14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

EXTRACTION OF LOCAL FEATURES FROM IMAGE SEQUENCES USING SPATIO-TEMPORAL INDEPENDENT COMPONENT ANALYSIS

Victor C. Chen

US Naval Research Laboratory
4555 Overlook Ave. SW Washington DC 20375 USA
phone: 202-404-1973, email: vchen@radar.nrl.navy.mil

ABSTRACT

Independent component analysis (ICA) has been used to extract salient features from natural video images. The purpose of this paper is to study how the ICA can be applied to extracting features from a sequence of images of an object taken at different time, from different view angles, or with different spatial structures. Local features are considered as basic building blocks of images. To extract localized features in image sequences, spatio-temporal ICA, which maximizes the degree of independence over space and time, is a suitable method for analyzing such image sequences and extracting local features.

1. INTRODUCTION

Independent component analysis (ICA) [1-4] was originally developed to separate mixed signals into independent components for blind source separation. It decomposes mixed signals into independent physical attributes or a basis whose components are statistically independent. As is well-known, independent events must be uncorrelated, but uncorrelated events may not be independent. Principal component analysis (PCA) only requires the components to be uncorrelated. But, ICA accounts for higher order statistics and thus represents local features.

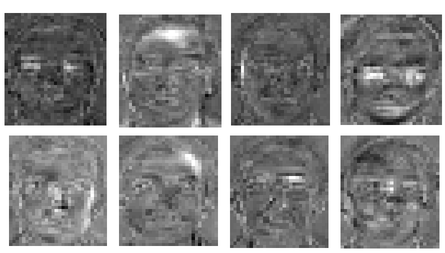

Fig.1 ICs derived from a face database showing enhanced features.

Local features are useful because they are more robust when a sequence of images is changing with time. To show the ability of ICA for extracting local features, as an example, independent components of 2D face images, which have enhanced features of nose, mouth, eye, forehead, jaw-line or cheekbone, are shown in Fig. 1. Local features are considered as basic building blocks of images and are more robust than global features.

ICA, which is usually applied to the 2-D spatial images, can also be applied temporally to a sequence of images to find the independent components over time. Joint spatial and temporal ICA, which maximizes the degree of independence over space and time, is a suitable method for analyzing image sequences. By using spatio-temporal ICA, independent feature space in both spatial and temporal domains can be extracted. The best features extracted from the spatio-temporal ICA are used to improve the overall performance of pattern recognition.

A local spatio-temporal independent component (IC) feature subspace derived from natural video image sequences was proposed in [2]. Recently, a unifying framework for analyzing natural image sequences in [3] suggested a spatio-temporal bubble model, which defines a region of activity localized both in time and in space. It was shown that the performance of ICA on natural image sequences is similar to simple cells in the primary visual cortex process. In this paper, we apply the idea proposed in [2, 3] to extracting local spatio-temporal IC features from image sequences beyond natural video image sequences, such as a sequence of images of an object taken at different time, from different view angles, or with different spatial structures. The goal of this paper is to extract local spatio-temporal features from sequences of images. With the best local features, the overall performance of pattern recognition may be improved. For example, combining global features of a face with localized features in nose, mouth, eye, forehead, jaw-line or cheekbone, the overall performance of face recognition can be improved.

2. ARCHITECTURES OF INDEPENDENT COMPONENT ANALYSIS

ICA was originally developed to separate mixed signals into independent components for blind source separation. Because it is useful to find independent physical attributes, an extended application of ICA is feature extraction. It decomposes a set of features into a basis whose components are statistically independent. ICA minimizes the statistical dependence between basis feature vectors. It searches for a linear transformation W_{ICA} to transform or decompose the feature vectors $X = (X_1, X_2, \ldots X_M)$ into statistically independent vectors $IC = (IC_1, IC_2, \ldots, IC_N)$, so that the transformed components of

$$IC = W_{ICA} X \tag{1}$$

are independent. Because there is no closed-form solution for finding the matrix W_{ICA}, iterative algorithms have been used to search for this matrix. Because the ICA accounts for higher order statistics, it provides a more powerful feature representation than the PCA.

ICA can also be applied to 2-D image analysis. A 2-D image is treated as a mixture of independent components. An architecture of ICA (architecture I [5]) for a set of 2-D images $X = (X_1, X_2, \ldots, X_M)$ is shown in Fig. 2(a). To apply ICA algorithm, a 2-D image $X_k(x,y)$ must be concatenated into a 1-D vector. The ICA algorithm learned the N by M weighting matrix W_{ICA} such that a set of N independent basis functions is obtained by $IC = W_{ICA} X$. Thus, 2-D images can be formed by mixing the independent components together with a mixing matrix A:

$$X = A \cdot IC \tag{2}$$

where $A = W_{ICA}^{-1}$ and is a permuted version of W_{ICA}. Fig. 2(b) illustrates how a 2-D image is decomposed into a set of independent basis components

$$X_k = a_1 IC_1 + a_2 IC_2 + \ldots + a_N IC_N \,. \tag{3}$$

In fact, ICA finds a new projection space $(IC_1, IC_2, ..IC_N)$. Although the IC space derived here is not a orthogonal space, the coefficients $(a_1, a_2, \ldots, a_N)$ of the linear combination of the independent components shown in (3) can be used to represent the 2-D image.

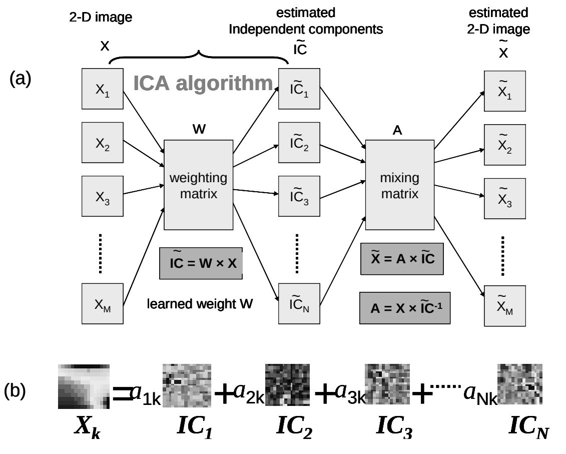

Figure 2 – (a) Architecture I of ICA; (b) IC decomposition.

Most applications of the ICA are in spatial x-y domains. Spatial ICA finds independent components of a 2-D feature over space. Thus, a 2-D spatial feature can be represented as a sum of the independent components ICs:

$$X(x,y) = \sum_{k=1}^{N} a_k \cdot IC_k(x,y) \tag{4}$$

where a_k is the amplitude of independent component $IC_k(x,y)$. The $IC_k(x,y)$ $(k=1,2,\ldots N)$ is regarded as the basic building blocks of the image. The amplitude a_k, which defines the strength of each building block, is extracted by a corresponding IC filter $F_k(x,y)$:

$$a_k = \sum_{x,y} F_k(x,y) \cdot X(x,y) \,. \tag{5}$$

where the $[IC_k]$ matrix and $[F_k]$ matrix are inverse to each other [2].

3. SPATIO-TEMPORAL ICA OF IMAGE SEQUENCES

If a 2-D feature is varying over time, then temporal ICA may be used to find independent components over time. For many applications where 2-D features can be either spatially dependent, temporally dependent, or both, ICA correspondingly can be applied to the spatial feature decomposition, temporal feature decomposition, or the joint spatiotemporal feature decomposition. The best features extracted from the joint spatial and temporal ICA may be used to improve the overall performance of pattern recognition.

To decompose a sequence of 2-D images over time, spatio-temporal ICA can maximize the degree of independence over space as well as time. Instead of using a set of 2-D spatial images, a set of 3-D spatio-temporal image cubes, that is stacked 2-D spatial images over time, are used to find a set of space-time independent components, as shown in Fig. 3.

$$X(x,y,t) = \sum_{k=1}^{N} a_k \cdot IC_k(x,y,t) \tag{6}$$

The independent component method assumes that each particular image sequence is a superposition of elementary components that are independent of each other.

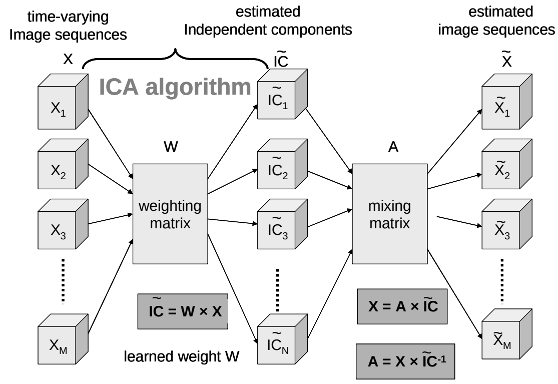

Figure 3 – Spatiotemporal ICA.

The corresponding IC filter $F_k(x,y,t)$ is localized in space and time, band-passed in spatial and temporal frequencies, tuned in orientation, and selective for the direction of movement [2]. With the IC filter, any sequence of images will generate a set of coefficients, which represent the feature of the image sequence.

4. SPATIOTEMPORAL ICA FOR EXTRACTING LOCAL FEATURES

Local topographic features extracted by local feature analysis (LFA) of images are sparse-distributed and low-dimensional. LFA provides a description of image feature in terms of localized characteristics and has useful applications in head segmentation and face recognition [5-8]. ICA can also be used for extracting local features by applying spatial ICA to randomly selected image patches from the image. Local spatial ICs derived from a face image are 2-D Gabor-like functions (sinusoidal function modulated by a Gaussian envelope) as shown in Fig. 4. The properties of IC filters actually resemble Gabor-type of receptive fields in simple visual cortex cells [2]. Spatio-temporal IC filters should resemble 3-D Gabor-type functions as discussed below.

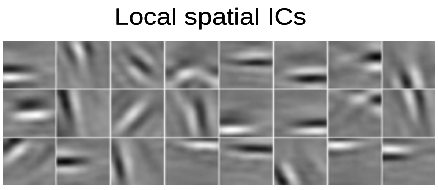

Fig.4 Local spatial ICs derived from a face image.

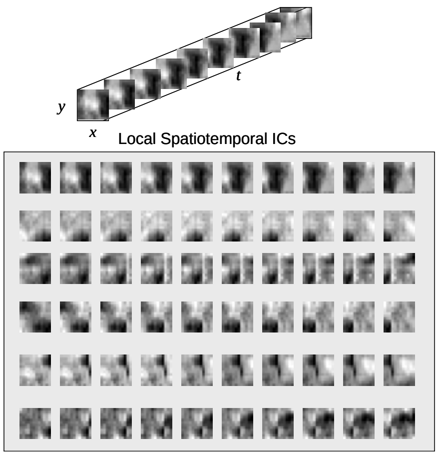

Fig.5 Local spatiotemporal ICs derived from a face database.

To extract local features from image sequences, spatiotemporal ICA should be applied to patch image sequences that are randomly selected from the entire image sequences. A weighting matrix learned by the fast ICA algorithm provided in [9] is derived from the ORL face dataset downloaded from [10]. The database includes face images from 40 different persons, where each person has 10 face images taken at different times, from different view angles, or with different local spatial structures or movements. Thus, each image sequence consists of 10 images. By using the spatio-temporal ICA algorithm, a set of spatio-temporal IC subspace can be learned. These ICs serve as basic building blocks of the face image sequences. These building blocks are 3-D Gabor-like functions [2], or bubbles [3], which have different orientations and spatial frequencies as shown in Fig. 5. Compared to a spatial IC of an image, the spatiot-temporal IC changes its phase from time to time and represents a more invariant local feature. Any patch sequence in the image sequences can be decomposed into these spatio-temporal ICs as shown in Fig. 6. ICs calculated from image sequences resemble edge features moving temporally. These local spatio-temporal features of relative movement, as indicated in Fig. 7, are important. They add a new dimension to the feature vector and, thus, are useful for improving the overall performance of pattern recognition systems.

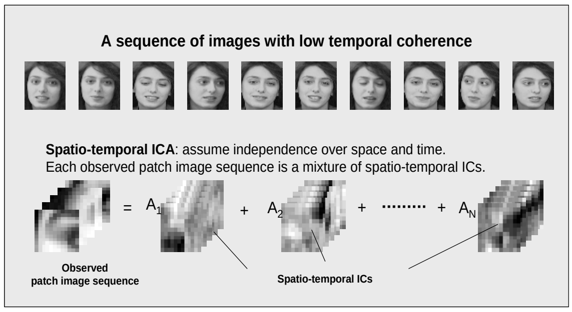

Fig.6 A patch image sequence is decomposed into the spatio-temporal ICs derived low temporal coherent face image sequences.

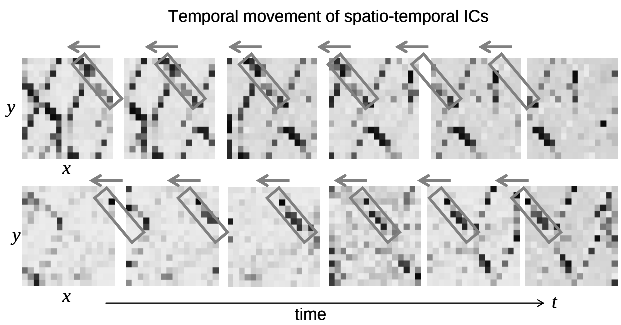

Fig.7 Temporal movement of local spatiotemporal IC features.

5. SUMMARY

In this paper, we use a face image database to study how to use ICA to extract local features from a sequence of images that taken at different times, from different view angles, or with different spatial structures. This analysis can also be extended to image sequences other than facial images. Local features extracted by ICA algorithm are basic building blocks of images and are more robust than global features. The best local features extracted from the spatio-temporal ICA may be used to improve the overall performance of pattern recognition.

6. ACKNOWLEDGEMENT

This work was supported by the Office of Naval Research, through the NRL base program.

REFERENCES

[1] A. J. Bell and T. J. Sejnowski, “An information-maximization approach to blind separation and blind deconvolution,” *Neural Comput.*, vol. 7, pp. 1129–1159, 1995.

[2] J.H. van Hateren and D.L. Ruderman, “Independent component analysis of natural image sequence yields spatio-temporal filters similar to simple cells in promary visual cortex,” *Proc. R. Soc. Lond.* B. *Biol. Sci.* Dec. 7, vol.265, no.1412, pp.2315-2320, 1998.

[3] A. Hyvarinen, J. Hurri, and J. Vayrynen, “Bubbles: a unifying framework for low-level statistical properties of natural image sequences,” *J. Opt. Soc. Am.* A. vol.20, no.7, pp.1237-1255, 2003.

[4] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, Wiley-Interscience, 2001.

[5] M.S. Barlett, J.R. Movellan, and T.J. Sejnowski, “Face recognition by independent component analysis”, *IEEE Trans. on Neural Networs*, vol.13, no.6, pp.1450-1464, 2002.

[6] A. Hyvärinen and P. O. Hoyer, “Emergence of phase and shift invariant features by decomposition of natural images into independent feature subspaces,” *Neural Computation*, vol.12, no.7, pp.1705–1720, 2000.

[7] A. Hyvarinen and E. Oja, “Independent component analysis: Algorithms and Applications,” *Neural Networks*, vol.13 no.4-5, pp.411-430, 2000.

[8] P. Penev and J. Atick, “Local feature analysis: a general statistical theory for object representation”, *Network: Computation in Neural Systems*, vol. 7, pp.477–500, 1996.

[9] FastICA : http://www.cis.hut.fi/projects/fastica.

[10] The AT&T (Olivetti) Research Laboratory ORL face database.
http://www.uk.research.att.com/facedatabase.html.