International Journal of Emerging Engineering Research and Technology
Volume 2, Issue 7, October 2014, PP 267-273
ISSN 2349-4395 (Print) & ISSN 2349-4409 (Online)

Comparative Analysis of ICA Based Features

Sunita Kumari[1], Anita Gunjal[2], Swati Dhopte[3]

Department of Computer Engineering, MITCOE, Pune, India (Asst. Prof.)

Abstract: *Reduction of search space during real-time identification of an individual using any biometric trait has become an important issue. This requires some kind of filtering approaches to prune the database. In this thesis, a detailed investigation has been made to choose ICA features that provides maximum separability. To study the gender separability of various ICA features, two algorithms viz. InfoMax and Fast ICA are used considering two different architectures i.e., ICA-1 and ICA-2 for each algorithm. Hence, four different ICA features are extracted namely, InfoMax ICA-1, InfoMax ICA-2, Fast ICA-1, and Fast ICA-2 using faces from FERET database. The separability potential of each feature is evaluated using cd ratio and dprime index. InfoMax ICA-1 outperforms others.*

Keywords: *InfoMax , Fast , ICA-1 , ICA-2, cd, dprime*

1. Introduction

Suitable representation of a multivariate data by means of an appropriate transformation is an important issue in the field of pattern recognition, data compression, denoising, visualization, etc. The process of finding such representation of data is known as feature extraction[1,2]. The goal of feature extraction is to discover compact and informative components that perfectly represent the dataset. The basic approach adopted in this process is to transform a given set of measurements to a new set of features. Most of the classification-related information is squeezed in a relatively small number of features, leading to the reduction of feature space. The principal components of the image give highly uncorrelated coefficients. Thus, using the principal component as features seems to be a reasonable choice which is achieved using Principal Component Analysis (PCA). Although PCA removes the second order dependencies from the data, still there exists the higher order dependency in the data. Higher order dependencies among the data can be removed using Independent Component Analysis and Selection of ICA Features Analysis (ICA) which is an extension of PCA

2. Independent Component Analysis (ICA)

ICA algorithm was initially proposed to solve the Blind Source Separation (BSS) problem i.e., given a mixture of signals generated from a set of underlying sources; the task is to separate the mixed signals and retrieve the original sources [3, 4].

To represent the ICA model mathematically, let us denote M-dimensional random variable by X; the problem is to find a linear transformation of the observed variable which can be found as,

$$S = WX \tag{1}$$

Where, transformed vector S is a set of N sources, $S = S_1, S_2, \cdot \cdot \cdot, S_N$ and W is a matrix to be determined.

2.1 Contrast Functions for ICA

The ICA data model can be estimated by formulating an objective function and then by maximizing or minimizing it. The objective function is also known as contrast or cost function of ICA. The quantitative measures for non-Gaussianity are kurtosis and negentropy which are discussed below.

2.1.1. Kurtosis

Kurtosis is the fourth order moment of a random data. For a given random data y, the kurtosis *kurt(y)* is defined as,

$$\text{kurt}(y) = E\{y^4\} - 3E\{y^2\} \tag{2}$$

where, $E\{\cdot\}$ is the statistical expectation operator. For normalized y, it is assumed that the variance of y is equal to unity i.e. $E\{y^2\} = 1$.

Hence,

$$\text{kurt}(y) = E\{y^4\} - 3 \tag{3}$$

For a Gaussian y,

$$E\{y^4\} = 3(E\{y^2\})^2 \tag{4}$$

Thus, for a Gaussian random variable, the kurtosis value is zero and for non-Gaussian random variable it is non-zero. For positive kurtosis value, the random variable is called super-Gaussian and for negative kurtosis value, it is known as sub-Gaussian. Super-Gaussian random variables have a spiky PDF with heavy tails and sub-Gaussian random variables have a flat PDF.

2.1.2. Negentropy

A second measure of non-Gaussianity is negentropy and is based on the information theoretic differential entropy. Entropy is the measure of the uncertainty associated with a random variable. The more random and unpredictable the data, the larger will be its entropy[6]. The entropy H of a random variable y is given as,

$$H(y) = E(I(y)) \tag{5}$$

Where, $E(\cdot)$ is the expected value and I is the information content of y. If p denotes the probability mass function of y then the entropy can explicitly be written as,

$$H(y) = -\sum_{i=1}^{n} p(y_i)\log_b p(y_i) \tag{6}$$

Among all random variables with unit variance, a Gaussian variable has the largest entropy value. Differential entropy normalized with respect to Gaussian variables result in negentropy. Negentropy of y denoted by J(y) is defined as,

$$J(y) = H(y_{Gauss}) - H(y) \tag{7}$$

Where, y_{Gauss} is the Gaussian random variable.

Approximations to Negentropy: Negentropy can be approximated in terms of a non-linear function f as,

$$J(y) = K[E\{f(y)\} - E\{f(v)\}]^2 \tag{8}$$

Where, K is a constant and v is a Gaussian random variable of zero mean and unit variance. Optimization of contrast function $J(y)$ for ICA depends on good choice of f. The frequent choices of f that have been proven to be useful are,

$$f1(y) = (1/a_1)\log\cosh(a_1 y) \tag{9}$$

$$f2(y) = -(1/a_2)\exp(-a_2 y^2/2) \tag{10}$$

$f3(y) = (¼)y^4$ (11)

Where, a_1 and a_2 are constants.

2.2 Mutual Information

Mutual information is the measure of mutual dependence between two variables. If y is a n-dimensional random variable and $p_y(y)$ is its PDF, then y has mutually independent components (ICs) iff—

$$p_y(y) = p_{y1}(y_1) + p_{y2}(y_2) + \cdot \cdot \cdot + p_{yn}(y_n)) \quad (12)$$

A common way of checking whether y has ICs is to measure the distance between both sides of Equation (12), which results the following,

$$I(p_y) = \delta(p_y, \Pi\, p_{yi}) \quad (13)$$

Average mutual information of y given by Comon [2] is defined as,

$$I(p_y) = \int p_y(y) \log(p_y(y)/\Pi\, p_y(y))dy \quad (14)$$

Average mutual information vanishes if and only if the variables are mutually independent and strictly positive otherwise. In terms of negentropy, mutual information is written as,

$$I(y_1, y_2, \ldots, y_n) = H(y) - \sum H(y_i) \quad (15)$$

The contrast function of mutual information discussed above requires the estimation of density function and thus has severely restricted the use of these contrast functions [7].

Algorithm1. *Information Maximization ICA*

Result: W

1. Initialize W_t=0 randomly ;
2. Update $W_{t+1} = W_t + (I + Y' S^T) W_t$;
3. if $W_{t+1} - W_t \approx 0 + \epsilon$ then
4. stop ;
5. else
6. go back to step 2 ;

Algorithm2. *Fast Fixed Point ICA*

Result: W

1. Initialize W_i (column vector of W) randomly;
2. $W_i^+ = E\{X\, \Phi(W_i^T X)\} - E\{\Phi'(W_i^T X)\}W_i$;
3. $W_i^+ = W_i^+ / \| W_i^+ \|$;
4. if $W_i^+ - Wi \approx 0 + \epsilon$ then
5. stop ;
6. else
7. $W_i = W_i^+$;
8. go back to step 2;

2.3 Simulation Results and Discussions

2.1.3. Basis Images

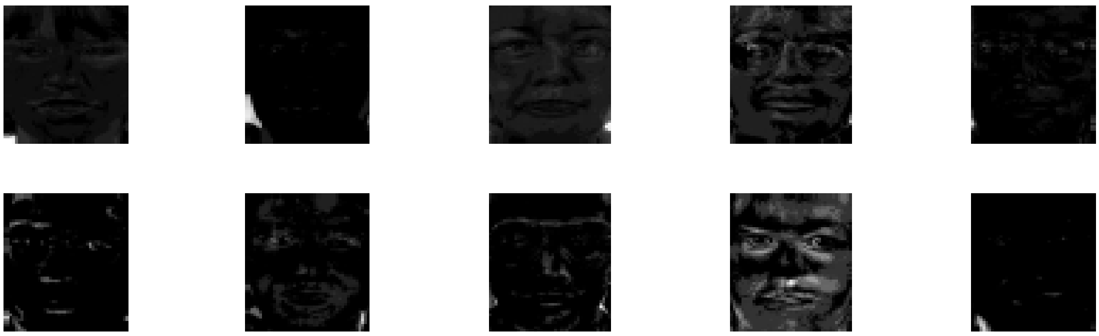
The goal of ICA is to find a better set of basis images so that it can easily estimate an unknown dataset. In this subsection, thrust has been laid for choosing the best set of basis images among InfoMax ICA-1, InfoMax ICA-2, Fast ICA-1, and Fast ICA-2. The basis images are obtained using InfoMax algorithm with architecture-1 and 2 as shown in Figures(i) and (ii) respectively. The basis images are also obtained using Fast ICA-1 and Fast ICA–2 as shown in Figures (iii) and (iv) respectively.

Fig(i). *Basis images from InfoMax ICA Architecture-1*

2.1.4. Choosing the Best Set of Features

If the separation between female and male class for each point is known then one can easily chose the optimum set of features. This is achieved using a technique

Fig(ii). *Basis images from InfoMax ICA Architecture-2*



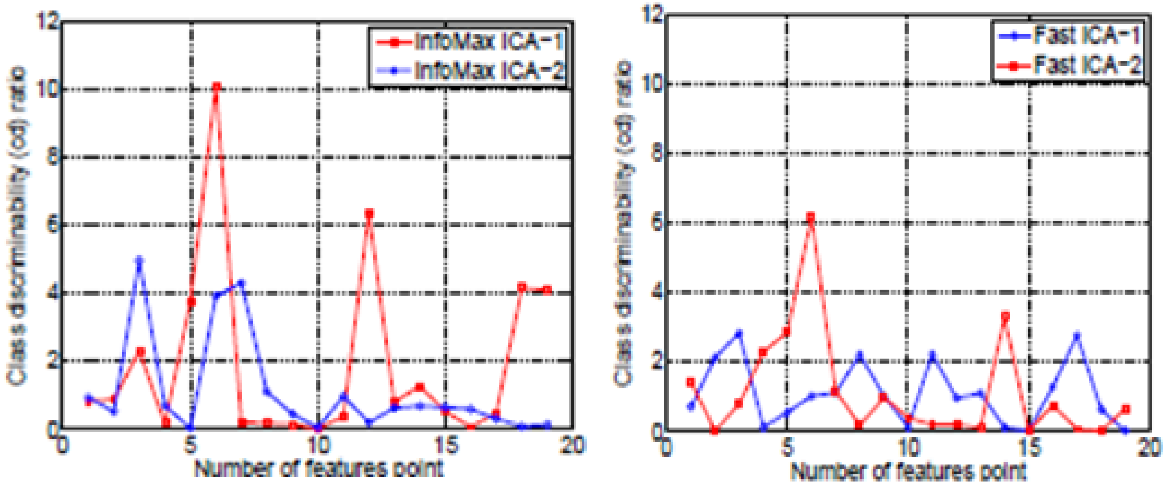
Fig(iii). *Basis images from Fast ICA Architecture-1*

Fig(iv). *Basis images from Fast ICA Architecture-2*

Based on the class discriminability (*cd*) which is given as,

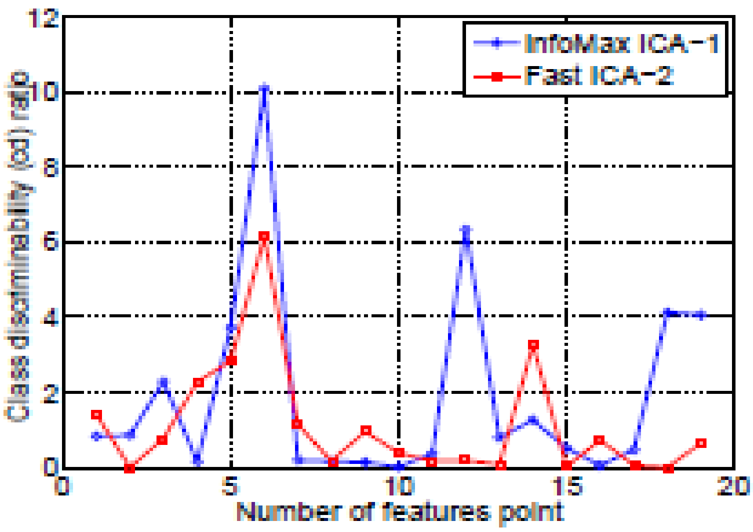
$$cd = \sigma_{between} / \sigma_{within} \quad (16)$$

Where $\sigma_{between}$ is the variance between male and female classes and σ_{within} is the sum of variances within each class.



(a) cd ratio distribution for InfoMax ICA (b) cd ratio distribution for Fast ICA

Figv. *Separation between female and male classes in terms of cd ratio*



Fig(vi). *Comparison of cd ratio between InfoMax ICA-1 and Fast ICA-2*

The experiments are carried on 19 independent coefficients which are obtained from the set of 20 face images (first 10 are of female and last 10 are of male). Thus, the size of the dataset becomes 20×19. Now, the distribution of cd ratio for each feature point is calculated. Such distributions of cd ratio for InfoMax ICA-1, InfoMax ICA-2, Fast ICA-1 and Fast ICA-2 are shown in Figure (v). From this

graphical representation, it can be approximated that separation between male and female for InfoMax ICA is more in architecture-1 than architecture-2 and reverse for Fast ICA. Thus, it can be concluded that InfoMax ICA-1 performs better than InfoMax ICA-2 and Fast ICA-2 performs better than Fast ICA-1.This makes a selection more concise. Figure vi depicts the distribution of cd ratio between InfoMax ICA-1 and Fast ICA-2. This approximates that InfoMax ICA-1 have more separability than Fast ICA-2. To make this approximation more accurate, dprime index is used as discussed in the following subsection.

2.1.5. DPrime Index

The dprime index (d') measures the separation between the mean of the male and female probability distributions in standard deviation units [8]. It is defined as,

$$d' = \sqrt{2}\, |\mu_{Female} - \mu_{Male}| / (\sigma^2_{\ Female} + \sigma^2_{\ Male}) \tag{17}$$

Where, μ and σ are the means and standard deviations of means of male and female classes respectively. A higher dprime index value indicates better separability. The dprime index value is

Table1. *dprime index value observed by different approaches of ICA*

	Infomax ICA-1	**Infomax ICA-2**	**Fast ICA-1**	**Fast ICA-2**
dprime index	0.6347	0.2294	0.0995	0.6123

measured using different approaches of ICA and is listed in Table 1. InfoMax ICA-1 has the highest dprime index value which indicates that it has the better separability than other approaches which are also supported by cd ratio distribution. Thus, InfoMax ICA-1 is chosen as an optimum candidate for feature selection for the purpose of gender classification.

3. CONCLUSION

In this chapter, an endeavour has been made to develop an efficient feature extraction technique based on ICA for gender classification. The two algorithms: InfoMax and Fast, and two architectures of ICA are used to find the gender features. In order to have the maximum separation between male and female class, class discriminability (cd) is used which is further evaluated using dprime index. Through experiments it has been found that Infomax ICA-1 has the highest dprime index value than other approaches and hence it is used for extraction of gender features from faces.

REFERENCES

[1] S. Theodoridis and K. Koutroumbas. Pattern recognition. Springer-Verlag, 2006.

[2] P. Comon. Independent component analysis: a new concept? Signal Processing, 36:287–314, 1994.

[3] D. P. Acharya. Development of novel independent component analysis techniques and their applications. PhD thesis, National Institute of Technology Rourkela.

[4] A. Hyvrinen. Survey on independent component analysis. Neural Computing Surveys, 2:94–128, 1999.

[5] A. J. Bell and T. J. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. Neural Computation, 7:1129–1159, 1995.

[6] M.S. Bartlett, J.R. Movellan, and T.J. Sejnowski. Face recognition by independent component analysis. IEEE Transactions on Neural Networks, 13(6):1450–1464, 2002.

[7] A. Hyvrinen and E. Oja. Independent component analysis: algorithms and applications. Neural Networks: The Official Journal of the International Neural Network Society, 13:411–430, 2000. PMID: 10946390.

[8] A. A. Ross, K. Nandakumar, and A. K. Jain. Handbook of multibiometrics (International Series on Biometrics). Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.