

Structural Geodesic-Tchebychev Transform: An Image Similarity Measure for Face Recognition

¹Ali Nadhim Razzaq, ^{2,3}Zahir M. Hussain and ²Hind Rustum Mohammed

^{1,2}Faculty of Computer Science and Math, University of Kufa, Iraq

³School of Engineering, ECU, Australia

Article history

Received: 7-04-2016

Revised: 7-11-2016

Accepted: 28-11-2016

Corresponding Author:

Zahir M. Hussain

Faculty of Computer Science

and Math, University of Kufa,

Iraq

Email: zmhussain@ieec.org

Abstract: This work presents a new holistic measure for face recognition. Face recognition involves three steps: Face Detection, Feature Extraction and Matching. In the face detection process to identify the face area in face images, Viola-Jones algorithm has been used. Feature extraction is based on performing double-transformation, where discrete Tchebychev transform is performed on the geodesic distance transform of the grayscale image. Structural Similarity (SSIM) is applied to the resulting image double-transform to find matching factor with other image faces in the FEI (Brazilian) database. Performance is measured using a confidence criterion based on the similarity distance between the recognized person (best match) and the next possible ambiguity (second-best match). Simulation results showed that the proposed approach handles the face recognition efficiently as compared with SSIM.

Keywords: Discrete Tchebychev Moments, Generalized Geodesy via Geodesic Time, Structural Similarity (SSIM), Viola-Jones, Face Recognition, Image Processing

Introduction

Face image is a complex, varied, high-dimensional pattern. Although people recognize familiar faces easily, using the machine to accurately identify the face is still a difficult task, where correct recognition can fall below 60% (McCormick, 2013; Hashim and Hussain, 2014). However, face recognition in the areas of authentication and security systems has been widely used. Generally, face recognition is an unresolved problem that attracted many researchers, and continued to be an important area of research in computer vision and pattern recognition (Wang *et al.*, 2015).

Different methods have been proposed for face recognition. These methods generally involve a feature extractor like Principal Component Analysis (PCA) or wavelet decomposer to reduce the size of input, and a classifier like neural networks, support vector machines, and nearest distance classifiers to find the features which are most likely to be found. In the field of face recognition, the accuracy and robustness of the algorithm largely depend on the facial feature description (Zhu and Can, 2012). In general, the facial features include global and local features. Global features mainly describe the whole properties of the face, such as color, shape and are used for general

matching; while local features are used for confirming accuracy, mainly used to describe the fine details on human face, such as scars and dimples (Zhu and Can, 2012).

Moments invariants were firstly introduced by Hu (1962), who proposed a method of deriving moment invariants using algebraic methods. He used geometric moments to generate a set of invariants. However, geometric moments are not derived from a family of orthogonal functions, and are sensitive to noise, especially for higher order moments (Li, 1995). Thus, Hu's moment invariants have limited applications. Many works in the literature have presented approaches to construct moment invariants, such as Zernike moment (Chong *et al.*, 2003a), pseudo-Zernike moment (Chong *et al.*, 2003b) and Legendre moment (Chong *et al.*, 2004). However, in these approaches the accuracy of recognition descends due to the discrete approximation of the continuous integrals (Mukundan, 2004). To resolve these problems, Mukundan proposed discrete orthogonal Tchebychev moments (Mukundan *et al.*, 2001).

Face recognition in the visible spectrum is a well known field of research over many years of study using low-cost cameras. Many publications are available in

this direction, containing smart and sophisticated algorithms. The field of face recognition has undergone a steady growth due to its security applications; also, other applications such as access permission and even identity control. However, this field is still facing unresolved problems like the diminished recognition ability of algorithms under variations in the intensity of illumination in the image (Hermosilla *et al.*, 2015). Still failure of recognition is the main problem. This work focuses on accuracy of recognition where a double transform is found to optimize revealing hidden distinctive features of a particular face.

There are basically three main directions for face recognition as follows (Pandya *et al.*, 2013):

Feature-Based Approach

In feature-based approach, the local features like nose, eyes are segmented and can be used as input data in face detection to support the task of face recognition.

Holistic Approach

In holistic approach the whole face taken as the input in the face detection system to perform face recognition.

Hybrid Approach

Hybrid approach is combination of feature-based and holistic approach. In this approach both local and whole face is used as the input to face detection system.

In this work the holistic approach has been considered. The paper is organized as follows. Section 2 presents a brief on theoretical and practical basics. Section 3 presents the proposed method. Section 4 presents the results of applying proposed method for face recognition. Section 5 presents the performance of the proposed method under noise. Section 6 presents the discussions, and Section 7 presents a conclusion.

Background

In this section we present the different techniques used on the method proposed. There are three main steps in the proposed method:

- Pre-processing
- Feature extraction
- Matching
- Pre-Processing

The following pre-processing operations are needed before applying the proposed method:

Modifying Image Scales

All images must be square and have even dimensions.

Face Detection

Face detection is a necessary preprocessing stage. Viola-Jones algorithm (Viola and Jones, 2001) is applied to detect the face from image. Face detection is the process that determines the locations and sizes of human faces in digital images.

It detects facial features and ignores anything else. MATLAB vision function CascadeObjectDetector uses the Viola-Jones Algorithm (VJA) to detect objects in an image. This detector uses Haar-like features and a cascade of classifiers. The cascade object detector is pre-trained to detect faces, noses and other objects. The detection of face from image is represented in Fig. 1.

The algorithm has four stages:

- Haar Feature Selection
- Creating an Integral Image
- Adaboost Training
- Cascading Classifiers

After this step, we get a square image of size $N \times N$ containing face only $x(i,j)$.

Feature Extraction

Generalized Geodesy via Geodesic Time Transform

The time necessary to cover a path on a grey-scale image is the sum of the grey-level values along the path. The geodesic time between two points in a grey-scale image is defined as the smallest amount of time allowing to link these points. The geodesic time allows the definition of generalized geodesic distances and dilations. It will be shown that the application of minimal path extraction on greyscale images gives highly distinctive features if used in conjunction with Tchebychev moments.

The geodesic time along a path P with length l connecting a pixel $P_0 = x(i,j)$ [in a gray-scale image x] and a seed pixel $q = P_l = x(r, c)$ in the same image is given by (Soille, 1994):

$$D_x(P) = \frac{P_0}{2} + \frac{P_1}{2} + \sum_{i=1}^{l-1} P_i \quad (1)$$

The geodesic distance transform denoted by $d_x(i,j) = d_x(x(i,j))$ is the smallest amount of geodesic time D_x needed to link the general image pixel $p = x(i,j)$ to the specific seed image pixel $q = x(r,c)$ as follows:

$$d_x(i,j) = \inf\{D_x(P) \mid P \text{ links } x(i,j) \text{ to } q\} \quad (2)$$

where, \inf denotes the infimum operator.

The result in Fig. 2 has been simulated as per Equation 2 links between p and q .

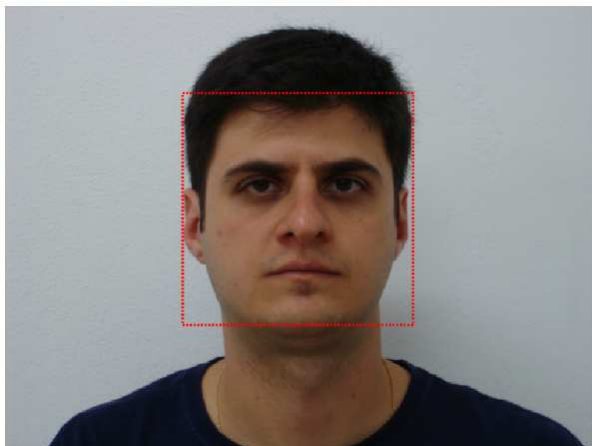


Fig. 1. Face detection with VJA using a face image from FEI Face Database

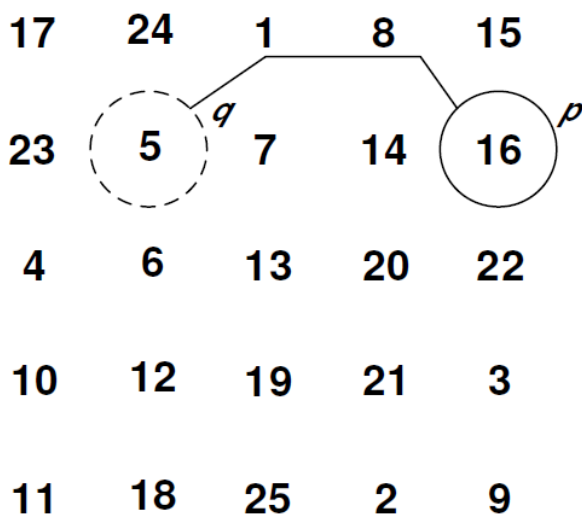


Fig. 2. Geodesic or shortest path between point p and seed q on a grey scale image. The time necessary to cover this path equals (19.5)

Discrete Tchebychev Moments

Discrete orthogonal moments such as the Tchebychev (Chebyshev) moments. The use of discrete orthogonal Tchebychev polynomials as basis function for image moments eliminates the discrete approximation associated with the continuous moments. In this work we will use Tchebychev moments for the image geodesic transform instead of the image itself. The Normalized Discrete Tchebychev polynomials are given by the following recurrence formula (Mukundan, 2004):

$$p_0(r) = \frac{1}{\sqrt{N}} \quad (3)$$

$$p_1(r) = (2r+1-N) \sqrt{\frac{3}{N(N^2-1)}} \quad (4)$$

$$p_k(r) = (a_k r + b_k) p_{k-1}(r) - c_k p_{k-2}(r) \quad (5)$$

with $k, r = 0, \dots, N-1$

Where:

$$a_k = \frac{2}{k} \sqrt{\frac{4k^2-1}{N^2-k^2}} \quad (6)$$

$$b_k = \frac{1-N}{k} \sqrt{\frac{4k^2-1}{N^2-k^2}} \quad (7)$$

$$c_k = \frac{k-1}{k} \sqrt{\frac{2k+1}{2k-3} \cdot \frac{N^2-(k-1)^2}{N^2-k^2}} \quad (8)$$

Now the 2D normalized discrete Tchebychev moments T_{mn} for an $N \times N$ generalized geodesic transform $d_x = \tau(x) = [d_x(i, j)]$ of the image $x(i, j)$ are given by:

$$g_x(m, n, d_x) = T_{mn}\{t(x)\} = T_{mn}(d_x) \quad (9)$$

$$= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} p_m(i) \cdot p_n(j) d_x(i, j)$$

With $m, n = 0, \dots, N-1$. $p_m(i)$: The m^{th} normalized discrete Tchebychev polynomials for the variable i .

Image Structural Similarity

In this study we will extend the application of the well-known similarity measure SSIM in (Wang *et al.*, 2004) to extract features from geodesic-Tchebychev transforms of images as defined in the next section. The Structural Similarity (SSIM) index is a novel method for measuring the similarity between two images. It can be viewed as a quality measure of one of the images being compared, provided the other reference image is regarded as of perfect quality.

To find similarity between two images x and y of common size $\times N$, the SSIM index is calculated as the average of the following local similarity:

$$SSIM(x, y) = \frac{(2M_x M_y + c_1)(2\sigma_{xy} + c_2)}{(2M_x^2 M_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (10)$$

with:

- M_x = The local average of x over a window
- M_y = The local average of y
- σ_x^2 = The local variance of x
- σ_y^2 = The local variance of y
- σ_{xy} = The local covariance of x and y
- c_1, c_2 = Two variables to stabilize the division in case of weak denominator

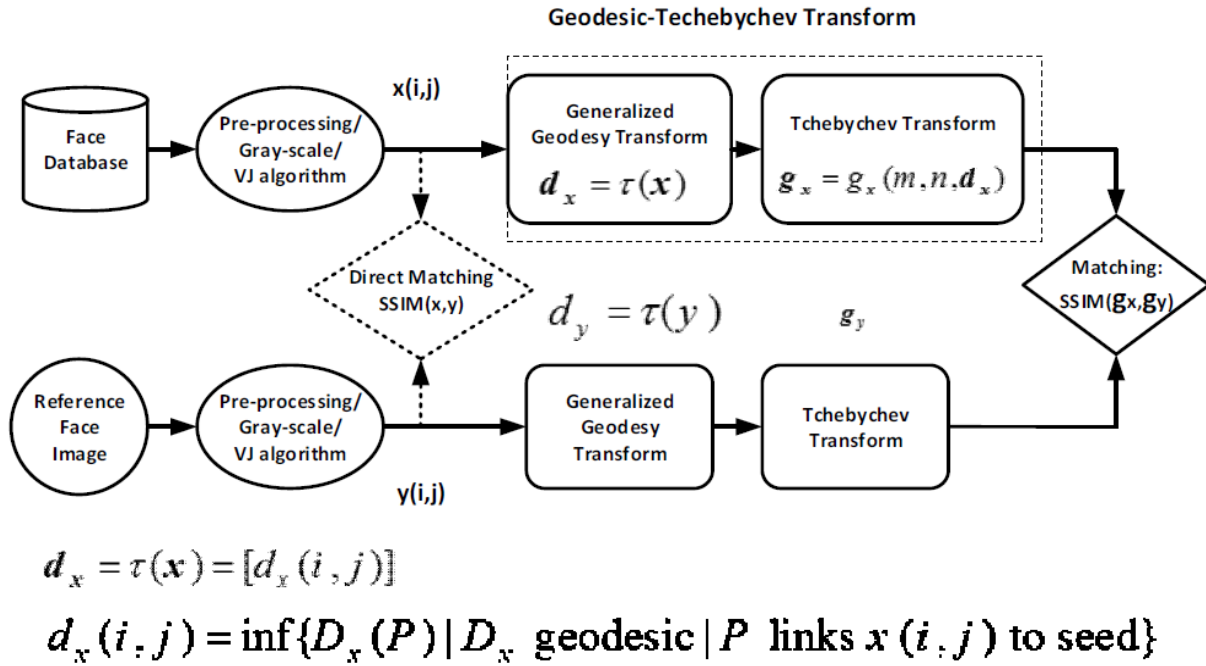


Fig. 3. Face recognition based on geodesic Tchebychev transform

Geodesic-Tchebychev Similarity

Both Tchebychev moments and geodesic distance are good tools to extract image features; however, face recognition involves a lot of failures. This could be due to the fact the different approaches may not be able to show their power when used separately. In this work a double transform that is based on the interaction between the powerful Tchebychev and geodesic transforms is designed to reveal an outstanding performance as compared to using each transform alone.

The geodesic distance is an optimized way to handle the distances between pixels in the sense of the nearest paths between pixels and a specified center, hence reducing the effects of varying illumination and noise. Applying Tchebychev moments to geodesic transform will get the distinctive geodesic features from the original image. Now applying a statistical approach like SSIM based on the mean, variance, and covariance will sieve the best of these distinctive features, giving superior performance in face recognition.

The overall diagram of the proposed system is shown as Fig. 3.

The Geodesic-Tchebychev similarity measure (called SSIMGT) is the average of the following local measure performed on smaller windows of the original images:

$$R(x, y) = SSIM[g_x, g_y] \quad (11)$$

with:

$x(i, j)$ = The original face image

$g_x = [g_x(i, j, d_x)]$ = A discrete Tchebychev moment for the geodesic transform $d_x(i, j)$ of original image x

SSIM = The similarity index as per equation (10)

Another High-Performance Structural

Measure in the Tchebychev Domain In this section, another high-performance similarity measure that is directly based on the statistical structure in the Tchebychev domain (called SSIMT) is proposed as follows:

$$S(x, y) = SSIM[T(x), T(y)] \quad (12)$$

where, $T(x) = [T_{mn}(x)]; T(y) = [T_{mn}(y)]$.

For most cases, SSIMT performance is inferior by a small amount to that of SSIMGT. For a few cases, the situation is reversed. Hence, SSIMT remains a strong competitor for SSIMGT.

MSE-Based Similarity Measure in the Tchebychev Domain

For the purpose of performance comparison, the following similarity measure based on the meansquared error in the Tchebychev domain (called MSETS) is defined by Equation 14 as follows:

$$E(x, y) = MSE[T(x), T(y)] \quad (12)$$

$$r(x, y) = 1 - E(x, y) / \max\{E(x, y)\} \quad (13)$$

$$p(x, y) = r(x, y) / \max\{r(x, y)\} \quad (14)$$

with $T(x) = [T_{mn}(x)]$; $T(y) = [T_{mn}(y)]$ and $MSE(x, y)$ is define as:

$$MSE(x, y) = \frac{1}{MN} \sum_i \sum_j [x(i, j) - y(i, j)]^2$$

This measure is weaker than SSIMGT and SSIMT; giving less recognition confidence.

Results and Discussion

Tests to examine SSIMGT performance can be summarized in the following steps:

- Face area should be identified in the original face image within the *pre-processing* phase. First, the image is converted from a color image to a gray-scale image. The face area of the image is extracted using Viola-Jones algorithm. It should be square of unified size $N \times N$. In this work, FEI Face Database

(Brazilian) (Carlos Eduardo Thomaz (Caru)) has been used. It consists of 50 people, each with two poses. The size of the face is 250×250

- The resulting image $x(i, j)$ from step (a) passes through distance transform as per Equations 1 and 2. The seed row and columns vectors are chosen as random integers within the face image size of length 3 each
- The resulting transform in step (b), $d_x(i, j)$, passes through a discrete orthogonal Tchebychev transform to get Tchebychev moments $g_x(i, j)$. The low-pass area of size 30-by-30 is chosen from the Tchebychev transform matrix
- The last step is to apply structure similarity (SSIM) to the selected area of the double transform above to reveal exceptional amount of similarity between the reference image transform $g_x(i, j)$ and database image transform $g_y(i, j)$

Other measures are based on steps related to those mentioned above. For SSIM, the constants $c_1 = (K_1L)^2$ and $c_2 = (K_2L)^2$ (K_1 and K_2 being small constants, $L=255$) where chosen with $K_1 = 0.01$ and $K_2 = 0.003$ as per (Wang *et al.*, 2004). The performance of SSIM is insensitive to these constants.



Fig. 4. A sample from FEI Face Database (Brazilian)

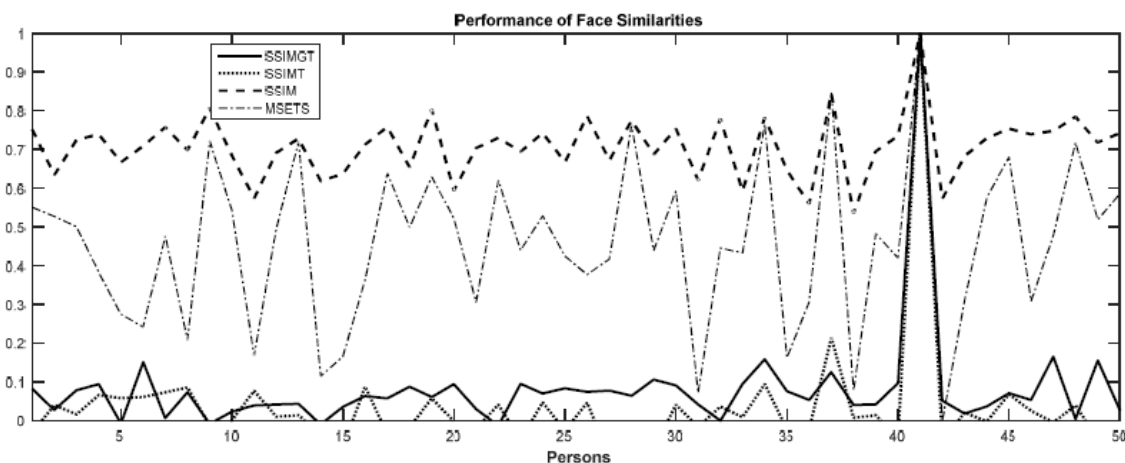


Fig. 5. Similarity measures using 50-face sub-database from FEI Face Database (Brazilian). Confidence in recognition for SSIMGT, SSIMT, SSIM, and MSETS is 83.6%, 78.6%, 15.4% and 15.3%, respectively

Confidence in Recognition

To test performance in recognition, a differential performance measure $d(r)$ for a similarity measure r is defined using the difference between the maximum value of the measure r (should be 1 in normalized measures), which is supposed to occur on the point (integer count) representing the correct person (in case of correct recognition), and the second-largest value, considered as the most serious confusion in the recognition process. The measure for recognition confidence is defined as follows:

$$c(r) = [1 - d(r)] \times 100\% \quad (15)$$

Figure 4 shows a sample of face the images used. In this test 50 images are selected as a database. Figure 5 shows the results of one case of applying the above measures of face recognition. Performance comparison based on confidence defined above for different measures in 50 different cases using 50 different reference face images showed that the geodesic-Tchebychev measure SSIMGT displays better performance for most of the cases. The only competitor is SSIMT in some cases only. It is clear that Tchebychev transform alone cannot reveal the distinctive face features, however, if used with a structural and geodesic transforms, it goes deeper into image analysis.

Conclusion

This paper presents three methods for face recognition. The process is to identify a reference face in a database which contains a different pose image for the same reference person. These methods, called SSIMGT, SSIMT and MSETS are based on Tchebychev transform, along with a definition for confidence in recognition. Simulation results show that the most efficient among these methods (in terms of recognition confidence) is the similarity measure SSIMGT, obtained by applying the well-known structural similarity SSIM on the double geodesic -Tchebychev transform of the grayscale version of the face images.

Funding Information

The Authors would like to thank the Ministry of Higher Education & Scientific Research (Iraq) and Edith Cown University (Australia) for financial support of this MSc (Research) project.

Author's Contribution

Ali Nadhim Razzaq: Analysis of proposed method, simulation, and writing.

Zahir M. Hussain: Mathematical formulation, simulation and writing.

Hind Rustum Mohammed: Simulation and data analysis.

Ethics

The Authors would like to confirm that there are no ethical issues associated with this work.

References

- Chong, C.W., P. Raveendran and R. Mukundan, 2003a. The scale invariants of pseudo-Zernike moments. *Pattern Analysis Applications*, 6: 176-184. DOI: 10.1007/s10044-002-0183-5
- Chong, C.W., P. Raveendran and R. Mukundan, 2003b. Translation invariants of Zernike moments. *Pattern Recognition*, 36: 1765-1773. DOI: 10.1016/S0031-3203(02)00353-9
- Chong, C.W., P. Raveendran and R. Mukundan, 2004. Translation and scale invariants of Legendre moments. *Pattern Recognition*, 37: 119-129. DOI: 10.1016/j.patcog.2003.06.003
- Hashim, A.N. and Z.M. Hussain, 2014. Local and semi-global feature-correlative techniques for face recognition. *Int. J. Advanced Comput. Sci. Applications*, 5: 157-167. DOI: 10.14569/IJACSA.2014.050824
- Hermosilla, G., F. Gallardo, G. Farias and C.S. Martin, 2015. Fusion of visible and thermal descriptors using genetic algorithms for face recognition systems. *Sensors*, 1: 17944-17962.
- Hu, M.K., 1962. Visual pattern recognition by moment invariants. *IRE Trans. Information Theory*, 8: 179-187. DOI: 10.1109/TIT.1962.1057692
- Li, B., 1995. High-order moment computation of gray-level images. *IEEE Trans. Image Processing*, 4: 502-505. DOI: 10.1109/83.370680
- McCormick, D., 2013. Face Recognition Failed to Find Boston Bombers. *IEEE Spectrum*.
- Mukundan, R., 2004. Some computational aspects of discrete orthonormal moments. *IEEE Trans. Image Processing*, 13: 1055-1059. DOI: 10.1109/TIP.2004.828430
- Mukundan, R., S.H. Ong and P.A. Lee, 2001. Image analysis by Tchebichef moments. *IEEE Trans. Image Processing*, 10: 1357-1364. DOI: 10.1109/83.941859
- Pandya, J.M., D. Rathod and J.J. Jadav, 2013. A survey of face recognition approach. *Int. J. Eng. Res. Applications*, 3: 632-635.
- Personal Homepage of Dr. Carlos Eduardo Thomaz (Caru).
- Soille, P., 1994. Generalized geodesy via geodesic time. *Pattern Recognition Letters*, 15: 1235-1240. DOI: 10.1016/0167-8655(94)90113-9

- Viola, P. and M. Jones, 2001. Rapid object detection using a boosted cascade of simple features. Proceedings of the Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Dec. 8-14, IEEE Xplore press, Cambridge, MA, USA.
DOI: 10.1109/CVPR.2001.990517
- Wang, Y., C. Wang and L. Liang, 2015. Sparse representation theory and its application for face recognition. *Int. J. Smart Sensing Intelligent Systems*, 8: 107-124.
- Wang, Z., A.C. Bovik, H.R. Sheikh and E.P. Simon, 2004. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Processing*, 13: 600-612.
DOI: 10.1109/TIP.2003.819861
- Zhu, Y. and S. Can, 2012. Sub-image method based on features sampling and feature fusion for face recognition. *J. Software*, 23: 3210-3220.