fig. 3. The dendrogram for the m th measurement D^m is pruned in order to get the clustered set Q^m . This is accomplished by simply obtaining the cutset of the dendrogram that yields the target number of C clusters. In other words, the dendrogram is cut, as shown in Fig. 3 at a distance value, that is the ordinate, to yield the desired number of clusters. Within each one of the C clusters, the merged bands are similar to each other according to the chosen correlation coefficient metric, while across clusters they are dissimilar. The leaves of the dendrogram, that is the singleton clusters, which correspond to the initial bands, become thus grouped into $C = 3$ larger bands.

Once the agglomerative clustering is accomplished we obtain M dendrograms, one for each measurement. To extract a single set of subbands from the M clusterings we resort to a voting scheme. At this stage there

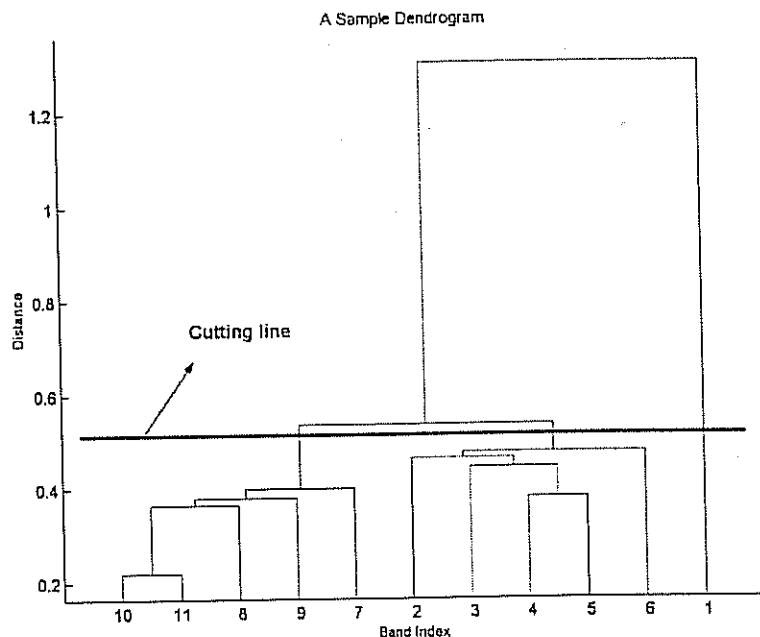


Figure 3. A typical dendrogram: The horizontal axis indexes the initial bands, vertical axis indicates with pairwise cluster distances.

are $C \times M = 3 \times 72 = 216$ candidate bands with possibly differing bandwidths, out of which we try to determine the most frequently occurring ones. We therefore rank these band patterns based on their frequency of occurrence. To make this point clear, let's consider again the sample dendrogram in Fig. 3, which results in the following subbands: {0–30 mHz, 30–180 mHz, 180–330 mHz}. We determine how many times each of these subband formations are generated by the clustering of all the R^m measurements, each occurrence counting as a vote. Selecting the subband patterns that have received the highest number of votes (frequency of occurrence), we achieve a partitioning of fNIRS frequency spectrum, where the resulting bands are non-overlapping and exhaustively cover the frequency interval 0–330 mHz.

3.2. Results of Band Grouping

As a result of clustering and voting, we obtained 9 candidate bands that shared 216 votes, as shown in Table 2.

Based on the results of Table 2, there are many alternative ways to partition the spectrum. Nevertheless, there is one obvious choice that exhibits the strongest evidence with 186 votes (86.1% of the total): this partition is 0–30 mHz, 30–60 mHz, 60–330 mHz. We believe that this partitioning is characteristic of fNIRS power profiles. Hereafter we call them as "canonical frequency bands" and denote them by letters *A*, *B*, *C* as in Table 3. Similarly when we say that we used the *AB*-band, *BC* band or *ABC* band, or that the signal was prefiltered in the *AB*-band etc we mean the band stretching, respectively, in the 0–60, 30–330, 0–330 ranges.

Table 2. Candidate bands (total number of occurrences is 216).

Band	(mHz)	Votes	Percentage (%)
0	30	72	33
30	60	57	26
30	90	11	5
30	180	3	1
30	120	1	0
60	330	57	26
90	330	11	5
120	330	1	0
180	330	3	1

Table 3. Canonical fNIRS spectrum bands.

Bands (mHz)	0–30 <i>A</i>	30–60 <i>B</i>	60–330 <i>C</i>
Votes	33%	26%	26%

4. Characteristics of the Canonical Band Signals

Once the canonical bands are determined, we set ourselves three goals:

- (i) To compare the cross-correlations of the relative power profiles in different bands and to discuss the physiological meaning of these bands.
- (ii) To explore the existence of any periodicity in the temporal patterns within the bands. This is relevant because the stimuli are quasi-periodic and in some of the bands, we expect that periodicity subsists more heavily than in the others. The intent of this periodicity search is to corroborate the evidence that the fNIRS does indeed measure cognitive activity, as will be demonstrated in Sections 4.2 and 4.3.
- (iii) Finally, to correlate and to fit fNIRS waveform excerpts collected right after target onsets and the Gamma function, a model for brain hemodynamic response function popular in fMRI studies.

4.1. Interpretation of the Canonical Bands

In several other studies, three main frequency bands of interest have been identified for cerebral hemodynamics: a very low frequency VLF (8–33 mHz), a low frequency LF (around 100 mHz) named as the Mayer waves or V-signal (Mayhew et al., 1999; Obrig et al., 2000a), and a high frequency component HF (around 250 mHz), the latter being definitely synchronous with breathing rate (Kuo et al., 1998). Similarly, we conjecture that each of the canonical bands is associated with one or more of the physiological activities related to hemoglobin concentrations. The lowest frequency *A* band (0–0.03 Hz) is mainly responsible for the slow baseline signal that is thought to be reflecting the very slow vasomotor activity due to heart rate fluctuations and thermoregulation (Toronov et al., 2000; Francheschini et al., 2000; Giller et al., 1999). In fact, reports on the frequency content of such fluctuations have identified this signal as being the phasic dilation and contractions of "the small regulating arteries, and these vasomotor waves produce fluctuations in cerebral

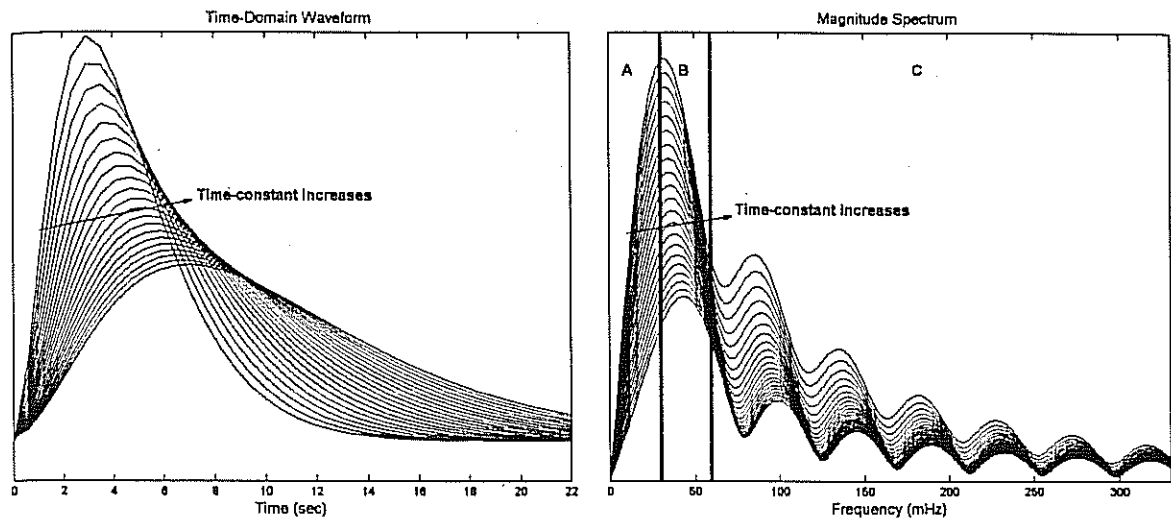


Figure 4. A family of Gamma waveforms with various time-constants (left) and their magnitude spectra (right). (The spectral sidelobes are an artifact due to centering operation).

blood volume, which are eventually reflected in the intracranial pressure" (Kuo et al., 1998). Similarly, in many fMRI studies, the Gamma function model has been used to parameterize the brain hemodynamic response. In Fig. 4, we display a family of Gamma waveforms (see Eq. (13)) with different time-constants, their magnitude spectra, on which we have superimposed the limits of the *B*-band. Since the brain hemodynamic model response, e.g. the "centered" Gamma function, has its peak located within the 30–50 mHz range (see Fig. 4), the narrow *B*-band should be related to task-related cognitive activity of the subject (Cordes et al., 2001; Toronov et al., 2000). The larger *C*-band is also assumed to carry cognitive activity related information, most probably due to the periodicity of the target stimuli sequence as we will explore more in detail in the next section. Moreover, vasomotion and breathing rate are two other physiological factors that are responsible for the emergence of the *C*-band. The *C*-band in our findings is compatible with the components of cerebral autoregulation identified by others (Cordes et al., 2001; Zhang et al., 1998; Strik et al., 2002; Giller et al., 1999).

When we consider the three bands found by our algorithm and the spectra of the BHR model as in Fig. 4, we observe that most of the energy of the Gamma waveform resides in the *B* band, followed, in order, by the bands *A* and *C*. It is of interest to consider the cross-correlations between the *R*-series within the bands. To this end, we consider the relative power time series

$R_n^m(t)$, $t = 1, \dots, T$, where m ranges over the M measurements, but n now ranges over the 3 canonical bands *A*, *B*, *C*; hence for band *A* we will use the notation $R_A^m(t)$ and similarly for the others. Cross-correlations between these series are given in matrix form below:

$$\begin{bmatrix} & A & B & C \\ A & 1 & -0.65 & -0.79 \\ B & -0.65 & 1 & 0.21 \\ C & -0.79 & 0.21 & 1 \end{bmatrix}$$

We can observe that there exists a negative correlation (correlation coefficient of minus 0.79) between the *A*-band and *C*-band. This implies that an increase in power in one band (say *A*) causes depletion of power in the other band (*C*), and vice versa. In this sense the time series $R_A^m(t)$ and $R_C^m(t)$ are almost "antipodal". A representative triad of *R*-series in these bands is shown in Fig. 5, where the antipodality can be observed when curves in the *A* and *C* bands are superposed. The correlation coefficients between these time-series suggest a negative feedback control between the cerebral autoregulatory activities (*C*-band) and the slow frequency components (*A*-band). One conjectures that the *A*-band signals are generated partly by the cognitive activity and partly by oscillations of the cerebral spinal fluid (CSF), which has been associated with intracranial pressure oscillations in the 8–40 mHz band and are dubbed as the *B*-waves (Strik et al., 2002).

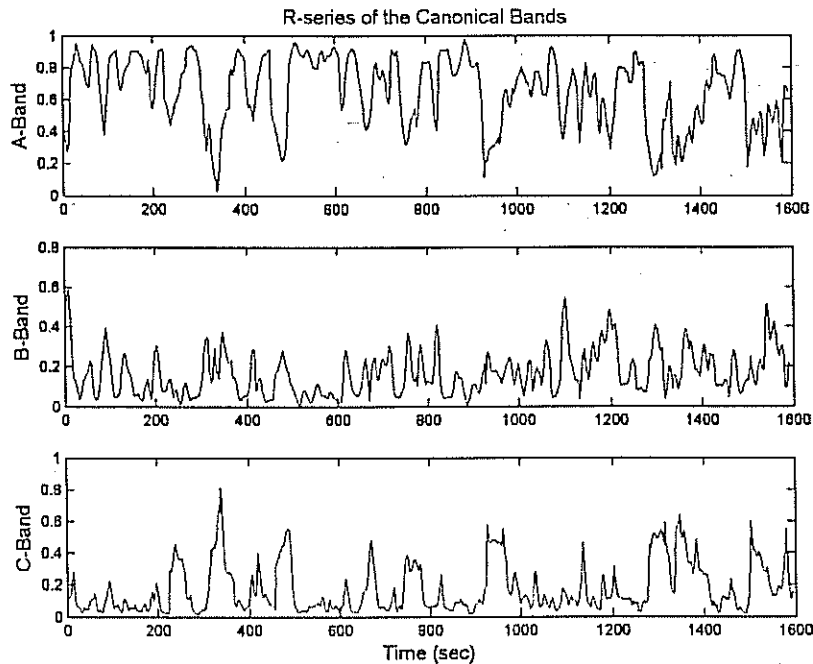


Figure 5. A representative set of R -series in the canonical bands.

Due to the probe geometry and physics of photon migration in tissues, it is most likely that the detectors are picking up a weighted sum of all the hemodynamic activities lying underneath each detector, namely the skin dynamics, cranium and CSF dynamics, and lastly, the cortex dynamics.

4.2. Search for Periodicities

Since the cognitive stimuli are quasi-periodic, whose target presentation intervals are uniformly distributed between 18–29 seconds, we can expect some sort of periodic behavior in the cognitive activity signals. The frequency bands in which such periodic response emerges more distinctly can be thought to better reflect cognitive activity or the “brain hemodynamic response.” Recall that the brain displays continuous activity (baseline activity) patterns even in the absence of any cognitive task. The cognitive activity waveforms, if any, will be in general immersed in the baseline activity. In fact, cognitive activity responses are very difficult to discern by simple observation of the waveforms. It follows that classical Fourier spectrum, correlation and peak picking techniques are not suitable for hunting these responses and more sophisticated statistical

methods must be invoked to detect and estimate these hidden periodicities (Toronov et al., 2000).

We analyze the bands $\{A, B, C\}$ and search for periodicities. The admissible periods should be in the vicinity of target exposition periods of the experimental protocol, that is, in the 18–29 seconds range or its harmonic/subharmonic multiples. Recall that, any periodicity in the time domain will emerge as discreteness (line spectrum) in the frequency domain. Thus, if a time-domain signal is periodic with P_0 (18 to 29) seconds, the corresponding spectrum should exhibit spectral samples $1/P_0$ Hz (34 to 59 mHz) apart. Since the B -band can at most accommodate one such spectral line, we merge the two bands B and C , to cover the 30–330 mHz band range. Previous researchers (Kuo et al., 1998; Giller et al., 1999; Hu et al., 1999) have also excluded the A -band since it was mostly mirroring the fNIRS baseline activity. Another plausible argument for excluding the A -band is that it is inherently a nonstationary process, which obviates signal-processing tools requiring stationarity. To this effect, we used a zero-phase finite impulse response (FIR) high-pass filter with unit gain and a 3 dB transition bandwidth of 1 mHz at 30 mHz. We will denote the corresponding band-pass filtered signals for simplicity as $x(t) = s_{BC}(t)$.

The periodicity measure we adopted is based on a classical method to estimate the pitch period in speech signals: the least-square periodicity estimation (LSPE) (Friedman, 1977). It is simply based on the minimization of the weighted mean-squared error (MSE) between the observed signal $x(t)$ and an estimated signal $x_0(t)$ that satisfies $x_0(t) = x_0(t + kP_0)$, $t = 1, 2, \dots, T$ and $k = 0, 1, \dots, K = \lfloor T/P_0 \rfloor - 1$ where $\lfloor \cdot \rfloor$ denotes the floor operation. The optimal $x_0(t)$, for a given P_0 , is

$$x_0(t) = \frac{\sum_k x(t + kP_0)w(t + kP_0)}{\sum_k w(t + kP_0)} \quad (7)$$

where $w(t)$ is the weight sequence of length T . Observe that (7) reduces to the following if all signal samples are equally weighted, i.e. $w(t)$ is a rectangular window,

$$x_0(t) = \frac{1}{K} \sum_k x(t + kP_0) \quad (8)$$

It has been argued that the weight sequence should be selected so that it has the maximum value of unity at the center of its support and that it decays smoothly down to zero towards the extremes since the period deviates more heavily at the extremes than at the center. It has been shown that P_0 that minimizes the weighted MSE is, equivalently, the one that maximizes the functional:

$$J_1(P_0) = \frac{I_0 - I_1}{E - I_1} \quad (9)$$

where I_0 stands for the weighted energy of the estimate $x_0(t)$ and E for the weighted energy of the original signal $x(t)$, and where $I_1(P_0) = \sum_t \frac{x^2(t)w^2(t)}{\sum_k w(t+kP_0)}$. Note that the LSPE with J_1 -index is also called as the pseudo-maximum likelihood estimation of periodicities (Friedman, 1977).

In (9), we search for the dominant period in the signal, \hat{P}_0 , that maximizes the $J_1(P_0)$ functional provided the periodicity index $J_1(P_0)$ is adequately high. In fact, the index function can be interpreted as a confidence score that becomes 1 only for a truly periodic signal. Since some maximizing value of \hat{P}_0 can always be found, for this estimate to correspond to a genuine periodicity, the confidence score should exceed a minimum threshold. Based on the protocol parameters we allow \hat{P}_0 to take values between P_{\min} and P_{\max} . Since the cognitive stimuli are not exactly periodic and since, furthermore, the cognitive activity signals are heavily embedded in baseline signals, we do not expect $J_1(\hat{P}_0)$

scores to be high, and hence we avoid heavy thresholding. A final confounding factor to be taken into account is that the cognitive response may not be fired exactly after the target presentation, but some variable delay may occur.

In order to illustrate the viability of the LSPE algorithm, we use a simulated data sequence that consists of the superposition of a Gamma waveform train embedded in white noise (the signal-to-noise ratio is 10 dB) and of an actual signal copied from the A-band. The average period of Gamma waveforms is 40 samples with a uniform random jitter between $(-10, 10)$ samples in order to simulate our experimental protocol (see Fig. 6).

Figure 7 illustrates the advantage of filtering out the A-band signal and to run the periodicity detector only in the BC band. It can be seen that the index values become higher when the band A is removed from the signal.

The results of the periodicity estimation on real data are presented in Fig. 8. In this particular case without prefiltering (ABC bands), the periodicity could not even be detected with any reasonable confidence, while in the filtered case (BC bands) we observe that a periodicity is detected in the expected range with acceptable confidence (Fig. 8).

In order to prospect the actual fNIRS-HbO₂ data for periodicities, we run the LSPE algorithm session by session, since the experimental protocol consists of eight identical sessions in succession. That is, we consider each of the eight sessions of the experiment separately and apply Eqs. (7)–(9). We use for each session a superscript $j = 1, \dots, 8$, so that $\{x^j(t)\}$ denotes the j th experimental session in the 30–330 mHz band-pass filtered fNIRS signal from whatever subject/detector. The session-wise processing of the fNIRS signal, together with prefiltering within the BC-band, helps also to mitigate the nonstationarity of these signals. In fact, we may view the signal portions in different sessions as independent realizations of the target categorization experiment if we neglect any correlation due to the baseline signal, which is independent of the cognitive activity. The hunting for periodicity maxima proceeds as follows:

- (i) The range in which we search for periodicities is (20, 60) samples.
- (ii) We look for local maxima of the $J_1(\hat{P}_0)$ functional; furthermore, once a peak is found, no further peak is searched within its neighborhood of $(-3, 3)$ samples.

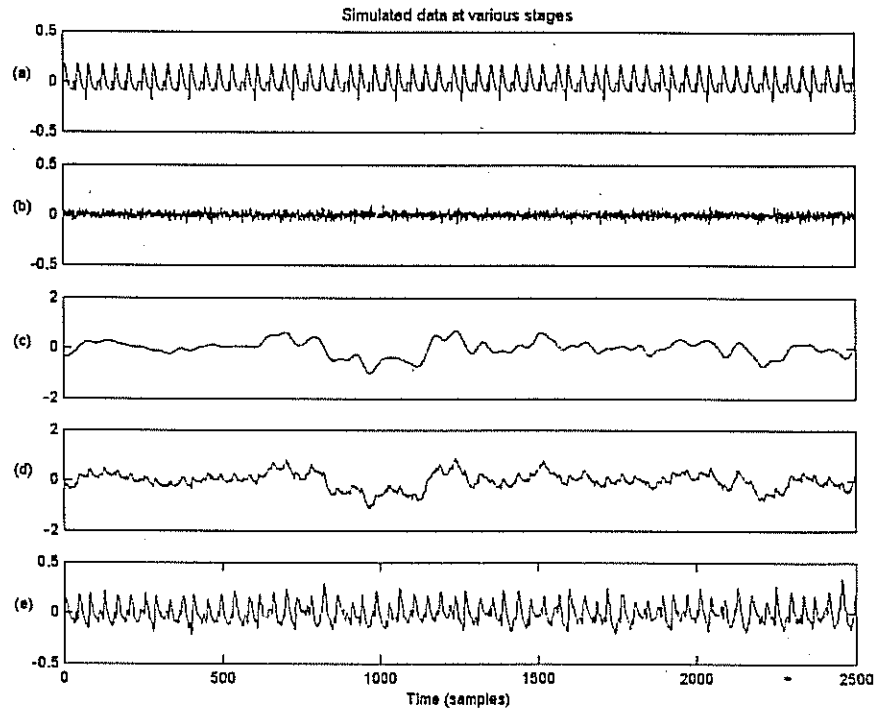


Figure 6. (a) Simulated quasi-periodic sequence of cognitive activity waveform; (b) white noise sequence (SNR = 10 dB); (c) an actual A-band signal; (d) superposition of the signals in (a), (b) and (c); (e) Band-pass filtered version of (d) in the BC-band: this signal is then used for periodicity detection.

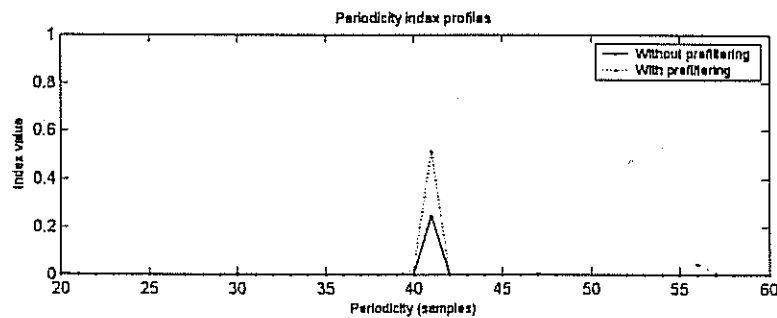


Figure 7. Periodicity index profiles for simulated data without prefiltering (solid line) and with prefiltering (dotted line), after local maxima selection and thresholding.

- (iii) We set a threshold of 0.1 on the periodicity belief value $J_1(\hat{P}_0)$.
- (iv) For each signal portion, we let the algorithm return the periodicity estimate with largest J_1 -index.

Those \hat{P}_0 values that fall within the (30, 50) samples interval are thought to belong to the single-trial cognitive activity in the experiment. Those falling outside are considered as fortuitous values, indicative of the fact

that that detector is not capturing properly any cognitive activity signal during that session. Since there are 8 session signals $x^j(t)$ per detector (each in turn possessing 8 target stimuli), each $\{x(t)\}$ signal returns eight period estimates, \hat{P}_0^j , $j = 1, \dots, 8$ along with their confidence scores. We accumulate separately the scores of the periodicities falling, respectively, inside and outside the expected range. We define the cumulative score of inside periodicities S_{in} and the count of

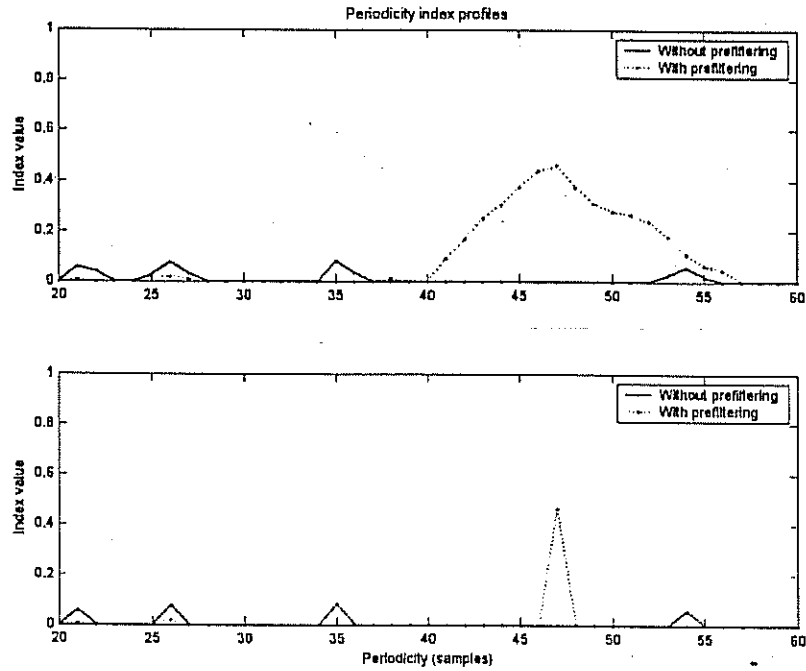


Figure 8. Periodicity index profiles with (dotted line) and without (solid line) prefiltering (top), after local maxima selection and thresholding (bottom): thanks to prefiltering out band A, a periodicity with a good confidence emerges within the expected range.

inside periodicities C_{in} for a given detector and target categorization experiment as:

$$S_{in} = \sum_{j=1}^8 J_1(\hat{P}_0^j) \delta(\hat{P}_0^j) \quad (10)$$

$$C_{in} = \sum_{j=1}^8 \delta(\hat{P}_0^j) \quad (11)$$

with

$$\delta(\hat{P}_0^j) = \begin{cases} 1 & \text{if } \hat{P}_0^j \text{ is inside the expected range} \\ 0 & \text{if } \hat{P}_0^j \text{ is outside the expected range} \end{cases} \quad (12)$$

where $j = 1, \dots, 8$ is the session index. Similarly we define corresponding expressions for the outside periodicities S_{out} and C_{out} . Furthermore, in order to investigate intersubject and interdetector variability of periodicities, we compute two quantities: periodicities falling in the proper range averaged over all subjects for a given photodetector, denoted as $\bar{P}_{subject}(k)$, $k = 1, \dots, 16$, and inside periodicity averaged over all photodetectors for a given subject, denoted as $\bar{P}_{detector}(l)$, $l = 1, \dots, 5$. The error bar plots corresponding to these two quantities, $\bar{P}_{subject}(k)$, $k = 1, \dots, 16$ and $\bar{P}_{detector}(l)$, $l =$

$1, \dots, 5$, are displayed in Fig. 9. We also show the bar plots of the scores and the scatter plots of the detected periodicities with respect to the photodetector number in Fig. 9.

Several conclusions can be drawn from these results:

- (i) The averaged period estimates match the expected value of P_0 , whether the average is computed over detectors or subjects, as illustrated in Fig. 9.
- (ii) For any detector or subject there is some dispersion of estimated periodicity values. The large spread, of the order of 10% in each sense, may be due to the jitter of target instances, to the presence of context activity, and to the limited observation interval containing, at most, eight target stimuli.
- (iii) We have also a method to classify detectors as responsive of BHR periodicity and the non-responsive ones, that is, those detectors that do not show any evidence of periodicity within the expected range. The discrimination method is based on the not-in-the-range periodicity score, as illustrated in Fig. 10, over subjects 1 to 5. The reason why some detectors do not yield conjectured periodicity could be due to the lateralization effect for

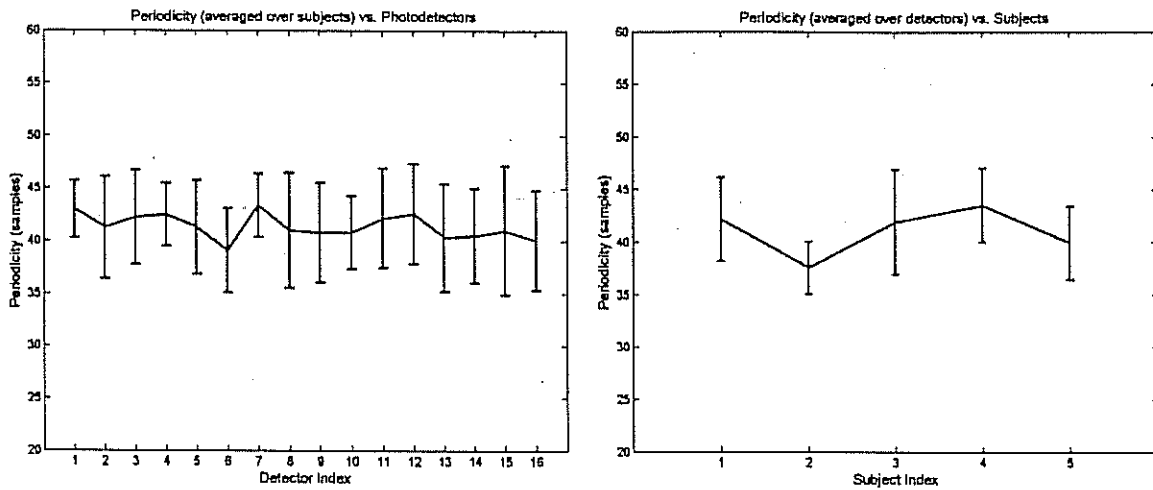


Figure 9. Plots of \bar{P}_{subject} (left) and $\bar{P}_{\text{detector}}$ (right), the bars indicate the inter-quartile range of data, computed over admissible (that is, within-range) cases.

that subject or simply to corrupted measurements. One argument that supports lateralization conjecture is that adjacent detectors all succeed or fail in a group. For example, for Subject 1, detectors 3–12 are “good”, while detectors 1–2 and 13–16 are “bad” (see Table 4). One other reason could be due to the location of the optodes with respect to the light sources. Since the source-detector distribution determines the volume of brain being sampled, millimeter range shifts and alignments in the probe will result in a significant change in the brain volume being monitored. Finally, the corruption observed could be due to the skin effects (larger arteries on the skin surface right underneath the optode) dominating the signal.

- (iv) There are also marked differences between subjects. For example, Subjects 1, 3 and 4 (especially Subject 4) yield high periodicity scores consis-

tent across all their detectors, while Subjects 2 and 5 are dubious. Although inter-subject variability is always expected in such studies, there is no standard procedure to isolate corrupted data from statistical analysis for fNIRS signals. The periodicity analysis method provided in this paper might be used as a rule of thumb in identifying the corrupted data or the patient that is not cooperating.

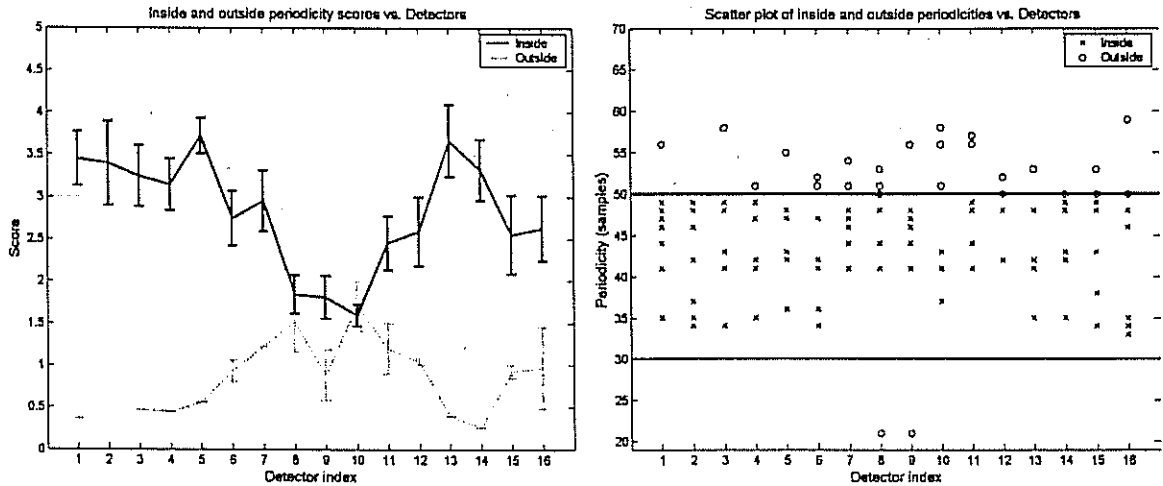
4.3. Correlation Analysis

We have seen in Section 4.2 that fNIRS possesses evidence of cognitive activity as reflected in the periodical patterns of target stimuli. We can further explore such activity by correlating with our data a brain hemodynamic response model, as borrowed from fMRI

Table 4. Photodetectors with S_{in} higher than S_{out} .

Subject	Photodetector quadruples			
	Left (1–4)	Mid-left (5–8)	Mid-right (9–12)	Right (13–16)
1	3 and 4	5 to 8 (all)	10, 11 and 12	16
2	-not any-	8	9, 11 and 12	13 to 16 (all)
3	4	5 to 8 (all)	9 to 12 (all)	15 and 16
4	1 to 4 (all)	5 to 8 (all)	9, 11 and 12	13 to 16 (all)
5	1 to 4 (all)	5 and 7	9, 11 and 12	13 to 16 (all)

Subject 4



Subject 5

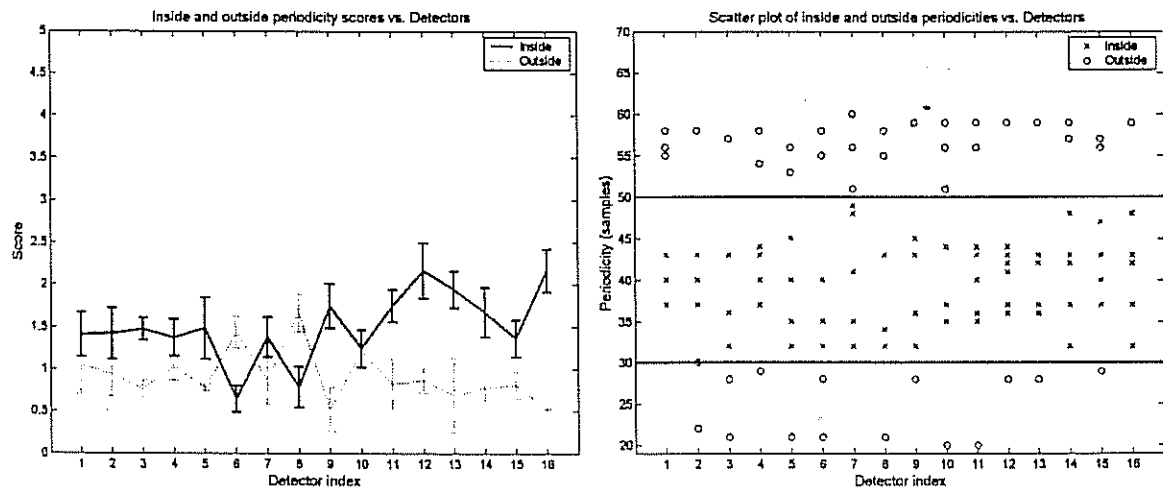


Figure 10. Bar plots of the scores and the scatter plots of the detected periodicities with respect to photodetectors. On the left, the bars indicate standard deviation of data. On the right, thick black bars enclose the expected range of periodicities. In the presented cases, good periodicity is observed in Subject 4 while lesser periodicity exists in Subject 5.

analysis. One popular model is the Gamma function expressed as

$$h(t) = \begin{cases} A(t - T)^2 e^{-(t-T)/\tau} & \text{for } t \geq T \\ 0 & \text{for } t < T \end{cases} \quad (13)$$

where τ is the time-constant that characterizes the response, A is the strength, and T is the delay of the response to the target stimulus. Let's denote the sampled Gamma waveform by the m -dimensional vector \mathbf{h} (where $m = 40$) sampled at the instants

$t = 0, T_s, \dots, (m - 1)T_s, T_s$ being the sampling period (1.7 Hz in the experiment). Implicit in each \mathbf{h} vector is the set of parameters τ and T , where τ is the time-constant that characterizes the response, A is the strength, and T is the delay of the response to the target stimulus. Let's denote the sampled Gamma waveform by the m -dimensional vector \mathbf{h} , constituted of the samples of $h(t)$ in (13) at the instants $t = 0, T_s, \dots, (m - 1)T_s$ and $m = 40$. Consider the fNIRS signal $x(t)$, detrended and, in addition, possibly prefiltered to leave the BC-band, i.e., 30–330 mHz.

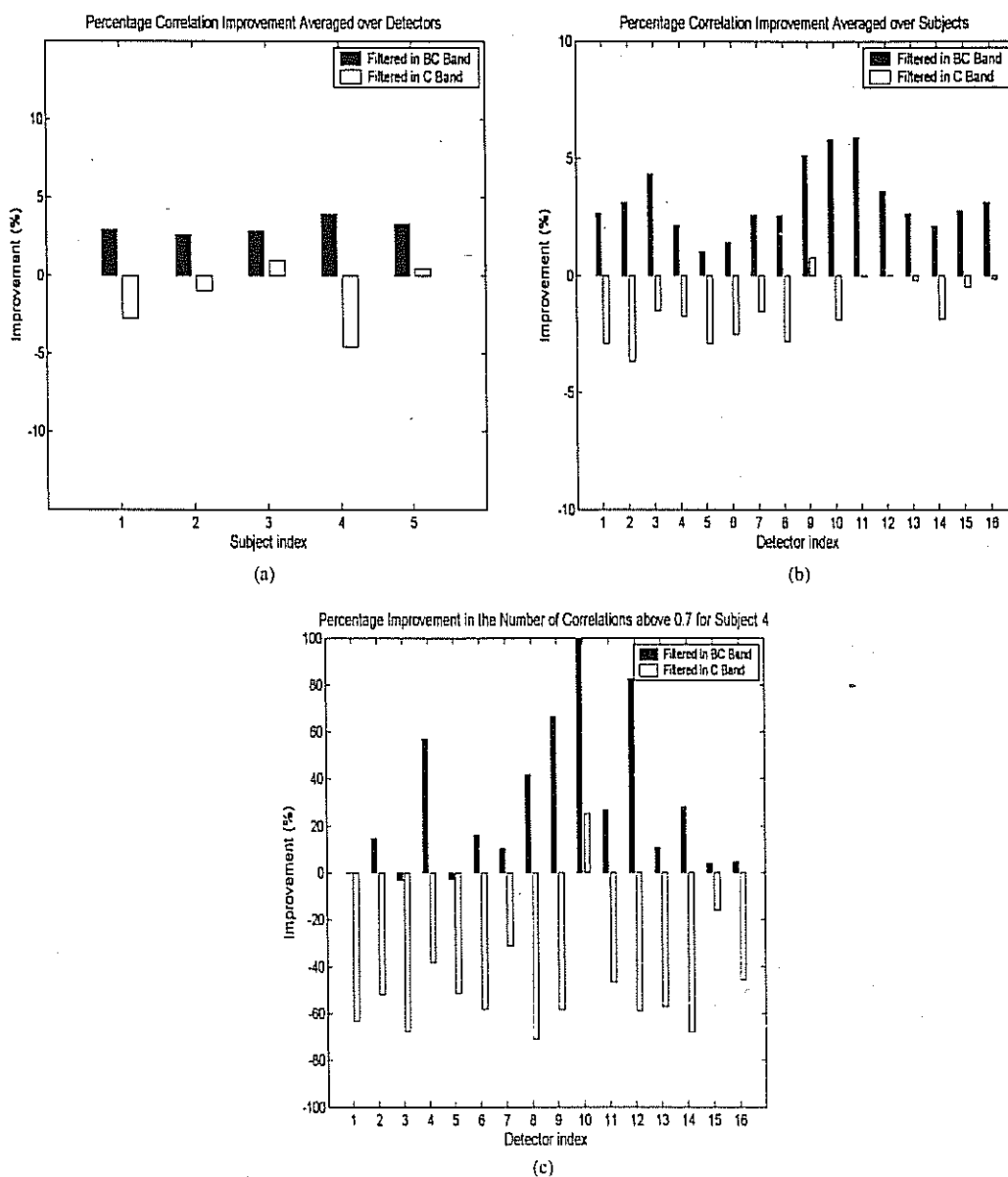


Figure 11. Plots of percentage improvements in the correlation scores with Gamma response model (clockwise): (a) Improvement averaged over all sessions and detectors; (b) Improvements averaged over all sessions and subjects, and (c) Improvement for Subject 4 averaged over all sessions.

- (i) Put into a vector form the 40 samples after each target presentation instant to form one of the z_k vectors. Recall that there are 64 stimuli per experiment, hence the vector index runs as $k = 1, \dots, 64$.
- (ii) Constrain the time-constant τ of the Gamma waveform in the 1.5–3.5 second range, and the response delay T in the range 0–10 samples. We

search exhaustively for the best match in this parameter space.

- (iii) For each $x(k)$, find the best match that maximizes the normalized correlation coefficient between $x(k)$ and h , i.e.,

$$\rho(k) = \max_{\tau, T} (h, z_k)$$

where the samples of z_k that precede the delay T are set to zero.

- (iv) If any prefiltering is used, say we focus on bands B and C , then we denote the correlation as $\rho(k) \equiv \rho_{BC}(k)$, and similarly for other band choices.
- (v) The $\rho(k)$, $k = 1, \dots, 64$ series for all detectors and subjects are pruned out, by deleting low-correlation sessions, such that we keep only those correlations that exceed the threshold of 0.7. Furthermore we calculate the mean of the correlations per detector, $m_\rho(l)$, where the detector index runs over $l = 1, \dots, 16$.

We have computed the mean correlation profiles for the original signals (no prefiltering) and for signals pre-filtered in the BC (30–330 mHz) and C (60–330 mHz) bands. The prefiltering of the AB band (0–60 mHz) was specifically avoided to preclude any resonance effect with the Gamma response model. In Fig. 10, we display some of these mean correlation profiles and bar plots for cases where correlations exceed the threshold of 0.7. As to be expected in the light of the assumed Gamma response, we observe that:

Prefiltering helps in general to improve the correlation scores. In fact, the scores have definitely increased in the BC band, whether considered over subjects or detectors. On the other hand, excluding the B band, that is considering only the C band, we find much lower scores. Another useful note is that the inter-detector variability in the number of encountered Gamma matches decreases in the BC -band. These observations point to the fact that the cognitive activity is mainly localized above 0.03 Hz, that is, in the 30–300 mHz range (Toronov et al., 2000; Prince et al., 2003). The results are illustrated in Fig. 11, where we contrast the percentage change in the number of sessions exceeding the correlation threshold of 0.7 when 30–330 mHz BC -band is considered vis-à-vis the 60–330 C -band or the full ABC -band. One can observe that the correlation scores improve whether observed over subjects or detectors by 10 to 50%.

5. Conclusions

Identification of fNIRS frequency bands where cognitive activity predominates in single-event trials is an important problem. Based on the clustering tendency of time-frequency power spectra, we have determined that there exist three bands of interest: A -band (0–30 mHz), B -band (30–60 mHz) and C -band (60–330 mHz). The

A -band is conjectured to represent the context activity and some cognitive activity. The B -band is predominantly cognitive-activity related, while the C -band accounts for various physiological activities as well as protocol-induced periodicity. Our analysis has covered up to 330 mHz range since this was the cut-off frequency of the built-in RC-filter in the fNIRS device. These bands are found to be very similar to those in related studies (Hu et al., 1999; Kuo et al., 1998; Obrig et al., 2000a; Franceschini et al., 2000; Toronov et al., 2000).

Correlation between the response model waveforms (Gamma waveform) and the fNIRS signals becomes maximum in the joint AB -band (0–60 mHz). On the other hand, protocol-induced periodicity is best observed in the BC -band (60–330 mHz). Since each subject was tested only once, we cannot comment on intra subject variability. We believe that evidence of periodicity in detectors provides a guideline for selecting responsive photodetectors and subjects and discarding non-responsive ones.

Acknowledgments

This work has been sponsored in part by funds from the Defense Advanced Research Projects Agency (DARPA) Augmented Cognition Program and the Office of Naval Research (ONR), under agreement numbers N00014-02-1-0524 and N00014-01-1-0986, and in part by Boğaziçi University Research Fund, BURF 02S101 and Tübitak Project 102E027.

References

- Boas DA, Gaudette T, Strangman G, Cheng X, Marota JJA, Mandeville JB (2001). The accuracy of near infrared spectroscopy and imaging during focal changes in cerebral hemodynamics. *NeuroImage* 13: 76–90.
- Blanco S, Garcia H, Quiroga RQ, Romanelli L, Rosso OA (1995) Stationarity of the EEG Series. *IEEE Engineering in Medicine and Biology Magazine* 14: 395–399.
- Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human v1. *J. Neuroscience* 16: 4207–4221.
- Chance B, Anday E, Nioka S, Zhou S, Long H, Worden K, Li C, Turray T, Ovetsky Y, Pidikiti D, Thomas R (1998) A novel method for fast imaging of brain function, noninvasively, with light. *Optics Express* 2: 411–423.
- Cordes D, Haughton VM, Arfanakis K, Carew JD, Turski PA, Moritz CH, Quigley MA, Meyerand ME (2001) Frequencies contributing to functional connectivity in the cerebral cortex in resting-state data. *Am. J. Neuroradiol.* 22: 1326–1333.

- Firbank M, Okada E, Delpy DT (1998) A theoretical study of the signal contribution of regions of the adult head to near-infrared spectroscopy studies of visual evoked responses. *Neuro Image* 8: 69-78.
- Franceschini MA, Fantini S, Toronov V, Filiaci ME, Gratton E (2000) Cerebral hemodynamics measured by near-infrared spectroscopy at rest and during motor activation. *Proc of the Optical Society of America In Vivo Optical Imaging Workshop Washington*. Optical Society of America, DC, pp. 73-80.
- Friedman DH (1977) Pseudo-maximum-likelihood speech pitch extraction. *IEEE Trans. on ASSP* 25: 213-221.
- Giller AC, Hatab MR, Giller AM (1999) Oscillations in cerebral blood flow detected with a transcranial doppler index. *J. Cereb. Blood Flow Metab.* 19: 452-459.
- Hu H, Kuo T, Wong W, Luk Y, Chern C, Hsu L, Sheng W (1999) Transfer function analysis of cerebral hemodynamics in patients with carotid stenosis. *J. Cereb. Blood Flow Metab.* 19: 460-465.
- Kim D, Ugurbil K (2002) Bridging the gap between neuroimaging and neuronal physiology. *Image Anal. Stereol.* 21: 97-105.
- Kuo T, Chern C, Sheng W, Wong W, Hu H (1998) Frequency domain analysis of cerebral blood flow velocity and its correlation with arterial blood pressure. *J. Cereb. Blood Flow Metab.* 18: 311-318.
- Mayhew J, Zheng Y, Hou Y, Vuksanovic B, Berwick J, Askew S, Coffey P (1999) Spectroscopic analysis of changes in remitted illumination: The response to increased neural activity in brain. *Neuroimage* 10: 304-326.
- McCarthy G, Luby M, Gore J, Goldman-Rakic P (1997) Infrequent events transiently activate human prefrontal and parietal cortex as measured by functional MRI. *J. Neurophysiology* 77: 1630-1634.
- Obrig H, Neufang M, Wenzel R, Kohl M, Steinbrink J, Einhaupl K, Villringer A (2000a) Spontaneous low frequency oscillations of cerebral hemodynamics and metabolism in human adults. *Neuroimage* 12: 623-639.
- Obrig H, Wenzel R, Kohl M, Horst S, Wobst P, Steinbrink J, Thomas F, Villringer A (2000b) Near-infrared spectroscopy: Does it function in functional activation studies of the adult brain? *Intr. J. Psychophysiol.* 35: 125-142.
- Prince S, Kolehtainen V, Kaipio JP, Franceschini MA, Boas David, Arridge SR (2003) Time series estimation of biological factors in optical diffusion tomography. *Phys. Med. Biol.* 48: 1491-1504.
- Stoica P, Moses RL (1997) *Introduction to spectral analysis*. Prentice Hall, NJ, p. 43.
- Strangman G, Culver JP, Thompson JH, Boas DA (2002) A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation. *Neuroimage* 17: 719-731.
- Strik C, Klose U, Erb M, Strik H, Grodd W (2002) Intracranial oscillations of cerebrospinal fluid and blood flows: Analysis with magnetic resonance imaging. *J. Magn. Res. Imag.* 15: 251-258.
- Svensen M, Kruggel F, von Cramon DY (2000) Probabilistic modeling of single-trial fMRI data. *IEEE Trans. on Med. Imaging* 19: 25-35.
- Toronov V, Franceschini MA, Filiaci ME, Wolf M, Fantini S, Gratton E (2000) Near-infrared study of fluctuations in cerebral hemodynamics during rest and motor stimulation: Spatial mapping and temporal analysis. *Med. Phys.* 27: 801-815.
- Villringer A, Chance B (1997) Non-invasive optical spectroscopy and imaging of human brain function. *Trends in Neuroscience* 20: 4435-4442.
- Wobst P, Wenzel RM, Kohl HO, Villringer A (2001) Linear aspects of changes in deoxygenated hemoglobin concentration and cytochrome oxidase oxidation during brain activation. *Neuroimage* 13: 520-530.
- Zhang R, Zuckerman JH, Giller CA, Levine BD (1998) Transfer function analysis of dynamic cerebral autoregulation in humans. *Am. J. Physiol. (Heart Circ. Physiol.)* 274: H233-H241.

Independent Factor Analysis for Estimating Fading Channels in CDMA Communication

M.A. Altinkaya*, O. Kalkan* and E.E. Kuruoğlu**

*Izmir Institute of Technology, Department of Electrical & Electronics Engineering, Turkey

**Istituto di Scienza e Tecnologie dell' Informazione Consiglio Nazionale delle Ricerche, Italy

Abstract

In a CDMA downlink scenario, the signals coming from the base station are a mixture of a desired signal for a specific user and other signals corresponding to other users of the system. Since the desired user usually cannot know the spreading code sequences of other users this can be considered as a blind source separation (BSS) problem. In this work, we applied independent factor analysis (IFA) which offers a flexible analytic source modelling by mixtures of Gaussians and which also accounts for the additive noise in solving this BSS problem. The simulation studies with the basic CDMA downlink model and with a fast fading CDMA channel showed the superiority of IFA with respect to regular independent component analysis (ICA). We will also utilize the factorized variational approximation method in IFA to mitigate the computational complexity which is exponential in the product of number of users and the number of multipaths considering the fact that in a typical CDMA system the number of users is a multiple of tens and the number of multipaths is at least three.

Keywords: independent factor analysis, independent component analysis, code division multiple access, blind source separation

1. Introduction

Code Division Multiple Access (CDMA) is a multiple access technique that uses spread spectrum modulation by each user with its own unique spreading code where all users share the same spectrum [1]. The increasing demands in wireless personal and mobile communication systems both to provide and accommodate high quality voice services and other multirate services such as internet access from hand-held mobile terminals, also increased the interest in CDMA because it provides high-frequency usage and it is suitable for multimedia and multirate services.

The easiest way of detecting the desired signal in CDMA communication is a single user detection scheme where the signals corresponding to other users are considered as white noise just contributing to the noise level presumably. However, a multiuser detection (MUD) scheme where also the multiple access interference caused by the signals of other users are tried to be eliminated in the detection process, will result in better performance. In such a CDMA system, specifically in a CDMA downlink scenario, the signals coming from the base station are a mixture of a desired signal for a specific user and other signals corresponding to other users of the system. Since the desired user usually cannot know the spreading code sequences of other users this can be considered as a blind source separation (BSS) problem.

In BSS problem, data measured by sensors arise from source signals that are mixed together by some linear transformation corrupted by noise. The task is to obtain those source signals. However, the sources are not observable and there is no information about their properties beyond their statistical independence nor about the mixing process and the noise [2]. These sources can be found using independent component analysis (ICA) provided that all sources are independent and non-Gaussian [3]. Most of the work in the field of BSS with ICA considered the idealized noise-free case until a satisfactory solution was found in the mid 90's. [4,5]. In real applications we usually encounter noise and the number of the sources differ from the number of sensors which fact is again not solved by ICA methodology. Although there have been some efforts to include noise into the analysis [6,7] we observed some deteriorating performance when the noise level increases.

Independent factor analysis (IFA) proposed by Attias [2] and Moulines et al. [8] solve the noisy BSS problem which can also model the CDMA downlink communication. Additionally, IFA can handle the mismatch in the numbers of sources and sensors. The sources are modelled as mixtures of Gaussians (MOG) whose parameters are estimated jointly with the mixing matrix. The

model also contains the noise covariance matrix which is another parameter to be estimated.

In this paper we use the IFA model to solve the blind MUD problem of CDMA downlink communication. The first results that we present in this summary includes application of IFA to a basic CDMA downlink model with comparison to existing methods and estimating the fading channels which may have an impulsive nature. A problem of using IFA method in CDMA systems for BSS arises in the computational complexity which is exponential in the product of number of users and the number of multipaths considering the fact that in a typical CDMA system the number of users is a multiple of tens and the number of multipaths is at least three. In order to solve this problem of IFA for CDMA systems, we also propose applying the factorized variational approximation suggested by Attias. The suggested solution utilizes a posterior density different from the exact posterior which is used in the E-step of the EM algorithm. This posterior has it's own set of parameters which are learned separately from the generative parameters of the IF model and makes the E-step easier providing a reasonable approximation of the exact posterior.

2. Independent Factor Analysis (IFA)

IFA is a new algorithm for blind separation of noisy mixtures with a possibly non-square mixing matrix [2, 8]. It is performed in two steps:

- Learning the IF model that is the mixing matrix, the noise covariance and source density parameters from the data where each source density is modeled by a mixture of one-dimensional Gaussian density.
- Recovering the sources from the sensor signals using the posterior density of the sources given the data.

The IFA model is a linear probabilistic model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{u}$$

where \mathbf{y} and \mathbf{u} are the $L' \times 1$ observation and noise vectors, respectively, \mathbf{H} is the $L' \times L$ mixing matrix and \mathbf{x} is the $L \times 1$ source vector. Each source density is modeled by a mixture of one-dimensional Gaussian (MOG) density given by:

$$p(x_i | \theta_i) = \sum_{q_i=1}^{n_i} w_{i,q_i} \zeta(x_i - \mu_{i,q_i}, \nu_{i,q_i}) \quad \theta_i = \{w_{i,q_i}, \mu_{i,q_i}, \nu_{i,q_i}\}$$

and the noise is Gaussian with zero mean and covariance Λ : $p(\mathbf{u}) = \zeta(\mathbf{u}, \Lambda)$.

The resulting model sensor density becomes:

$$\begin{aligned} p(\mathbf{y} | \mathbf{W}) &= \int p(\mathbf{y} | \mathbf{x}) p(\mathbf{x}) d\mathbf{x} \\ &= \int G(\mathbf{y} - \mathbf{H}\mathbf{x}, \Lambda) \prod_{i=1}^L p(x_i | \theta_i) d\mathbf{x} \end{aligned}$$

where $\mathbf{W} = (\mathbf{H}, \Lambda, \theta)$ denotes the collective parameter set. \mathbf{W} parameters are found iteratively by minimizing the Kullback-Leibler (KL) divergence of the true density and the density conditioned on the observation given by:

$$\begin{aligned} \xi(\mathbf{W}) &= \int p^\circ(\mathbf{y}) \log \frac{p^\circ(\mathbf{y})}{p(\mathbf{y} | \mathbf{W})} d\mathbf{y} \\ &= -E[\log p(\mathbf{y} | \mathbf{W})] - H_{p^\circ} \end{aligned}$$

For minimizing this error function the expectation maximization (EM) algorithm is used which is formed by the two steps:

- Forming the expected value of the complete-data likelihood, given the observed data and the current model:

$$F(\mathbf{W}', \mathbf{W}) = -E[\log p(\mathbf{q}, \mathbf{x}, \mathbf{y} | \mathbf{W})]$$
- Obtain the new parameters as the solution of

$$\mathbf{W} = \arg \min_{\mathbf{W}'} F(\mathbf{W}', \mathbf{W}'')$$

After the IF model parameters are estimated the sources can be reconstructed from the sensor signals using any parametric estimator such as the least mean squares or maximum a-posteriori probability estimators.

Next we will describe how the CDMA downlink problem fits to an IFA formulation.

3. CDMA Downlink Communication

The basic CDMA one shot downlink model is

given as $r(t) = \sum_{k=1}^K A_k b_k s_k(t) + \sigma n(t)$ where

$r(t)$ is the received signal, A_k is the k 'th users received signal amplitude, b_k is the antipodal data bit of the k 'th user, $s_k(t)$ is the k 'th user's signature waveform with normalized energy, σ^2 is the noise power with unit power spectral density. After chip matched filtering and

sampling the received signal \mathbf{r} can be expressed in $C \times 1$ vector form:

$$\mathbf{r} = \sum_{k=1}^K A_k b_k s_k + \sigma \mathbf{n}$$

where s_k is the $C \times 1$ signature vector of the k 'th user and \mathbf{n} is the $C \times 1$ AWGN vector. This equation can be written as:

$$\mathbf{r} = \mathbf{G}\mathbf{b} + \mathbf{n}$$

where $\mathbf{G}=\mathbf{S}\mathbf{A}$ is the $C \times K$ mixing matrix with $\mathbf{S}=[s_1 s_2 \dots s_K]$ and $\mathbf{A}=\text{diag}(A_1 \dots A_K)$ and \mathbf{b} is the $K \times 1$ source vector. Note that the above equation has the form of the IFA generative model. In the fast fading case the CDMA downlink model is given by

$$r(t) = \sum_{m=1}^N \sum_{k=1}^K b_{km} \sum_{l=1}^L a_{lm} s_k (t - mT - d_l) + n(t)$$

After chip matched filtering C -vectors \mathbf{r}_m of the form

$$\mathbf{r}_m = [r(mC) r(mC+1) \dots r((m+1)C-1)]^T$$

can be expressed as

$$\mathbf{r}_m = \sum_{k=1}^K \sum_{l=1}^L [a_{l,m-1} b_{k,m-1} \mathbf{g}_{kl}^E + a_{l,m} b_{k,m} \mathbf{g}_{kl}^L] + \mathbf{n}_m$$

where

$$\mathbf{g}_{kl}^E = [s_k(C-d_l+1) \dots s_k(C) \ 0 \dots 0]^T$$

and

$$\mathbf{g}_{kl}^L = [0 \dots 0 \ s_k(1) \dots s_k(C-d_l)]^T$$

represent the early and late parts of the code vectors. Thus, one obtains

$$\mathbf{R} = \mathbf{G}\mathbf{F} + \mathbf{N}$$

where $\mathbf{R} = [\mathbf{r}_1 \dots \mathbf{r}_N]$ represents N

transmitted symbols in compact form and

$\mathbf{F} = [\mathbf{f}_1 \dots \mathbf{f}_N]_{2KL \times N}$ contains the symbol

and fading terms:

$$\mathbf{f}_m = [a_{1,m-1} b_{1,m-1} \ a_{1,m} b_{1,m} \ \dots \ a_{L,m-1} b_{L,m-1} \ a_{L,m} b_{L,m}]$$

$$\mathbf{G} = [\mathbf{g}_{11}^E \ \mathbf{g}_{11}^L \ \dots \ \mathbf{g}_{KL}^E \ \mathbf{g}_{KL}^L]_{C \times 2KL}$$

contains the basis vectors and

$\mathbf{N} = [\mathbf{n}_1 \dots \mathbf{n}_N]$ is the noise matrix.

4. Results

We applied IFA to the basic CDMA model defined in the previous section. In the simulations, a CDMA downlink communication

that is the communication from the base station to a user with $K=2$ users and with $C=31$ Gold codes is considered. As the signal power of the desired user is fixed to 12 dB, the power of the interfering user is fixed to 30 dB. IFA, regularized ICA [9] and single user matched filter detectors are applied for 10 independent simulations and 1000 transmitted symbols. In Figure 1, the number of correctly estimated symbols are given for the three detectors. The figure depicts that IFA can estimate all transmitted symbols correctly even though the noise power is very high given by the variance = 0.7 and the desired user's signal power is low. On the other hand regularized ICA can estimate all symbols correctly when the signal power of the desired user is greater than 6 dB.

We also applied IFA to the fast fading CDMA model to estimate the fading coefficients. In the experiments, $K=2$ and $L=1$ are chosen to limit the number of sources to 4 to overcome the convergence problem of IFA. Fading coefficients are modelled by 1.85-stable distribution. In Figures 2 and 3, it is seen that IFA estimator tracks channel coefficients for both noise variances 0.1 and 0.01 successfully.

In the full paper, we will also present the results of a more practical CDMA scenario with multiple tens of users and multipath channels where the factorized variational approximation method is utilized to reduce computational complexity of IFA.

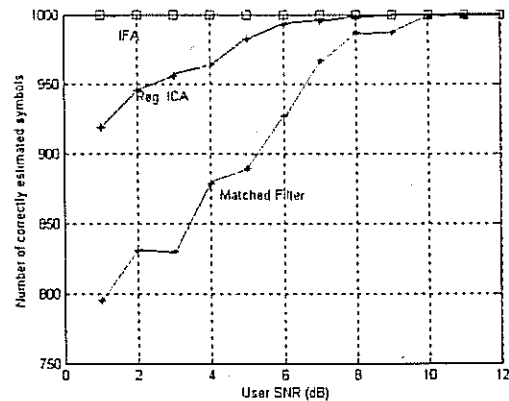


Figure 1 Number of correctly estimated symbols as a function of User SNR for IFA, Reg. ICA and MF detectors. ($K=2$, $C=31$ Gold codes)

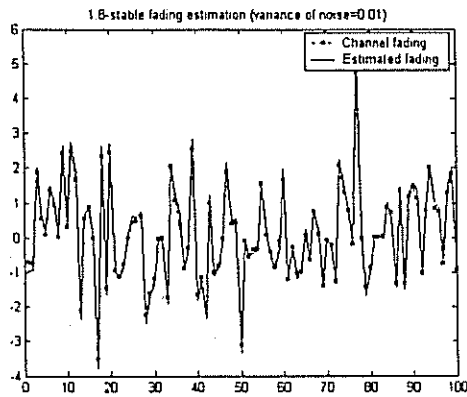


Figure 2 Fading channel estimation (variance of noise=0.01)

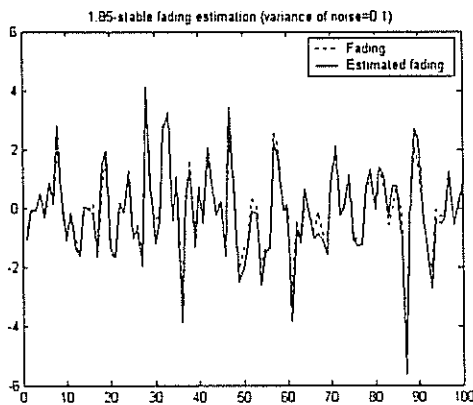


Figure 3 Fading channel estimation (variance of noise=0.1)

References

- [1] J.S. Lee and L.E. Miller, "CDMA Systems Engineering Handbook", Artech House: Boston, 1998.
- [2] H. Attias, "Independent Factor Analysis", *Neural Computation*, 11:803-855, 1999.
- [3] K.N. Leach, "A survey paper on independent component analysis", *System Theory*:pp.239-242, March 2002.
- [4] A.J. Bell and T.J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution", *Neural Computation*, 7:1129-1159, 1995.
- [5] J.F. Cardoso and B.H. Laheld, "Equivariant adaptive source separation", *IEEE Trans. on Signal Processing*, 44:3017-3030, 1996.
- [6] A. Hyvarinen, "Independent component analysis in the presence of Gaussian noise by maximizing joint likelihood", *Neurocomputing*, 22:49-67, 1998.
- [7] A. Hyvarinen, "Gaussian moments for noisy independent component analysis", *IEEE Signal*

Processing Letters, 6:145-147, 1999.

- [8] E. Moulines, J.F. Cardoso and E.Gassiat, "Maximum likelihood for blind separation and deconvolution of noisy signals using mixture models", *Proceedings of ICASSP'97*, 5:3617-3620, 1997.
- [9] K-H. Yap, L. Guan and J. Evans, "Blind adaptive detection for CDMA systems based on regularized independent component analysis", *GLOBECOM 2001*, 1:249-253, 2001.

İşlevsel Optik İşaretlerde Bilişsel Etkinliğin Varlığı

Evidence of Cognitive Activity in Functional Optical Signals

Ceyhun Burak Akgül¹, Bülent Sankur¹, Ata Akın²

¹ Elektrik ve Elektronik Mühendisliği Bölümü

² Biyomedikal Mühendisliği Enstitüsü

Boğaziçi Üniversitesi, İstanbul

{ceyhun.akgul, sankur, ataakin}@boun.edu.tr

Özetçe

Bu çalışmada, işlevsel yakın kızıl ötesi spektroskopisi işaretlerinin beyin bilişsel görevlerle hemodinamik tepkisini gözlemekte kullanılabileceği gösterilmektedir. İnsan deneklerin bir ekranda gösterilen hedefleri sınıfladığı deneylerde uyaran dizileri hemen hemen önemlidir. Önemlilikten sapmalara ve başka işaretlerin girişimlerine karşın, tüm foto-algılayıcılar ve denekler üzerinde incelenen işaretlerde tutarlı bir biçimde hemodinamik süreç döneminin kestirilebildiği görülmüştür.

Abstract

We show that functional near infrared spectroscopy signals can monitor task-related cerebral hemodynamic response. We detect periodicities in the near infrared spectroscopy signals and show that the estimated periods are consistent across multiple detectors and subjects, despite the jittered periodicity of the stimuli sequence that is used in a target categorization experiment.

1. Giriş

İşlevsel yakın kızıl ötesi spektroskopisi (fNIRS), bilişsel görevler sırasında beyin hemodinamik tepkisinin gözlemlenmesine elveren, kafatasının açılmasını gerektirmeyen, hızlı bir teknik olarak önerilmiştir [5]. Gerçekten de sinir hücrelerinin tepkilerinin, fNIRS yardımıyla ölçülen oksijensiz hemoglobinin (Hb) yoğunluk değişimlerine yansıdığı bilinmektedir. fNIRS'ın rakip işlevsel gözlem yöntemlerine bir diğer üstünlüğü de oksijenli hemoglobin (HbO₂) derişimindeki artmanın (ve beraberinde Hb derişimindeki azalmanın) milisaniye mertebesinde izlenebilmesidir. Araştırmacılar, beyin salınım içeren etkinliğinin bilişsel etkinlikten bağımsız olarak gerçekleştiğini ve bu iki etkinliğin farklı spektral davranışlarının olduğunu gözlemlemiştir [3]. Bunlara ek olarak, nefes alma, kalp atışı, yorgunluk gibi olguların her biri spektrumun farklı alanlarına yansımaktadır [1]. Bu varsayımlara uygun, kesin matematiksel modeller bulunmamakla birlikte, beyin çeşitli uyarılara karşı yanıtı için birim dürtü tepkisi temelli modeller önerilmiştir [4, 5, 6]. Bu çalışmada, fNIRS işaretlerinin gerçekten de beyindeki bilişsel etkinlikleri yansıttığı varsayımının geçerliliğini göstermeyi amaçladık. Bir fNIRS işaretinin uyaran işaretleriyle ilintili bir önemlilik içermesi, bilişsel etkinliğe dair bilgi taşıdığına en güçlü kanıttır. Bu nedenle, fNIRS işaretlerinin bilişsel etkinliği yansıtması en olası alt

bantlarında önemlilik kanıtını aradık. Alt bant seçimini [2]'de önerdiğimiz uyarıları topaklandırma temelli algoritma yardımıyla gerçekleştirdik. Önemliliklerin kestirimlerinde gözlenen tutarlılık, bunların arkaplan etkinliğinden kaynaklanan yapaylıklar olmadığını göstermektedir.

2. Veri Toplama

Deneklerden ölçülen veriler, alna yerleştirilmiş toplam 16 algılayıcı üzerinden alınmıştır. Deneklere ekranda rasgele biçimde "OOOOO" ve "XXXXX" örüntüleri gösterilmektedir. Birinci örüntü karşısında denek sol fare tuşuna, daha seyrek oluşan ve "hedef uyaran" diye nitelendirdiğimiz ikinci uyaran durumunda ise sağ tuşa basmaktadır.

Protokolde kullanılan uyaran dizisinin özellikleri şöyledir:

(i) Peş peşe yapılan 8 ayrı oturum halinde toplam 64 hedef uyaran sunulur; her bir oturumda ise hedeflerin zaman konumları beklenen periyodik konumlarından rasgele seçilen 8 ayrı hedef uyaran vardır; (ii) İki uyaran arasındaki süre 30 ilâ 50 örnek kadar değişebilen ve birbiciimli dağılmış rasgele bir değişkendir. (iii) Her iki tür uyaran da 500 ms boyunca ekranda kalır ve ardışık iki uyaran arasında bir saniyelik boşluk vardır. (iv) Sayısal kayıtlar 1.7 Hz örnekleme hızında yapılmıştır, buna göre Nyquist bant genişliği 850 mHz'dir ve her bir kayıt da 2700 örnekten oluşur. (v) Deneylerde, 22-50 yaş aralığında 5 denek kullanılmıştır.

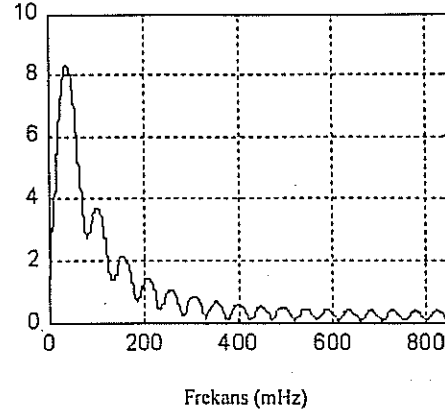
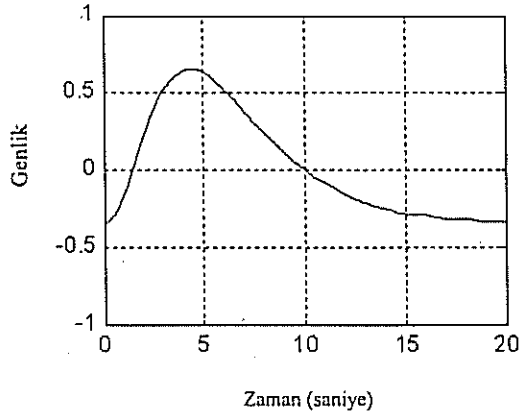
Foto-algılayıcılardan normal olarak 5 denek üzerinden $5 \times 16 = 80$ ölçüm elde etmeyi beklerken, bazı işaretlerin ciddi devinim yapaylıkları ya da algılayıcı hataları nedeniyle kullanılamaz durumda olduğunu gözlemledik. Kusurlu ölçümlerin elenmesinden sonra, 72 işaretlik bir veri kümesi elde ettik. Yaklaşık olarak 2700 örnekten oluşan işaretlerin her birinden 3 mHz'in altına ilişkin yönseme eğrisini çıkardık.

3. Yöntemler

3.1. Alt bant seçimi

Alt frekans bantlarını seçmek için, sıradüzensel bir toplama yöntemi kullandık [2]. Öncelikle frekans spektrumu için 10 mHz'lik 25 dar bant ve 250-850 mHz arasını kaplayan geniş bir yüksek frekans bandından oluşan bir başlangıç bölüntülemesi düşündük. Yüksek frekans bandını diğerlerine kıyasla oldukça geniş tutmamızın nedeni, tipik bir işaretin enerjisinin yalnızca yüzde 9'unun 250-850

TÜBİTAK 102E027 ve Boğaziçi Üniversitesi BURF 02S102 projeleri ile desteklenmiştir.



Şekil 1. Sıfır eksenine çekilmiş Gamma fonksiyonu (sol) ve Fourier spektrum büyüklüğü (sağ), yüksek frekans çıkıntıları işaretin ortalama değerinin çıkarılmış olmasından kaynaklanmaktadır.

mHz aralığında bulunmasıdır. Daha sonra veri kümesindeki her $s^m(t)$, $m=1, \dots, 72$, işareti için bu bantlarda zamanla değişen güç profilini, ya da R-dizisini, sıradaki gibi hesapladık:

$$R_n^m(t) = \frac{\int_{f_{n,l}}^{f_{n,h}} |S^m(t, f)|^2 df}{\int_0^{f_{Nyquist}} |S^m(t, f)|^2 df} \quad (1)$$

(1) ifadesindeki $S^m(t, f)$ terimi, $s^m(t)$ işaretinin kısa dönemli Fourier dönüşümüdür, n^{inci} bandın alt ve üst limitleri sırasıyla $f_{n,l}$ ve $f_{n,h}$ 'dir. Son olarak, bilgi taşıyan alt bantların seçimini aşağıda anlatıldığı üzere topaklandırma ve çoğunluk oylama kullanarak gerçekleştirdik.

İlk aşamada her bir işareten hesaplanan 26 tane R-dizileri arasında ikiserli olarak ilinti katsayılarını hesapladık. Ardından bu katsayıları temel alarak, birbirlerine en çok benzeyen R-dizilerini sıradüzensel bir yöntemle üç grup kalana dek topaklandırdık. Son tahlilde bir alçak, bir yüksek, bir de "yararlı" bir orta frekans bandı elde etmeyi umduğumuzdan son topak sayısının üç olarak belirledik. Bu işlemlerin sonucunda elde edilen her bir topaktaki R-dizilerinin kapsadığı bandı, bulmayı amaçladığımız bilgi taşıyan bantlardan birisi olarak atadık. Sonuç olarak her bir algılayıcı işareti için üçer tane olmak üzere, tüm denek ve algılayıcılardan $72 \times 3 = 216$ bant elde ettik. Enerji profillerine

Tablo 1. Aday bantlar

Bantlar (mHz)	Oy sayıları
250-850	44
0-30	35
0-40	35
40-250	35
30-40	28
40-850	28
30-250	7
0-50	2
50-250	2

ve oylamaya dayalı topaklandırmanın başarılı olduğunu, az sayıda bant adayının oyların çoğunu topladığına bakarak söyleyebiliriz.

İkinci aşama ise, genel bir fNIRS işareti için geçerli olabilecek bir spektrum bölüntülemesi oluşturmak için elde edilen bantlar üzerinden çoğunluk oylaması yürütmektir. Buna göre, elimizdeki 216 bandı, tekrarlanma sıklığına göre sıralandırdık. Sıralama sonucunda, 216 oyu paylaşan ve son bölüntülemeye yer alabilecek 9 ayrı bant bulduk. Bu bantlar ve aldıkları oy sayısı Tablo 1'de gösterilmiştir.

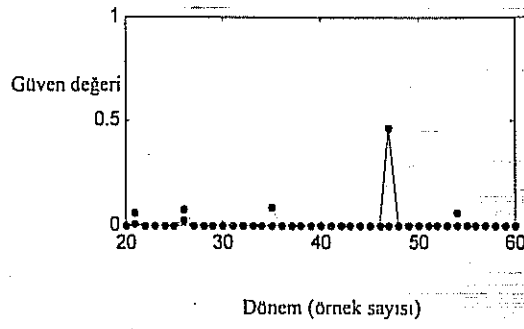
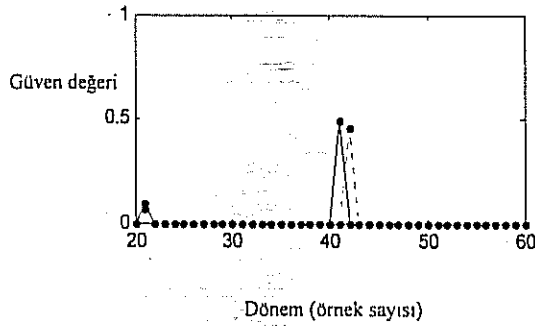
Tablo 1'de verilen sonuçlara göre, birlikte bütün spektrumu kaplayan ve birbirleriyle örtüşmeyen bantlar içeren birçok bölüntüleme oluşturmak olasıdır. Bu bölüntülemelerden bantlarının oyları toplandığında en çok oyu alan (216 oyun 142'si, % 65.7) 0-30 mHz, 30-40 mHz, 40-250 mHz ve 250-850 mHz bantlarından oluşan bölüntüleme değildir. Bu bantları, ifade kolaylığı açısından Tablo 2'deki gibi harflerle gösteriyoruz.

Tablo 2. fNIRS Bantları

Belirteç	Bantlar (mHz)	Oy sayıları
A	0-30	35
B	30-40	28
C	40-250	35
D	250-850	44

Özellikle transkraniyel doppler sonografinin kullanıldığı birçok çalışmada, beyin hemodinamiğine ilişkin üç frekans bandı saptanmıştır: (i) çok alçak frekans bandı (8-33 mHz), (ii) Mayer dalgaları ya da V-ışaretinin bulunduğu alçak frekans bandı (100 mHz çevresi) [8], (iii) ve bir yüksek frekans bandı (>250 mHz). Bu sonuncusunun solunum sıklığı ve kalp atışı işaretiyle ilişkili olduğu rapor edilmiştir [9].

Benzer bir biçimde, biz de fNIRS işaretlerinden yukarıda anlatılan yöntemle seçilen bantların, oksihemoglobin derişimleriyle ölçülen fizyolojik etkinliklerle birebir ilişkili olduğunu düşünüyoruz. En alçak frekans bandı olan A-bandı, çok yavaş damar içi salınımları yansıtan arkaplan etkinliğinden sorumludur. Bu bandın uyarana bağlı bilişsel etkinlikten bağımsız olduğunu vaysıyoruz. Diğer yandan, fMRI'da kullanılan modellerden biri olan sıfır-ortalama



Şekil 2. (a) Seçilmiş bir fNIRS oturumu için önsüzgeçlemesiz (noktalı çizgi) ve önsüzgeçlemeli (katı çizgi) güven değerleri profili: Önsüzgeçleme kestirilen dönem değerinin güven değerini artırıyor; (b) bir başka fNIRS oturumu için önsüzgeçlemesiz (katı çizgi) ve önsüzgeçlemeli (noktalı çizgi) güven değerleri profili: Bu durumda önsüzgeçleme sayesinde (30,50) aralığında yüksek güven değerli bir dönemlilik ortaya çıkıyor, oysa ki önsüzgeçleme yapıldığında hiçbir dönemlilik bulunamıyordu.

Gamma fonksiyonunun 30-50 mHz arasında yoğun olduğu gözlemlenirken hareketle (Şekil 1), B-bandının uyarana bağlı bilişsel etkinlikle ilişkili olması gerekir [10]. Daha geniş olan C-bandının da, büyük olasılıkla uyarın dizisindeki dönemlilikten kaynaklanan, bilişsel etkinlikle ilişkili bilgi taşıdığı düşünülmektedir. Bununla birlikte, damar içi salınımlar ve solunum işareti de C-bandının oluşmasında etkilidir. Son olarak, yüksek frekans bandı olan D-bandının birtakım zayıf, rasgele salınımlarla kalp atışı işaretinin (~1.1 Hz) katlanmış halini içerdiğini düşünüyoruz.

Son tahlilde, yukarıdaki varsayımlardan hareketle, B ve C bantlarını kapsayan 30-250 mHz aralığının bilişsel etkinliğin fNIRS işaretlerinde en yoğun olarak bulunduğu bant olduğunu savhıyoruz.

3.2. Dönemlilik kestirimi

Bilişsel uyarılar neredeyse dönemli olduklarından (uyaran arası zaman farkları 18-29 saniye arasında bir örnek dağılmıştır), bilişsel etkinliğe ilişkin işaret bölümlerinde bir dönemlilik davranışının varlığını düşünebiliriz. Bilişsel etkinlik dalga biçimleri, eğer varsa, genelde arkaplan etkinliğinin içine gömülmüş olacaktır. Gerçekten de, deneyler göstermektedir ki bilişsel etkinlik tepkilerini gözle ayırt etmek son derece güçtür. İşaretteki dönemlilikleri sezme için kullandığımız yöntem, ilkin konuşma işaretlerinde kullanılan en küçük kareler dönemlilik kestirimi ("least-squares periodicity estimation", LSPE) algoritmasıdır [7]. Bu yöntemde, gözlemlenen işaret $x(t)$ ile kestirilen işaret $x_0(t)$ arasındaki ağırlıklı ortalama karesel hata en küçük kılınır. Kestirilen işaret $x_0(t)$ şu sıradaki dönemlilik ilişkisini sağlar: $x_0(t) = x_0(t + kP_0)$, $t = 1, 2, \dots, T$ ve $k = 1, \dots, K$. En küçük ağırlıklı ortalama karesel hatayı veren P_0 dönemliliğinin aynı zamanda aşağıdaki ifadeyi enbüyüttüğü [7]'de gösterilmiştir:

$$J_1(P_0) = \frac{I_0 - I_1}{E - I_1} \quad (2)$$

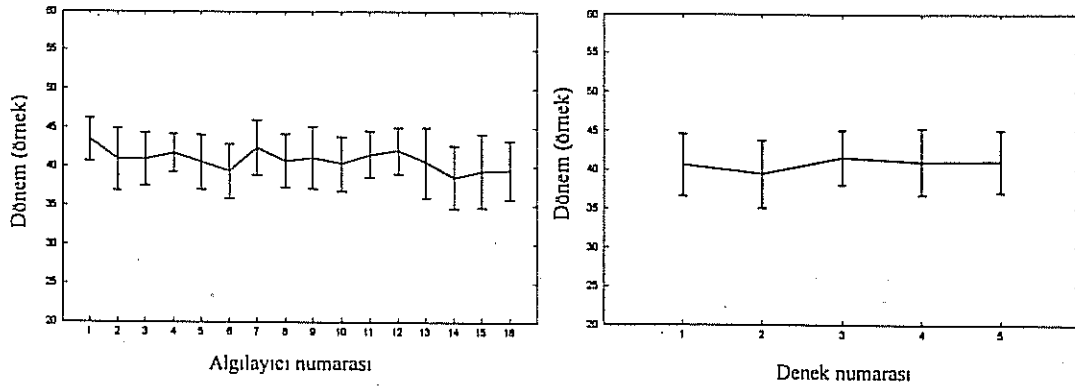
(2) ifadesinde, I_0 terimi $x_0(t)$ 'nin ağırlıklı enerjisi, E de $x(t)$ 'nin ağırlıklı enerjisi yerine geçer. Bunlara ek olarak, yüksek P_0 değerlerine doğru olan yanlılığı düzeltmek için konulan I_1 terimi de aşağıdaki hesaplanabilir:

$$J_1(P_0) = \frac{\sum_t x^2(t)w^2(t)}{\sum_k w(t+kP_0)} \quad (3)$$

(3) ifadesinde, $w(t)$ işaret örnekleri ağırlıklandırılmakta kullanılan, tipik bir penceredir. (2)'deki fonksiyonel en büyük kılan P_0 değeri, eğer $J_1(P_0)$ yeterince yüksekse, işarettaki en baskın dönemlilik olarak alınır. $J_1(P_0)$ işlevsel, bütünüyle dönemli bir işaret için 1 değeri alan güven değeri olarak yorumlanabilir. (2)'yi enbüyüten herhangi bir P_0 her zaman bulunabileceğinden, kestirimin gerçek bir dönemlilikle ilişkili olabilmesi için, güven değerinin bir eşiği geçmesi gerekir. Bu düşüncelerden hareketle fNIRS işaretlerindeki bilişsel etkinlikle ilişkili dönemlilikleri ortaya çıkarabilmek için, örnek sayısı olarak $P_{min} = 20$ and $P_{max} = 60$ değerleri arasında değişen her P_0 için $J_1(P_0)$ değerini hesapladık. Bilişsel uyarının tam anlamıyla dönemli olmamasından ve de işlevsel etkinlik işaretinin arkaplana iyice gömülmüş olmasından dolayı, $J_1(P_0)$ değerlerinin çok yüksek olması beklenmemelidir. Üçüncü bir seçirme, yani periyodiklikten sapma kaynağı da bilişsel etkinlik işaretinin rasgele bir süre içinde tepki vermekte gecikmesidir. Bütün bu bozucu etkenleri dikkate alarak, aykırı değerleri kestirilen dönemlerden elemeye karar verdik ve (30,50) örnek arasında düşen kestirimleri kabul edilebilir olarak nitelendirdik.

Bir önceki bölümde, fNIRS işaretlerinde farklı frekans bantlarının bir ya da birden çok özel fizyolojik etkinlikle ilişkili olduğunu göstermiştik. Buna göre, dönemlilik kestiriminin yapılacağı en uygun bant aralığı 30-250 mHz olarak saptanmıştır. Böylece ilkin fNIRS işaretlerini 30-250 mHz aralığında önsüzgeçledik. Önsüzgeçlemenin iyileştirici etkisini, Şekil 2'de, kestirilen periyoda ait hesaplanan güven değeri profilinin artmasında gözlemlemekteyiz.

DeneySEL protokol ardışık sekiz oturumdan oluştuğundan, LSPE algoritmasını her bir fNIRS işaretinin her bir oturumu için ayrı ayrı uygulamaya karar verdik. Buna göre, önsüzgeçlenmiş bir işaret için sekiz aday dönem değeri



Şekil 3. $\bar{P}_{algilayıcı}$ (sol) ve \bar{P}_{denek} (sağ) çizgeleri.

bulunabilir. Diğer bir deyişle, verili bir deneğe ait her bir algılayıcıdan sekiz dönem kestirimi hesaplanabilir. Denekler-arası ve algılayıcılar-arası değişimleri araştırmak için, iki ayrı nicelik hesapladık: (i) Belirli bir algılayıcı için kabul edilebilir bütün kestirimler ve bütün denekler üzerinden ortalama dönem $\bar{P}_{algilayıcı}(k)$, $k=1, \dots, 16$; (ii) Belirli bir denek için kabul edilebilir bütün kestirimler ve bütün algılayıcılar üzerinden ortalama dönem $\bar{P}_{denek}(l)$, $l=1, \dots, 5$. Bu niceliklerin grafikleri çeyrek kütleli gösteren çubuklarla birlikte Şekil 3'te gösterilmiştir.

4. Tartışma ve Sonuçlar

Bu çalışmada, fNIRS yardımıyla beyindeki bilişsel etkinliğin gözlemlenebileceği savının sağlanmasını yaptık. Deneysel protokollerin çoğu, neredeyse dönemli uyaran içeren bilişsel görevler dizisinden oluşur. Benzer bir protokol altında alınan fNIRS ölçümlerinin protokoldeki dönemliliği yansıttığını gözlemledik. Gerçekten de kestirilen dönemler, hedef sunum aralıklarıyla aynı erimde, (35,45) örnek ya da (20.5,26.5) saniye eriminde çıkmaktadır. Aykırı kestirimler elendikten sonra dönem kestirimlerinin bütün algılayıcılar ve denekler için tutarlı olduğu Şekil 3'ten görülebilir. Algılayıcıların geniş beyin alanlarından ölçüm almalarından ve buna karşıt olarak tepkilerin son derece yerel olabileceğinden dolayı, bütün algılayıcıların hemodinamik yanıtı aynı ölçüde doğru vermesi beklenmemelidir. Örneğin, arteriyel bir damarın yakınındaki bir algılayıcıdan alınan bir ölçüm daha güvenilir ve belirgin olabilir. Algılayıcıdan algılayıcıya ve denekten deneye değişen kabul edilemez kestirimlerin varlığı, bu kestirimlerle ilişkili algılayıcılarda ve oturumlarda yalnızca arkaplanın ölçüldüğünü göstermektedir. Bu bağlamda, fNIRS işaretlerine uygulanan dönemlilik kestirimi algılayıcı ve deneklerin elenmesinde kullanılabilir [2]. Düşük dönemlilik güven değerleri veren algılayıcı/deneklerin bilişsel etkinliğe dair bilgi sağlamadığı düşünülmektedir.

5. Kaynakça

[1] D. Kim and K. Ugurbil, "Bridging the gap between neuroimaging and neuronal physiology", *Image Anal. Stereol.*, 21:97-105, 2002.

[2] C. B. Akgül, B. Sankur, and A. Akın, "Selection of Relevant Frequency Bands in fNIRS", *to be submitted to J. Comp. Neuroscience*.

[3] H. Obrig, M. Neufang, R. Wenzel, M. Kohl, J. Steinbrink, K. Einhaupl, and A. Villringer, "Spontaneous low frequency oscillations of cerebral hemodynamics and metabolism in human adults", *Neuroimage*, 12:623-639, 2000.

[4] M. Svensen, F. Kruggel, and D. Y. von Cramon, "Probabilistic modeling of single-trial fMRI data", *IEEE Trans. Med Imaging.*, 19(1):25-35, 2000.

[5] A. Villringer and B. Chance, "Non-invasive optical spectroscopy and imaging of human brain function", *Trends in Neuroscience*, 20(10):4435-4442, 1997.

[6] P. Wobst, R. Wenzel, H. Obrig, M. Kohl, and Arno Villringer, "Linear aspects of changes in deoxygenated hemoglobin concentration and cytochrome oxidase oxidation during brain activation", *Neuroimage*, 13:520-530, 2001.

[7] David H. Friedmann, "Pseudo Maximum Likelihood Speech Pitch Extraction", *IEEE Trans. on ACSSP*, 25: 213-221, 1977.

[8] J. Mayhew, Y. Zheng, Y. Hou, B. Vuksanovic, J. Berwick, S. Askew, P. Coffey, "Spectroscopic analysis of changes in remitted illumination: The response to increased neural activity in brain", *Neuroimage*, 10:304-326, 1999.

[9] T. Kuo, C. Chern, W. Sheng, W. Wong, H. Hu, "Frequency domain analysis of cerebral blood flow velocity and its correlation with arterial blood pressure", *J Cereb. Blood Flow Metab.*, 18: 311-318, 1998.

[10] D. Cordes, V.M. Haughton, K. Arfanakis, J. D. Carew, P. A. Turski, C. H. Moritz, M. A. Quigley, M. E. Meyerand, "Frequencies contributing to functional connectivity in the cerebral cortex in resting-state data", *Am J Neuroradiol*, 22: 1326-1333, 2001.

El Şekillerine Dayalı Kişi Tanıma

Person Authentication Based On Hand Shape

Erdem Yörük¹, Ender Konukoğlu¹, Bülent Sankur¹, Jérôme Darbon²

¹ Elektrik ve Elektronik Mühendisliği Bölümü, Boğaziçi Üniversitesi, Bebek, İstanbul, Türkiye

² EPITA (Ecole Polytechnique Industrielle des Technologies Appliquées), Paris, France

{erdem.yoruk, konukoge, sankur}@boun.edu.tr, jerome.darbon@lrde.epita.fr

Özetçe

El imgelerine dayalı bir kimlik tanıma sistemi geliştirilmiştir. Kişilerin sol elleri herhangi bir poz kısıtı uygulanmaksızın bir renkli tarayıcı aracılığıyla elde edilmiş olup ve çıkarılan el şekilleri tüm ele ve teker teker parmaklara uygulanan öteleme ve döndürme işlemleriyle standart konumlara çakıştırılmıştır. Tanılama için şekil imgelerinin bağımsız bileşen öznitelikleri kullanılmış ve sınıflama başarımları en az yüz örneklik gruplarda uygulanabilecek kadar yüksek bulunmuştur.

Abstract

A system has been developed for person identification based on hand images. The images of the left hand of the subjects are captured by a flatbed scanner in an unconstrained pose. The silhouettes of hands are registered to a fixed pose, which involves both rotation and translation of the hand and, separately, of the individual fingers. Independent component features of the hand silhouette images are used for recognition. The classification performance is found to be very satisfactory and it is shown that, at least for groups of one hundred subjects, hand-based recognition becomes a viable and secure access control scheme.

1. Giriş

Biyometri teknolojileri, kişileri tanımak için deneklerin fiziksel ve davranışsal özelliklerini kullanır. Biyometrik tanılama sistemlerinde yaygınca kullanılan fiziksel öznitelikler yüz öznitelikleri, parmak izleri, iris ve retina taramaları, el ve parmak geometrileridir; ses izleri, imza, vücut hareketleri gibi özellikler ise davranışsal olarak sınıflandırılabilir.

Biz, bu çalışmada, kişiye özgü el şekillerinin, çeşitli giriş kontrolü sistemlerinde kolayca uygulanabilecek, basit ve tutarlı bir seçenek olduğunu varsayarak sadece el şekillerine dayalı bir tanılama yöntemi geliştirdik. Yüz işlemede sıkça karşılaşılan poz, ışıklandırma ve ifade değişimleri; veya sese dayalı tanıma uygulamalarındaki çevresel etkenler ve sinyal karışımları gibi sorunlar tanıma başarımlarını olumsuz yönde etkilemektedir. Oysa el imgeleri, standart bir tarayıcı aracılığıyla, bu tarz kısıtlamalara maruz kalmadan elde edilebilmekte, bu yönüyle düşük maliyet ve bellek gereksinimli, basit arayüzlü bir biyometrik seçenek olarak karşımıza çıkmaktadır.

Literatürde yer alan, bundan önceki ele dayalı tanılama yöntemleri çeşitli el özniteliklerini ve/veya aya izi bilgisini kullanmışlardır. Örneğin, [10], [8] ve [6]'daki yazarlar, el çevritlerinden çıkardıkları, parmak uzunluğu, genişliği, aya büyüklüğü, ayanın parmaklara göre genişlik oranı gibi belli bazı geometrik ölçüleri kullanıp, bunları tanıma evresinde Öklit veya Hamming uzaklıklarıyla karşılaştırmışlardır. [4] ve [11] gibi çalışmalarda ise özellikle aya izlerine önem verilmiş, izler, hayat ve kalp çizgilerine göre kaydedilip, düz çizgi yaklaşımlarına dayanılarak kıyaslanmıştır.

Biz ise bu çalışmada, kişi tanıma için sadece elin şekil bilgisini kullanıyoruz. Elde edilen ham imgeleri ilk önce kapsamlı bir ön işlemeyle bölütleyip ve standart bir duruşa getiriyoruz. Bu evrede, tüm eli, ve özellikle ayrı ayrı parmakları önceden belirlenmiş yönlere getirdikten sonra, elde edilen çakıştırılmış şekil imgelerinden ayırt edici öznitelikleri çıkarıyor, ve bunları tanıma evresinde karşılaştırıyoruz.

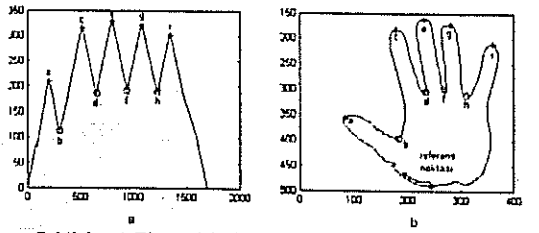
Bildirimiz şu şekilde düzenlenmiştir. Bölüm 2'de el imgelerinin bölütlenmesi, bölüm 3'te ise çakıştırılması anlatılmaktadır. Çıkarılan öznitelikler, Deney düzeneği ve tanıma başarımları bölüm 4 ve 5'te, çıkarılan vargılar da bölüm 6'da bahsedilmiştir.

2. El İmgelerinin Bölütlenmesi

Ham tarayıcı imgelerine bakıldığında, ellerin bölütlenmesi basit bir işlem gibi görünse de, elde edilen bölütlerin doğruluğu, yüzük, kol saati, bilekten taşan kıyafetler ve elin tarayıcıya gereğinden fazla veya az bastırılması gibi faktörlerden etkilenabiliyor. Dahası, farklı eller arasındaki şekilsel farkların genellikle küçük olması, uygulamanın başarımlarını, çıkarılması hedeflenen el çevritlerinin asıl el şeklinden yapacağı sapmalara büyük ölçüde duyarlı kılıyor. Bu noktada, renk kümelemesinin ardından uygulanan biçimsel düzeltmeler ve susunırı dönüşümüne dayalı bölütleme kullanarak, şekil doğruluğunu yüksek tutacak şekilde, iki farklı yöntem geliştirdik. Her iki işlemin sonunda da yüzük gibi aksesuarlardan kaynaklanan hataları gidererek, el çevritlerini detaylı bir şekilde elde ettik.

3. El Çevritlerinin Çakıştırılması

Bölütlenmiş ellerin çakıştırılması, şekle dayalı tanıma söz konusu olduğunda, hedeflenen biyometri uygulamasının en hassas aşamasını teşkil ediyor. Bu işlem, asıl şekillerini bozmadan, el örneklerinin döndürme ve ötelemeleri



Şekil 1: a) El çevritlerinde radyal uzaklık eğrisi; b) Parmak uçları ve vadileri

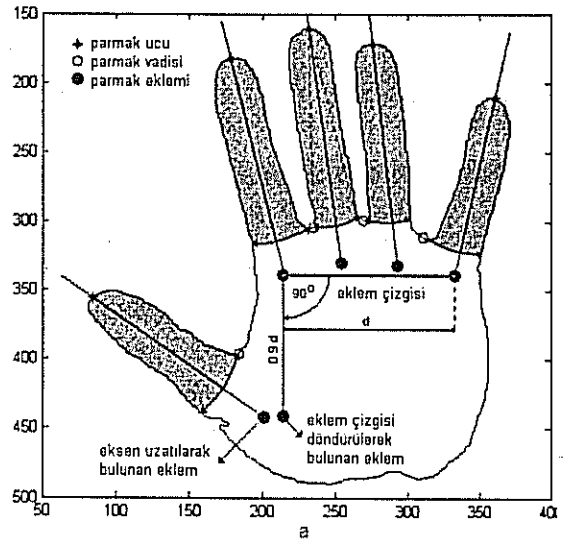
kaydedilmesini ve standart yönler esas alınarak parmak uzanışlarının tekrar düzenlenmesini içermektedir. Şekil 1 b.'de, aynı kişiye ait çevritlerin, çakıştırılma öncesi ve sonrası üstüste görüntüsüne bakıldığında bu anlamda bir işlemin gereği daha iyi anlaşılmaktadır. Hareketli bir organ olması dolayısıyla, parmak duruşlarından kaynaklanan bu tür sınıf içi farklılıklar, sınıflar arası farklara baskın çıkarılmakta ve tanımayı başarısız kılabilmektedir, dolayısıyla, öz niteliklerin çıkarımından önce parmakların aynı uzanışlara getirilmesi büyük önem taşımaktadır.

3.1. Parmakların konumlandırılması

Parmak uçlarını ve parmak arası vadileri belirlemek için, aya bölgesinin içinden, bileğe olabildiğince yakın seçilen bir referans noktası esas alınmakta ve bu noktaya göre diğer çevrit noktalarının radyal uzaklıkları kaydedilmektedir. Buradaki referans noktası, elin ana atalet ekseninin bilek çizgisiyle yaptığı kesişme olarak seçilmiş olup ortaya çıkan uzaklık eğrisi, aranan dokuz anatomik noktayı yerel uç değerler olarak sergilemektedir. Şekil 1'de görüldüğü gibi uzaklık eğrisi, çevritteki gürültüden etkilenmediği için bulunan bu uç değerler kararlı bir şekilde parmakları konumlandırılmayı sağlamaktadır.

3.2. Parmakların kaydedilmesi

Parmaklar, iki yanındaki vadi noktalarından uç noktasına daha yakın olanı hizasından kesilerek ayadan ayrılır. Sabit nirengi noktaları olmaları bakımından, parmak-aya bağlantı eklemleri, parmak ana eksenini üzerinde, uçtan başlayarak parmak uzunluğunun % 120'si kadar aya içinde konumlandırılır. Başparmak haricindeki dört parmak için bulunan bu eklem noktaları, elin bütününe büyüklüğü ve uzanışı hakkında önemli bir ölçüt teşkil etmektedir. İşaret parmağı ve serçe parmak eklem noktalarını birleştiren, eklem çizgisi olarak adlandırdığımız doğru parçası bundan sonraki çakıştırma işlemlerinde ana rolü oynamaktadır. Parmakların ve tüm elin uzanış açıları, bu çizginin uzanış açısı esas alınarak belirlenmekte ve gerekli öteleme ve döndürme işlemleri yine bu çizginin uzunluğu ve uzanışı ışığında gerçekleştirilmektedir. Parmak uzanış açıları, ayadan ayırdıktan sonra bulunan atalet eksenleriyle bulunur ve parmaklar, uzanış açıları, önceden belirlenen, parmağa özel ortalama uzanış açısıyla örtüşecek şekilde, eklem noktası etrafında döndürülür. Bu aşamada başparmak, tarayıcı düzlemi üzerinde iki farklı ekleme göre daha dinamik bir hareketlilik sergilemesi bakımından ayrı olarak ele alınmaktadır. Buradaki iki eklemlilik dönme hareketliliği, tek bir dönme ve ötelemenin katışımı olarak modellenir. Dolayısıyla, başparmaklar



Şekil 2: a) Parmakların ayadan ayırılışı, parmak eksenleri, parmak eklem noktaları ve eklem çizgisinin tanımı; Aynı kişiye ait, üstüste çizilmiş el çevritleri: b) Çakıştırmadan önce, c) Çakıştırmadan sonra

uzunluklarıyla bulunan eklem noktaları etrafında bir kez döndürüldükten sonra, yine eklem çizgisi referans alınarak saptanan, sabit ikinci bir noktaya doğru ötelenir. Bu şekilde parmak uzanış çakıştırması tamamlanmış olur. Bundan sonra, elin tamamı, eklem çizgi merkezi ve açısı, imge düzlemi üzerinde önceden saptanmış bir nokta ve uzanış açısıyla örtüşecek şekilde ötelenip döndürülür ve böylece elin konumsal ve uzanışsal çakıştırması tamamlanmış olur.

Bundan sonra ikinci bir düzeltme bilek kısmında yapılır, çünkü bölütlenmiş imgelerde kıyafetlerin ayağa doğru taşması, bilekten itibaren kolun uzanışının ve tarayıcıya uygulanan baskının farklı olabilmesi imgelerde değişik bilek yapıları ortaya çıkarabilmektedir. Bu anlamdaki düzensizlikleri gidermek için, çevritteki bilek eğrisini Euler yay eğrilerini kullanarak tekrar üretmek veya aya bölgesini bilek kısmına yakın bir hizadan kesmek gibi yöntemler kullandık.

4. Çıkarılan Öz nitelikler ve Tanıma

Bir biyometri uygulaması için elde edilen çakıştırılmış el şekillerinden farklı yöntemlerle öz nitelikler çıkarılabilir. Yakın zamanda, Bağımsız Bileşenler Analizi (ICA), imge işleme

$$x_i = A(i,1) \times \hat{s}_1 + A(i,2) \times \hat{s}_2 + \dots + A(i,N) \times \hat{s}_N$$

ICA1 metodunda i. elin öznelik vektörü: $a_i = [A(i,1), A(i,2), \dots, A(i,N)]$

$$x_i = a_i \times \hat{p}(1,i) + a_2 \times \hat{p}(2,i) + \dots + a_N \times \hat{p}(N,i)$$

ICA2 metodunda i. elin öznelik vektörü: $\hat{p}_i = [\hat{p}(1,i), \hat{p}(2,i), \dots, \hat{p}(N,i)]$

Şekil 3: ICA el örüntüleri: a) ICA1 elleri; b) ICA2 elleri.

güvenilir bir öznelik çıkarma tekniği olarak, yüz tanıma gibi problemlerde kullanılmaktadır. Bizim durumumuzda her ne kadar kullanılacak şekil imgeleri, ikili el silüetleri şeklindeyse de, ICA ile tanımda gri seviyesi de devreye sokulabilir.

Bağımsız Bileşenler Analizi herhangi bir karışım içinden istatistiksel anlamda bağımsız değişkenleri çıkarmaya yarayan bir tekniktir. Bu metotta, gözlenen K uzunluğundaki karışım sinyalleri $x_i(k)$ ($i = 1, \dots, N$, $k = 1, \dots, K$), aynı uzunlukta ve sayıdaki $s_i(k)$ ($i = 1, \dots, N$, $k = 1, \dots, K$) kaynak bağımsız bileşenlerinin $N \times N$ 'lik bir A matrisiyle doğrusal katışımı olarak kabul edilir. Burada, x_i ve s_i sinyalleri ait oldukları $N \times K$ büyüklüğündeki X ve S matrislerinin satırlarında yer alacak şekilde, kullanılan model

$$X = AS \quad (1)$$

şeklinde dir. Bu modele göre S'deki bağımsız kaynaklar ve onları katıştıran A'daki katsayılar, $Y=WX$ matrisinin satırları arasındaki istatistiksel bağımsızlığı enbüyüten bir W dönüşümü bulunarak kestirilir. Böyle bir dönüşümü bulmak için, Y'deki çıkımlar arası bağımsızlığı, onların negentropisi cinsinden modelleyen fastICA yordamı kullanılmakta ve sonuçta kaynaklar ve katışım matrisi şu şekilde kestirilir:

$$\hat{S} = Y = WX \quad (2)$$

$$\hat{A} = W^{-1} \quad (3)$$

Aynı model esas alınarak, çıkarılması hedeflenen özneliklerin seçimine göre ICA1 ve ICA2 diye adlandırılan iki farklı uygulama yapısı göze çarpmaktadır. Bunlardan ilkinde her x_i sinyali yani X matrisinin her satırı, i. el imgesinin bir boyutta dizilmiş piksel değerleri şeklinde düzenlenir. Bu durumda, 200×200 'lük bir imge düşünüldüğünde bu sinyaller, ayrıştırma sonunda ortaya çıkan \hat{s}_i kaynak imgeleri gibi $K = 40000$ uzunluğunda olacaktır. Bu durumda her bir x_i 'yi, bağımsız kaynakların farklı bir katışımı olarak ifade eden, i. kişiye ait N boyutlu öznelik vektörü, katışım katsayıları halinde A'nın i. satırında yer alacaktır. Tanıma evresinde gelen $(1 \times K)$ boyutundaki yeni bir x_{test} imgesi aynı katışım modelini izleyeceği varsayımıyla önceden belirlenmiş bağımsız bileşenlere izdüşürülür ve

$$a_{test} = x_{test} \hat{S}^T (\hat{S} \hat{S}^T)^{-1} \quad (4)$$

şeklinde elde edilen izdüşüm katsayıları vektörü eğitim aşamasında belirlenmiş özneliklerle karşılaştırılır. L1

uzaklığına göre en yakın olan öznelik vektörünün ait olduğu kişi, tanınacak olan test imgesinin kimliği olarak belirlenir:

$$i^* = \arg \min_i \left(\sum_{j=1}^N |a_{i,j} - a_{test,j}| \right) \quad (5)$$

İkinci yapıda ise izlenen model aynı olmakla beraber, çıkarılacak özneliklerin katışım matrisi A yerine kaynak matrisi S'de olması ve dolayısıyla birbirlerinden bağımsız olmaları gözetilmektedir, bu durumda x_i karışım sinyalleri, i. imgenin piksel değerleri şeklinde değil, i. pikselin farklı imgelerde aldığı değerler şeklinde düşünülmektedir. Yani, önceki yapıda kullanılan X matrisi yerine, onun devriği olan X^T , katışım matrisi olarak sisteme sokulur. Fakat, bu noktada aynı imge boyutu ($K=200 \times 200$) düşünüldüğünde ayrıştırılacak karışımın sayısı çok fazla olacağından, bir ön işleme olarak, Ana Bileşenler Analizi (PCA) ile boyut indirgeme gerekmektedir, yani X^T yerine, X^T 'nin, kendi kovaryans matrisinin ilk N ($N \ll K$) özvektörü üzerine izdüşümü, yeni katışım matrisi olarak kabul edilir. $K \times K$ boyutundaki $C = X^T X$ 'ye ait ilk N özdeğere $\{\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N\}$ karşılık gelen bu özvektörler $\{u_1, u_2, \dots, u_N\}$ ise Tekil Değer Ayrıştırma (SVD) Teoremine göre daha küçük olan $N \times N$ boyutundaki $D = X X^T$ matrisinin $\{v_1, v_2, \dots, v_N\}$ özvektörlerinden

$$u_i = 1 / \sqrt{\lambda_i} X^T v_i \quad (6)$$

şeklinde bulunabilir. Bu yöntemle çıkarılan K boyutlu N tane $\{u_1, u_2, \dots, u_N\}$ özvektörleri U'nun sütunlarında yer alacak şekilde

$$\tilde{X} = U^T X^T = V^T X X^T = R X^T \quad (7)$$

dönüşümüyle elde edilen boyut indirgenmiş $N \times N$ 'lik yeni katışım matrisi önceki yapıdaki gibi bağımsız bileşenlerine ayrıştırılır. Karışıklık yaratmaması için piksel kaynaklarını barındıran bağımsız bileşenler matrisi bu yapıda \hat{P} diye ifade edilmekte ve bu matrisin i. sütunu i. kişinin öznelik vektörünü teşkil etmektedir. Tanıma evresinde gelen $(1 \times K)$ boyutundaki yeni bir x_{test} imgesi ilk önce $R x_{test}^T$ olarak boyut indirgendikten sonra eğitim aşamasında bulunan katışım matrisinin tersiyle çarpılır, bu şekilde bulunan

$$\hat{p}_{test} = W R x_{test}^T \quad (8)$$

vektörü önceden belirlenmiş öznelik vektörleriyle karşılaştırılır ve aradaki açının kosinüsünü enbüyüten vektörün ait olduğu kişi, tanınacak olan test imgesinin kimliği olarak belirlenir:

Tablo 1: Küme büyüklüğüne göre doğru tanıma başarımları

Küme büyüklüğü	ICA1			ICA2		
	20	50	91	20	50	91
Pc: tekli eğitim	96.89	94.85	93.77	97.92	97.33	96.89
Pc: çiftli eğitim	97.92	96.00	94.87	98.54	98.81	98.53

$$i^* = \arg \max_i \left(\frac{\hat{p}_i \cdot \hat{p}_{test}}{\|\hat{p}_i\| \|\hat{p}_{test}\|} \right) \quad (9)$$

Her iki yapının da şematik gösterimi ve çıkarılan ICA el örüntüleri Şekil 3'teki gibidir.

5. Deney Sonuçları

Kullandığımız, el veri tabanı, 91 farklı kişinin sol ellerinden alınmış toplamda 273 adet imgeden oluşmaktadır [2]. 1754x1276 boyutundaki bu imgeler 150 dpi çözünürlükte, bir HP Scanjet 5300c tarayıcı ile elde edilmiş olup, tarama esnasında örneklerin parmak uzunluklarına veya kullandıkları aksesuarlara herhangi bir kısıt getirilmemiştir. Bu şekilde, kişiler, üç farklı zamanda alınmış el imgeleri ile 91'lik A, B, C diye adlandırdığımız üç kümeyi oluşturmaktadır.

Kişi tanıma deneyleri, ilk önce, çakıştırılmış el imgeleri kullanılarak, 91 insan içinden 20, 50 ve 91'lik alt kümeler şeklinde, üç farklı nüfustaki gruplar üzerinde yapılmıştır. Buradaki farklı küme nüfusları rasgele seçilen altkümelerin kullanımıyla, tanıma başarımının veri tabanındaki kişi sayısı ile nasıl değiştiğini görmemizi sağlamaktadır.

İkinci olarak, sistem, öğrenme kümesinin niceliğine göre başarımı test etmek amacıyla, her kişiye ait tekli ve ikili kayıtlar üzerinde ayrı ayrı eğitilmiştir. Tekli küme deneylerinde kullanılan test ve öğrenme kümeleri ((A,B), (B,A), (A,C), (C,A), (B,C), (C,B)) şeklinde, çiftlilerde ise {(A, BC), (B, AC), (C, AB)} şeklindedir. Sonuç olarak, Tablo 1'de verilen tanıma başarımları farklı eğitim ve test kümesi deneyleri üzerinden ortalama alınarak elde edilmiştir. Bu sonuçlara göre, 100 kişiyi bulan durumlarda, el şekline dayalı tanılama işlemi gayet başarılı gözükmektedir.

6. Vargılar

Bu çalışmada göstermiş olduk ki, el şekilleri, bir kişi tanıma uygulaması için, yüz tanımda olduğu gibi aydınlatma ifade ve aksesuar gibi faktörlere karşı bağıtsız olmasıyla, nüfusun 100'ün üstüne çıktığı durumlarda bile çok uygun bir seçenek olmaktadır. Fakat buradaki başarımlar büyük ölçüde el şekillerinin doğru belirlenip çakıştırılmasına bağlıdır.

Anlatılanların dışında, şekilleri temsil etmek için Eksenal Radyal Dönüşüm (Axial Radial Transform), özneliklerin çıkarımı için Çekirdeksele Ana Bileşenler ve Doğrusal Ayırtaç Analizleri, el çakıştırması içinse Aktif Şekil Modelleri gibi başka

seçenekler denenmektedir. Çalışmada kullanılan örneklerin sol ellerinin yanı sıra sağ ellerin başarıma etkisi araştırılacak, çok daha fazla nüfuslu kümelere deneyler yapılacaktır. Ayrıca şekil bilgisinin yanında doku bilgisinin de kullanılması planlanmaktadır.

7. Kaynakça

- [1] M. S. Bartlett, H. M. Lades, and T. J. Sejnowski, "Independent component representations for face recognition," in *Conference on Human Vision and Electronic Imaging III*, San Jose, California, 1998.
- [2] S. Garcia-Salicetti, C. Beumier, G. Chollet, B. Dorizzi, J. L. les Jardins, J. Lunter, Y. Ni, and D. Petrovska-Delacretaz, "BIOMET: a multimodal person authentication database including face, voice, fingerprint, hand and signature modalities," in *International Conference on Audio- and Video-based Biometric Person Authentication*, 2003.
- [3] B. A. Draper, K. Baek, M. S. Bartlett, and J. R. Beveridge, "Recognizing faces with PCA and ICA," *Computer Vision and Image Understanding* 91 (1-2), 115-137, 2003.
- [4] C.C. Han, H. L. Cheng, C. L. Lin, and K. C. Fan, "Personal Authentication using Palm Print Features," *Pattern Recognition* 36(2003), 371-381.
- [5] A. Hyvarinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Networks* 13 (4-5), 411-430, 2000.
- [6] A. K. Jain and S. Pankanti, "Biometric Systems: Anatomy of Performance," *IEICE Trans. Fundamentals*, Vol. E84-D, No. 7, pp. 788-799, July 2001.
- [7] B. B. Kimia, I. Frankel, and A. M. Popescu, "Euler Spiral for Shape Completion," *International Journal of Computer Vision* 54(1/2), 157-180, 2003.
- [8] Y. L. Lay, "Hand shape recognition," *Optics and Laser Technology*, 32(1), 1-5, Feb. 2000.
- [9] E. Konukoğlu, E. Yörük, J. Darbon, B. Sankur, "Shape-Based Hand Recognition," submitted to *IEEE Image Processing*, 2003.
- [10] R. Sanches-Reillo, C. Sanchez-Avila, and A. Gonzalez-Marcos, "Biometric Identification through Hand Geometry Measurements," *IEEE Transactions of Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, October 2000.
- [11] W. Shu, G. Rong, and Z. Bian, "Automatic palm print verification," *International Journal of Image and Graphics*, 1(1):135-151, 2001.

Hand Biometrics

Erdem Yörük, Helin Dutağacı, Bülent Sankur

¹ Electrical and Electronic Engineering Department, Boğaziçi University, Bebek, İstanbul, Turkey
[erdem.yoruk, dutagach, sankur]@boun.edu.tr

Telephone: (90) 212 359 6414, Fax: (90) 212 287 2465

Corresponding author: Bülent Sankur

Abstract:

The potential of hand-shape and hand-texture based biometry is investigated and algorithms are developed. Feature extraction stage is preceded by meticulous registration of the deformable shape of the hand. Alternative features addressing hand shape and hand texture are investigated. Independent component analysis features prove to be the best performing in the identification and verification tasks. It is shown that hand biometric devices can be built that perform reliably for a population of at least one thousand.

Keywords: *Biometry. Identification and verification. Principal component analysis and independent component analysis. Registration for deformable shapes. Distance transform.*

1. INTRODUCTION

Among a number of biometric features that are used for frequent human identification tasks, hand shape and hand shape-based biometry is an emerging new technique, with certain advantages over the more established competitor techniques. Human hand data acquisition is less cumbersome, user-friendlier. Furthermore it is much less susceptible to intrinsic variations and environmental artifacts [1]. In contrast, fingerprint acquisition is usually associated for criminal identification, and is therefore psychologically disturbing. Iris and retinal scans require sophisticated, expensive and more intrusive acquisition systems. Though these alternative techniques have considerable discriminative power, users may not be too keen employ them for daily access control or verification in e-commerce applications. Systems relying on human face or voice identification are more established at present and seem user-friendlier, but suffer from intra-class variations and background/noise elimination problems. Pose, expression and illumination artifacts affect face recognition, while voice recognition is prey to health and/or emotional state of the speaker. However, both face and voice recognition still remains as a very active research area. All in all, hand-based identification/verification systems provide an attractive and growing alternative biometric scheme [2].

There have been a number of previous studies investigating hand-based biometry. Some of these schemes rely solely on geometrical features, others use hand silhouette shape with or without geometric features, still others use extracted palm print lines, and finally there are approaches that combine palm lines with hand shape.

¹ This work was supported by the Bogazici University Research Fund 03A203 and TUBITAK Project 102E027.

Schemes that utilize *geometrical features* of the hand focus on such characteristics as widths of fingers at articulations, finger and palm lengths, finger deviations and the angles of the inter-finger valleys with the horizontal [3, 4, 5, 6]. The number of features varied typically, in the range of 20 to 30, and the acquisition styles also differ, in that some necessitate pegs to position fingers while others were peg-free. These schemes have the advantage that the features are local, and hence accurate hand registration and normalization are not needed. Oden et al. [7] used jointly finger *shape information and geometric features*. The shape information of the individual fingers (but not the whole hand) was extracted via implicit polynomials and the geometric features were joined at the feature fusion stage.

There is also increasing research on *palm print-based* identification/verification systems. One challenging problem with the palm print is the extraction of features such as line structures from the palm print image. For that reason, most of the proposed algorithms [8], [9], [10] require ink markings for obtaining enhanced palm print images. On the other hand, Zhan [8, 11] uses a special scanning system to automatically extract the palm curves from high-quality and well-aligned palm print images. The verification algorithm proposed by Han et al. [12] does not require peg or ink usage, and the palm print features. The extracted features via Sobel and some morphological operators are used with template matching and a back-propagation neural network to determine palm print similarity between hands. In Kong et al. [13] 2-D Gabor filters are implemented to obtain texture information and two palm print images are compared in terms of their hamming distance.

It is possible to use both *hand geometry and palm print information*. Kumar et al. [14] have proposed an approach based on the integration of hand geometry features to hand-based verification. They developed a system where the palm gray level information is fused with the geometrical features of the hand.

In this work we propose a novel hand-based biometric system. The originality of our work resides in two aspects. Our first contribution is the development of a very accurate hand normalization scheme, whereby hand images captured in arbitrary postures are brought to standard finger orientations and overall pose. Such a hand normalization allows the consideration of global features and increases the capacity of the system significantly, that is the size of the subject population correctly identified. As a second contribution, we consider global appearance-based features of the hands for identification and verification tasks. Two of these features are data-driven statistical features, such as principal component analysis and independent component analysis features, and two of them are general features, such as axial radial transform and distance transform. We extract these features either from pure shape information or from the hand texture plus shape information, and we investigate their biometric impact. Finally the size of the enrollment in our work exceeds the populations used in the literature by almost an order of magnitudes. As a byproduct, we have analyzed the similarities between the right and left hands of people, since we capture both of them.

The paper is organized as follows. In Section 2, the segmentation of hand images from its background and the various normalization steps for the deformable hand images are given. Section 3 describes the features we extract from the normalized hand silhouettes and textures. The experimental setup and the classification results are discussed in Section 4 and conclusions are drawn in Section 5.

2. HAND NORMALIZATION

The normalization of hands is of paramount importance for any verification and identification task. This is because the difference between appearance-based features of a hand in different postures, as in Fig. 1, by far exceeds the difference between those features between hands belonging to different subjects. The normalization task involves several consecutive processing steps, namely, segmentation of the hand image from the background, hand rotation and translation, finding the finger axes and tips, removal of ring artifacts, completion of the wrist, estimation of finger pivots (metacarpal-phalanx joints), rotation and translation of fingers to standard orientations. These steps are pictured in the block diagram in Fig. 2 and are described below.

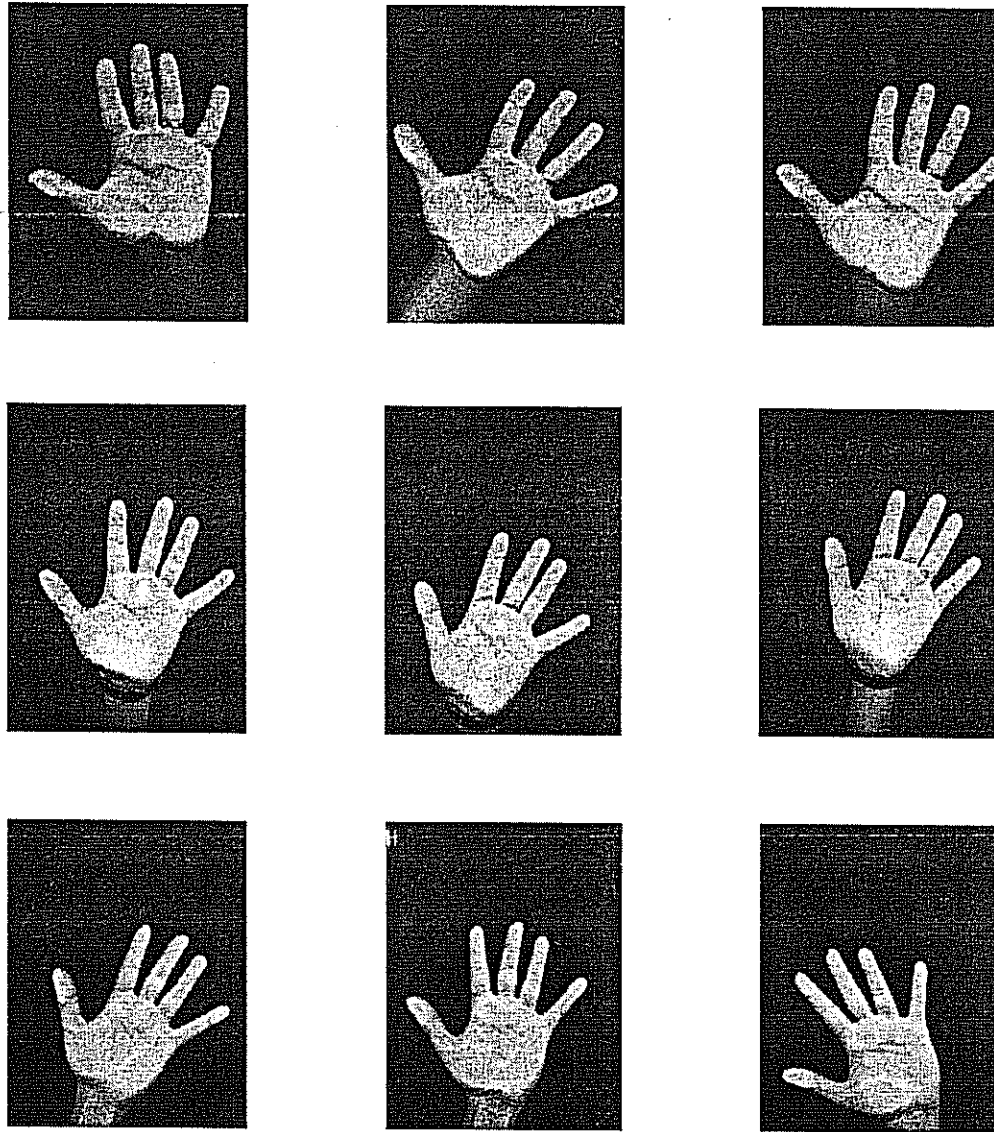


Fig. 1: Images of the left hand of a subject in different postures

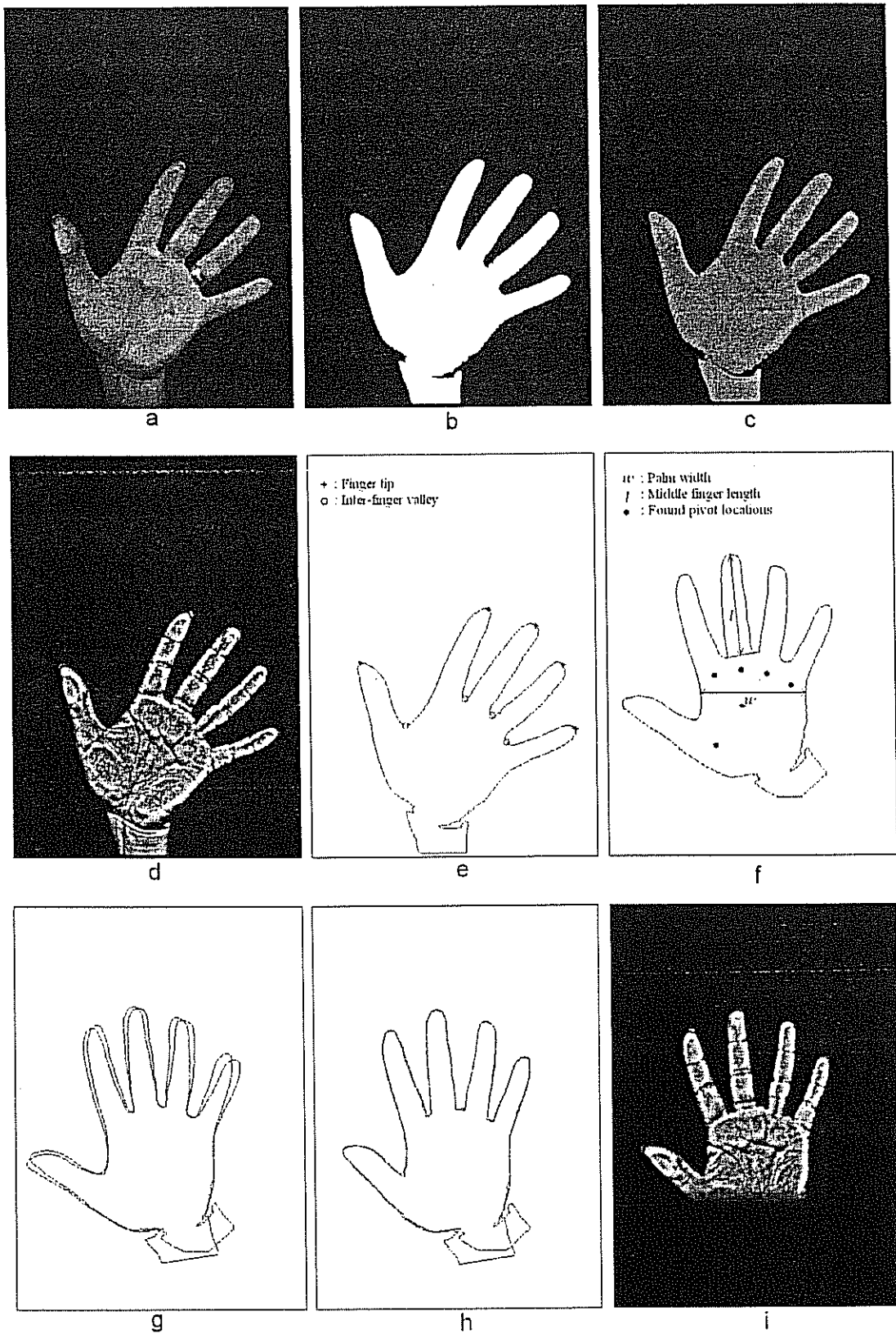


Fig. 2: Processing steps for hand normalization: a) Original hand image; b) Segmented hand image; c) Illumination corrected hand image (ring removed); d) Gray-scale, texture enhanced

hand image; e) Determination of finger tips and valleys; f) Initial global registration by translation and rotation: Middle finger length and palm width for hand image scaling and derivation of the metacarpal pivots; g) Superposed contours taken from different sessions of the same individual with rigid hand registration only; h) Superposed contours taken from different sessions of the same individual after finger orientation normalization; i) Final gray-scale, normalized hand with cosine-attenuated wrist.

2.1 Hand segmentation

Image capturing devices (scanner or digital camera) yield basically a two-class image, with hand texture in the foreground and a darker background. We start with the two-class K-means clustering algorithm, followed by morphological operators to fill in holes and remove isolated foreground debris [15]. Thus size filtering is applied on the connected components of the image to remove spurious components in the background, and then repeat this operation on the video reverse image to remove background holes in the hand region. These two steps satisfactorily separate and extract the hand from the background. Finally we apply a "ring artifact removal" algorithm [16] to correct for any straights or isthmuses caused by the presence of rings. The outcome is a binary image corresponding to the silhouette of the hand. (Fig. 2.a to 2.d).

2.2 Initial hand registration

Hand images are first subjected to a global rotation and translation. This coarse registration involves translation of the centroid of the binary hand mass and its rotation in the direction of the larger eigenvector of the inertia matrix [17]. The inertia matrix can be envisioned as an ellipse fitted to the connected component of the hand object, where the larger eigenvalue determines the hand orientation and corresponds to the major axis of the ellipse. The sign ambiguity in the resulting eigenvector, that is the problem of the sense of rotation, is resolved by considering the relative shift in the centroid of the eroded hand image, since loss of mass would be more rapid in the portion of fingers.

2.3 Finger tips and valleys

Hand extremities, that is the finger tips and the finger valleys form fiduciary landmarks. A robust method to extract these contour extremities consists in computing the radial distances with respect to a reference point around the wrist region. This reference point can be taken as the first intersection point of the major axis (the larger eigenvector of the inertial matrix) with the wrist line. The resulting sequence of radial distances yields minima and maxima corresponding to the sought extremum points. Since the resulting extrema are very stable, the definition of the five maxima (fingertips) and of the four minima are not easily affected by segmentation artifacts on the contour. The radial distance function and a typical hand contour with extremities marked on it are given in Fig. 2.e.

2.4 Wrist completion

The hand contours in the wrist region can be irregular and noisy due to clothing occlusion or due to the different angles that the forearm can make with the platen, and due to the varying pressure exerted on the imaging device. These factors generate different wrist contours in every session, which can adversely affect the recognition rate. In order to create consistently

a smooth wrist contour for every hand image we tried two approaches to synthesize a wrist boundary. The first approach is a curve completion algorithm called the Euler spiral [18, 16]. In the second approach, the hand was guillotined at some latitude, that is, a straight line connected the two sides of the palm. It turned out that tapering off the wrist region with a cosinusoidal window starting from the half distance between the pivot line and the wrist line was the most effective approach in appearance-based recognition and verification tasks (Fig. 2.i). In this operation the wrist line is defined as the horizontal line passing through the estimated thumb pivot on the globally rotated hand image.

2.5 Finger Pivots

As a deformable organ, hand can exhibit a large range of variations due different orientations of fingers. Thus, even after the re-positioning of the hand along its principal orientation axis and on its center of mass, matching scores between hands of the same person captured in different sessions still remains low if the fingers are not exactly aligned. In this context, the posture normalization of the hands consists of reorientation of fingers along predetermined directions and around their metacarpal-phalanx pivots or finger pivots.

Notice that the finger pivots are located inside the palm at the knuckle positions, and are not directly observable. The accurate estimate of these pivots is crucial for reliable finger pose registration. For this purpose we use two pieces of information: first, we know that the pivot must lie on the finger principal axis, which in turn is given by the major eigenvector of the finger's inertial matrix [17]. Secondly, hand anatomy suggests that, once a prototypical hand with correct pivot locations has been established, the pivots of other hands can be approximated fairly accurately by a scaling transformation.

Let $x_{i, proto}$, $i = 1, \dots, 5$ denote the accurately known pivot coordinates of the fingers of a hand, called the proto-hand. When these coordinates are subjected to a scaling transformation based on the actual size measurement of the hand, we obtain $x_{i, scale}$. Finally, each finger pivot is separately corrected as detailed below, yielding $x_{i, actual}$ and hopefully, $x_{i, actual} \approx x_{i, true}$ for each finger index i . Thus the pivot updates go through the modifications $x_{proto} \rightarrow x_{scale} \rightarrow x_{actual}$, which are successive steps to approximate x_{true} , and where the finger indices i have been omitted for simplicity. We use the finger valley between middle and ring fingers as a reliable and easy computable coordinate origin of the actual hand, and denote it as, x_{ref} . In Fig. 3 we mark generically a finger tip and a true (but hidden) pivot point with F and $T = x_{true}$, respectively. The two stages for the estimation of finger pivots are as follows:

Scaling transformation of prototypical pivots:

The positions of the proto-pivots are adapted using the size information of the actual hand and this is realized by a linear transformation $x_{scaled} = Sx_{proto} + x_{ref}$, where S is a scaling transformation. The most reliable hand scale data, with the smallest intra-class spread, is obtained as the length of middle finger (tip-to-valley) and the mean width of the palm, as illustrated in Fig. 2.g. Both of these dimensions can be measured accurately, and they give the notion of the longitudinal and latitudinal scale of the hand. The transform coefficients in the diagonal of the matrix, S , are given by the ratio of the lengths, l , of the actually measured

and prototypical middle fingers, that is, $S_{11} = \frac{l_{actual}}{l_{proto}}$, and the ratio of the respective palm

widths, w , $S_{22} = \frac{w_{actual}}{w_{proto}}$.

Per finger correction of pivots:

The estimated pivots found after scaling transformation may not fall right on the actual finger axis, that is, $x_{scale} \neq x_{true}$. Though the true pivot is not directly observable, an indication of a possible mismatch is given by the fact that x_{scale} does not lie on the finger axis. A trivial solution would be to project x_{scale} orthogonally onto the finger axis, as illustrated in Fig 3.a. Figure 3.a shows the projections of x_{scale} onto fingers axis in two different orientations, yielding respectively, estimates $P_1 = x_{actual}^1$ and $P_2 = x_{actual}^2$ (Notice that the position x_{scale} in this figure is shown grossly in error vis-à-vis x_{true} for illustration purposes). The point we want to make is that this first order approximation via orthogonal projection does not always lead to satisfactory solution especially when there exist large finger orientation variations among different imaging sessions. What we are seeking is a projection from x_{scale} onto the finger axis that causes the least variation in the tip-to-pivot (F-to- x_{actual}) length estimate of the finger. Under the assumption that the fingers can sway between the extreme angles $\alpha_1 = \alpha_{right}$ and $\alpha_2 = \alpha_{left}$ [19] with uniform distribution (angles are measured with respect to the horizontal line as shown in Fig. 3, we can find the projection angle β (not necessarily orthogonal) (see Figure 3.b) that minimizes the variance in the estimate of the *length* when the finger is allowed to sway between its extreme postures.

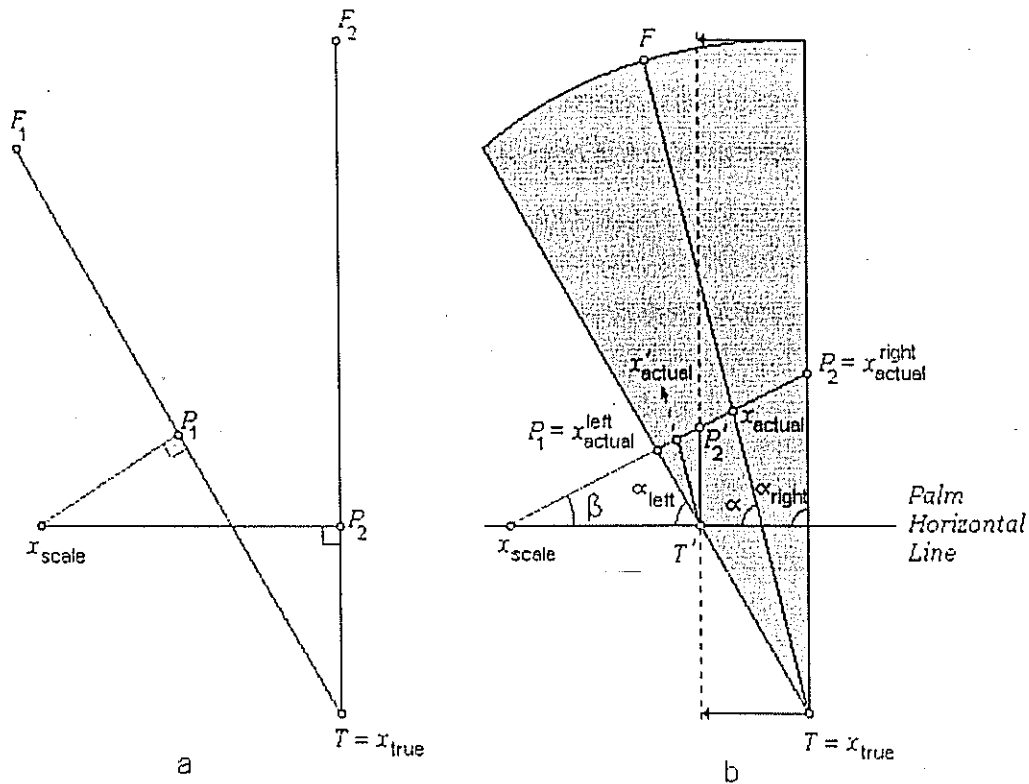


Fig. 3: a) x_{scale} projected orthogonally onto two instances of finger axes resulting in different amounts of error; b) Geometry to determine projection angle β to minimize tip-to-pivot length estimation error. The palm horizontal line becomes a reference for the angle measurement, and it is the horizontal line passing through the scaled pivot on the globally rotated hand image.

In Figure 3.b the sector (α_{left} , α_{right}), within which the finger can sway with respect to the palm, is indicated with shading. We seek a projection angle β for x_{scale} onto the observed finger axis, such that the variance of the estimated finger length can be minimized. The finger length is estimated to be $|Fx_{actual}| \approx |Fx_{true}|$, as measured from the projected pivot x_{actual} to the finger tip F . In Appendix A we show that the variance-minimizing projection angle is

$$\beta = \frac{\pi - \alpha_{right} - \alpha_{left}}{2}.$$

2.6 Enhancement of Texture Component

The appearance data has two components: hand shape, as given by its silhouette, and hand texture. We intend to base hand biometric information not only on the global shape of the hand, but also on its texture. The hand texture is constituted by the wrinkles and crevices in the palm as well as at and around the finger articulations. The inclusion of the hand texture information demands a number of preprocessing steps. First, we render the hand monochromatic by choosing the principal component color with the largest variance. Second, we compensate for the artifacts due to the nonuniform pressure of the hand applied on the platen of the imaging device, which causes spurious gray-level variations due to withdrawal

of the blood from capillaries. An example of this pressure artifact can be observed around the thumb metacarpal mount on the palm. These artifacts are typically low-pass in that they form gross whitish blobs, hence can be easily removed by high-pass filtering. In particular, since on any scan line we search for jump features (palm lines) at relatively higher frequencies, we can extract them by high-pass filtering (Fig. 2.c and 2.d). The image was smoothed with a Gaussian kernel (kernel window size is 20x20 and filter aperture, standard deviation of Gaussian, is $\sigma = 5$) and subtracted from the original image. Thus these two steps of choosing principal component color and then its high-pass filtering constitute the color normalization of the hand.

2.7. Texture Blending with Finger Registration

If texture is to be used as a discriminating feature, one must compensate for the plastic deformation of the knuckle region of the palm ensuing from each finger's rotation around its metacarpal joint, the x_{actual} pivot discussed in Section 2.5. In other words, the texture around the pivot of a rotated finger should be carefully blended with the palm region, from which it was extracted. Since the pivot is some 20% of the finger's length inside the palm (Fig. 4), the whole 120% of the finger segment is extracted (100% visible finger + 20% inside the palm), rotated and re-attached using texture blending. Fig. 4 illustrates attachments of the finger to the palm without texture blending and with texture blending.

The texture blending is effected by a linear combination of the pixel value of the rotated finger and the palm pixel at the corresponding position. The combiner ratio favors palm pixel if the position is deep inside the palm as it will be less subject to rotation. In contrast, it favors finger pixel if it is close to the phalanx. The weights are constituted as distance proportions, where the distances are the corresponding closest distances to the boundaries (Geometry illustrated in Fig. 4). A pixel position, (i, j) , common to both finger and palm regions, has the blended intensity:

$$I_{blend}(i, j) = w_{finger}(i, j)I_{finger}(i, j) + w_{palm}(i, j)I_{palm}(i, j)$$

where the weights are defined as:

$$w_{finger}(i, j) = \frac{d_{finger}(i, j)}{d_{finger}(i, j) + d_{palm}(i, j)} \quad \text{and} \quad w_{palm}(i, j) = \frac{d_{palm}(i, j)}{d_{finger}(i, j) + d_{palm}(i, j)}$$

such that $d_{finger}(i, j)$ and $d_{palm}(i, j)$ are the closest distances of (i, j) 'th pixel to the boundary of the finger and palm segments, respectively. (See inset in Fig. 4).

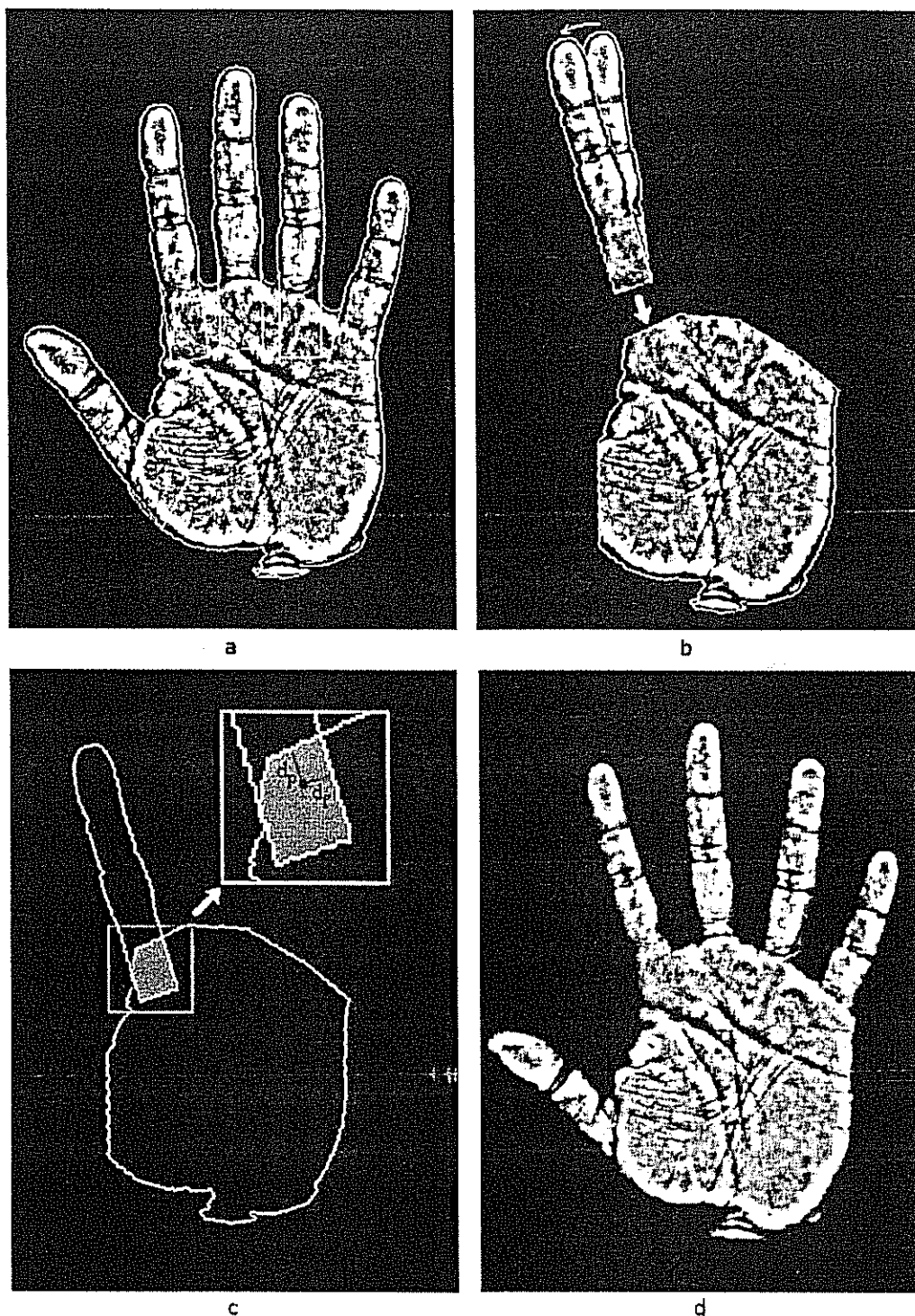


Fig. 4: a) Color normalized and high-pass filtered hand with original pose; thin white borderlines indicate extracted fingers with their extensions inwards the palm. b) Palm and a rotated finger (index) before blending. c) Superposed contours of the palm and of the index finger. Their common region is shown shaded. d) Finger registered Hand with finger registered to the standard orientation and attached with texture blending.

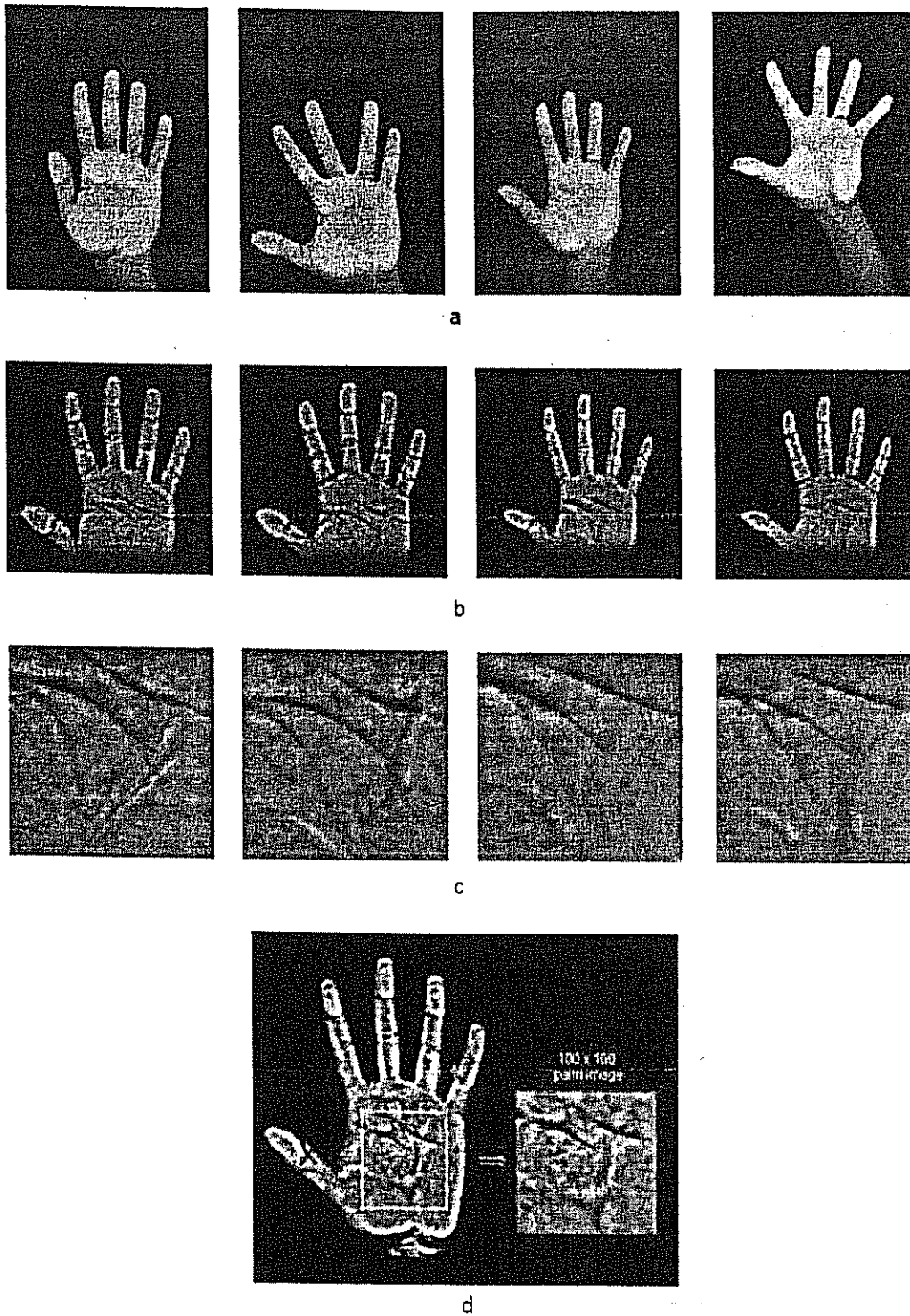


Fig. 5: a) Various hands in the database, some with rings. b) The registered hands with color normalized texture. c) Details of the central portion of the palm texture. d) The 100x100 pixel excerpt from the palm region.

In order to blend judiciously silhouette-based shape information with the texture, we scale the texture information with a weighting factor. In other words, we adjust the contribution of the texture component by tuning its standard deviation. If this tuning parameter is set to 1, we have the original hand appearance data, i.e., texture and shape (texture after color normalization). As this parameter is reduced toward zero, we use more of the silhouette and less of the texture. When the tuning parameter is set to zero, we rely solely on the silhouette data and exclude all texture information.

3. FEATURE EXTRACTION

We have considered comparatively several features that extract either pure shape information or shape information in addition to the texture information. The features that apply to shape only are independent component analysis (ICA) features, principal component analysis (PCA) features, axial radial transform (ART) and distance transform (DT). In a previous study, we had considered weighted Hausdorff distance as well, but we have excluded it from the competition as its performance falls short of ICA [16]. On the other hand, the features extracting information from both shape and texture are the ICA, PCA and ART schemes.

3.1 Principal Component Analysis Features [20, 21]

Let us represent the hand contour vector of length $2n$ as $\mathbf{z} = (c_x(1), \dots, c_x(n), c_y(1), \dots, c_y(n))^T$ where n is the number of points of the hand contour and $(c_x(i), c_y(i))$ are the 2D coordinates of the i th point on the contour. We first establish the nine fiduciary reference points, We first establish eleven fiduciary reference points, consisting of the first and last contour elements, of the five finger tips and the four finger valleys, and then resample the contour data in order to guarantee correspondence between contour elements of all hands. The number of samples between two landmark points is kept equal for all hands; hence the sampling step sizes differ proportionally to the hand size and shape. Fig 6 gives the number of contour elements chosen between landmarks of the hand. Notice that we exclude from the contour the horizontal line above the wrist.

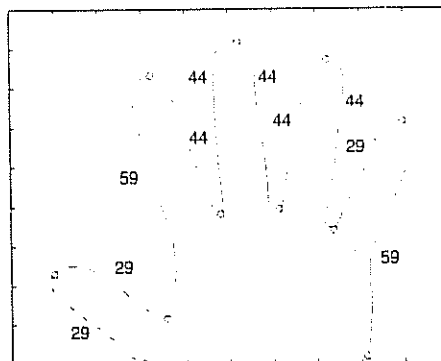


Fig. 6: The number of contour elements chosen between landmarks of the hand

The covariance matrix C of the contour vectors is constructed as $C = \frac{1}{s-1} \sum_{i=1}^s (z_i - \bar{z})(z_i - \bar{z})^T$ using the s sample hands, and where \bar{z} is the mean contour vector.

The eigenvectors, $\{v_i\}$ of the covariance matrix sorted in decreasing order with respect to the corresponding eigenvalues, $\{\lambda_i\}$ model the variations in the training set. If V contains the M eigenvectors corresponding to the largest eigenvalues, then any shape vector in the training set can be approximated as $z \approx \bar{z} + Vb$, where $V = [v_1 \ v_2 \ \dots \ v_M]$ is the selected eigenspace basis set and b is the projection of shape z to this eigenspace, that is, $b = V^T(z - \bar{z})$. The vector b serves as the feature vector of length M of a hand contour in the matching stage.

Fig. 7 shows the effect of varying the first ten modes of b , one at a time. The shapes in this figure are obtained by summing a perturbed n^{th} eigenvector with the mean shape vector. The perturbations are exaggerated intentionally to make the effect of the corresponding mode more visible. A comment is added on top of each figure inset related to the major visible effect of eigenvalue perturbation, though especially for higher eigenvalues multiple effects can occur.




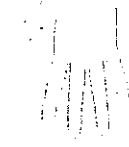






Range of i^{th} coefficient	Modes of variation	Range of i^{th} coefficient	Modes of variation
1 st mode $\pm 2\sqrt{\lambda_1}$	 Scale of hands	6 th mode $\pm 6\sqrt{\lambda_6}$	 Shape of the thumb valley
2 nd mode $\pm 4\sqrt{\lambda_2}$	 Elongation of hands	7 th mode $\pm 6\sqrt{\lambda_7}$	 Shape of the thumb
3 rd mode $\pm 6\sqrt{\lambda_3}$	 Relative lengths of the fingers	8 th mode $\pm 7\sqrt{\lambda_8}$	 Finger widths relative to the palm width
4 th mode $\pm 8\sqrt{\lambda_4}$	 Lengths of index and little fingers and sharpness of the finger tips	9 th mode $\pm 8\sqrt{\lambda_9}$	 Relative positions of little finger with respect to the palm
5 th mode $\pm 6\sqrt{\lambda_5}$	 Sharpness of the finger valleys and thinning of fingers.	10 th mode $\pm 8\sqrt{\lambda_{10}}$	 Shape of the thumb

Fig. 7: Effect of varying the weights of the first ten eigenvectors.

3.1.b. PCA of Hand Appearance:

The hand texture information can also be expressed via the principal component analysis. We have followed Coote's method [21] to decouple texture information from shape. To this effect each image is warped to make its landmarks match with those of some mean shape. Thin-plate splines are used for image warping as in Bookstein [22]. The resulting warped texture information is then expressed as a one-dimensional vector. Finally PCA is applied to the texture vectors of the training hand examples to obtain modes of variation of the texture.

Let b_s the projection of a hand to the shape eigenspace and b_g the projection of the warped hand to the texture eigenspace. The vector $b = [b_s \ b_g]^T$ serves as the feature vector of the hand. The dimensions of both shape and texture eigenspaces are important parameters and are optimized through experimental work. The distance between two hands are computed using a weighted sum of squared differences of feature vector components. When matching is performed using only shape information the distance between two feature vectors, b^k and b^l , is:

$$D(k, l) = \sum_{i=1}^M \frac{1}{\sqrt{\lambda_i}} (b_i^k - b_i^l)^2. \quad (1a)$$

When matching is performed using shape and texture information together, the distance is

$$D(k, l) = \sum_{i=1}^M \frac{1}{\sqrt{\lambda_{si}}} (b_{si}^k - b_{si}^l)^2 + \sum_{i=1}^N \frac{1}{\sqrt{\lambda_{gi}}} (b_{gi}^k - b_{gi}^l)^2. \quad (1b)$$

where $\{b_{si}^k\}_{i=1}^M$ are the M-dimensional shape features of the k^{th} hand, $\{b_{gi}^l\}_{i=1}^N$ are the N-dimensional texture features of the l^{th} hand, and λ_{si} and λ_{gi} are the i^{th} eigenvalues obtained from PCA of shape and texture vectors. The squared difference of each feature is divided by the square root of the feature variance as observed in the training set.

3.2 Independent Component Analysis Features

The Independent Component Analysis (ICA) is a technique for extracting statistically independent variables from a mixture of them and it has found several applications in feature extraction and person authentication tasks [23, 24]. We apply the ICA analysis tool alternatively on binary silhouette images to extract and summarize prototypical shape information as well as on the appearance data, which is shape plus texture.

ICA assumes that each observed hand image, $\{x(k), k = 1, \dots, K\}$ is a mixture of a set of N unknown independent source signals s_i ($i = 1, \dots, N$). Here $\{x(k), k = 1, \dots, K\}$ results from the lexicographic ordering of the image $I(i, j)$ in the scene, which has a total of K pixels. Notice also that, while in the PCA analysis we had considered images resting only within the contours of the hand, in the case of ICA, we consider the total scene image, consisting of its foreground and background. With x_i and s_i , ($i = 1, \dots, N$) forming the rows of the $N \times K$ matrices X and S , respectively, we have the following mixture model:

$$X = AS \quad (2a)$$

$$\hat{S} = Y = WX \quad (2b)$$

where A is the matrix of mixing coefficients. The ICA algorithm finds a linear transformation $\hat{S} = Y = WX$ that minimizes the statistical dependence between the hypothesized independent sources s_i , ($i = 1, \dots, N$).

There exist two possible architectures for ICA, called ICA1 and ICA2 [23], depending on whether one aims for independent basis images or for independent mixing coefficients [23]. In a previous study [16] we found that the ICA2 architecture yielded superior performance. In the ICA2 architecture, the superposition coefficients are assumed to be independent, but not the basis images. In this model we start with the transpose of the data matrix, X^T , and reduce its large dimensionality (typically number of pixels \gg number of images or $K \gg N$) via a PCA stage. Thus we proceed by computing the eigenvectors of the $K \times K$ covariance matrix $C = \frac{1}{N} X^T X$, (actually it suffices to consider the eigenvalues of the much smaller $N \times N$ matrix XX^T) we project the data vectors onto the M ($M \leq N$) largest eigenvalues and obtain the reduced data matrix $X_{reduced}^T$. The source and mixing coefficients are then obtained using the FastICA algorithm [25] using $X_{reduced}^T$ in Eq. 2b. The synthesis of a hand in the data set from the superposition of hand "basis images" is illustrated in Fig. 8. Notice that the columns of the estimated A matrix are the basis in of this architecture, whereas the coefficients in the corresponding column of the estimated source matrix are the independent weights, constituting the feature vector to be extracted.

$$x_i = a_1 \times \hat{S}(1,i) + a_2 \times \hat{S}(2,i) \dots + a_N \times \hat{S}(N,i)$$

$$\text{ICA2 representation for } i^{\text{th}} \text{ hand: } \hat{S}_i = [\hat{S}(1,i), \hat{S}(2,i), \dots, \hat{S}(N,i)]$$

Fig. 8: Hand pattern synthesis using ICA2 basis functions. a_i , $i = 1, \dots, N$ denote the N basis images, while the weighting coefficients $S(n,i)$, $i = 1, \dots, N$ for the hand i are statistically independent.

Similarly, whenever we want to take into account the texture of images in the ICA formalism, we consider the image $I(i, j)$ (foreground + background), $I_{shape}(i, j)$, the binary hand silhouette (foreground set to 1 and background to 0), and finally $I_{texture}(i, j)$, the textured image, also normalized to (0,1) interval. The image fed into the ICA2 algorithm is $I(i, j) = I_{shape}(i, j) + \alpha I_{texture}(i, j)$, where α is the tuning factor $0 \leq \alpha \leq 1$. Matching between hands are performed by comparing ICA basis vectors, for example, the mean-square distance between hands k and l becomes: $D(k, l) = \sum_{i=1}^M (p_i^k - p_i^l)^2$. We have observed that increasing the texture-to-shape ratio α , from zero to 0.2, makes the performances significantly better as expected, since discriminative texture starts also to play a role together with the pure shape information. However, for α beyond 0.5 we see a certain decrease in the overall

identification rate. This can be attributed to the irrelevant skin texture other than palm prints, which becomes to appear more for larger values of α .

The ICA2 algorithm parameters were as follows: the number of pixels in the hand images was $K = 40,000$, the number of subjects was $N = 458$, and finally the number of features, M , used in the ICA2 architecture was 200 as it yielded the best classification result. The texture-shape power ratio was taken as $0.2 \leq \alpha \leq 0.5$.

3.3 Angular Radial Transform Features [26]

Angular radial transform (ART) is a complex transform defined on the unit disk. The basis functions $V_{nm}(\rho, \theta)$ are defined in polar coordinates as a product of two separable functions along the angular and radial directions:

$$V_{nm}(\rho, \theta) = A_m(\theta)R_n(\rho), \text{ where } A_m(\theta) = \frac{1}{2\pi} \exp(jm\theta) \text{ and } R_n(\rho) = \begin{cases} 1 & n = 0 \\ 2 \cos(\pi n \rho) & n \neq 0 \end{cases}$$

Fig. 9 shows real parts of the ART basis functions. As can be observed from this figure, with increasing order n , the basis functions vary more rapidly in the radial direction, whereas the order m expresses the variation in the angular direction.

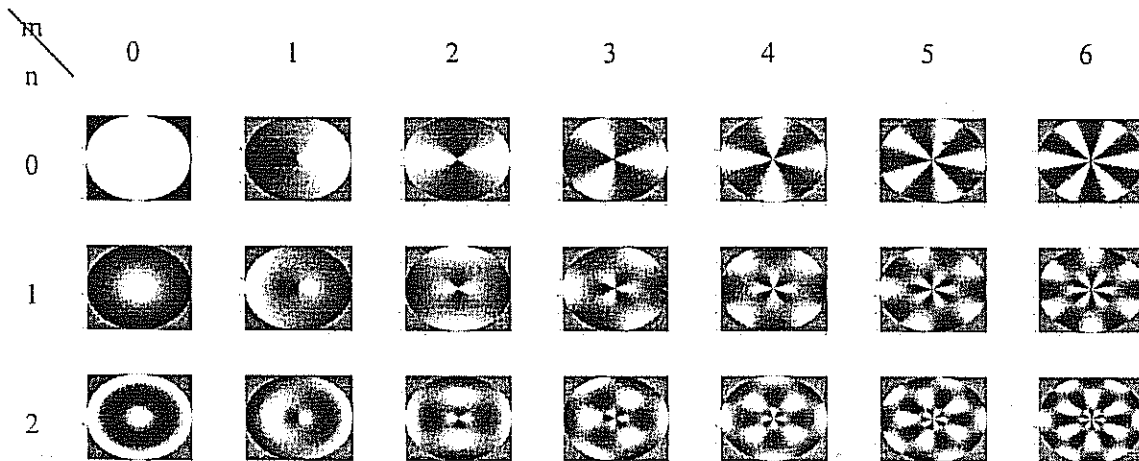


Fig. 9: Real parts of ART basis functions.

The angular radial transform of an image $f(\rho, \theta)$ in polar coordinates is a set of ART coefficients $\{F_{nm}\}$ of order n and m . These ART coefficients can be derived as follows:

$$F_{nm} = \int_0^{2\pi} \int_0^1 V_{nm}^*(\rho, \theta) f(\rho, \theta) \rho d\rho d\theta$$

and a set of $N \times M$ ART magnitude coefficients can be used as features. Notice that, while in shape recognition, the ART coefficients are normalized to $|F_{00}|$ in order to achieve scale invariance; in our work we specifically make use of this coefficient for discriminatory size information. After aligning the hand images and placing them in a fixed-size image plane, we

take the center of the plane as the center of the unit disk. Furthermore, each pixel location is converted to polar coordinates and the radial coordinate is normalized with the image size to have a value between 0 and 1.

We compute the ART coefficients both for the silhouette (binary) hands as well as for the shape plus texture appearance data, which includes palm and finger gray-level details.

3.4 Distance Transform Features [28]

In the shape-based retrieval of objects based on their 2D views, as proposed by Funkhouser et al. [28] first, the distance transform (DT) on the planar shape is calculated, and this is followed by sampling of the DT surface with concentric circles (Fig. 10). The one-dimensional periodic mass (say, 1 for hand region, 0 for background) on the circles is subjected to the discrete Fourier transform (DFT) and a shape signature is obtained by considering a selected number of low-order DFT magnitude coefficients. Thus these features are indexed both by the circle number and DFT coefficient number. As in the case of ART features, the center of the circles is positioned on the center of the plane. The span of radii is constant for all hands. This feature applies obviously only to the shape information, and not to the texture.

Figure 10.a and 10.b show the contour of a hand image and its distance transform. Figure 10.c and 10.d show, respectively, the concentric circles drawn and the resulting profiles. Finally Fig. 10.e illustrates the feature extraction scheme.

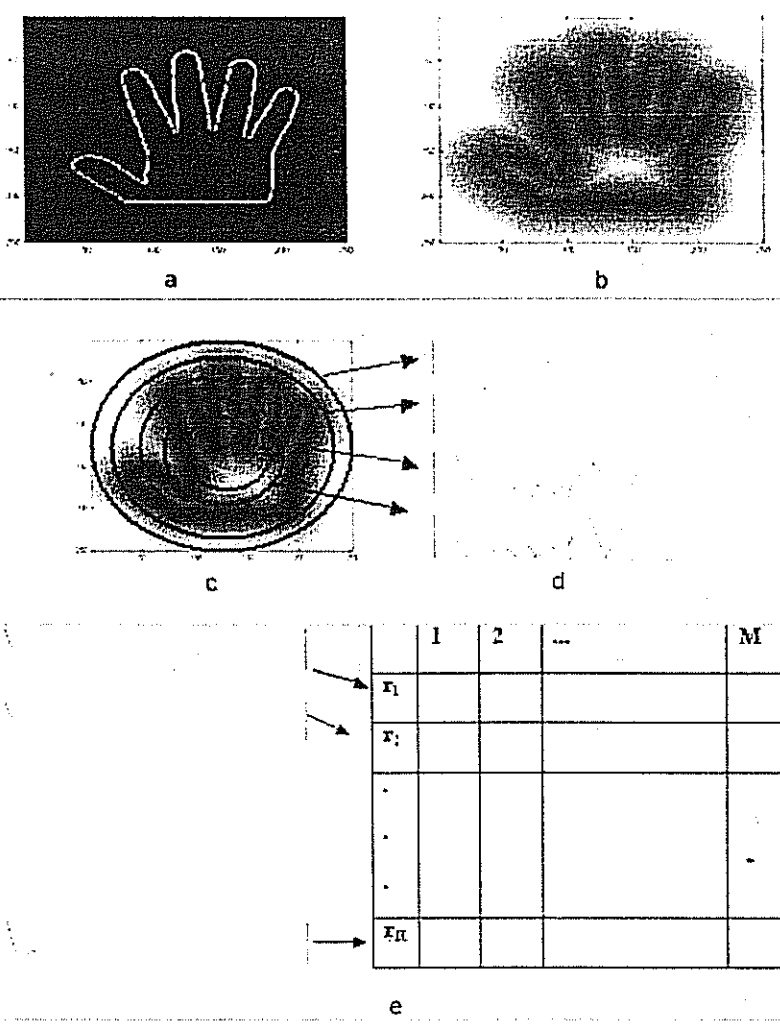


Figure 10: a, b) Contour of a hand and its distance transform defined on the plane. c, d) Concentric spheres on the distance transform and extracted profiles on circles. e) Feature extraction: DFTs of the circular profile of the distance transform function and the selected coefficients.

4. HAND-BASED PERSON RECOGNITION and VERIFICATION PERFORMANCE

Let f_i and f_j be two hand images, and $F_i = \{F_{i,k}\}_{k=1}^K$, $F_j = \{F_{j,k}\}_{k=1}^K$ be generic hand K -dimensional feature vectors. The following hand distance metrics can be used: The L1 norm of the difference of the feature vectors, the L2 norm of the difference of the feature vectors, and the arc-cosine of the angle between feature vectors, respectively calculated as follows:

$$d_1(f_i, f_j) = \sum_{k=1}^K |F_{i,k} - F_{j,k}|, \quad d_2(f_i, f_j) = \sum_{k=1}^K |F_{i,k} - F_{j,k}|^2, \quad d_{\cos}(f_i, f_j) = 1 - \frac{F_i \cdot F_j}{\|F_i\| \|F_j\|},$$

where \cdot

is the inner product notation. Thus hand distances, whether shape-based or shape plus texture-based, are measured in terms of some norm of their feature vectors.

In identification mode, the user does not provide any identity claim, but the system must find out the user's identity from a database of enrolled users. For person identification task, we measure the distance between the test feature vector, F_{test} and all the feature vectors $F_i, i = 1, \dots, N$ in the database belonging to N different subjects. The index of the hand giving the minimum distance is selected as the identity of the input hand, that is, the person i^* is identified if $i^* = \underset{(i)}{\arg \min} \{d(F_{test}, F_i)\}$. Notice that there could be more than one hand image

stored per person, consequently the number of comparisons amounts to the number of subjects times the number of images per subject.

For a person verification task, one must differentiate the "genuine hand" from the "impostor hands" as the user provides her hand image in support of her claimed identity. For this purpose, the distances between the hand of the applicant and all the hands in the database are calculated and the scores compared against a threshold. As we lower the acceptance threshold on $d(F_{test}, F_i)$, the probability of detection increases at the risk of increased false alarm, that is, of accepting an impostor as a genuine.

Table 1 and Fig. 11 show the experimental identification performance of various feature types, as the number of feature components grows, while Table 2 gives the variation of performance scores as the population size increases. Several observations can be made. First, all feature types benefit from an increase in dimensionality, as they start from a modest size of 40. For the population size of 458, all of them seem to have a broad peak at an around dimension 200. Second, ICA features applied on the hand appearance data are the best among all. This holds both in the competition between shape-only recognition and shape-plus-texture recognition. Thirdly, as expected, the addition of texture information improves the recognition performance. However, the improvement remains around 1% for the ICA and PCA features and 2% for the ART feature. Finally, as in Table 2 the population size grows an order of magnitude, from 40 to 458, all features suffer a performance drop ranging from 1 to 3 percent. The only exception is the ICA features on appearance data, where the performance drop is only a meager 0.2%, which again points out to the robustness of the ICA features. These results are obtained by averaging the performance over several randomly chosen subsets of populations, for example, over 30 different choices of 40-tuple subsets etc., as shown in the first row of Table 2.

Table 1: Identification performance of feature types with different number of selected features (population size: 458)

Feature type	Number of features			
	40	100	200	400
ICA_shape	94.32	97.67	98.4	97.31
ICA_appearance	95.41	98.25	99.49	99.36
PCA_shape	96.01	96.97	97.19	97.16
PCA_appearance	96.27	97.91	97.99	97.91
ART_shape	94.18	95.78	95.63	95.05
ART_appearance	95.92	97.38	97.67	97.60
Distance Transform	93.38	95.49	95.71	95.99

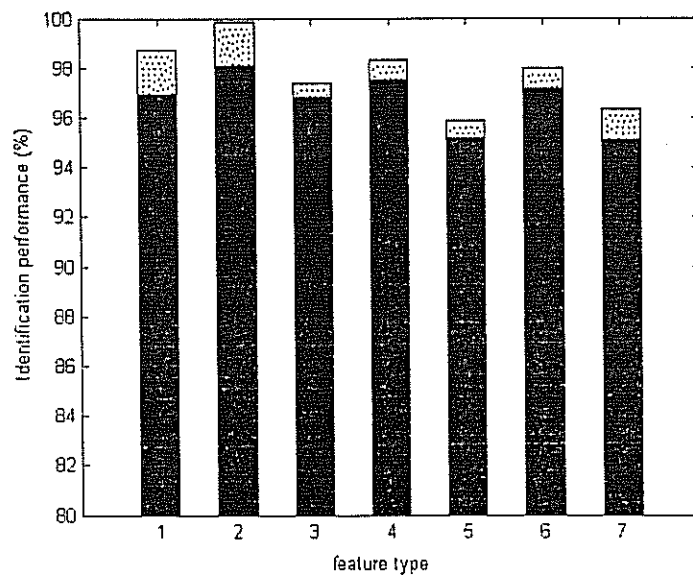


Figure 11: Bar graph showing the performance attainable with every feature set. Every feature set is optimized with respect to the feature dimensions. Maximum population size is used, i.e., 458. Gray whiskers show the region between the best performance and average performance with each feature set.

Table 2: Identification performance of feature sets with increasing population size. (Best feature dimension selected for each feature type)

Feature type	Population size			
	40 (30 random experiments)	100 (12 random experiments)	200 (6 random experiments)	458
ICA_shape	99.19	99.09	98.55	98.40
ICA_appearance	99.68	99.65	99.58	99.49
PCA_shape	98.67	98.69	98.56	97.19

PCA_appearance	99.14	98.89	98.72	97.99
ART_shape	98.72	97.78	97.00	95.78
ART_appearance	99.28	98.72	98.06	97.67
DT	99.17	98.22	96.22	95.99

Table 3 and Fig. 12 present the hand-based verification results. Table 3 gives the correct recognition percentages as a function of population size. The main observations are parallel to those for the identification case, namely: i) the slow performance drop with increasing population size, ii) the superiority of the ICA features, iii) the less than 1% contribution of the texture component to the shape only recognition. Notice that the verification performances seem to increase in row 2 of the Table 3. Actually, while the number of impostors that pass through increases, the population size increase is more rapid than the growth in the false alarm rate.

Table 3: Verification performance with feature sets with increasing population size. (Every feature set optimized with respect to the number of selected features. (Equal Error Rate results))

Feature type	Population size			
	40	100	200	458
ICA_shape	97.49 1 person	98.97 1 person	99.41 1 person	99.45 2.5 person
ICA_appearance	97.94	98.93	99.49	99.74
PCA_shape	98.27	97.80	97.83	97.78
PCA_appearance	98.61	98.50	98.73	98.49
ART_shape	98.29	97.89	97.95	97.91
ART_appearance	97.50	97.28	97.36	97.51
DT	98.31	98.03	98.08	98.34

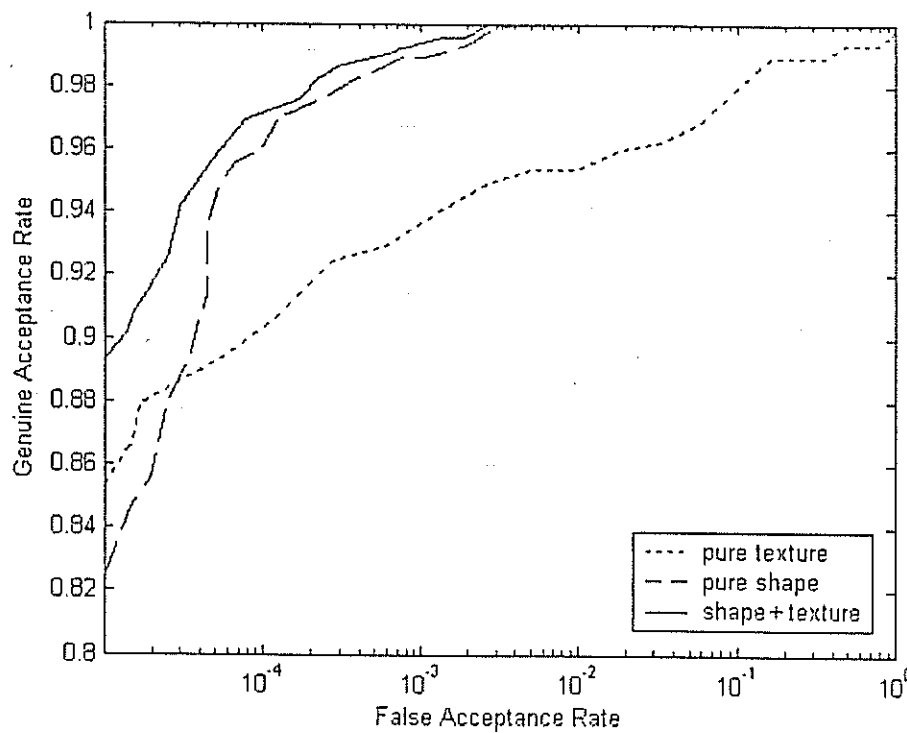


Fig. 12: ROC curves of ICA2 feature sets in hand-based person verification experiments.

Since both right and left hands were measured, we computed the distribution of the distance between the ICA feature vectors between the two hands of the same person. In Fig. 13 we show three distributions: The plots on the left side are the ICA distances between the same side hand (right-right or left-left) of the same subject as observed in different sessions, that is intra-personal distances. The distribution in the middle shows the ICA distances across the two hands of the same subject, that is right-left or left-right. Finally, the distribution on the right is the distances between the same-side hands of different subjects, that is, inter-personal distances. One can observe that there is considerable difference between the two hands of people. The ensuing differences, although less than inter-personal differences, nevertheless is still sufficiently large as not to enable mixed hand biometry, that is identify a subject with the left (right) hand while enrolled with the right (left) hand.

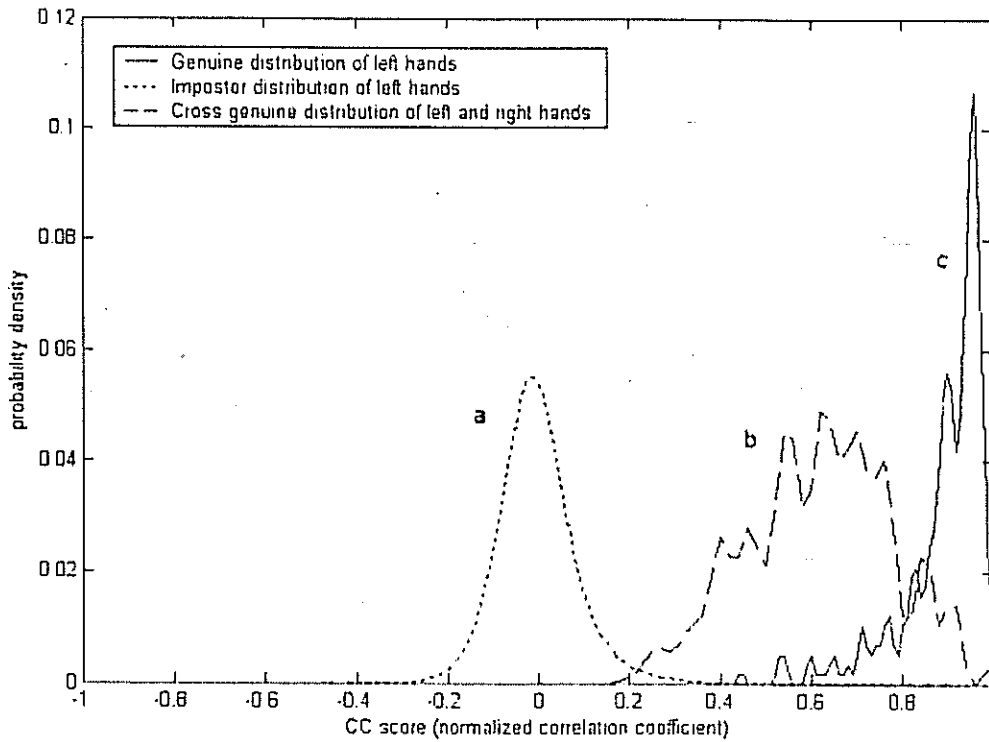


Fig. 13: Distance histograms: a) Intra-distances (left-to-left of the same person). b) Distances between left and right hands of the same person. c) Inter-distances (between left hands of different persons)

Since we have established that the Independent Component Analysis features yield superior performance compared to all others, we have conducted experiments on enlarged database solely with ICA features. Table 4 and 5 give the identification and verification performances obtained by ICA features, when the databases is scaled up to sizes 500, 600 and 756. It is encouraging to observe that results show that the algorithm can handle even larger databases without any noticeable performance degradation.

Table 4: Identification performance of ICA features with larger populations. (Best feature dimension selected for each feature type)

Feature type	Population size		
	500	600	756
ICA_shape	98,73	98,89	98,85
ICA_appearance	99,53	99,61	99,65

Table 5: Verification performance of ICA features with larger populations. (Best feature dimension selected for each feature type)

Feature type	Population size		
	500	600	756
ICA_shape	99,69	99,68	99,72
ICA_appearance	99,86	99,83	99,86

Finally, we compared the performance of our algorithm with the available hand identification results in the literature. In Table 6, we describe briefly the types of features used, the population size and the comparative performance figures. A more detailed summary of alternative algorithms is given in a companion paper [16]. We can observe that our algorithm outperforms all the available algorithms in the literature and its performance remains stable over a large plateau of population sizes.

Table 6: Comparison of the identification performance of our algorithm vis-à-vis the performance of competitor algorithms in the literature.

Reference	Algorithmic features	Population size	Performance of competitor algorithms	Performance of our algorithm
Reillo, Avila, Marcos [5]	Finger widths at different latitudes, finger and palm heights, finger deviations and angles of the inter-finger valleys.	20	97.0	99.70
Öden, Erçil and Büke [7]	Hand geometry via finger widths at various positions, palm size and finger geometry via fourth degree implicit polynomials.	35	95.0	99.68
Bulatov, Jambawalikar, Kumar, Sethia [6]		70	98.5	99.67
Kumar, Wong, Shen, Jain [14]	Hand geometry and palmprint information with decision fusion.	100	99.1	99.65
-		756	-	99.65

5. CONCLUSION

An algorithm has been presented for hand-based biometry in identification and recognition tasks. The very critical importance of a proper registration that takes into account the deformations not only in the shape but also in the texture of the hand is shown. Note that although the hand undergoes strong processing for geometrical correction, the texture is not affected since the transformations are rigid. The only instance when the texture might be affected is when the fingers are rotated around their metacarpal joints. But for this case, we use the texture interpolation and correction scheme, as detailed in Fig. 4. Several feature schemes are comparatively evaluated, and the Independent Component Analysis features are found to perform uniformly superior to all other features considered. The attained

performance of 99,65% correct identification, and of 99,86% EER verification for a population of 756 subjects is very encouraging and it indicates that hand-biometric devices can respond to the security requirements for populations of several hundreds.

The work is continuing to assess the performance in a time lapse of the order of months as well as for even larger populations. The complementary role of hand biometry in a multimodal is also being investigated.

APPENDIX

In this appendix we derive the optimum angle β that minimizes the variance of the positions of x_{actual} . We can simplify the geometry of the problem by shifting one of the bounding axes, say the right one, towards the left, such that they intersect at the point T' as shown in Fig. 5.b. If we work with the triangle $P_1TP_2 = x_{actual}^{left} x_{true} x_{actual}^{right}$, the length estimation error becomes $|x_{actual} x_{true}|$ for any given projection angle β . On the other hand, if we work with the triangle $P_1T'x'_{actual}$, the length estimation error becomes proportional to $|T'x'_{actual}|$. Notice that the line $T'x'_{actual}$ is parallel to the line Tx_{actual} and also the two triangles are similar, hence minimizing $\overline{x_{actual}T}$ is tantamount to minimizing $\overline{x'_{actual}T'}$, since errors calculated from similar triangles are proportional and we are only interested in minimizing the relative error.

If we focus on the 'random' triangle $x_{scale}x'_{actual}T'$, from the sine rule we have:

$$\frac{\overline{x'_{actual}T'}}{\sin \beta} = \frac{\overline{x_{scale}T'}}{\sin(2\pi - \alpha - \beta)} = \frac{\overline{x_{scale}T'}}{\sin(\alpha + \beta)}$$

where $c = \overline{x_{scale}T'}$ is a deterministic constant, which is the distance of the scaled pivot, x_{scale} , to the leftmost position of the finger axis. With this relation we can formulate the variance of $x'_{actual}T' = \frac{|x_{scale}T'| \sin \beta}{\sin(\alpha + \beta)}$, as a function of β . Taking the expectations with respect to the random variable α , we have:

$$Var\{|x'_{actual}T'|\} = \frac{c^2 \sin^2 \beta}{(\alpha_2 - \alpha_1)} \int_{\alpha_1}^{\alpha_2} \frac{1}{\sin^2(\alpha + \beta)} d\alpha - \left[\frac{c \sin \beta}{(\alpha_2 - \alpha_1)} \int_{\alpha_1}^{\alpha_2} \frac{1}{\sin(\alpha + \beta)} d\alpha \right]^2$$

This expression does not have a closed form minimum, except for the trivial solution $\beta = 0$. However, gradient descent schemes have shown the minimum is found at around the mean finger orientation, i.e. $\beta = \frac{\pi - \alpha_1 - \alpha_2}{2}$. Thus we use this value of β to update the pivot position to x_{actual} , and then the finger is rotated to its predetermined orientation angle with respect to this pivot.

REFERENCES

- [1] A.K. Jain, A. Ross, S. Prabhakar, "An Introduction to Biometric Recognition", *IEEE Trans. Circuits and Systems for Video Technology*, 14(2004), 4-20.

- [2] R.L. Zunkel, "Hand Geometry Based Verification", pp. 87-101, in Biometrics, Eds. A. Jain, R. Bolle, S. Pankanti, Kluwer Academic Publishers, 1999.
- [3] A.K. Jain, A. Ross and S. Pankanti, "A prototype hand geometry based verification system", Proc. of 2nd Int. Conference on Audio- and Video-Based Biometric Person Authentication, 1999, pp. 166-171.
- [4] A.K. Jain and N. Duta, "Deformable matching of hand shapes for verification, Proc. of Int. Conf. on Image Processing, 1999, Kobe, Japan, 857-861.
- [5] R. Sanches-Reillo, C. Sanchez-Avila, and A. Gonzalez-Marcos, "Biometric Identification through Hand Geometry Measurements," *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 22 (2000), 1178-1171.
- [6] Y. Bulatov, S. Jambawalikar, P. Kumar, and S. Sethia. Hand Recognition Using Geometric Classifiers. In proc. of the *1st International Conference on Biometric Authentication (ICBA)*, 2004, Hong Kong, China.
- [7] C. Öden, A. Erçil and B. Büke, "Combining implicit polynomials and geometric features for hand recognition", *Pattern Recognition Letters*, 24 (2003), 2145-2152.
- [8] D. Zhang and W. Shu, "Two novel characteristics in palmprint verification: datum point invariance and line feature matching", *Pattern Recognition*, 32 (1999), 691-702.
- [9] J. You, W. Li and D. Zhang, "Hierarchical palmprint identification via multiple feature extraction", *Pattern Recognition* 35(2002), 847-859.
- [10] N. Duta, Anil K. Jain and K. V. Mardia, "Matching of palmprints", *Pattern Recognition Letters* 23 (2002), 447-485.
- [11] D. Zhang, W. K. Kong, J. You, and M. Wong, "Biometrics - Online palmprint identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2003), 1041-1050.
- [12] C.C. Han, H. L. Cheng, C. L. Lin, and K. C. Fan, "Personal authentication using palm print features," *Pattern Recognition* 36(2003), 371-381.
- [13] A. W. Kong, D. Zhang and W. Li, "Palmprint feature extraction using 2-D Gabor filters", *Pattern Recognition* 36(2003), 2339-2347.
- [14] A. Kumar, D. C. M. Wong, H. C. Shen and A. K. Jain, "Personal Verification Using Palmprint and Hand Geometry Biometric", Proc. of Audio-Visual Based Person Authentication, 2003, Guildford, UK, 668-678.
- [15] P. Soille, "Morphological Image Analysis – Principles and Applications," *Springer-Verlag*, 1999.
- [16] E. Yoruk, E. Konukoğlu, B. Sankur, J. Darbon, "Shape-Based Hand Recognition," to appear in *IEEE Image Processing*, 2006. Available at <http://busim.ee.boun.edu.tr/sankur/>
- [17] A. R. Weeks, *Fundamentals of Electronic Image Processing*, pp. 466-467, SPIE Press, 1996.
- [18] B. B. Kimia, I. Frankel, and A. M. Popescu, "Euler spiral for shape completion," *International Journal of Computer Vision* 54 (2003), 157-180, 2003.
- [19] Lin, J., Y. Wu, and T. Huang, "Modeling Human Hand Constraints", *In Proceedings of the Workshop on Human Motion (Humo2000)*, Austin, TX, December 2000.

- [20] T. F. Cootes and C. J. Taylor, "Statistical models of appearance for computer vision", Technical Report, University of Manchester, 2000.
- [21] T. F. Cootes, G. J. Edwards, and C.J. Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23 (2001) 681-685.
- [22] F.L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(1989), 567-585.
- [23] H. K. Ekenel, B. Sankur, Feature selection in the independent component subspace for face recognition, *Pattern Recognition Letters*, 25 (2004), 1377-1388.
- [24] B. A. Draper, K. Baek, M. S. Bartlett, and J. R. Beveridge, "Recognizing faces with PCA and ICA," *Computer Vision and Image Understanding*, 91 (2003), 115-137.
- [25] A. Hyvarinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Networks* 13 (2000), 411-430.
- [26] S. Jeannin, "Mpeg-7 Visual part of eXperimentation Model Version 9.0", in ISO/IEC JTC1/SC29/WG11/N3914, 55th Mpeg Meeting, Pisa, Italia, Jan. 2001.
- [27] W.-Y. Kim and Y.-S. Kim, "A new region-based shape descriptor", in TR 15-01, Pisa, Dec. 1999.
- [28] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman and D. Dobkin, "A Search Engine for 3D Models", *ACM Transactions on Graphics*, vol. 22, No. 1, Jan. 2003.