# Gradient based Fusion of Infrared and Visual Face Images using Support Vector Machine for Human Face Identification

Priya Saha[a], Mrinal K. Bhowmik*[a], Debotosh Bhattacharjee[b], Barin K. De[c], Mita Nasipuri[b]

[a]Department of Computer Science and Engineering, Tripura University (A Central University), Suryamaninagar-799022, Tripura, India; [b]Department of Computer Science and Engineering, Jadavpur University, Kolkata-700032, India; [c]Department of Physics, Tripura University (A Central University), Suryamaninagar-799022, Tripura, India.

## ABSTRACT

Pose and illumination invariant face recognition problem is now-a-days an emergent problem in the field of information security. In this paper, gradient based fusion method of gradient visual and corresponding infrared face images have been proposed to overcome the problem of illumination varying conditions. This technique mainly extracts illumination insensitive features under different conditions for effective face recognition purpose. The gradient image is computed from a visible light image. Information fusion is performed in the gradient map domain. The image fusion of infrared image and corresponding visual gradient image is done in wavelet domain by taking the maximum information of approximation and detailed coefficients. These fused images have been taken for dimension reduction using Independent Component Analysis (ICA). The reduced face images are taken for training and testing purposes from different classes of different datasets of IRIS face database. SVM multiclass strategy 'one-vs.-all' have been taken in the experiment. For training support vector machine, Sequential Minimal Optimization (SMO) algorithm has been used. Linear kernel and Polynomial kernel with degree 3 are used in SVM kernel functions. The experiment results show that the proposed approach generates good classification accuracies for the face images under different lighting conditions.

**Keywords:** Face recognition, gradient based fusion, wavelet domain, independent component analysis, support vector machine, sequential minimal optimization, IRIS face database

## 1. INTRODUCTION

Pose and illumination invariant face recognition problem is now-a-days an emergent problem in the field of information security. Changes in pose and illumination in human face images has made face recognition to be more complex task. Illumination can change the appearance of an object significantly. It makes difficult the recognition task when the differences induced by illumination are larger than differences between individuals. It is also valid for pose variation. Even though a plenty of methods has been developed to achieve illumination independent recognition, still many lack in the applicability to practical application domains. Jobson et al.[1] theoretically demonstrated that face images sampled with varying lighting directions form an illumination cone, thus allowing to interpolate novel illuminated views from a sampled space. Ramamoorthi[2], Hanrahan[3,4] and Basri and Jacobs[5] independently developed the spherical harmonic representation. This representation explained why images of an object under different lighting conditions can be described by a low dimensional subspace[5-8]. Raviv and Shashua[9] developed a "class-based" recognition and image synthesis method under varying illumination. The set of images were all having the same shape, which were generated by varying lighting conditions, and surface texture. The reported results have shown that they can be characterized analytically using images of a prototype object and a (illumination invariant) "signature" image per object of the class. Shim[10] proposed a technique for re-lighting by jointly estimating the pose, reflectance functions, and lighting. In this paper, feature level fusion, in which the visual gradient face images are fused with corresponding infrared images, has been introduced to obtain illumination insensitive features for face recognition under varying illumination. Here, Daubechies wavelet transforms, termed as db4, coefficients from visual and corresponding coefficients computed in the

*mkb_cse@yahoo.co.in; phone 0 94361 29933; fax 0381 237 4802

Multimedia Content and Mobile Devices, edited by David Akopian, et al., Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 8667, 86670Z · © 2013 SPIE-IS&T · CCC code: 0277-786X/13/$18 · doi: 10.1117/12.2001976

same manner from infrared (IR) images are used for the fusion method. Apart from this, we have also focused on efficient feature extraction method Independent Component Analysis (ICA). Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Independent Component Analysis (ICA) are three powerful tools commonly used for data reduction and feature extraction in the holistic approaches[11, 12, 13]. F. M. Pop et al.[14] investigated face recognition problem under varying pose and expression varying condition and found that in case of frontal images with low intra-personal variation, PCA-based techniques in the visible spectrum achieves higher performance. But, in practical face recognition application, the performance of the classical PCA-based approaches is highly affected because acquisition's conditions cannot be always controlled. One of the major shortcomings of LDA is that the data samples of different classes will be collapsed into one single cluster, and each class is modeled as a unimodal Gaussian. As pose and illumination produces different facial images of the same person, LDA cannot perform well in these situations of face recognition. The use of independent component analysis is preferred by some authors for feature extraction and dimensionality reduction in face recognition research [8, 15]. In fact, in some cases it is preferred better in terms of recognition and performance as ICA uses the higher-order statistical information generated from a set of basis vectors that acquires maximum independence for representation and recognition. It is also useful in maximizing the information transmission in the presence of noise which may be robust in different expression and illumination variations of image[13].

In this paper, fusion in the wavelet domain is performed using the gray level gradient visual image and the gray level thermal image. The experimental results show the effectiveness of the proposed technique even with considerable variations of the subjects due to different illumination with pose changes. The paper is organized as follows: System Overview has been given in section 2, and experimental results with discussions on them are given in section 3, and the conclusions are drawn in section 4.

## 2. SYSTEM OVERVIEW

The system flow is given in Figure 1. All the processing steps to generate the gradient fused images of optical and infrared images and classification of these fused images using SVM are shown in the block diagram. The visual images are smoothed using Gaussian Kernel Function and then the gradients $I_X$, and $I_Y$ in the direction X and Y respectively have been calculated. At each image point, the gradient vector points in the directions of the largest increase in the image brightness, while the length of the gradient vector corresponds to the rate of change in that direction. The gradient image is generated from $I_X$ and $I_Y$ by taking their inverse tangent. The gradient image is obtained from these two coefficients. Then the image fusion of thermal image and corresponding visual gradient image is done in the wavelet domain by taking the maximum information of approximation and detailed coefficients. The decomposition of thermal and visual gradient images has been completed at level-4.

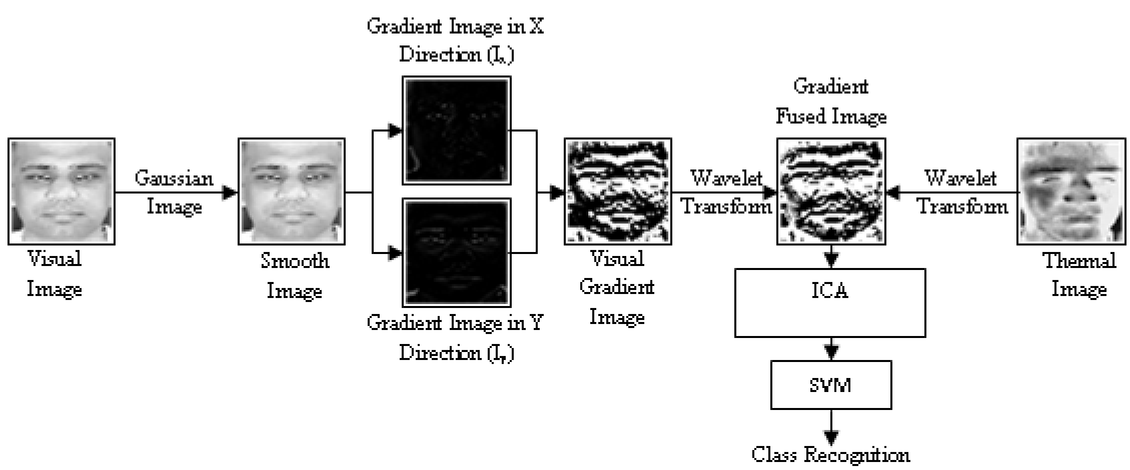


Figure 1. Block diagram of the face recognition system using gradient fused images

### 2.1 Generation of Registered Images

Image Registration is the process of transforming two images into the similar coordinate system. In this work, image registration is performed taking image intensity as a parameter estimator of a transformation between two images using an approach involving all pixels of the image. Images are taken from different illumination types of each person and registered the thermal and visual images separately with respect to the frontal image of each illumination and full dataset. Out of the 11 images of different rotations, the frontal image is considered as the base image and registration is done to the other 10 images with respect to the frontal image. Some sample images of visual and thermal base images, unregistered images, and their corresponding registered images are shown in Figure 2. Affine transformation is used in this work as it has six degrees of freedom and is equivalent to combine effect of translation, rotation, isotropic scaling, and shear (non-uniform scaling in some direction). Properties like parallelism, ratio of lengths of collinear or parallel segments (e.g. mid-points), ratio of areas, and linear combination of vectors are invariant under affine transformation.

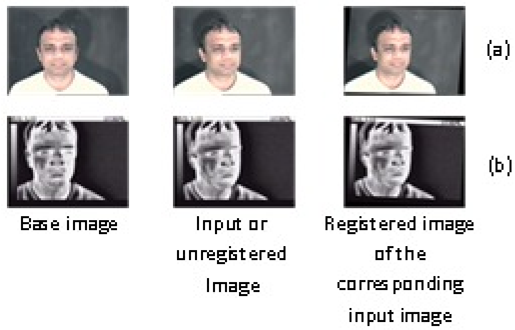


Figure 2: Sample of base, unregistered and corresponding registered image of (a) visual and (b) thermal image.

### 2.2 Gaussian filtering and implementation of visual gradient image

The gradient images are generated using Gaussian filtering[16]. According to the Tikonov regularization theory, Gaussian smoothing is a good approximation of the optimal polynomials to regularize a signal[17]. Performing a regularization step before the actual computation of partial derivatives of the image intensity, the effects of image noise are considerably reduced, still preserving the integrity of the differential. If, I be an image under variable lighting conditions, then gradient image (GI) of image I is defined as:

$$G = \tan^{-1}\left(\frac{I_y}{I_x}\right), G \in [0, 2\pi] \tag{1}$$

Where $I_x$ and $I_y$ are the gradients of image I in the X and Y direction respectively[18]. The implementation of visual gradient images is presented here. This visual image has been smoothed first with Gaussian filter to compute the gradient stably. Convolution type smoothing has been used to produce the smoothing image. In the next step, smooth image has been computed to find out the gradient in the x and y direction. By taking the inverse tangent of the image gradients in Y and X direction respectively, the gradient images have been generated. The algorithm to generate the gradient images is given below.

Input: Image I

Steps:

1. Creation of Gaussian Filter G.
2. To generate the smooth image $I_G$ by convolving the input image I with Gaussian Filter.

$$I_G = I * G, \text{ where, } * \text{ is the convolution operator.}$$

3. Find the gradients $G_X$ and $G_Y$ of image $I_G$ in X and Y direction respectively.

$$I_X = I_G * Grad_X(x, y, \sigma) \text{ and } I_Y = I_G * Grad_Y(x, y, \sigma)$$

Where, $Grad_X(x, y, \sigma)$ and $Grad_Y(x, y, \sigma)$ are the derivative of Gaussian Kernel function in X and Y direction respectively. σ is the standard deviation.

4. Calculating the inverse tangent of the image component $I_{Grad} = \tan^{-1}\left(\frac{Iy}{Ix}\right)$

Output: Gradient Image $I_{Grad}$.

### 2.3 Daubechies Wavelet Transform

Daubechies wavelet generates notable results in image analysis due to its mathematical properties. In the experiment, Daubechies transform (db4) has been used to perform decomposition and reconstruction of visual images and corresponding infrared (IR) images. Daubechies wavelets are the family of orthogonal wavelets and have obtained the highest number of vanishing moments for defined support width. Daubechies wavelet maintains the information of the original data and preserves spectral information[19]. The basic idea of wavelet image fusion is to unite the wavelet decompositions of the two original images with the help of fusion methods applied to approximations coefficients and details coefficients. The main idea behind the fusion algorithm are as follows: the two images are to be resampled to the one with the same size after processing, after that using forward wavelet transform they are respectively decomposed into the sub-images having same resolution at the same levels , and different resolution among different levels and information fusion is done based on the high-frequency sub-images of decomposed images; and finally inverse wavelet transform is performed to obtain the resultant image[20].

### 2.4 Independent Component Analysis

Independent Component Analysis (ICA) is a computational method from statistics and signal processing which is a special case of blind source separation. It is a generalization of PCA technique that assigns data from a high-dimensional space to a lower-dimensional space and decorrelates the higher-order statistics[21, 8]. ICA offers a set of basis vectors with maximum statistical independence, whereas in case of PCA basis vectors are determined by the eigenvectors. The ICA method can find a linear transform for the observed data using a set of basis functions, but it is insufficient to explain the complex, nonlinear variations[22]. A number of popular ICA algorithms include FastICA[8, 22], Infomax[23], Common's algorithm[24], and Kernel ICA[25]. The FastICA has been used here for implementation of independent component analysis. The FastICA algorithm has many desirable properties compared to other independent component analysis (ICA) algorithms. FastICA algorithm directly calculates independent components (ICs) of an image. The independent components can be approximated one by one, which is nearly a counterpart of doing projection pursuit.

To define ICA[24], we can use a statistical "latent variables" model. Assume that we observe n linear mixtures of n independent components

$$x_j = a_{j1}s_1 + a_{j2}s_2 + \cdots + a_{jn}s_n \text{ for all } j \tag{2}$$

It is convenient to use the vector-matrix notation instead of the sums like in Eq. (2). Let us denote by $x$ the random vector whose elements are the mixtures, and likewise by $s$ the random vector with elements $s_1, \cdots, s_n$ . Let us denote by $A$ matrix with elements. All vectors are understood as column vectors; thus, or the transpose of, is a row vector. The above mixing model is written as

$$x = As \tag{3}$$

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

$$x = \sum_{i=1}^{n} a_i s_i \tag{4}$$

The statistical model in Eq. (4) is called independent component analysis, or ICA model. The achievement of ICA for a given data set may require some application-dependent pre-processing steps.

The most basic and necessary pre-processing is to center $x$ i.e., subtract its mean vector $m = E\{x\}$ so as to make $x$ as a zero-mean variable. This implies that its zero-mean can be seen by taking expectations on both sides of (1). This preprocessing is made solely to simplify the ICA algorithm. After estimation of the mixing matrix $A$ with centered data, estimation can be completed by adding the mean vector of $s$ back to the centered estimates of $s$ . The mean vector of $s$ is given by, where $m$ is the mean that has subtracted in the preprocessing.

The algorithmic Steps for Centering are as follows: first create an array with all zeros of image size I, then, find the mean value of loaded image I, and finally, subtract the mean value from I. Second part of dimension reduction using independent component analysis is to find whiten of the observed matrix. The transformation of the observed vector has been performed linearly after centering so that a new vector $\tilde{x}$ which is white can be obtained, i.e., its components are not correlated, and their variances equal unity. In another word, the covariance matrix of $\tilde{x}$ equals to the identity matrix:

$$E\{\tilde{x}\tilde{x}^T\} = I \tag{5}$$

To calculate the whitening matrix, we use the Eigenvalues Decomposition (EVD) method of the covariance matrix, where $E$ is the orthogonal matrix of Eigenvectors of $E\{xx^T\}$ and $D$ is the diagonal matrix of its eigenvalues, $D = diag(d_1, \dots d_n)$. The main advantage of whitening is to reduce the number of parameters to be estimated. Instead of estimating $n^2$ parameters of the original matrix $A$, the new, orthogonal matrix $\tilde{A}$ have to be estimated. An orthogonal matrix contains n (n −1)/2 degrees of freedom. Thus, whitening solves half of the problem of ICA as whitening is a simple, standard procedure, and much simpler than any ICA algorithms. So, it helps to reduce the complexity of the problem using whitening. It may also be quite useful to reduce the dimension of the data at the same time as we do whitening. This has the effect of reducing noise[21].

Algorithmic steps for Whitening are as follows: centered image and original image I are taken as inputs; then, calculate co-variance for dimensionality reduction by applying PCA, the covariance between two random variables is $E[(x_1 - \mu_1)^*(x_2 - \mu_2)]$ where E is the mathematical expectation and $\mu_i = Ex_i$; calculate the eigenvalues and eigenvectors of the covariance matrix; finally, calculate the whitening and de-whitening matrices (these handle dimensionality simultaneously). These data are used for further pre-processing.

Some application-dependent pre-processing steps are needed to be performed to achieve success in ICA for a given dataset. Some band-pass filtering may be very useful, if the data consists of time-signals. If we filter linearly the observed signals $x_i(t)$ to acquire new signals, say $x_i^*(t)$, the ICA model still holds for $\mathbf{x_i^*(t)}$, with the same mixing matrix. Now, time filtering of **X** (where **X** is a matrix that contains the observations **x(1),...,x(T)** as its columns, and similarly for **S**) corresponds to multiplying $X$ from the right by a matrix, **M**. This gives

$$\mathbf{X^* = XM = ASM = AS^*} \tag{6}$$

which shows that the ICA model still remains valid[21]. Data can be further pre-processed using in different algorithms, FastICA and KernalICA. Fixed-point algorithm has been implemented in this work for ICA.

### 2.5 Support Vector Machine

Support Vector Machine (SVM), proposed by Vapnik[26] is a tool used for binary class classification that performs classification tasks by constructing optimal separating hyper planes in multidimensional space[27]. SVM is used to classify nonlinear separable data with the help of the kernel function. Boser et al.[28] discusses about the kernel trick to accomplish the generalization of the methods used for linear classification in nonlinear classification. M. Alizadeh and M. M. Ebadzadeh[29] have been defined Support Vector Classification as a kernel-based classification learning technique. We have focused on 2 basic kernels: Polynomial and Linear. In this paper, we have used multiclass SVM to carry out recognition on face images, and Sequential Minimal Optimization (SMO) algorithm to train the SVM. The kernels used in SVM are expressed in the following way:

Linear: $k\left(x_i, x_j\right) = (x_i^T x_j + c)$

Polynomial: $k(x_i, x_j) = (1 + x_i^T x_j)^p$

## 3. EXPERIMENTAL DISCUSSIONS

### 3.1 Database descriptions

The benchmark database OTCBVS (Object Tracking and Classification Beyond Visible Spectrum) contains videos and images recorded in and beyond the visible spectrum; which contains different sets of data like OSU Infrared (IR) Pedestrian Database, IRIS Infrared (IR)/Visible Face Database, OSU Color-Infrared (IR) Database, Terravic Facial IR Database, Terravic Weapon IR Database, and CBSR NIR Face Dataset. Among all of these different datasets, IRIS (Imaging, Robotics and Intelligent System) Infrared (IR)/Visible face dataset has only been considered in our experiments. This database contains simultaneously acquired unregistered infrared (IR) and visible face images under variable illuminations, expressions, and poses. Two different sensors are used to capture the infrared (IR) and visual images of this database. Database images are 320×240 pixels color images in bmp format. The subjects were recorded in 3 different expressions Ex1 (Surprised), Ex2 (laughing), Ex3 (Anger) and mainly 5 different illuminations Lon (left light on), Ron (right light on), 2on (both lights on), dark (dark room), off (left and right lights off) with varying poses. Here, a question may be arise about the absence of other expressions(pain, sad/disgust, delight, threat); in this context, we have to mention that this database contains only three expressions of visual and their corresponding thermal images. The total size of the database is 1.83 GB[30]. There is another database of visual and their corresponding thermal images named Equinox, but it is not freely available now-a-days.

### 3.2 Experimental Results

In this section, results are given from a set of conducted experiments. The overall experiment is carried out with the help of illumination and full database (which comprises expression, illumination and others without dark and eyeglass face images). The illumination and full dataset consist of 17 and 28 classes with total images of 748 and 1626 respectively where illumination dataset comprises 44 images per class. The experiment is started with image registration process of visual and their corresponding thermal images. The registered visual and thermal images are then cropped into 50×50 dimensions. The cropped visual and thermal images are fused using gradient fusion method. After that, all the fused images are taken for dimension reduction using ICA. After reduction, the dimension of the newly generated images is 1×49. These reduced images are taken for training and testing purposes from different classes of different datasets. In illumination dataset, total 524 images are taken as training images and 224 images are considered as testing images. 1139 images and 487 images are used as training and testing images in full dataset respectively. SVM multiclass strategy 'one-vs.-all' has been taken in the experiment. So, n numbers of classification functions have been constructed where n is the number of classes[31]. For training support vector machine, Sequential Minimal Optimization **(**SMO) algorithm has been used where tolerance is 1.0000e-003 with which KKT conditions are checked, maximum number of iterations of the algorithm is 50,000 and size of the kernel matrix cache is 1000. Linear kernel and Polynomial kernel with degree 3 are used in our SVM kernel functions. The accuracy of a classification process defined as the portion of true positives, and true negatives in the population of all instances, classification accuracy A = (TP+TN) / (TP+TN+ FP+FN). TP=True Positives, TN=True Negatives, FP=False positives and FN=False Negatives. All these values have been taken from the confusion matrix acquired from the classifier performance. The classification accuracy of 2 different datasets is listed in table 1.

Table 1. Classification accuracy of 2 different dataset

| | **Recognition Rate** | |
|---|---|---|
| | **SVM(linear)** | **SVM(polynomial)** |
| **Illumination** | 65.34% | 94.15% |
| **Fulldata** | 69.05% | 96.48% |

Experiment results show that polynomial kernel generates more accurate result than linear kernel. Linear kernel produces 28.81% less accuracy than polynomial kernel. Likewise, polynomial kernel performs better than linear in full dataset. So, it is clearly observed that polynomial kernel outperforms linear kernel for both datasets.

## 4. CONCLUSION

An approach of face recognition using gradient fused image is presented here. The efficiency of our scheme has been demonstrated on IRIS database, which contain images gathered with varying lighting, facial expression, pose, and facial details. In this paper, a novel approach has been presented to recognize the human faces of varying expressions and illumination using feature level fusion with the help of Independent Component Analysis where support vector machines are used as classifiers.

## ACKNOWLEDGEMENT

The research has been supported by the grant from DIT, MCIT, Govt. of India, Vide No. 12(2)/2011-ESD, dated 29/03/2011.

## REFERENCES

[1] Jobson, D. J., Rahman, Z. –U. and Woodell, G. A., “Properties and performance of a center/surround retinex,” IEEE Trans. Image Processing 6(3), 451-462 (1997).

[2] Ramamoorthi, R., “Analytic PCA construction for theoretical analysis of lighting variability in images of a Lambertian object,” IEEE Trans. Pattern Analysis and Machine Intelligence 24(10), 1322-1333 (2002).

[3] Ramamoorthi, R. and Hanrahan, P., “An efficient representation for irradiance environment maps,” Proc. Computer graphics and interactive techniques, 497-500 (2001).

[4] Ramamoorthi, R., and Hanrahan, P., “On the relationship between radiance and irradiance: determining the illumination from images of a convex Lambertian object,” J. Opt. Soc. Am. 18(10), 2448-2459 (2001).

[5] Basri, R., and Jacobs, D., “Lambertian Reflectance and Linear Subspaces,” IEEE Trans. PAMI, 25(2), 218-233 (2003).

[6] Epstein, R., Hallanan, P. and Yuille, A. L., “5±2 Eigenimages suffice: An empirical investigation of low-dimensional lighting models,” Proc. Physics- Based Modeling in Computer Vision, 108-116 (1995).

[7] Hallanan, P. “A low-dimensional representation of human faces for arbitrary lighting conditions,” Proc. CVPR, 995-999 (1994).

[8] Hyvarinen, A. “Fast and robust Fixed-Point algorithms for Independent Component Analysis,” IEEE Trans. Neural Networks 10(3), 626-634 (1999).

[9] Raviv, T. R. and Shashua, A."The Quotient image: class based recognition and synthesis under varying illumination conditions," Proc. CVPR 2, 566-571 (1999).

[10] Shim, H., Luo, J. and Chen, T. “A subspace model-based approach to face relighting under unknown lighting and poses,” IEEE trans. image processing 17(8), 1331-1341 (2008).

[11] Turk, M. A. and Pentland, A. P. “Eigenfaces for recognition,” J. Cognitive Neuroscience 3(1), 71–86 (1991).

[12] Belhumeur, P. N., Hespanha, J. P. and Kriegman, D. J. “Eigenfaces vs. Fisherfaces: recognition using class specific linear projection,” IEEE Trans. PAMI 19(7), 711–720 (1997).

[13] Bartlett, M. S., Movellan, J. R. and Sejnowski, T. J. “Face Recognition by Independent Component Analysis,” IEEE Trans. Neural Networks 13 (6), 1450-1464 (2002).

[14] Pop, F. M., Gordan, M., Florea, C. and Vlaicu, A. “Face recognition under pose and expressivity variation using thermal and visible images,” J. UbiCC. 6, http://www.ubicc.org/files/pdf/paper_id_528_528.pdf.

[15] Liu, Z., Xu, Z. and Hu, L. A. “Comparative Study of Distance Metrics in Face Recognition,” J. theoretical and applied information technology 5(2), 206-211 (2009).

[16] Haddad, R. A. and Akansu, A. N. “A class of fast Gaussian Binomial Filters for speech and image processing,” IEEE Trans. Acoustics, Speech and Signal Processing 39(3), 723-727 (1991).

[17] Poggio, T., Torre, V. and Koch, C. “Computational vision and regularization theory,” J. Nature 317(6035), 314-319 (1985).

[18] Zhang, T., Tang, Y. Y., Fang, B., Shang, Z. and Liu, X. “Face recognition under varying illumination using gradient faces,” IEEE Trans. Image Processing 18 (11), 2599-2606 (2009).

[19] Wahid, K. A., Dimitrov, V. S., Jullien, G. A. and Badawy, W. “An Algebraic Integer Based Encoding Scheme for Implementing Daubechies Discrete Wavelet Transforms,” Proc. Signals, Systems and Computers 1, 967–971 (2002).

[20] Shu-long, Z. “Image Fusion using Wavelet Transform,” Int. Symp. Geospatial Theory, Processing and Applications, 5-9 (2004).

[21] Hyvarinen, A. and Oja, E. “Independent Component Analysis: Algorithms and Applications,” J. Neural Networks, 13 (4-5), 411-430 (2000).

[22] Li, L. "Flexible Kernel Independent Component Analysis Algorithm and Its Local Stability on Feature Space," Proc. Machine Learning and Cybernetics, 2990-2994 (2006).
[23] Lee, T. W. Girolami, M. and Sejnowski, T. J. "Independent Component Analysis Using an Extended Infomax Algorithm for Mixed Sub-Gaussian and Super-Gaussian Sources," J. Neural Computation 11 (2), 417-441 (1999).
[24] Comon, P. "Independent Component Analysis, a new concept," J. Signal Processing 36 (3), 287-314 (1994).
[25] Bach, F. R. and Jordan, M. I. "Kernel Independent Component Analysis," J. Machine Learning Research 3, 1-48 (2002).
[26] He, L. -M., Yang, X. -B. and Lu, H.-J. "A comparison of Support vector machine ensemble for classification," Proc. Machine Learning and Cybernetics, 3613–3617 (2007).
[27] http://www.statsoft.com/textbook/support-vector-machines/.
[28] Boser, B. E., Guyon, I. M. and Vapnik, V. "A training algorithm for maximum margin classifiers," Workshop on Computational Learning Theory, 144-152 (1992).
[29] Alizadeh, M. and Ebadzadeh, M. M. "Kernel Evolution for Support Vector Classification," Workshop EAIS, 93–99 2001.
[30] http://www.cse.ohio-state.edu/otcbvs-bench/
[31] Lihong, Z., Ying, S., Yushi, Z. Cheng, Z. and Yi, Z. "Face recognition based on multi-class SVM," Proc. CCDC, 5871 – 5873 (2009).