Signature Texture Features Extraction Using GLCM Approach in Android Studio

HUSSEIN AWAD DWAICH, HUDA ABDULAALI ABDULBAQI
1St. in Department of Computer Science, College of Science, Mustansiriyah University/ Baghdad.
2Dr. in Department of Computer Science, College of Science, Mustansiriyah University/ Baghdad

husseinaljabiri2@gmail.com
huda.it@uomustansiriyah.edu.iq

Abstract. Signature Features Extraction is a method for deriving informative values of image signature for indexing and signature identification. It is a dimensionality reduction of image data to be manageable for processing. Image texture is a key spatial attribute used for feature extraction and image coding. However, Android environment lack of efficient algorithm for automatic extracting texture features which may cause serious security issues and unreliability problem in the Android application (app). The challenge in Signature Features Extraction of a mobile app is to be as robust and stable as possible. For that, the present paper presents an efficient algorithm for extracting signature texture features using a gray level co-occurrence matrix (GLCM). The image signature is quantized into five texture features of energy feature, entropy feature, contrast feature, dissimilarity feature, and homogeneity feature. The processes developed in the android studio environment to be used in mobile phone cell applications as a tool for signature detection. The results show that the present method provides significant results due to obtained average values and acceptable computational time.

Keywords. GLCM, Signature Texture Features, Android process, Texture, Pattern recognition, Features, Frames.

1. Introduction
The biometrics area for security application is became an interested field in the past few years. It is used in various application such as banks, education, airport security and public office. Biometrics technologies are automated methods for identification based on behavioral and biological characteristics of individuals. It cannot forge - meaning accurate identification and authentication of the specific individual [1]. Biometric systems are increasingly applied in security and specific individual recognition. They represent the most important aspects of designing a biometric system. Biometric systems are valid on different platforms such as mobile platforms for the use of biometrics [2]. Authentication of the user based on biometrics on mobile devices has a lot of concern from the
researchers recently due to its importance and effectiveness. They developed and investigated many useful techniques, modalities and features that can improve software performance.

Developing new sensors and exploring new features on smartphones has led to create new technologies to enhance user authentication. This enhancement involves the infusion of several features and modalities for more effective authentication system. IoT devices such as smartphones, robots, and autonomous vehicles enable continuous authentication using biometric specifications [3]. A smartphone is one of the IoT application is provided with many system devices such as BlackBerry, iOS Android. Among these operating systems, Android became the fastest growing one. In the mobile app, the inaccurate extracted features may lead to misclassifying the predefined categories [4]. For that, the continuous development in biometrics technologies of human specifications characteristic, biometrics, are commonly used for personal identification [5]. It covers two main streams: behavioral biometrics and physiological biometrics. The behavioral biometrics involves typing rhythm, voice and face recognition, finger movements on touch surface, gait recognition, mouse dynamics and device vibration during strolling. The behavioral biometrics generally do not required special hardware on mobile devices. Its built in in devices as a features and sensors. The physiological biometrics developed based on personality checking such as DNA, iris and unique mark. These features required exceptional equipment to analyze the physiological issues using related equipment. For instance, equipment breakdown and the special devices for tuning and upkeep [6].

2.  Gray level co-occurrence matrix (GLCM)

Gray level co-occurrence matrix (GLCM) defined as one of the best methods for extraction texture features. This method proposed by Haralick in 1973 with cooperation of team of researchers [8]. GLCM structured by two dimensionless histogram using gray levels for pixels pair and separated by spatial relationship. In general, the texture discrimination problem is to analyze a set of co-occurrence matrices. It consists a statistical approach which involves an index of raw and columns. The indexed data represent the given rang of the gray level image content and the value of \( P(i,j) \) in certain position represents the frequency of gray levels i and j occur with a specific distance and direction [9]. GLCM is computed using vector \( d \) which is displaced by the radius \( \delta \) and orientation \( \theta \) [10]. Haralick used GLCM to extract thirteen texture features based on image representation [11, 12]. The main useful features selected in the present work discussed in below:

2.1.  contrast

Contrast is a measure of gray-level variations or intensity. It measure the differences between the pixel point and neighbor points. In real applications, the contrast determined based on the differences in color brightness. Also, it concerns about the object brightness and other objects in the same field of view. The main formula used in contrast calculations is [13,14]:

\[
F_1 = \sum_{n=0}^{N-1} \left( \sum_{j=0}^{N-1} p_{d,g}(i,j)^2 \right) \quad \text{………(1)}
\]

\( p_{d} \): gray level co-occurrence matrix.
\( d \): displacement vector.
\( \theta \): the direction.

\((i,j)\) of \( p_{d} \) is the number of occurrences used to compare a pair of gray levels i and j which represent the distance from point I to point j based on the direction \( \theta \) and distance d. and \( n = |i - j| \)

2.2.  energy

Energy is defined as a measurement of uniformity of the image. The main concept of this measurement is to find the differences between the similar points. When the points are similar, the formula of angular second moment (ASM) will be used to investigate the differences. Also, this method called uniformity or energy due to maximizing the pixel value magnitude [15,16]:

\[
F_2 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} p_{d,g}(i,j)^2 \quad \text{………(2)}
\]

\[
\mu = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} i \cdot p_{d,g}(i,j) \quad \text{………(3)}
\]
\[ \sigma_x = \sqrt{\mu = \Sigma_{i=0}^{N-1} \Sigma_{j=0}^{N-1} (i - \mu)^2 \cdot p_{d,\theta}(i,j)} \]

\[ \sigma_y = \sqrt{\mu = \Sigma_{i=0}^{N-1} \Sigma_{j=0}^{N-1} (j - \mu)^2 \cdot p_{d,\theta}(i,j)} \] \hspace{1cm} \text{(4)}

2.3. Homogeneity

Homogeneity (or the inverse difference moment IDM) is a process of measuring distance from one cluster to another cluster in the domain of color features. In this case, the color space should be less than the threshold of features. IDM features provide the closeness measurements of the GLCM distribution of elements to the GLCM diagonal [17, 18]. The general formula is:

\[ F_3 = \Sigma_{i=0}^{N-1} \Sigma_{j=0}^{N-1} \frac{1}{1+(i-j)^2} \cdot p_{d,\theta}(i,j) \] \hspace{1cm} \text{(5)}

2.4. Entropy

The degree of turbulence in the image is expressed, and the value of randomness is large when all elements in the GLCM are similar and small in the opposite case and are expressed by the following relationship [19].

\[ F_4 = -\Sigma_{i=0}^{N-1} \Sigma_{j=0}^{N-1} p_{d,\theta}(i,j) \log_2(p_{d}(i,j)) \] \hspace{1cm} \text{(6)}

Dissimilarity:

Measures the differences in the distribution of elements in the GLCM to the GLCM diagonal [20].

\[ F_5 = \Sigma_{i=0}^{N-1} \Sigma_{j=0}^{N-1} |i - j| \times p_{d,\theta}(i,j) \] \hspace{1cm} \text{(7)}

The researchers found that the \( \delta \) values range from 1 to 10. This conclusion is justified as a pixel is more likely to be correlated to other similar located pixels than the one which is located far away. Also, displacement value equal to the size of the texture element improves classification [21]. In this method, every pixel has eight neighboring pixels. This arrangement allowing to have eight choices for \( \theta \), they are 0°, 45°, 90°, 135°, 180°, 270°, 315° [22].

3. Methodology

The used methodology in the present study addresses the difficulties involved in IoT mobile application. The application characterized by the specification of the android environment. The scenarios are developed in two main stages. The goal of the first stage is to increase the features extraction reliability to improve the application processes in the next steps. In the second stage, the gray level co-occurrence matrix (GLCM) applied to find the texture features. The flow chart of the system methodology is shown in Figure 1.
To implement the algorithm, the following steps are used:

Step 1: obtain co-occurrence matrix of each image

In this step, the calculation of independent co-occurrence matrix will be built and the result will be a matrix of 16x16 integer element for each image.

Step 2: in this step, the values of each chosen descriptor will be obtained. The individual co-occurrence matrix will be calculated and for each descriptor, the results values will be stored as a matrix.

Step 3: in this step, the image signature will be generated and the values of descriptor matrix will be organized with different categories.

Step 4: the image in this step will be compared based on signature. The distance of two images will be calculated by applying the signature element in Euclidian distance function.

The details of algorithm steps will be discussed in the next sections.

3.1. Image Acquisition

Signature images are select from the studio. These images are from a Kaggle dataset, each person has five signatures, as shown in figure (2).
3.2. **pre-processing**

In the present step, the signature standard will be developed and created in order to investigate and extract the features. Its crucial step to improve the image quality and makes it suitable for finding the unique features that can be easily extracted. This preprocessing stage involves the following processes.

3.2.1. **Median filter.** Median filter is a nonlinear class filter that is used to improve image quality by removing the image noise. It is used salt and pepper noise reduction process and, in some applications, suppress speckle noise. In the present step, a 3x3 median filter is applied as shown in figure 3.

![Figure 3. Median filter process](image)

3.2.2. **Grayscale Image.** The color image consists of three bands red, green, and blue (RGB). Every one of these three has an interval from (0 to 255), and take 24 bit for all. While the grayscale image has one byte, just one interval (0 to 255). So for more flexible and high-speed performance, it is necessary to convert the color image into a grayscale image. There are many methods for converting into grayscale, here will use a simple method which applies the following equation:

\[
\text{Grayscale} = \frac{R + G + B}{3}
\]

3.2.3. **Contrast Stretching.** Contrast stretching is the process of improving the low-contrast of image. This phenomenon occurs when the image has poor illumination due to wrong lens setting or weakness in dynamic range of the sensor. The operation of this process is done by stretching the pixel intensity range to occupy the dynamic range in output image [23]. The formula of this process is:

\[
N(x, y) = \frac{o(x,y) - o_{\text{min}}}{o_{\text{max}} - o_{\text{min}}} \times 255 + 255
\]

3.2.4. **Otsu's Binarization.** Otsu's Binarization used in the present work to enhance image and minimize the overlapping of the class distribution [24]. It converts the gray image into binary image and considered as a successful global thresholding method. The main process of this method is the automatic production of shape histogram of the image. The algorithm proposes two classes of pixels
as an image content. They represent the foreground and the background which calculate the optimal threshold of the two classes. The appropriate threshold T can be found by using the formula below:

$$N(x,y) = \begin{cases} 
1, & \text{if } g(x,y) > T \\
0, & \text{if } g(x,y) \leq T 
\end{cases} \quad \text{(10)}$$

where N(x, y) represents the binary image.

$$N(x,y) = \begin{cases} 
1, & \text{if } g(x,y) > T \\
0, & \text{if } g(x,y) \leq T 
\end{cases} \quad \text{(10)}$$

3.3. Feature Extraction

The process of GLCM is normalized and each element considered the joint probability occurrence of spatial relationship based on pixel pairs. The extracted features were 16 and stored as a vector which contains the \( \theta \) values (0, 45, 90, 135) degrees for the features Energy, Entropy, Contrast, Dissimilarity, and Homogeneity respectively. The extracted features shown in figure 5.

![Preprocessing steps](image-url)
4. Results and discussion

Conducting the same experiment using the Haralick features Contrast, Correlation, Energy, Entropy and Homogeneity we obtained an average values of (contrast = 1.763, homogeneity = 0.991, entropy = 0.087, energy = 0.917 and dissimilarity = 0.117). The results obtained from a five-dimensional feature vector for querying the database. This is a clear indication of the improvement of performance using the proposed features as shown in table 1.

| Image No. | Contrast | Homogeneity | Entropy | Energy | Dissimilarity | Time |
|-----------|----------|-------------|---------|--------|---------------|------|
| 1         | 1.837    | 0.991       | 0.089   | 0.915  | 0.122         | 15s  |
| 2         | 1.724    | 0.992       | 0.082   | 0.923  | 0.114         | 16s  |
| 3         | 1.784    | 0.992       | 0.089   | 0.915  | 0.118         | 16s  |
| 4         | 1.756    | 0.992       | 0.088   | 0.916  | 0.117         | 18s  |
| 5         | 1.713    | 0.992       | 0.087   | 0.917  | 0.114         | 17s  |
| Avg       | 1.762    | 0.991       | 0.087   | 0.917  | 0.117         | 16s  |

Since contrast is the measure of intensity, it is determined by the difference in the color of the object's brightness and other objects within the same field of view. Therefore, we note that he has more values than other features because signature pictures always contain a difference in lighting, so a greater value appears to him. If the image is black, the contrast is zero.

The reason for extracting signature features successfully in Android environment using GLCM method is special relationship of texture features. The GLCM set of features are on second order statistics, for that they have the ability to reflect the degree of correlations in all pixel pairs in different aspects. This process facilitate the definition of pixel of interest and apply the specific relationship on the pair of pixels.

5. Conclusions

In this paper, an effective Android application has been proposed for image processing and extracting signature features. A gray level co-occurrence matrix (GLCM) have been applied based on texture features. In addition, a series of operations used starting with Median filter, Grayscale Image, Contrast Stretching, and Otsu. Testing on several signature images of one person provides the string values of
Energy, Entropy, Contrast, Dissimilarity, and Homogeneity features. The results observe an improvement performance of signature features detection in android environment.

Acknowledgment
The author thankful Department of Computer Science, College of Science, Mustansiriyah University/Baghdad, Iraq, for supporting this work.

References
[1] Abdullayeva, F., Imamverdiyev, Y., Musayev, V., & Wayman, J. (2008). Analysis of security vulnerabilities in biometric systems. In The second international conference: problems of cybernetics and informatics.
[2] Yang, W., Wang, S., Hu, J., Zheng, G., & Valli, C. (2019). Security and accuracy of fingerprint-based biometrics: A review. Symmetry, 11(2), 141.
[3] Liang, Y., Samtani, S., Guo, B., & Yu, Z. (2020). Behavioral Biometrics for Continuous Authentication in the Internet of Things Era: An Artificial Intelligence Perspective. IEEE Internet of Things Journal.
[4] Reyhani Hamedani, M., Shin, D., Lee, M., Cho, S. J., & Hwang, C. (2018). AndroClass: An effective method to classify Android applications by applying deep neural networks to comprehensive features. Wireless Communications and Mobile Computing, 2018.
[5] Bhunia, A. K., Alaei, A., & Roy, P. P. (2019). Signature verification approach using fusion of hybrid texture features. Neural Computing and Applications, 31(12), 8737-8748.
[6] Al-Khafaji, S. S. M. (2018). An Anomaly Detection Model for Signature Authentication on Mobile Devices (Doctoral dissertation, Middle East University).
[7] Rani, S., & Dalal, P. (2014). A review on offline signature recognition and verification techniques.
[8] Sebastian V, B., Unnikrishnan, A., & Balakrishnan, K. (2012). Gray level co-occurrence matrices: generalisation and some new features. arXiv preprint arXiv:1205.4831.
[9] Felipe, J. C., Traina, A. J., & Traina, C. (2003, June). Retrieval by content of medical images using texture for tissue identification. In 16th IEEE Symposium Computer-Based Medical Systems. Proceedings. (pp. 175-180). IEEE.
[10] Jabur, Z. F., & Ali, S. K. (2014). Off line Handwritten Signature Recognition based on Fusion of Global and GLCM Features Using Fuzzy Logic. JOURNAL OF THI-QAR SCIENCE, 6(1), 83-88.
[11] Zakaria, A. K., Khadra, Y., & Al-Abboud, E. (2019). Improving the process of recognition the treated teeth in the panoramic images based on the optimal features selection: تحسين عملية التعرف على الأسنان المعالجة في الصور البانورامية بالاعتماد على الاختيار الأمثل للأسنان. مجلة العلوم الهندسية، 3(3).
[12] Jabr, Z. F., Saleh, S. R., & Fasial, A. N. (2016). A Hybrid Features for Signature Recognition Using Neural Network. journal of thi-qar science, 6(1), 83-88
[13] Wang, Q., & Ward, R. (2006, October). Fast Image/Video Contrast Enhancement Based on WTIE. In 2006 IEEE Workshop on Multimedia Signal Processing (pp. 338-343). IEEE.
[14] Tripathi, A. K., Mukhopadhyay, S., & Dhara, A. K. (2011, November). Performance metrics for image contrast. In 2011 International Conference on Image Information Processing (pp. 1-4). IEEE.
[15] Er. Kanchan Sharma et al, (2015). GLCM and its Features.
[16] Mohanaiah, P., Sathyarayaraya, P., & GuruKumar, L. (2013). Image texture feature extraction using GLCM approach. International journal of scientific and research publications, 3(5), 1.
[17] Noel, G., & Jourlin, M. (2019). Region homogeneity in the Logarithmic Image Processing framework: application to region growing algorithms. arXiv preprint arXiv:1904.12597.
[18] Knauer, U., & Meffert, B. (2010, June). Fast computation of region homogeneity with application in a surveillance task. In Proceedings of ISPRS Commission V Mid-Term Symposium Close Range Image Measurement Techniques (pp. 337-342).
[19] Razlighi, Q. R., & Kehtarnavaz, N. (2009, January). A comparison study of image spatial entropy. In Visual Communications and Image Processing 2009 (Vol. 7257, p. 72571X). International Society for Optics and Photonics.

[20] Amelio, A. (2016). Approximate matching in ACSM dissimilarity measure. Procedia Computer Science, 96, 1479-1488.

[21] Pham, T. A. (2010). Optimization of texture feature extraction algorithm.

[22] Gadkari, D. (2004). Image quality analysis using GLCM.

[23] Jriash¹, H. J., & Abdullah, N. A. (2015). OFFLINE HANDWRITTEN SIGNATURE VERIFICATION SYSTEM USING NEURAL NETWORK.

[24] Yousefi, J. (2011). Image binarization using Otsu thresholding algorithm. University of Guelph, Ontario, Canada.