

# Destek Vektör Makinesi Kullanarak Bağımsız Bileşen Tabanlı 3B Nesne Tanıma

## Independent Component Based 3D Object Recognition Using Support Vector Machines

O. G. Sezer, A. Erçil and M. Keskinöz

Mühendislik ve Doğa Bilimleri Fakültesi, Sabancı Üniversitesi  
ogsezer@su.sabanciuniv.edu, aytulercil@sabanciuniv.edu, keskinoz@sabanciuniv.edu

### Özetçe

Bu makalede, zaman ve belleğin bileşimsel (kombinezon) patlaması olmaksızın yüksek dereceden istatistikleri kullanan bir nesne tanıma tekniği önerilmektedir. Önerilen yöntem literatürdeki iki gözde yöntem olan Bağımsız Bileşen Analizi (BBA) ve Destek Vektör Makinesi (DVM)'nin kaynaşımıdır. İmgelerdeki artıklığı gidermek ve her imge için daha düşük boyutlu öznitelik vektörleri elde etmek için BBA'yı ve sonrasında BBA adımından gelen bu öznitelik vektörlerini sınıflandırmak için DVM'nin kullanılması önerilmektedir. Coil-20 veritabanı ve kendi ürettiğimiz bir 2B üretim nesnelere veritabanı için deney sonuçları verilmiştir.

### Abstract

In this paper, we propose an object recognition technique using higher order statistics without the combinatorial explosion of time and memory complexity. The proposed technique is a fusion of two popular algorithms in the literature, Independent Component Analysis (ICA) and Support Vector Machines (SVM). We propose to use ICA to reduce the redundancy in the images and obtain some feature vectors for every image which has lower dimensions and then make use of SVM to classify these feature vectors coming from the ICA step. Experimental results are shown for Coil-20 and an internally created database of 2D manufacturing objects.

### 1- Giriş

İnsan gözlemciler için basit bir işlem olan nesne tanıma, Yapay Görme alanında birçok araştırmanın konusunu oluşturmuş ve makine tabanlı birçok nesne tanıma sisteminin temel bileşeni olmuştur. Literatürde çok sayıda değişik yaklaşım önerilmiştir. Yapay görme alanında biçim eşleme üstüne ayrıntılı bir inceleme [1, 2, 3]'te bulunabilir. Genelleme yapmak gerekirse, iki tür yaklaşım vardır: 1) Kenar elemanları veya eklem gibi çıkarılan özniteliklerin uzaysal düzenlemelerinin kullanımını içeren, öznitelik tabanlı yaklaşımlar, 2) piksel parlaklığının daha doğrudan kullanımını yapan, parlaklık tabanlı yaklaşımlar.

Parlaklık tabanlı (veya görünüm tabanlı) yöntemler kapatan çevritin biçimine veya çıkarılan diğer özniteliklere odaklanmak yerine, nesnenin görünen kısmındaki gri değerlerini kullanır. BBA veya Temel Bileşen Analizi (TBA)

gibi altuzay yöntemleri yüz tanıma [4] ve robot görme sistemlerine başarıyla uygulanmıştır. Literatürde BBA ve TBA arasında karşılaştırma yapan birçok çalışma da vardır; BBA yüksek dereceden istatistikler kullandığı için ikinci dereceden istatistikleri kullanan TBA'dan daha başarılı olması beklenir. Bu çalışmalardan biri Sahambi ve arkadaşlarına ait [5], BBA ve TBA'nın başarımlarını nesne tanıma işinde karşılaştıran çalışmadır. Bağımsız ve temel bileşen analizini Coil-20 veritabanına değişik örnekleme açılarıyla (25° ve 50°), sınıma nesnelere bağımsız ve temel bileşenlerinin katsayılarıyla eğitim aşamasında kullanılan nesnelere bağımsız ve temel bileşenlerinin katsayılarının ortalamaları arasındaki farkı bulmak için Öklid uzaklığını kullanarak uygulamışlardır. Bu yöntem için Coil-20 veritabanında tanıma oranları %70-80 civarındadır.

Pontil ve meslektaşları tarafından [6] nesne tanıma için Destek Vektör Makinesi (DVM) kullanan başka bir yaklaşım önerilmiştir. DVM'nin yüksek boyutlu veriyi sınıflandırmadaki gücü bilindiği için, Pontil ve arkadaşları boyutluluğu azaltmak için herhangi bir öznitelik çıkarma tekniği kullanmamışlardır. Hesaplamaların hızını artırmak için özgün hali 128x128'lik olan imgeleri 32x32'ye ölçeklemler ve imgeleri yüksek boyutlu bir uzaydaki noktalar olarak düşünüp açıyı kestirmeksizin tanıma yapmışlardır. Ancak, algoritma 32x32'lik imgeleri veri noktaları olarak aldığı için hesap yükü fazladır.

Makalede önerilen yöntem, BBA'nın bir imgedeki piksel parlaklık değerleri arasındaki doğrusal olmayan ilişkileri de içeren yüksek dereceden bağımlılıkları modellemedeki üstünlüğü ile DVM'nin bir en iyi ayıran altdüzlem yaratmadaki üstünlüğünü birleştirmektedir.

Makalenin düzeni şu şekildedir: Bağımsız bileşen analizi (BBA) hakkında önbilgi Bölüm 2'de verilmiştir; destek vektör makinesi (DVM)'nin altında yatan fikir Bölüm 3'te kısaca açıklanmıştır; bu makalede kullanılan yaklaşım Bölüm 4'te özetlenmiştir ve deney sonuçları Bölüm 5'te verilmiş; hemen ardından Bölüm 6'da tartışma ve sonuçlar eklenmiştir.

### 2- Bağımsız Bileşen Analizi (BBA)

BBA kör kaynak ayırması yapabilen bir yöntemdir. Hem kaynak işaretlerinin hem de bunların nasıl karıştıklarının bilinmemesinden dolayı ayırıştırma kör olarak yapılmaktadır. BBA algoritması öyle bir doğrusal kordinat sistemi bulur ki elde edilen işaretler istatistiksel olarak birbirinden bağımsız olur. TBA gibi ilinti tabanlı dönüşümlerden farklı olarak, BBA

işareti sadece ilintisizleştirmekle kalmaz (2. derece istatistik) ayrıca yüksek derece istatistiksel bağımlılığı da azaltır.

BBA yöntemi birbirinden bağımsız ve Gauss dağılımına sahip olmayan faktörlerin bulunmasında klasik yöntemlere göre çok güçlü bir araçtır. BBA modelinde saklı faktörler ya da bağımsız bileşenler doğrusal veya doğrusal olmayan şekilde bilinmeyen bir karıştırma mekanizması ile karıştırılarak gözlem verilerini oluştururlar. Temel doğrusal karışım modeli matematiksel olarak şu şekilde ifade edilebilir [7]:

$$\mathbf{x}=\mathbf{A}\mathbf{s} \quad (1)$$

Burada  $\mathbf{x}$  gözlem verileri  $x_j$ 'lerden oluşan gözlem vektörü,  $\mathbf{s}$  kaynak vektörü;  $\mathbf{A}$  da karıştırma matrisidir. Amaç bilinmeyen  $\mathbf{A}$  ve  $\mathbf{s}$  değerlerinin gözlem vektörü  $\mathbf{x}$  kullanılarak kestirilmesidir. Kaynak dağılımları bilinmemekte sadece gauss olmadıkları ve birbirlerinden istatistiksel olarak bağımsız oldukları varsayılmaktadır. [7].

Modeldeki varsayımlar kullanılarak;  $\mathbf{A}$  matrisi kestirilir daha sonra da kaynaklar aşağıdaki gibi hesaplanır:

$$\mathbf{s}=\mathbf{W}\mathbf{x} \quad (2)$$

Burada  $\mathbf{W}$  matrisi  $\mathbf{A}$  matrisinin tersidir. Bizim uygulamamızda bu kaynaklar bağımsız bileşenlerin katsayılarını verir.

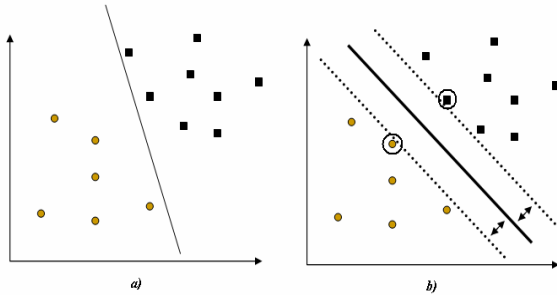
### 3- Destek Vektör Makinesi (DVM)

Destek vektör makinesi örüntü tanımda gittikçe gözde hale gelen bir genelleme yöntemidir. DVM iki sınıflı veriye ait bir nokta kümesini ayıran bir en iyi altdüzlem bulmaya çalışır.

Bu bölümde, ilk olarak doğrusalca ayrılabilir veri basit durumu açıklanmış, sonra destek vektör kavramı ve ayrılabilen veri için genel durum ayrıntılı olarak anlatılmıştır.

#### 3.1. Doğrusalca Ayrılabilir Veri

Doğrusalca ayrılabilir veri durumunda herbiri  $y_i \in \{-1,1\}$  ile gösterilen sınıflardan birine ait olan,  $\mathbb{R}^n$ 'in elemanı olan  $x_i$ 'ler,  $i=1,\dots,N$ , kümesi  $\mathcal{S}$  verilmiştir. Amaç, veri kümesini verilen etiketlere göre bir altdüzlemle ayırıp, aynı sınıfa ait bütün veri noktalarını altdüzlemin aynı tarafında bırakmaktır.



Şekil 1. İki sınıflı veriye ayıran bir altdüzlem (a), en iyi altdüzlem ve marjları (b)

Bir  $x_i$ 'ler veri kümesi, eğer  $i=1,\dots,N$  için

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \quad (3)$$

koşulunu sağlayan bir  $\mathbf{w}$  varsa doğrusalca ayrılabilir. Burada  $(\mathbf{w}, b)$  aşağıdaki denkleme sahip bir altdüzlem tanımlanmaktadır;

$$\mathbf{w} \cdot \mathbf{x} + b = 0 \quad (4)$$

ve ayıran altdüzlem olarak adlandırılır ve denklem (3)'teki çarpım, veri noktası ile etiketinin altdüzlemin aynı tarafında olmasını belirler. Şekil 1(a)'da iki sınıflı ayıran böyle altdüzlemlerden biri gösterilmiştir. Tabii ki, iki sınıflı ayıran sonsuz sayıda altdüzlem vardır. Şimdi altdüzlemden bir veri noktası  $x_i$ 'ye olan bir  $d_i$  uzaklık ölçüsü tanımlayalım:

$$d_i = \frac{\mathbf{w} \cdot \mathbf{x}_i + b}{\|\mathbf{w}\|} \quad (5)$$

Eğer denklem (3) ve (5) birleştirilirse aşağıdaki denklemi elde ederiz:

$$y_i d_i \geq \frac{1}{\|\mathbf{w}\|} \quad (6)$$

Burada  $1/\|\mathbf{w}\|$ 'nin  $(\mathbf{w}, b)$  altdüzlemi ile  $x_i$  veri noktası arasındaki uzaklığın bir alt sınırı olduğunu gözlemliyoruz. Dolayısıyla, eğer veri noktaları kümesi içinde bu alt sınırı eşitlikle sağlayan bir  $x_i$  noktası bulabilirsek bu, altdüzleme en yakın noktayı bulduğumuz anlamına gelir. Aynı zamanda, en iyi altdüzlemi elde etmek için altdüzlemle en yakın veri noktası arasındaki uzaklığı enbüyütmemiz gerekir ve bu marj gerçekten de  $1/\|\mathbf{w}\|$  değerine karşılık gelir. Dahası, bu marjı enbüyütmek  $\mathbf{w}$ 'nin düzgesi  $\|\mathbf{w}\|$ 'i enküçültmeye karşılık gelir.

Bu marjı enküçültmek ile elde edilen  $\underline{\mathbf{w}}$  ve  $\underline{b}$  kullanarak sınıflandırma problemi,  $\mathbf{x}$  gelen veri olmak üzere,

$$\underline{\mathbf{w}} \cdot \mathbf{x} + \underline{b} \quad (7)$$

yukardaki denklemin işaretini bulmaya indirgenir. Böylelikle, karar fonksiyonumuz;

$$f(\mathbf{x}) = \text{sign}(\underline{\mathbf{w}} \cdot \mathbf{x} + \underline{b}) = \text{sign}(\langle \underline{\mathbf{w}}, \mathbf{x} \rangle + \underline{b}) \quad (8)$$

olur.

#### 3.2. Doğrusalca Ayrılabilen Veri

Bilindiği gibi, gerçek hayatta gürültü ve veri sınıflarının doğrusal olmaması nedenleriyle doğrusalca ayrılabilir veri genelde bulunmaz. DVM'yi bu tip veriye uygulamak için, doğrusalca ayrılabilir veri durumunda sunulan kısıtları bir yapay  $\xi_i$  değişkeni tanımlayarak esnetebiliriz.

Şimdiki problemimiz iki sınıflı veriye, aynı sınıfa ait verinin mümkün olan en büyük kısmını altdüzlemin aynı tarafında bırakacak şekilde ayıran altdüzlemi bulmaktır. Yukarıdaki denklemin çözümü bize bu en iyi ayıran altdüzlemi verir. [6, 8]

Daha önce olduğu gibi ikiliği kullanarak,  $\mathbf{w}'$ yi  $y_i x_i$ 'lerin doğrusal bileşeni olarak gösterirsek karar fonksiyonu şuna dönüşür:

$$\begin{aligned} f(\mathbf{x}) &= \text{sign}(\langle \mathbf{w}, \mathbf{x} \rangle + b) \\ &= \text{sign}\left(\sum_{i=1}^N \alpha_i y_i \langle \mathbf{x}_i, \mathbf{x} \rangle + b\right) \end{aligned} \quad (9)$$

Karar fonksiyonumuzdaki iç çarpımın (çekirdeğin) yerine uygun bir doğrusal olmayan çekirdek konulup aşağıdaki şekilde gösterildiği gibi verinin daha yüksek boyutlu bir uzaya taşınabileceğini gözleyin. [6, 8]

Böylece  $\langle \mathbf{x}_i, \mathbf{x} \rangle$ 'in yerine  $\langle Q(\mathbf{x}_i), Q(\mathbf{x}) \rangle$  koyuyoruz; burada  $Q(\cdot)$  çekirdek fonksiyonudur. Literatürde değişik uygulamalar için çok çeşitli çekirdekler vardır ama bu çalışmada üç temel çekirdeğin başarımları karşılaştırılmıştır:

- Doğrusal:  $K(x_i, x_j) = x_i^T x_j$ .
- Çokterimli:  $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d$ ,  $\gamma > 0$ .
- Radyal taban fonksiyonu (RTF) :  $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$ ,  $\gamma > 0$

Sonuç olarak, karar fonksiyonu şöyle olur:

$$f(x) = \text{sign}\left(\sum_{i=1}^N \alpha_i y_i K(x_i, x) + b\right) \quad (10)$$

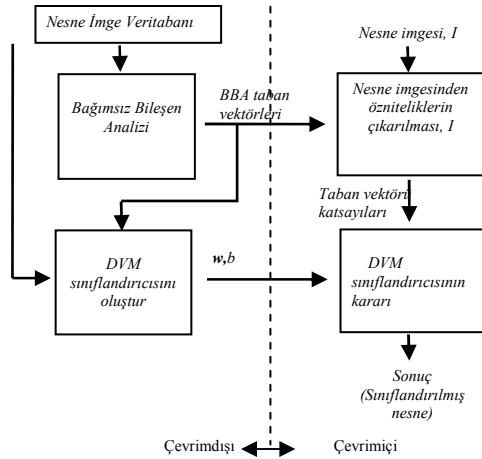
#### 4- Yöntem

Makalede kullanılan nesne tanıma yöntemi Şekil 2’te gösterildiği gibi özetlenebilir. Çevrimdışı aşamada, bağımsız bileşenler elde edilir ve DVM sınıflandırıcısı oluşturulur. Çevrimiçi aşamada, çevrimdışı aşamada üretilen bilgi, nesne imgesinden öznelik çıkarmakta kullanılır ve sonra nesnenin kimliğini belirlemek için DVM sınıflandırıcı parametreleriyle çok sınıflı karar verme uygulanır.

Çevrimdışı aşamada, BBA algoritması eğitim kümesinde bulunan her nesnenin imgeleri kullanılarak eğitilir. BBA’nın ilk aşamasında, imgelerdeki toplam enerjinin en az %90’ı korunacak şekilde temel bileşen analizi kullanılarak boyutluluk azaltılır. Böylece, boyutları azaltmak BBA algoritmasının daha hızlı çalışmasını sağlar.

Çevrimdışı aşamadaki DVM kısmı eğitim verisinden sınama verisi için genellemeleri elde etmeyi amaçlar. Buradaki eğitim verisi BBA aşamasından gelen ayrıştırma matrisi ( $W$ ) ile veritabanındaki nesne imgelerini çarpma ile elde edilen bağımsız bileşenlerin (veya taban vektörlerinin) katsayılarıdır. Son olarak, DVM sınıflandırıcısı eğitim kümesindeki nesne imgeleri ile eğitilir.

Çevrimiçi aşamada, ilk olarak gelen nesnenin imgesi sisteme verilir, sonra verilen nesnenin BBA öznelik vektörleri (bağımsız bileşenlerin (BB’ler) katsayıları) çıkarılır ve son olarak DVM sınıflandırıcısı nesnenin kimliğine karar verir.



Şekil 2. Bu projedeki nesne tanıma sisteminin blok çizeneği

#### 5- Deney Sonuçları

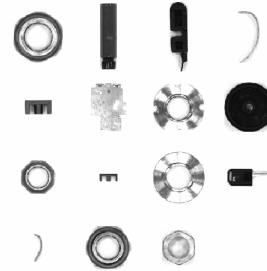
Önerdiğimiz yöntemi test etmek için iki tane ayrı veritabanı kullandık. Bunlardan ilki 3B nesne tanımda standart haline gelmiş olan COIL-20 veritabanıdır [9] (bkz Şekil 3). Bu

veritabanı, 20 nesnenin dikey eksene göre 5 derecelik döndürülmeleri sonucunda her nesne için 72 toplamda ise 1440 adet imgeden oluşur.



Şekil 3. COIL-20 veritabanı nesnelerinin önden görünümü

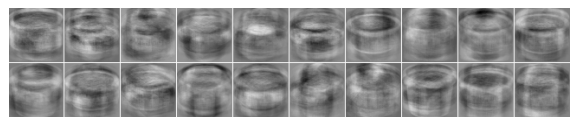
Kullandığımız bir diğer veritabanı ise kendi laboratuvarımızda oluşturduğumuz FESTO veritabanıdır (bkz Şekil 4). Bu veritabanı COIL-20’den farklı olarak 15 endüstriyel nesnenin sayfa içi eksene göre 10 derecelik döndürülmeleri ile elde edilmiştir. FESTO’da her nesne için 36 toplamda ise 540 imgeden oluşmuştur. Bu makalede FESTO veritabanının iki deneyi kullanılmıştır. İlk deney şekil 4’teki gibi beyaz arka planla hazırlanmışken ikinci deneyde arka plan siyahtır. Ayrıca deneylerimizde DVM için LibSVM yazılımını kullandık [10].



Şekil 4. FESTO veritabanı nesnelerinin görünümü

##### 5.1. COIL-20 Veritabanı Sonuçları

COIL-20 veritabanı için sonuçlarımızı oluştururken, karşılaştırmaya olanak vermesi açısından bu veritabanına uyguladığımız diğer yöntemlerin nesne tanıma başarımını da ekledik. Genel olarak sonuçları elde ederken veritabanını eğitim ve sınama kümeleri olarak ikiye ayırdık. Eğitim kümeleri nesnelerin 10 veya 30 derecelik döndürülme aralıklarından seçtik. TBA’de özdeğerleri yüksek ilk 20 temel bileşeni kullandık. Böylece imgelerdeki toplam enerjini %90’a yakını korumuş olduk. BBA’de ise  $\tanh(x)$  doğrusalsızlığı ile sabit-nokta yöntemini kullandık. Altta BBA ile elde ettiğimiz bağımsız bileşen süzgeçleri verilmiştir.



Şekil 5. COIL-20 veritabanının bağımsız bileşen süzgeçleri

Tablo 1’de BBA ve TBA’dan gelen öznelik vektörlerinin enyakın nesne ortalamasına göre sınıflandırınca elde ettiğimiz sonuçlar iki farklı örnekleme aralığı için verilmiştir. Burada BBA’nın TBA’dan daha üstün başarımlar sağladığı görülmektedir.

Tablo 1. BBA ve TBA ile nesne tanıma başarımları (%)

COIL-20	Örnekleme Aralığı	
	10 <sup>0</sup>	30 <sup>0</sup>
TBA	74.03	75.17
BBA	88.33	88.00

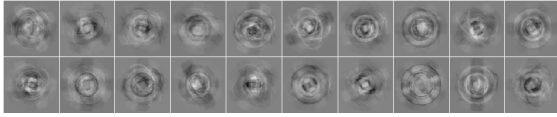
TBA ve BBA’dan gelen öznelik vektörleri DVM’in farklı çekirdekleri ile sınıflandırılması durumunda ise çokterimli ve rtf çekirdekleri için hatasız nesne tanıma başarılmıştır (bkz Tablo 2). Tablo 2 deki başarımlar değerleri DVM değişkenlerinin ızgara araması (grid-search) ile elde edilen en yüksek yüzdelerdir.

Tablo 2. TB Ave BBA’ya DVM eklenince tanıma başarımları

COIL	Doğrusal		Çokterimli		Rtf	
	10 <sup>0</sup>	30 <sup>0</sup>	10 <sup>0</sup>	30 <sup>0</sup>	10 <sup>0</sup>	30 <sup>0</sup>
TBA+ DVM	99.4	95.5	100	97.2	100	97.4
BBA + DVM	99.9	97.3	100	97.5	100	97.5

## 5.2. FESTO Veritabanı Sonuçları

COIL-20 veritabanı için yapılan işlemler bu veritabanının iki farklı deneyi içinde tekrarlanmıştır. Yalnız örnekleme aralığı olarak sadece 30 derece kullanılmıştır. Bunun sebebi ise 10 derece örnekleme aralığı eğitim için kullanılırsa sınama kümesinin oluşturulamayacağıdır.



Şekil 6. FESTO veritabanının bağımsız bileşen süzgeçleri

FESTO veritabanına BBA uygulanınca elde edilen bağımsız bileşen süzgeçleri Şekil 6’de verilmiştir.

Tablo 3. BBA ve TBA ile Nesne Tanıma Başarımları (%)

FESTO	Deney #1	Deney #2
TBA	57.22	47.50
BBA	98.33	100.00

BBA’nın iki deneyde de başarımlarının TBA’dan çok daha başarılı olduğu görülmektedir. Hatta ikinci deneyde sadece BBA uygulayarak bile hatasız nesne tanıma başarılmıştır. Daha önce COIL’de olduğu gibi enyakın nesne ortalamasına göre sınıflandırma yapılmıştır.

Tablo 4. TB Ave BBA’ya DVM eklenince nesne başarımları

FESTO	Doğrusal		Çokterimli		Rtf	
	#1	#2	#1	#2	#1	#2
TBA+ DVM	83.9	75	93.3	91.9	93.1	93.1

BBA + DVM	96.4	95.6	97.2	97.2	98.1	100
-----------	------	------	------	------	------	-----

## 6- Vargılar

Makalede TBA ve BBA’nın DVM ile kaynaşımıyla nesne tanıma başarımlarının arttığı gösterilmiştir. Ayrıca DVM’den önce TBA veya BBA uygulayarak öznelik çıkarımının işlem süresini kısaltması da yöntemin doğal bir sonucudur. Makalenin açığa vurduğu bir diğer durum ise FESTO veritabanında BBA ile elde ettiğimiz hatasız nesne tanıma başarımlarıdır. Buradan şu çıkarımı yapabiliriz ki; BBA ile elde edilen öznelikler özellikle sayfa içi eksenlerdeki dönmeye karşı gürbüzdür ve bir nesnenin imgelerinden elde edilen öznelik vektörleri o nesnenin ortalama değeri etrafında kümelenmişlerdir. Bunu BBA için DVM’in Rtf çekirdeğiyle elde ettiğimiz sonuçlar da desteklemektedir.

## 7- Kaynakça

- [1] M. Hagedoorn, Pattern Matching using Similarity Measures, PhD Thesis, University of Utrecht, 2000
- [2] S. Loncaric, A Survey of Shape Analysis Techniques, Pattern Recognition, Vol. 31, No. 8, pp. 983-1001, 1998
- [3] R.C. Velkamp, M. Hagedoorn, State of the Art in Shape Matching, Technical Report, UU-CS-1999-27, Utrecht, 1999.
- [4] M. S. Bartlett, H. M. Lades, and T. J. Sejnowski, Independent component representations for face recognition, Proc. SPIE Symp. Electron. Imaging: Science Technology; Conf. Human Vision Electronic Imaging III, CA, 1998, pp. 528-539.
- [5] H. S. Sahambi, K. Khorasani, A Neural-Network Appearance-Based 3-D Object Recognition Using Independent Component Analysis, IEEE Trans. on Neural Networks, vol. 14, no. 1, January 2003, pp. 138-149.
- [6] M. Pontil, A. Verri, Support Vector Machines for 3D Object Recognition, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 20, no. 6, June 1998, pp. 637-646.
- [7] A. Hyvarinen, Survey on independent component analysis, Neural Computing Surveys, 2,(99) 94-128.
- [8] N. Cristianini, J.Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge Univ. Press 2000.
- [9] S. A. Nene, S. K. Nayar and H. Murase, Columbia Object Image Library (COIL-20), Technical Report CUCS-005-96, February 1996.
- [10] Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin, A Practical Guide to Support Vector Classification, Department of Computer Science and Information Engineering National Taiwan University.