Texture recognition by using a non-linear kernel

Essam Al-Daoud*, Feras Al-Hanandeh,
Emad E. Abdallah and Alaa E. Abdallah

Department of Computer Science,
Faculty of Science and Information Technology,
Zarqa University,
Zarqa, Jordan
Email: essamdz@zu.edu.jo
Email: feras@hu.edu.jo
Email: Emad@hu.edu.jo
Email: aabdallah@hu.edu.jo
*Corresponding author

Abstract: This study proposes the use of features combination and a non-linear kernel to improve the classification rate of texture recognition. The feature vector concatenates three different sets of feature: the first set is extracted using grey-level cooccurrence matrix, the second set is collected from three different radii of local binary patterns, and the third set is generated using Gabor wavelet features. Gabor features are the mean, the standard deviation, and the skew of each scaling and orientation parameter. The aim of the new kernel is to incorporate the power of the kernel methods with the optimal balance derived from the features. To verify the effectiveness of the proposed method, numerous techniques are tested using the three data sets, which consist of various orientations, configurations and lighting conditions.

Keywords: texture recognition; cooccurrence matrix; Gabor wavelet; LBP; non-linear kerner.

Reference to this paper should be made as follows: Al-Daoud, E., Al-Hanandeh, F., Abdallah, E.E. and Abdallah, A.E. (2013) 'Texture recognition by using a non-linear kernel', *Int. J. Computer Applications in Technology*, Vol. 48, No. 3, pp.235–240.

Biographical notes: Essam Al-Daoud is currently an Associate Professor and the Chairman at Computer Science Department in Zarka Private University. He received his BSc from Mu'tah University, MSc from Al al-Bayt University and his PhD in Computer Science from University Putra Malaysia in 2002. His research interests include machine learning, soft-computing, cryptography, bioinformatics and quantum cryptography.

Feras Al-Hanandeh is currently an Assistant Professor at the Prince Al Hussein Bin Abdullah II, Faculty of Information Technology, Al-Hashemite University, Jordan. He obtained his PhD in Computer Science from the University Putra Malaysia, Malaysia, in 2006. His current research interests include distributed databases, parallel databases focusing on issues related to integrity maintenance, transaction processing and query processing.

Emad E. Abdallah is currently an Assistant Professor in the Department of Computer Information Systems at the Hashemite University (HU). He received his PhD in Computer Science from Concordia University in 2008, where he worked on multimedia security, pattern recognition and 3D object recognition. He received his BS in Computer Science from Yarmouk University, Jordan, and his MS in Computer Science from the University of Jordan in 2000 and 2004, respectively. His current research interests include computer graphics, multimedia security and pattern recognition.

Alaa E. Abdallah is an Assistant Professor in the Department of Computer Science of Hashemite University, having joined in 2011. He obtained his BSc in Computer Science from Yarmouk University in 2000, MSc in Computer Science University of Jordan in 2003, PhD in Computer Science from Concordia University in 2008, Montreal-Canada. His research interest includes the routing protocols for ad hoc networks, parallel and distributed systems, and multimedia security.

1 Introduction

In most industrial applications, texture analyses combined with image processing techniques are important recognition and classification methods (Nasirzadeh et al., 2010; Zhao and Wang, 2012). Basically, texture recognition tasks always involve two steps: the feature extraction step, where texture features are extracted from the image; and the recognition step, where texture class membership is assigned according to the extracted texture features (Prasetiyo et al., 2010; Varma and Zisserman, 2009). A large number of feature extraction methods have been proposed and tested in various pattern recognition tasks. These methods include, but are not limited to, statistical, structural, signal processing and model-based approaches (Kobayash et al., 2009; Xie and Mirmehdi, 2008). Vácha and Haindl (2010) proposed a statistical scheme for recognising three-dimensional textures shown in motion images. The texture characteristics emerged from the distinct movement of the motion images, and the dynamic cues are especially useful for recognising ambiguous texture patterns in noisy images. They applied cubic higher order auto-correlation (CHLAC) to extract features of both the textures and their movements. Liu et al. (2011) proposed two new sparse descriptors. The authors showed that more discriminative features can be extracted from a single image by combining sparse LBP and sparse HoG. Prabhu and Savvides (2011) used 3D generic elastic model to construct a 3D image for each subject from a 2D image. Their model can be applied to uncontrolled real-world images where variations in expression, illumination and pose are encountered. Wolf et al. (2011) captured local similarities by designing a family of novel descriptors. Moreover, they used unlabelled background samples to enhance classification performance.

2 Grey-level cooccurrence matrix

Another way to describe an image is by counting the pairs of pixels with grey-level i that occur diagonally, vertically or horizontally to adjacent pixels with grey-level j and at a given distance d. For example, if we have five grey-levels (0, 1, 2, 3 and 5) and an image represented as follows (Yong et al., 2008):

$$\text{Im age} = \begin{bmatrix} 0 & 0 & 1 & 4 \\ 3 & 1 & 1 & 4 \\ 4 & 0 & 2 & 3 \\ 4 & 1 & 3 & 2 \end{bmatrix}$$

then the vertical (right) cooccurrence matrix with distance 1 can be represented as follows:

$$C = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Several cooccurrence matrices can be incorporated using different orientations and distances. In this paper, seven features are extracted from matrix C for each of the orientations at 0°, 45°, 90°, 135° and the distance 1. The total number of features is 28. The seven features are as follows:

Entropy $c_1 = -\sum_{i,j=0}^{\text{n}} c(i,j) \;\log\; c(i,j)$

Intertia $c_2 = \sum_{i,j=0}^{n} (i-j)^2 c(i,j)$

Cluster shade $c_3 = \sum_{i,\text{j}=0}^{n} \left(i - M_x + j - M_y\right)^3 c\left(i,j\right)$

Local homogeneity $c_4 = \sum_{i,j=0}^{n} \frac{1}{1+(i-j)^2}\, c(i,j)$

Total energy $c_5 = \sum_{i,j=0}^{n} c^2(i,j)$

Measure of correlation $c_6 = \frac{c_1 - H_{xy}}{\max\left(H_x\;, H_y\right)}$

Cluster prominence $c_7 = \sum_{i,j=0}^{n} (i - \text{M}_\text{x} + j - M_y)^4 c(i,j)$

where

$$H_x = -\sum_{i=0}^{n} S_x(i) \log S_x(i),$$

$$H_y = -\sum_{i=0}^{n} S_y(i) \log S_y(i),$$

$$H_{xy} = -\sum_{i,j=0}^{n} c(i,j) \;\log S_x(i) S_y(i),$$

$$S_x(i) = \sum_{i=0}^{n} c(i,j),$$

$$S_y(i) = \sum_{j=0}^{n} c(i,j),$$

$$M_x(i) = \sum_{i,j=0}^{n} ic\left(i,j\right),$$

$$M_y(i) = \sum_{i,j=0}^{n} jc(i,j),$$

3 Two-dimensional Gabor wavelet

The frequency and orientation of the two-dimensional Gabor wavelet are similar to the human visual system; it contains very rich information about the structure of the image and the texture representation. It has been successfully used with numerous image processing techniques, such as fingerprint recognition, handwritten numerals recognition, face recognition and texture segmentation (Meyers and Wolf, 2008; Tenllado, et al., 2008). 2D Gabor function is defined as follows:

$$g(x,y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right) \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + 2\pi jWx\right) \quad (1)$$

The Gabor wavelet is

$$g_{m,n}(x,y) = a^{-m} g(x',y'),\ a > 1,$$
$$x' = a^{-m}\left(x\ \cos\ \theta + y\ \sin\ \theta\right)$$
$$y' = a^{-m}\left(-x\ \sin\ \theta + y\ \cos\ \theta\right)$$

where $j = \sqrt{-1}$, W is the radial frequency, σ_x and σ_y are the scaling parameters, m and n are integers, a^{-m} is to ensure that the energy is independent of m, $\theta = n\pi/K$ is the orientation of the filter and K is the number of orientations. Thus, the Gabor wavelet transformation of the image $I(x, y)$ is

$$GW_{m,n} = \iint I\left(x_1, y_1\right) \overline{g_{m,n}}\left(x - x_1, y - y_1\right) dx_1 dy_1 \tag{2}$$

In this paper, we use 48 Gabor wavelet features, which are concatenation of the values $\left\{\mu_{m,n}, \sigma_{m,n}, S_{m,n}\right\}$, where $m = 1, \ldots, 4$, $n = 1, \ldots, 4$ and

$$\mu_{m,n} = \iint \left|WG_{m,n}\left(x,y\right)\right| dx\ dy \tag{3}$$

$$\sigma_{m,n} = \sqrt{\iint \left(\left|WG_{m,n}(x,y)\right| - \mu_{m,n}\right)^2 dxdy} \tag{4}$$

$$S_{m,n} = \iint \left(\frac{\left|WG_{m,n}\left(x,y\right)\right| - \mu_{m,n}}{\sigma_{m,n}}\right)^3 \tag{5}$$

4 Local binary patterns

Local Binary Patterns (LBP) has achieved impressive classification results and has been extended into various fields. The most attractive advantage of LBP is its invariance to scale and rotation changes. The LBP process starts by comparing every pixel with its eight neighbours (Ahonen et al., 2006). The values of the LBP cells can be calculated as follows:

$$LBP\left(x,y\right) = \sum_{n=0}^{7} t\left(i_n - i_c\right) 2^n \tag{6}$$

where i_n is the grey values of the eight pixels and i_c is the grey value of the centre and

$$t\left(x\right) = \begin{cases} 1 & \text{if} \quad x \geq 0 \\ 0 & \text{if} \quad x < 0 \end{cases}$$

An illustration of this process is shown in Figure 1.

To achieve rotation invariant, LBP is extended by allowing different radii and samples. For example, $LBP_{16,2}^{u2}$ means 16 points on a circle of radius 2 and the superscript $u2$ indicates a uniform pattern. The general form can be expressed as follows (Prasetiyo et al., 2010):

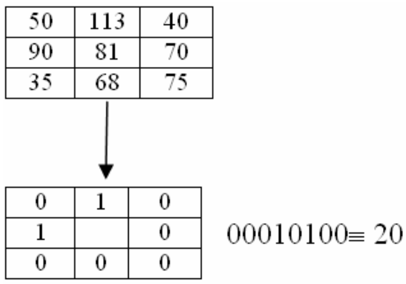
$$LBP_{N,R}^{u2} = \begin{cases} \sum_{n=1}^{N} t\left(i_n - i_c\right) & if \quad U\left(LBP_{N,R}\right) \geq 2 \\ N + 1 & \text{otherwise} \end{cases} \tag{7}$$

where

$$U\left(LBP_{N,R}\right) = \left|t\left(i_N - i_c\right) t\left(i_1 - i_c\right)\right| + \sum_{n=2}^{N} \left|t\left(i_n - i_c\right) - t\left(i_{n-1} - i_c\right)\right|$$

Therefore, the number of extracted features using $LBP_{N,R}$ and the histogram representation are at most $N + 2$. In this paper, concatenation of LBP_8, LBP_{16} and LBP_{24} are used and so the total number of features is 54.

Figure 1 The original LBP operator



5 Support Vector Machine (SVM)

The SVM is a classification method that tries to find the optimal hyperplanes between different classes. If the training data contain mislabelled points, then a slack variable should be added to the SVM optimisation problem (Zhang et al., 2006; Fan, 2012; Shi, 2012). Thus, an SVM problem can be expressed as follows:

$$\text{M ax} \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j X_i^T X_j \tag{8}$$

Subject to $0 \leq \alpha_i \leq C$

$$\sum_{i=1}^{N} \alpha_i y_i = 0$$

where C measures the degree of misclassification, αs are the support vectors, y is the target, and X is the feature vector. If the data are not linearly separable, then the optimisation problem must be extended to a non-linear case. Non-linear SVM can be written as follows:

$$\text{Max} \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j K\left(X_i, X_j\right) \tag{9}$$

Subject to $0 \leq \alpha_i \leq C$

$$\sum_{i=1}^{N} \alpha_i y_i = 0$$

where K is a suitable kernel function. The decision function is

$$F(x) = \sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b \tag{10}$$

Popular predefined kernel functions, such as:

Radial basis function (RBF): $K(x, y) = e^{\left(-\gamma\|x-y\|^2\right)}$

Polynomial: $K(x, y) = \left(x.y + 1\right)^d$

Sigmoid: $K(x,y) = \tan h(\kappa x.y + c)$

6 New weighted kernel

To ensure the solution of an SVM optimisation problem converges to a unique solution, the kernels must be positive semi-definite. A symmetric kernel K is called positive semi-definite if and only if for any complex numbers $c_1, c_2, \ldots, c_n$ we have

$$\sum_{i,j=1}^{n} c_i c_j K\left(x_i, x_j\right) \geq 0 \qquad (11)$$

The selection of a suitable kernel for a given application or for a set of features is still an open problem. For this reason we suggest, in the following proposition, a new kernel with free parameters:

Proposition: *Let A be a diagonal matrix of size n × n and its elements are non-negative elements, then the kernel* $K\left(X, Y\right) = \left(1 + X^t AY\right)^d$ *is positive semi-definite.*

Proof: Since any diagonal non-negative matrix is positive semi-definite, then from the definition, the matrix X^tAY is positive semi-definite; and since a matrix of positive constant is positive semi-definite, the summation and multiplication of a positive semi-definite matrix is again positive semi-definite, then $\left(1 + X^t AY\right)^d$ is positive semi-definite.

The matrix A can be used to optimise the weight of the features. Grid search or genetic algorithm can be applied to find the suitable values of the diagonal matrix A. Although this approach is time-consuming, the values need to be found only once and can be reapplied to several similar data sets and applications. If the feature vector is concatenation of features that are extracted using M methods, then the values of the diagonal can be grouped into M sets. In this study, the feature vector is concatenation of three sets: 28 features from grey-level cooccurrence matrix, 54 features from Gabor wavelet and 48 from LBP.

7 Experimental results

To test the suggested method, three databases ware used: KTH-TIPS database (Fritz et al., 2011), OUTEX database (Ojala at al., 2002) and UIUC database (Lazebnik et al., 2005). KTH is the abbreviation for Kungliga Tekniska Högskolan University, and TIPS stands for textures under varying illumination, pose and scale. Images were taken at nine different scales spanning two octaves. At the central scale, the distance between the camera and the target was 28 cm. In this study, we used nine different types, with 81 images for each type. The textures are sandpaper, crumpled aluminium foil, styrofoam, sponge, corduroy, linen, cotton, brown bread, orange peel and cracker. OUTEX database contains 320 surface textures that include both macrotextures and microtextures. Many textures have variations in local colour content, which results in challenging local greyscale variations in intensity images. Some of the source textures have a large tactile dimension, which can induce considerable local greyscale distortions. The textures in this database include barley, rice, canvas, carpet, chips, crushed stone, flakes, flour, foam and granite. UIUC database contains 25 texture classes, each class contains 40 samples. All images were in greyscale JPG format at 640 × 480 pixels. Figures 2–4 show samples from each database.

Figure 2 Textures from KTH-TIPS database

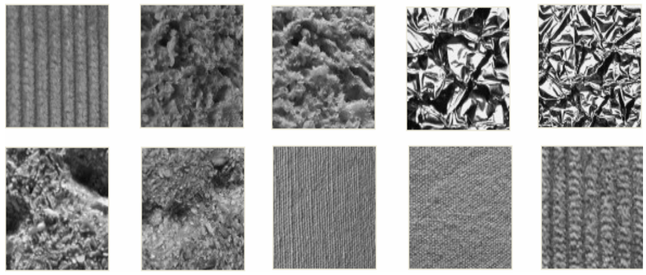
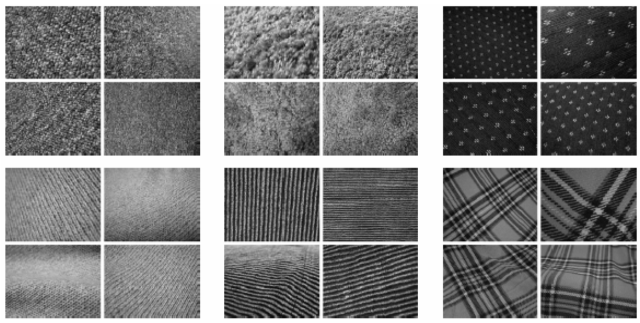


Figure 3 Textures from OUTEX database (see online version for colours)

Figure 4 Textures from UIUC database



Tables 1–3 show the classification rate using 10%, 30% and 70% of the data set for training and the reminder for testing. The proposed method is to compare six methods: Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Batch Linear Discriminant Analysis (Batch-ilda),

Independent Component Analysis (ICA), Gabor-SVM and LBP-SVM. In all experiments, Matlab 10.0 was used to implement the seven methods. The number of features that were extracted using grey-level cooccurrence matrix is 28, while the number of Gabor-SVM features was 48, and the number of LBP-SVM features was 54. Therefore, the total number of features was 130. The adopted diagonal matrix was $\{a_{11} = a, \ldots . a_{28,28} = a,\ a_{29,29} = b, \ldots, a_{76,76} = b,\ a_{77,77} = c, \ldots, a_{130,130} = c\}$, thus the parameters that must be optimised were a, b and c. In our experiments, the optimal values were $a = 0.3$, $b = 0.8$ and $b = 0.6$, where 10% of the data was selected randomly and grid search was used to find the optimal values.

It can be noted that the proposed kernel enhanced the classification rates for any testing set and for all databases. For example, if 50% of the data sets was used for training, then the best classification rates were 72.0, 81.1 and 79.3 which can be achieved using the suggested method.

Table 1 The classification rate using 10% for training

Method	*KTH-TIPS*	*OUTEX*	*UIUC*
LDA	55.1	73.1	71.5
PCA	47.7	68.0	67.2
BILDA	56.3	71.4	69.3
ICA	46.4	59.4	63.2
Gabor SVM	53.4	71.7	70.7
LBP SVM	55.0	69.3	70.8
Proposed	55.4	71.5	73.5

Table 2 The classification rate using 30% for training

Method	*KTH-TIPS*	*OUTEX*	*UIUC*
LDA	66.7	82.0	80.2
PCA	62.0	79.0	78.1
BILDA	71.0	83.1	79.1
ICA	61.7	70.8	75.2
Gabor SVM	68.1	80.5	78.6
LBP SVM	70.3	80.1	77.0
Proposed	72.0	81.1	79.3

Table 3 The classification rate using 70% for training

Method	*KTH-TIPS*	*OUTEX*	*UIUC*
LDA	81.6	92.1	90.1
PCA	77.5	89.0	88.0
BILDA	86.5	93.7	89.9
ICA	76.1	80.3	85.1
Gabor SVM	84.5	92.2	88.4
LBP SVM	85.6	89.0	87.2
Proposed	87.0	91.7	89.0

8 Conclusions

The performance of the proposed method was demonstrated on three databases that contained various images per texture, different texture orientations, and configurations. The results are quite promising. The proposed approach outperforms other methods and is suitable for many real-world scenarios. It can be concluded that combining several descriptors and weighted features boost texture recognition performance. An important direction for future work includes using the diversity of linear and non-linear kernels.

References

Ahonen, T., Hadid, A. and Pietikainen, M. (2006) 'Face description with local binary patterns: application to face recognition', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 12, pp.2037–2041.

Fan, J.S. (2012) 'Multi-region particle swarm optimisation algorithm', *International Journal of Computer Applications in Technology*, Vol. 44, No. 2, pp.117–123.

Fritz, M., Hayman, E., Caputo, B. and Eklundh. J-O. (2011) *The KTH-TIPS Database*. Available online at: www.nada.kth.se/cvap/databases/ (accessed on 10 November 2011).

Kobayash, T., Higuchi, T., Miyajima, T. and Otsu, N. (2009) 'recognition of dynamic texture patterns using CHLAC features', *Symposium on Bio-inspired Learning and Intelligent Systems for Security*, pp.58–60.

Lazebnik, S., Schmid, C. and Ponce, J. (2005) 'A sparse texture representation using local affine regions', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 8, pp.1265–1278.

Liu, N., Lai, J. and Qiu, H. (2011) 'Robust face recognition by sparse local features from a single image under occlusion', *6th International Conference on Image and Graphics*, pp.500–505.

Meyers, E. and Wolf, L. (2008) 'Using biologically inspired features for face processing', *International Journal in Computer Vision*, Vol. 76, No. 1, pp.93–104.

Nasirzadeh, M., Khazael A. and Bin Khalid, M. (2010) 'Woods recognition system based on local binary pattern', *2nd International Conference on Computational Intelligence, Communication Systems and Networks*, pp.308–313.

Ojala, T., Mäenpää, T., Pietikäinen, M., Viertola, J. Kyllönen, J. and Huovinen, S. (2002) 'Outex – new framework for empirical evaluation of texture analysis algorithms', *Proceedings of 16th International Conference on Pattern Recognition*, Quebec, Canada, pp.701–706.

Prabhu, U. and Savvides, M. (2011) 'Unconstrained pose invariant face recognition using 3D generic elastic models', *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 33, No. 10, pp.1952–1961.

Prasetiyo, M., Khalid, R. and Meriaudeau, F. (2010) 'A comparative study of feature extraction methods for wood texture classification', *6th International Conference on Signal-Image Technology and Internet Based Systems*, pp.23–29.

Shi, H. (2012) 'Particle swarm optimisation-based support vector machine for intelligent fault diagnosis', *International Journal of Computer Applications in Technology*, Vol. 44, No. 2, pp.159–164.

Tenllado, C., Gomez, J.I. and Setoain, J. (2008) 'Improving face recognition by combination of natural and Gabor faces', *Pattern Recognition Letters*, Vol. 31, No. 11, pp.1453–1460.

Vácha, P. and Haindl, M. (2010) 'Natural material recognition with illumination invariant textural features', *International Conference on Pattern Recognition*, pp.858–861.

Varma, M. and Zisserman, A. (2009) 'A statistical approach to material classification using image patches', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, No. 11, pp.2032–2047.

Wolf, L., Tal, H. and Yaniv, T. (2011) 'Effective unconstrained face recognition by combining multiple descriptors and learned background statistics', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 33, No. 10, pp.1978–1990.

Xie, X. and Mirmehdi, M. (2008) 'A galaxy of texture features', *Handbook of Texture Analysis*, Imperial College Press, pp.375–406.

Yong, H. Chun-Xia, Z. and Hong-Nan, W. (2008) 'Directional analysis of texture images using gray level co-occurrence matrix', *IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application*, pp.277–281.

Zhang, H., Berg, A., Maire, M. and Malik, J. (2006) 'SVM-KNN: discriminative nearest neighbor classification for visual category recognition', *IEEE Conference Computer Vision and Pattern Recognition*, pp.2126–2136.

Zhao, L. and Wang, L. (2012) 'A multiple camera coordination method based on genetic programming and material character', *International Journal of Computer Applications in Technology*, Vol. 43, No. 4, pp.351–358.