Detecting Occluded Objects Using Independent Component Analysis

Hitendra K. Dutt   Amar Makhija
Santanu Chaudhury *   S. D. Joshi
Department of Electrical Engg.
Indian Institute of Technology
New Delhi-110016
INDIA
e-mail: santanuc@ee.iitd.ernet.in

Abstract

In this paper, we present an appearance based recognition engine that is based on Independent Component Analysis(ICA). The recognition engine has the ability to tackle occlusion. The proposed approach does not use local features but makes use of the complete view of the object. Consequently, complexity of the recognition process is not increased. Using this technique we can recognise all the known objects in a scene provided the appearance models of the objects have distinct statistically independent components. In this paper we have also compared the performance of an ICA based recognition engine with a PCA based approach for occluded object recognition using a robust distance measure.

1 Introduction

Object recognition has been an active area of research in Computer Vision. Appearance based techniques are being widely used for this problem because of their inherent ability to exploit image based information. In this paper, we have proposed a novel method for appearance based recognition of occluded objects without using local features.

Appearance based object recognition methods make recognition systems easily trainable from visual data. These systems typically operate by comparing a two-dimensional, image-like representations of object appearance against prototypes stored in the memory, and finding the closest match. A class of appearance based methods make use of a lower dimensional subspace of the higher dimensional representation memory for the purpose of comparison. In [1], an approach based upon principal component analysis has been proposed. In this approach, appearance information is stored in the form of uncorrelated components and the object recognition is done by finding the nearest neighbor of the projections of unknown images using the Euclidean Distance norm. Pentland. et al. [15] have used view-based and modular eigenspaces for face recognition. In [16] similar approach has been used for parametric modeling of shape. These methods, in general, fail to recognise partially occluded objects because these approaches make use of the complete view of an object. In order to overcome this limitation, some local feature based approaches have been proposed [14, 19, 20]. However, the use of local features increases the time complexity of the recognition process. Black et al. [6] have proposed use of a robust distance norm for comparing global appearance information, which can take care of occlusion. However, a nearest neighbour based matching framework precludes the possibility of recognising all the known objects present in a scene using a robust distance measure.

The PCA approach stores appearance information in terms of uncorrelated components. Independent component analysis [13] is another linear decomposition technique which seeks statistically independent and non-Gaussian components, modeling observation data as a linear mixture of independent components. It has been used for face recognition [17], texture discrimination [8], analysis of EEG signals [18], etc. In this paper we have presented an ICA based approach for occluded object recognition. This technique makes use of the global view of the object and hence does not introduce additional complexity in the recognition process. Using this technique we can recognise all the known objects in a scene provided the appearance models of the objects have distinct statistically independent components. In this paper we have also compared the performance of an ICA based recognition engine with a PCA based approach for occluded object recognition using robust distance measure.

*Author for Correspondence

2 Recognition using ICA

Our approach is based on the Independent Component Analysis which is another way to encode the appearance based information content in an image. This technique was first developed to tackle problems of the kind of the cocktail party problem [10]. To rigorously define ICA, we use a statistical "latent variables" model. Assume that we observe n linear mixtures x_1,...,x_n of n independent components

$$x_j = a_{j1}s_1 + a_{j2}s_2 + ... + a_{jn}s_n \tag{1}$$

In the ICA model, we assume that each mixture x_j as well as each independent component s_k is a random variable, instead of a proper time signal. The observed values $x_j(\mathrm{t})$, are then a sample of this random variable. Without loss of generality, we can assume that both the mixture variables and the independent components have zero mean. If this is not true, then the observable variables x_i can always be centered by subtracting the sample mean, which makes the model, zero mean. It is convenient to use vector matrix notation instead of the sums like in the previous equation. Let us denote by $\mathbf{x}$ the random vector whose elements are the mixtures x_1,...,x_n , and likewise by s the random vector with elements s_1,...,s_n. Let us denote by $\mathbf{A}$ the matrix with elements a_{ij}.Using this vector matrix notation, the mixing model can be written as

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{2}$$

Sometimes we need the columns of matrix $\mathbf{A}$; denoting them by $\mathbf{a_j}$ the model can also be written as

$$\mathbf{x} = \sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{a_i s_i} \tag{3}$$

The statistical model in equation 2 above is called ICA model. The independent components are latent variables, meaning that they cannot be directly observed. All we observe is the random vector $\mathbf{x}$, and we must estimate both $\mathbf{A}$ and $\mathbf{s}$ using it. The ICA approach basically finds $\mathbf{s}$ for us. However, there are some constraints on this approach. One of them is that the random variables $\mathbf{s}$ must be non-gaussian.

2.1 Appearance based recognition using ICA

In this section we have described construction of a simple ICA based object recognition system. For each object in the database we have considered a set of images showing the unoccluded model object from a different viewpoint. We now obtain vectors ($N * N \times 1$ matrices) by reading off the pixels from each image(of size $N \times N$) in a raster scan manner. Using the fast ICA algorithm [10], we compute the independent components of these images. The number of independent components computed by this algorithm will be equal to or less than the number of input images. These ICs form a space based on the appearance characteristics of the model objects as seen in the images. Each view of the model objects correspond to a point in this appearance space. A test image, that is an image showing unoccluded view of an unknown model object, can now be projected on to this space. The nearest neighbour of this point in terms of Euclidean distance in the appearance space indicates the identity of the object in the test image.

The above approach cannot be used for detection of occluded objects in cluttered image. However by changing the euclidean measure to a robust distance measure [6], we may eliminate outliers in the test image and thereby recognize the occluded object in the image.

2.2 Using robust measure for occluded object detection

This section describes how the matching process can be made robust following the methodology proposed in [6].

Let $\mathbf{e}$ be an input image, written as a $nm \times 1$ vector, that we wish to match to the appearance space. The traditional methods construct an approximation, $\mathbf{e}^*$, to the input image, $\mathbf{e}$ as

$$\mathbf{e}^* = \sum_{i=1}^{t} c_i IC_i \tag{4}$$

where each c_i is computed by taking the dot product of $\mathbf{e}$ with IC_i.Here, IC are the various Independent Components. This approximation corresponds to the least-squares estimate of c_i. In other words, the c_i are those that give a reconstructed image that minimizes the squared error $E(\mathbf{c})$ between $\mathbf{e}$ and $\mathbf{e}^*$ summed over the entire image:

$$\begin{aligned} (\mathbf{c}) &= \sum_{j=1}^{n \times m} (\mathbf{e}_j - \mathbf{e}_j{}^*)^2, & (5) \\ &= \sum_{j=1}^{n \times m} \left(\mathbf{e}_j - \left(\sum_{i=1}^{t} c_i IC_{i,j} \right) \right)^2 & (6) \end{aligned}$$

The least squares approximation works well with the images which have clearly segmented objects which look roughly like those used to build the eigenspace. But it is commonly known that least squares is sensitive to gross errors or "outliers", and it is easy to construct situations in which the standard representation is a poor approximation of the input data.

To robustly estimate the coefficients $\mathbf{c}$ we replace the quadratic error norm in Equation 6 with a robust error norm ρ, and minimize

$$\mathbf{E}(\mathbf{c}) = \sum_{j=1}^{n\times m} \rho\left(\left(\mathbf{e}_j - \left(\sum_{i=1}^{t} c_i IC_{i,j}\right)\right), \sigma\right). \quad (7)$$

where σ is the scale parameter. We can take ρ to be

$$\rho(x,\sigma) \quad = \quad \frac{x^2}{\sigma^2 + x^2}, \quad (8)$$

$$\frac{\partial\rho(x,\sigma)}{\partial x} = \quad \psi(x,\sigma) \quad = \frac{2x\sigma^2}{(\sigma^2+x^2)^2}, \quad (9)$$

a robust error norm which has been used extensively for optical flow estimation. The shape of the function is such that it "rejects", or down-weights, large residual errors. The function $\psi(x,\sigma)$, is the derivative of ρ and characterizes the influence on the residuals. As the magnitudes of the residuals $(\mathbf{e}_j - \mathbf{e}_j{}^*)$ grow beyond a point, their influence begins to decrease and the value of $\rho(.)$ approaches a constant. The value σ is a scale parameter that affects the point at which the influence begins to decrease. By examining the $\psi-$function it can be seen that the outlier region begins where the second derivative of ρ is zero.

However, this minimisation based approach can identify only one of the known objects in the scene in both ICA and PCA based formulations. In the next section, we describe a more general technique for occluded object recognition using ICA.

3 Occluded object detection using ICA

In a cluttered test image there may be more than one object. Our aim is to detect all the known images in the clutter. Our aim is to use statistically independent properties of the objects to recognize the objects in an image.

3.1 Finding the Appearance Space

The image recognition engine should be robust to recognize images of a variety of objects from different viewpoints. As a result the images initially given will be representative images from all possible viewpoints. Our database has different unoccluded views of model objects. We take the average of all the images corresponding to a single object from all the given viewpoints. After repeating this for each object we have one representative image corresponding to each object. We now obtain vectors ($N*N\times 1$ matrices) by reading off the pixels from each image (of size $N\times N$) in a raster scan manner. These vectors may now be placed in a matrix of size $(N*N\times M)$. It is being assumed that there are M images corresponding to M objects in the database. We now construct a space whose basis vectors are the independent components obtained from these representative vectors. The fast ICA algorithm [10] is used to find the independent components. With these IC's we have with us a space on which other image vectors may be projected. We also normalize this appearance space using Gram-Schmidt normalization.

Each representative image may now be projected on to this space. Those independent components on which the projection of the representative image vector is greater than a given threshhold, are said to correspond to that object. If two or more objects get mapped on to the same set of independent components, then they have statistical properties which are common to both, and thus, these two objects cannot be distinguished by the recognition engine. When each object has statistically independent properties, the IC's will uniquely correspond to individual objects. Also we have to assume the number of images to be greater than the number of objects for the recognition to take place correctly.

3.2 Recognizing Occluded Objects

A typical test image will have in it more than one object for recognition.To recognize the objects in the test image, we take the following steps.

- Read the image in a raster scan manner to compute the pixel vector correspoding to the image

- Make the vector zero mean and also normalize it

- Project this vector on to the IC space and find the ICs on which the projection of this image is greater than a threshold. A general rule of thumb is to consider as many number of ICs as there are objects in the test image.

The objects present in the test image correspond to those objects which have their appearance characteristics concentrated on the ICs found in the previous step. These algorithm fails when objects do not have distinct indepepndent components.

4 Results

In this section we shall discuss the results of our recognition engine for unoccluded as well as occluded objects. The database has been procured from the University of Columbia website. Some of the model objects are shown in Figure 1-3.

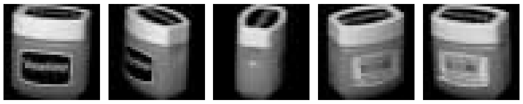


Figure 1: Images in the database of Object1

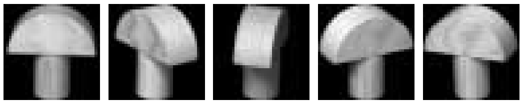


Figure 2: Images in the database of Object2

4.1 Results for recognition of unoccluded objects using ICA

The results for recognition of unoccluded objects are presented here. On comparison, performance of PCA based method was similar to that of ICA based approach. This is consistent with the results reported in [9]. Overall recognition rate for our experimental set was 95 percent. Some of the test images correctly recognised by the ICA based technique are presented in figures 4-8.

4.2 Using the Robust Measure along with ICA to tackle occlusion

The robust measure used by Black et.al.[6] has been used to tackle occlusion. Some of the test images are shown in fig. 9. Here, only one of the known objects is present in the scene.

Comparison between the results obtained using the robust measure with PCA and ICA is presented in table 4.2.

4.3 Results for recognition of occluded objects using ICA

The result for tackling occlusion from objects lying inside the database are presented in Tables 2-4 below. The results were good if the objects have statistically independent components. We shall now discuss some of the results of our experimentation. The database used was the same as presented in Figures 1-3. In Figure 10 we show the independent components that were obtained on using the mean of all poses of a single object as a representative image for the object. Some of the results that were obtained with 3 objects in the database are presented in Table 2 and 3 respectively.

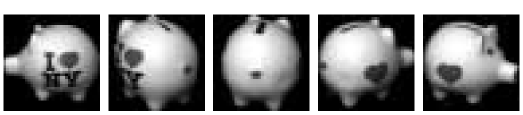


Figure 3: Images in the database of Object3

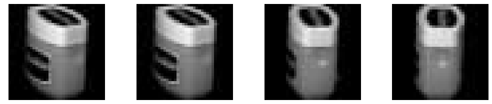


Figure 4: Some Test Images for Unoccluded object recognition using ICA,results-1

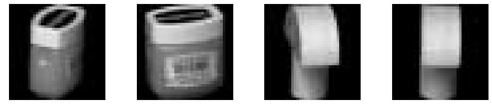


Figure 5: Some Test Images for Unoccluded object recognition using ICA,results-2

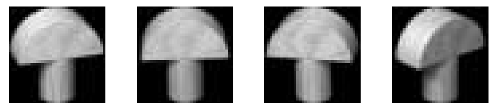


Figure 6: Some Test Images for Unoccluded object recognition using ICA,results-3

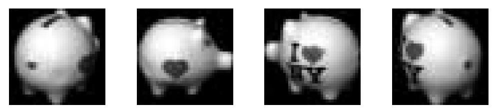


Figure 7: Some Test Images for Unoccluded object recognition using ICA,results-4

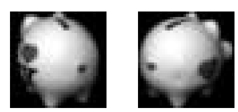


Figure 8: Some Test Images for Unoccluded object recognition using ICA,results-5

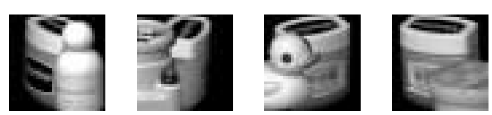


Figure 9: Some Test Images for Robust measure using ICA

Test name	PCA	ICA
No. of Test Images	30	30
Objects recognized correctly	28	28
Success percentage	93%	93%

Table 1: Recognition results with robust measure

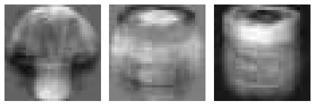
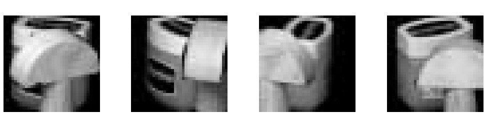


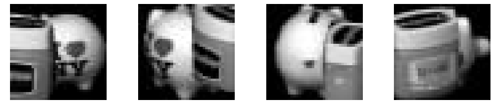
Figure 10: Independent Components of 3 Objects

Here, starred component indicate the one with least projection. Therefore, the objects present correspond to other projections(IC1 correspond to object1, IC2 to object2, IC3 to object3) As can be seen that the results for the three objects case were correct as the objects had statistically independent components. We also show an example, where the proposed approach does not provide expected results. We have used a model database of five objects presented in Figure 11(a). The corresponding Independent Components are shown in Figure 11(b). However, it is obvious from the figure that the objects did not have distinct statistically independent components. In this case, it is difficult to identify all the objects in a conglomerate. However, our experimental results show that if an image of an conglomerate of, say, two known objects occluding each other are projected onto this space, it consistently projects along two independent Components as shown in Table 4 (in this case, they are IC1 and IC4). However, the projections, though consistent, cannot clearly establish the identity of the objects because these components do not correspond to a single object in the database. We can only conclude existence of of objects belonging to known subsets of the modelbase.



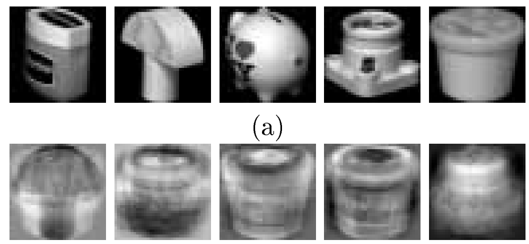
ic1 proj	0.049	0.111	0.007	0.008
ic2 proj	-0.55*	-0.142*	-0.055	-0.136*
ic3 proj	0.115	-0.015	-0.061*	-0.001

Table 2: IC occlusion results-1



ic1 proj	-0.086*	0.152	-0.08*	-0.140*
ic2 proj	0.159	0.069	0.082	0.129
ic3 proj	-0.020	-0.006*	0.055	-0.047

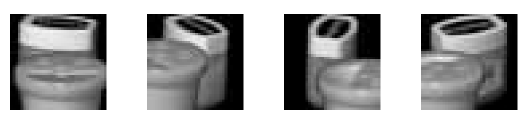
Table 3: IC occlusion results-2



(a)

(b)

Figure 11: (a) Original Objects (b) Independent Components of 5 Objects



ic1 proj	0.022354*	0.052805*	-0.034289	0.053654*
ic2 proj	-0.004320	-0.000246	-0.087359	0.003010
ic3 proj	-0.065584	-0.054623	-0.165289	-0.077402
ic4 proj	0.207848*	0.065804*	0.061921*	0.119362*
ic5 proj	0.010363	0.020174	0.015573*	0.016531

Table 4: IC occlusion results-3: With 5 ICs

5 Conclusion

We have seen that Independent Components offer us certain advantages over Principal Components in that it stores appearance information in the form of statistically independent components. The constraint on ICA that the independent components have to be Gaussian is normally true for images for their pixels have definite positive values. Also, ICA help us to recognize occluded objects even when all the objects present in a conglomerate belong to the database.

References

[1] Shree K. Nayar, Hiroshi Murase, "Visual Learning and Recognition of 3-D Objects from Appearance", International Journal of Computer Vision, 14, 5-24(1995)

[2] Erez Sali, Ullman Shimon, "Recognizing novel 3-D objects under new illumination and viewing position using a small number of example views or even a single view", Sixth International Conference on Computer Vision, January 4-7, 1998

[3] Ali Shokoufandeh, Ivan Marsic and Sven J. Dickinson, "View-Based Object Matching", Sixth International Conference on Computer Vision, January 4-7, 1998

[4] G. J. Edwards, C. J. Taylor, C. J. Cootes, "Learning to Identify and Track Faces in Image Sequences", Sixth International Conference On Computer Vision, January 4-7, 1998.

[5] Rogers David F. , Adams J. Alan, "Mathematical Elements for Computer Graphics", 2nd Edition, McGraw-Hill Publishing Company, 1990.

[6] Black Michael J., Jepson Allan D., "EigenTracking: Robust Matching and Tracking of Articulated Objects Using a View-Based Representation", Technical Report, Department of Computer Science, University of Toronto : RBCV-TR-96-50, October 1996.

[7] Fukunaga, K. 1990. Introduction to Statistical Pattern Recognition, Academic Press, London

[8] Manduchi R. ,Portilla J., " Independent Component Analysis of Textures", Fifth International Conference on Computer Vision, September 20-25, 1999, Corfu, Greece.

[9] Moghaddam Baback, "Principal Manifolds and Bayesian Subspaces for Visual Recognition", Fifth International Conference on Computer Vision, September 20-25, 1999, Corfu, Greece.

[10] Hyvarinen A., Oja E. , "Independent Component Analysis: A tutorial", http://www.cis.hut.fi/projects/ica/

[11] Tomar Anoop Singh, "Studies On the Karhunen Loeve Expansion of Stochastic Processes defined over the scale and rotation group", M.Tech Thesis, IITD, 1999

[12] http://www.cs.columbia.edu/CAVE/

[13] http://www.cnl.salk.edu/ tewon/ICA/intro.html

[14] Ikeuchi Katsushi, Ohba Kohtaro, "Detectability, Uniqueness and Reliability of Eigen Windows for Stable Verification of Partially Occluded Objects", IEEE transactions on Pattern Analysis and Machine Intelligence, Vol-19, September 1997.

[15] A.Pentland, B.Moghaddam and T.Starner. View-based and modular eigenspaces for face recognition. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, Seattle, WA, June 1994.

[16] T.F.Cootes and C.J.Taylor. Active Shape models: Smart snakes. In *Proc. British Machine Vision Conference*, pages 9-18. Springer-Verlag, 1992.

[17] M.S.Bartlett, H.M.Lades and T.J.Sejnowski. Independent component representations for face recognition. In *Proc. of the SPIE, Vol 2399: Conference on Human Vision and Electronic Imaging III*, pages 528-539, 1998.

[18] S.Makeig, A.J.Bell, T.Jung and T.J.Sejnowski. Independent component analysis of electroencephalographic data. In *Advances in Neural Information Processing Systems* 8, pages 145-151, 1996.

[19] C.Schmid and R.Mohr. Combining greyvalue invariants with local constraints for object recognition. In *Proc. CVPR96*, pages 872-877, San Francisco CA, June 1996.

[20] R.C.Nelson and A.Selinger. Large Scale tests of a Keyed, Appearance-Based 3-D Object Recognition System. In *Proc. of the Vision Science 1999.*