Analysis of Widely-used Descriptors for Finger-vein Recognition

Fariba Yousefi [1], Erdal Sivri[1], Ozgur Kaya[1,2], Selma Süloğlu[1,2] and Sinan Kalkan[1]
[1]*Middle East Technical University, Ankara, Turkey*
[2]*SoSoft Ltd. Şti., Ankara, Turkey*
fariba.yousefi@gmail.com,{erdal, skalkan}@ceng.metu.edu.tr, {okaya, selma}@ceng.metu.edu.tr

Keywords: Finger-vein Recognition, Histogram of Oriented Gradients, Fourier Descriptors, Zernike Moments, Global Binary Patterns.

Abstract: For finger-vein recognition, many successful methods, such as Line Tracking (LT), Maximum Curvature (MC) and Wide Line Detector (WL), have been proposed. Among these, LT has a very slow matching and feature-extraction phase, and LT, MC and WL are translation and rotation dependent. Moreover, we show in the paper, they are affected by noise. To overcome these drawbacks, we propose using popular feature descriptors widely used for several Computer Vision or Pattern Recognition (CVPR) problems in the literature. The CVPR descriptors we test include Histogram of Oriented Gradients (HOG), Fourier Descriptors (FD), Zernike Moments (ZM), Local Binary Patterns (LBP) and Global Binary Patterns (GBP), which have not been applied to the finger-vein recognition problem before. We compare these descriptors against LT, MC, and WL and evaluate their running times, performance and resilience against noise, rotation and translation. We report that the LT and WL methods accuracy are comparable to each other and WL gives the best accuracy, LT method's speed is the slowest. Our results indicate that WL can be used together with ZM and GBP in case of rotation and noise, respectively.

1 INTRODUCTION

Biometric human identification has become very important due to pitfalls even in extreme security systems, and alternative methods that can be used in place of or together with finger-print have been being sought during the last decade. It has been shown that finger-vein pattern is distinctive enough for human biometric identification (Yanagawa et al., 2007), and with the advent of technology, and especially since finger-vein is more difficult to manipulate compared to, e.g., fingerprints, we have seen many successful applications.

We can divide existing studies regarding finger-vein feature extraction into three main categories (without aiming to be comprehensive): (i) Those that use filtering or transformation methods: Such methods mainly use Gabor filters to extract filter responses using the following Gabor function at different scales and orientations (Yang et al., 2011; Yang et al., 2009), steerable filters (Yang and Li, 2010) or wavelets (Xueyan et al., 2007). After filtering, from the filter responses, using histograms, minimum-maximum or average values, features are extracted, and these features are usually compared using Euclidean or Cosine distance (Yang et al., 2011). Moreover, Fourier Transform (Mahri et al., 2010) can be applied to the finger vein images and the phase components of the images can be taken as feature vectors. (ii) Those that track finger veins, segment the pixels corresponding to the veins, and represent or directly compare these pixels (Miura et al., 2004; Qin et al., 2011). (iii) Those that use descriptive feature extraction methods such as local binary patterns (Lee et al., 2011; Yang et al., 2012) or local derivative patterns (Lee et al., 2011).

Among these methods, Line Tracking (LT - (Miura et al., 2004)), Maximum Curvature (MC - (Miura et al., 2007)) and Wide Line Detector (WL - (Huang et al., 2010)) are widely used and shown to perform very well. However, LT has a very slow matching and feature-extraction phase. Moreover, LT, MC and WL are rotation dependent, and they are affected by image noise. To overcome these drawbacks, we propose using some popular feature descriptors widely used for several Computer Vision or Pattern Recognition (CVPR). These descriptors include Histogram of Oriented Gradients (HOG) (Dalal and Triggs, 2005), Fourier Descriptors (FD) (Gonzalez and Woods, 2001), Zernike Moments (ZM) (Teh and Chin, 1988), Local Binary Patterns (LBP) (Ojala

et al., 1994) and Global Binary Patterns (GBP), (Sivri, 2013). Among these, HOG, FD, ZM and GBP have not been applied to finger-vein recognition before. We compare these descriptors against line tracking (LT), Maximum Curvature (MC), and Wide Line Detector (WL), which are most widely used methods for finger-vein recognition. The novelty of the paper is in (i) applying new feature extraction methods that have not been used for finger-vein recognition before and (ii) evaluating the performance of all these methods under translation, rotation and noise.

We focus on the "feature extraction" step, and the preprocessing step is kept as simple as possible. As for "matching", the matching method specific to LT, MC and WL which is called mismatch ratio is used and for all other descriptors, three different distance metrics called Euclidean distance, Chi-Square distance and Earth Mover's distance have been used and compared to each other. For performance evaluation, we use the SDUMLA-HMT finger-vein database that is publicly available[1].

2 Feature Extraction and Matching

In this section, we briefly describe the methods that we analyze in the article.

2.1 Line Tracking (LT)

LT-based method of Miura et al. (Miura et al., 2004) is one of the leading approaches used in finger-vein extraction. The method exploits the fact that, due to the light-absorbing nature of finger veins, finger veins appear darker in the image, making veins look like 'valley's in the infrared image. The LT method is based on randomly finding a pixel in a valley, and tracking the pixels along a valley as long as possible. To determine whether a pixel is on a valley (i.e., a part of the finger vein), the LT method checks whether the cross-section orthogonal s.p.t. to the center pixel form a valley in intensity values. The method tracks "valley pixels" and restarts randomly in another position in the image for tracking finger-veins. The output is a locus table that list how many times a pixel has been tracked, and this captures the finger-vein in the infrared image.

2.2 Maximum Curvature (MC)

Miura et al. (Miura et al., 2007) proposed a method that is based on calculating curvatures in cross-sectional profiles of a vein image. In each profile, the location of the maximum curvatures are found, and those maxima and their width are taken as the center and the width of the veins respectively.

2.3 Wide Line Detector (WL)

WL (Huang et al., 2010) uses a circular sliding region to detect pixels belonging to a finger vein. At each location, the intensity distribution of the pixels in the neighborhood is evaluated with respect to the intensity of the center. If a small proportion of the pixels have different intensities, then the center of the window is taken as belonging to a finger vein.

2.4 Local Binary Patterns (LBP)

LBP (Ojala et al., 1994) is one of the widely used methods for feature extraction from finger-vein images. In LBP, a window is placed on each pixel in the image, and within each window, the intensity of the center pixel is compared against the intensities of the neighboring pixels. During this comparison, bigger intensity values are taken as 1 and smaller values as 0. These numbers are interpreted as binary numbers and histogram of their corresponding decimal values are used as descriptors in LBP.

2.5 Fourier Descriptors (FD)

FD are constructed by applying a Fourier transform on a shape signature, which is a one dimensional function derived from the shape boundary (Zhang and Lu, 2002). Any shape signature can be used to obtain Fourier descriptors, such as complex coordinates, centroid distance and curvature function.

The first step to compute Fourier descriptors from the image is to extract points representing the finger-vein: $(x_0, y_0), (x_1, y_1), ..., (x_{N-1}, y_{N-1})$. For these points, we use the locus table that the LT method returns. Then, pixel coordinates (x_k, y_k) are converted to complex coordinates as $z_i = x_k + jy_k$, where $j^2 = -1$. Complex numbers, $z_0, z_1, ..., z_{N-1}$, in spatial domain are converted to frequency domain using the Discrete Fourier Transform (Proakis and Manolakis, 2006). After complex coordinates are transformed to the frequency domain, the coefficients are normalized by the first term, and the normalized first K (empirically determined as 100) coefficients are taken to be the descriptor.

2.6 Zernike Moments (ZM)

ZM (Teh and Chin, 1988) is one of the most popular geometric moments in the literature and defined as

[1] http://mla.sdu.edu.cn/sdumla-hmt.html

follows (Teague, 1980):

$$A_{nm} = [(n+1)/\pi] \int\int f(x,y)[V_{nm}(p,\theta)]\,dxdy, \quad (1)$$

where $f(\cdot,\cdot)$ is the function, i.e., image, n defines the order and V_{nm} is known as the Zernike polynomials: $V_{nm}(x,y) = V_{nm}(\rho sin\theta, \rho cos\theta) = R_{nm}(\rho)\exp^{il\theta}$. Rotation invariance is achieved by using the magnitudes of Zernike moments because rotation does not affect the magnitude (Khotanzad and Hong, 1990). Translation invariance is achived by shifting the origin of the image to the centroid of the finger vein image. In this study, we set $n = 40$ (determined empirically), $m = n$ and $n - |m|$ being even and extracted A_{nm} on the median thresholded locus table produced by the Line Tracking algorithm, thresholded Maximum Curvature map and Wide Line Detector map.

2.7 Histogram of Oriented Gradients (HOG)

HOG constructs a histogram of gradient occurrences in localized grid cells (Dalal and Triggs, 2005). HOG has been demonstrated to be very successful in human detection or tracking, and in this study, the performance of HOG on finger-vein recognition is analyzed and compared with other popular descriptors. HOG in an image patch P is defined as follows:

$$HOG(k) = \sum_{p \in P} \delta\left(\left\lfloor \frac{\theta^p}{L} \right\rfloor \right), \quad (2)$$

where $\delta(\cdot)$ is the Kronecker delta function, L is a normalizing constant and θ^p is the orientation at point p, which is equal to the image gradient at that point. $HOG(k)$ corresponds to the value of the kth bin in a K-bin histogram. The value of K used in the experiments is set to 9, and the value of the normalizing constant, L, is equal to $180/K = 20$ (Dalal and Triggs, 2005).

2.8 Global Binary Patterns (GBP)

GBP (Sivri, 2013) is a simple, yet efficient, shape descriptor similar to LBP. The method creates a set of bit strings for any direction of a thresholded binary image and interprets these bit strings as binary numbers to build a global descriptor (see Figure 1). In its simplest form, GBP of a row, r, of a binary image I is defined as follows:

$$GBP_h(r) = \sum_{j=1}^{R} I(r,j) \cdot 2^{j-1}, \quad (3)$$

where R is the number of columns in image I. GBP_h computes GBP along horizontal direction. Similarly,

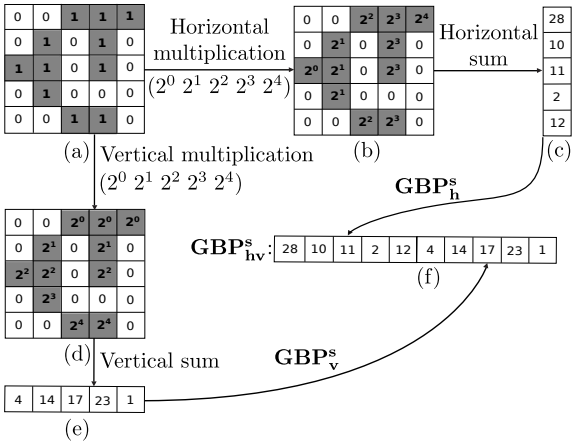


Figure 1: GBP computation. (a) The binary image. (b) After rows are multiplied by powers of two. (c) After each row is summed horizontally. (d) After columns are multiplied by powers of two. (e) After columns are summed vertically. (f) Resulting GBP descriptor. (Source: (Sivri, 2013))

GBP along vertical direction, denoted GBP_v, can be defined as follows:

$$GBP_v(c) = \sum_{i=1}^{C} I(i,c) \cdot 2^{i-1}, \quad (4)$$

where C is the number of rows in image I. See Figure 1 for an illustration of GBP computation along horizontal and vertical directions, denoted GBP_{hv}.

GBP_h and GBP_v are defined along horizontal and vertical directions. In fact, GBP can be constructed along any arbitrary direction, which may effect the performance of the descriptor (Sivri, 2013).

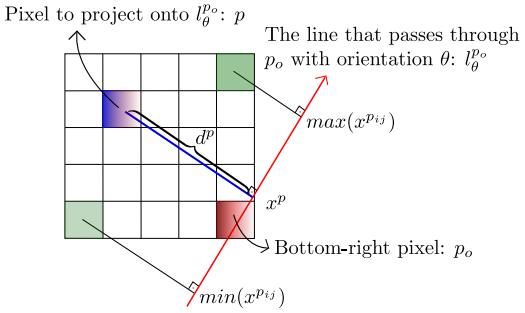


Figure 2: Illustration of the projection of a pixel onto a line that passes through the pixel p_o with orientation θ. (Source: (Sivri, 2013))

It is possible to use any number of projections to form the GBP descriptor. Analysis in this study is performed using horizontal and vertical projections denoted GBP_h and GBP_v respectively.

2.9 Matching

Different matching methods have been used for matching the features extracted from each method. For LT, MC and WL, the literature uses the so-called *mismatch ratio* in which a query map is compared against a stored map pixel by pixel (like in template matching). For FD, ZM, HOG, LBP and GBP, the literature uses many distance metrics, like Euclidean distance, Chi-Square distance and Earth Mover's distance. In this article, we use the mismatch ratio for LT, MC and WL, and Euclidean distance, χ^2 distance and Earth Mover's distance for FD, ZM, HOG, LBP and GBP.

3 Results

For performance evaluation, we use the SDUMLA-HMT finger-vein database that is publicly available[2]. SDUMLA-HMT database contains in total 3,816 images of index fingers, middle fingers and ring fingers of both hands. Moreover, for each finger of an individual, 6 different snapshots are included, making it ideal for extensive analysis. The methods are compared based on the False-Rejection Rate (FRR) and False-Acceptance Rate (FAR), as widely used by the finger-vein recognition literature (e.g., (Miura et al., 2004)). FAR is the rate of accepting imposters by mistake and calculated as follows:

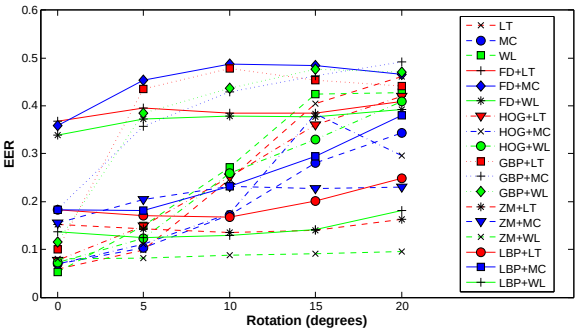
$$FAR = \frac{FA}{FA+CR}, \tag{5}$$

where FA denotes the number of false acceptances and CR the correct rejections. FRR is the rate of rejecting authorized users by mistake and calculated as follows:

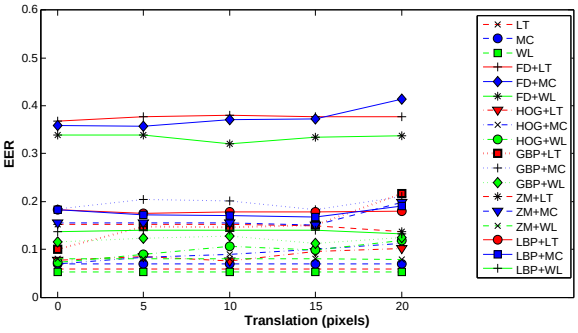
$$FRR = \frac{FR}{FR+CA}, \tag{6}$$

where FR denotes the number of false rejections and CA the correct acceptances. Another important metric for evaluating the methods is the EER rate, which is the point where FAR equals FRR. As we discussed before, our aim is to improve the running times and the resilience of LT, MC and WL using HOG, ZM, FD, LBP and GBP. We apply HOG, ZM, FD, LBP and GBP to the binarized locus table (from LT), the binarized Maximum Curvature map and the Wide Line Detector map. For the sake of simplicity and space, in the remainder of the article, we will use X+Y, where $X \in \{\text{FD, ZM, LBP, HOG, GBP}\}, Y \in \{\text{LT, MC, WL}\}$, to refer to the application of method X to the output of method Y.

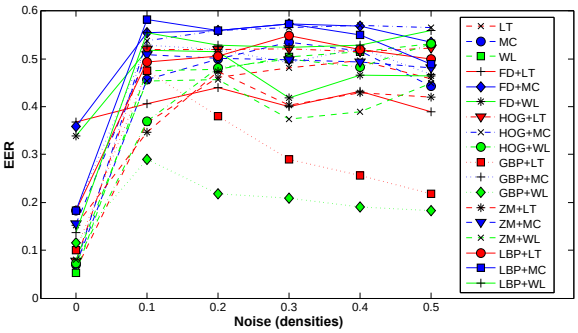
[2]http://mla.sdu.edu.cn/sdumla-hmt.html



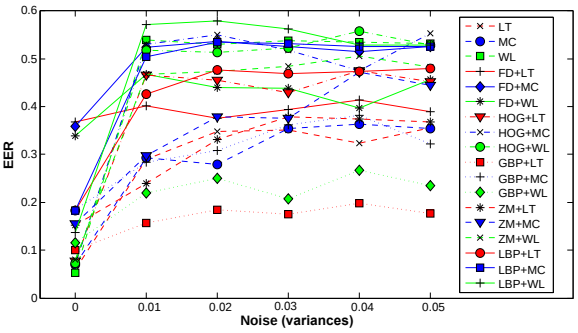
(a)



(b)



(c)



(d)

Figure 3: Change of EER values under rotation, translation and noise using χ^2 distance. (a) Rotation. (b) Translation. (c) Salt and Pepper Noise. (d) Gaussian Noise.

3.1 Rotation, Translation and Noise Resilience

We tested the methods on their resilience against translation, rotation and noise. This is crucial since in finger-vein image capturing systems, fingers tend to be not aligned and positioned the same each time they are scanned. We arbitrarily selected 20 images which are translated horizontally for different shift amounts ($0-20$ pixels) and rotated clockwise for different angles (0^o-20^o).

Figures 3(a), 4(a) and 5(a) show the changes in EER values for different horizontal translations. We see that HOG-based features are not affected by translation. The reason is that histograms that HOG extract from each block are not affected by shifts in positions. Moreover in terms of translation, all methods seem to be good enough to be translation invariant except GBP.

As for the rotation performance of the methods, which is shown in Figures 3(b), 4(b) and 5(b), we see that all the methods except for ZM-variants are severely affected by rotation. ZM is not affected by rotation due to its polar formulation.

Two different noises are added to the images called salt and pepper (SP) noise and Gaussian noise. The variables whose effect were investigated are the likelihood of changing a pixel's intensity in the case of SP noise and the variance in the case of Gaussian noise. In Gaussian noise, the mean was kept zero. The effect of SP noise can be seen in Figures 3(c), 5(c) and 4(c), and the effect of Gaussian noise in Figures 3(d), 5(d) and 4(d). From the figures, we see that GBP and its variants are affected less from noise while the others suffer significantly from the added noise.

3.2 Overall Performance Comparison

Figure 6 displays the FAR-FRR graph for best variant of each method (e.g., the best FD among FD+LT, FD+MC and FD+WL). We see that LT, MC and WL provide the best performance. However, HOG+WL provides better FRR rate. This is striking because HOG is usually used in human detection and tracking applications, and simple histograming of gradients in grids may provide better results than directly comparing images using mismatch ratios. ZM+WL result is comparable with HOG+WL since ZM is invariant to translation and rotation. GBP+WL also has comparable result with HOG+WL and ZM+WL. Surprisingly, FD, which has been successfully applied to similar problems such as hand-written digit recognition, yields the worst results as far as the FAR-FRR values are concerned.

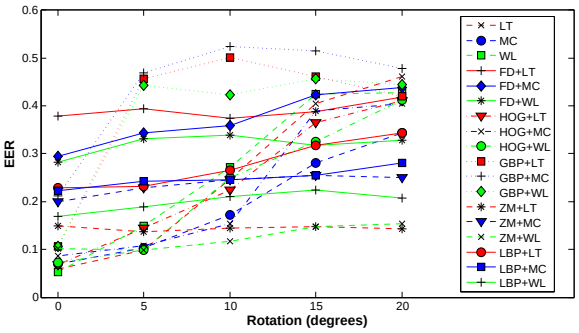
As we discussed in Section 2.9, original matching method of LT, MC and WL as well as Euclidean distance, Ch-Square distance and Earth Mover's distance are evaluated. The best results of each distance metric in Figure 6 are placed in Figure 7. As seen in Figure 7, EMD and χ^2 perform better than the Euclidean distance.

3.3 Running Time Comparison

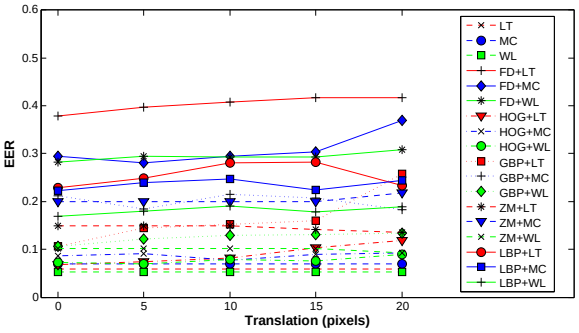
We compared the methods based on their running times as well. Each method is run on the same machine using MATLAB. We compared 100 images against a total of 3804 images in the database, and calculated the average times for both feature extraction and feature matching over 3804 images. As shown in Table 1, we see that FD, HOG and GBP are the fastest among the methods since their computational complexities are quite low; whereas LT, which provides best results when there is no translation, rotation or noise, is the slowest. The reason for LT's slowness is due to the fact that the feature extraction step requires tracing veins for several iterations. WL, the other method providing best results when there is no translation, rotation or noise, is among the best in terms of speed.

Table 1: Running-time comparison of all methods.

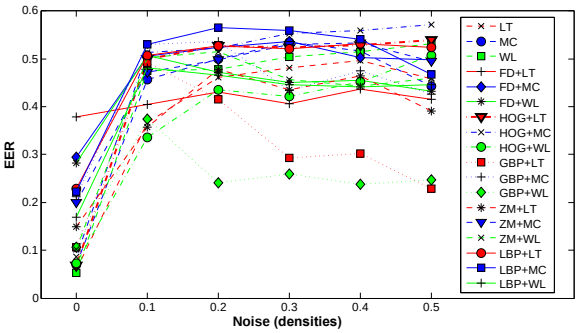
Method	**Time (seconds)**	**Speed Rank**
LT (feat. extraction only)	9.2831	15
MC (feat. extraction only)	0.0585	1
WL (feat. extraction only)	0.1286	8
LT	9.4675	19
MC	0.0595	3
WL	0.1335	10
FD + LT + Euc	9.2837	16
FD + MC + Euc	0.0586	2
FD + WL + Euc	0.1289	9
ZM + LT + Euc	9.6507	21
ZM + MC + Euc	0.5183	7
ZM + WL + Euc	0.5920	14
GBP + LT + Euc	9.2850	17
GBP + MC + Euc	0.0601	4
GBP + WL + Euc	0.1454	12
HOG + LT + Euc	9.2925	18
HOG + MC + Euc	0.0664	5
HOG + WL + Euc	0.1368	11
LBP + LT + Euc	9.5831	20
LBP + MC + Euc	0.0890	6
LBP + WL + Euc	0.1686	13



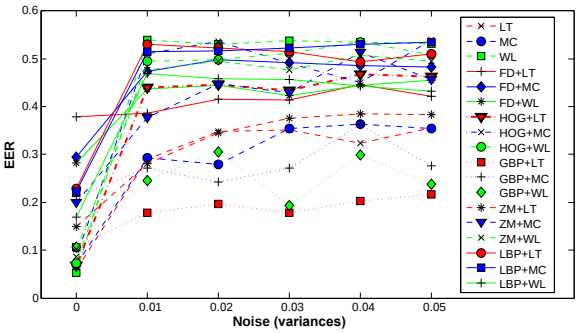
(a)



(b)

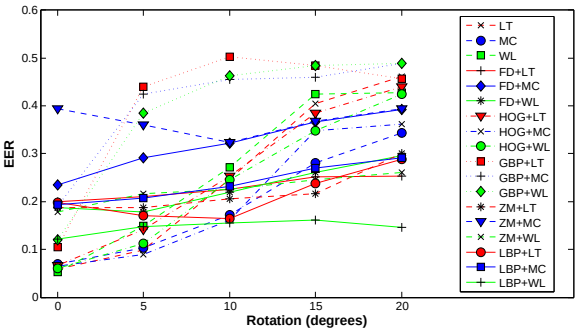


(c)

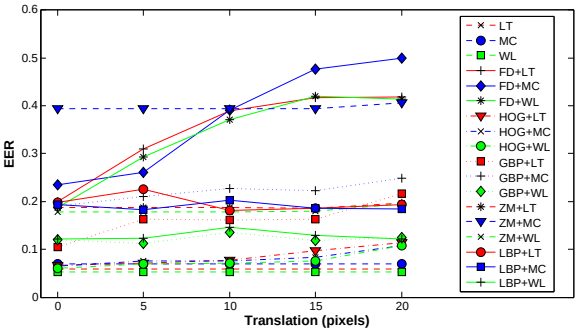


(d)

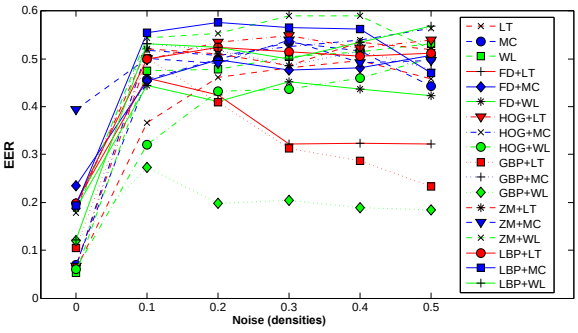
Figure 4: Change of EER values under rotation, translation and noise using Euclidean distance. (a) Rotation. (b) Translation. (c) Salt and Pepper Noise. (d) Gaussian Noise.



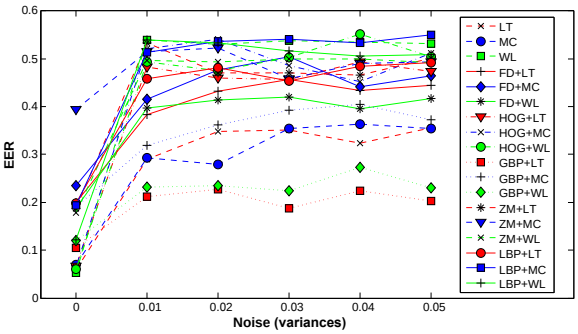
(a)



(b)

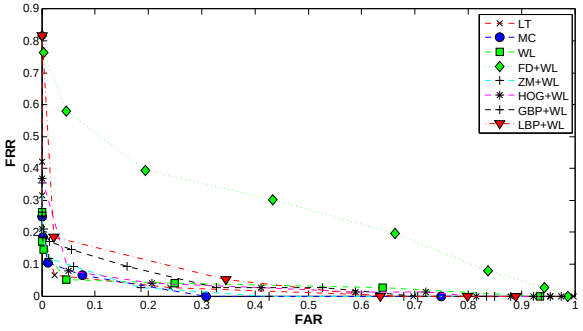


(c)

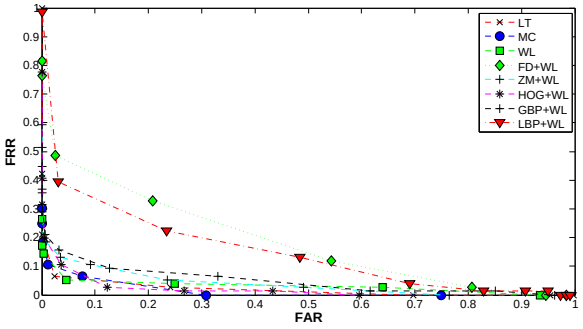


(d)

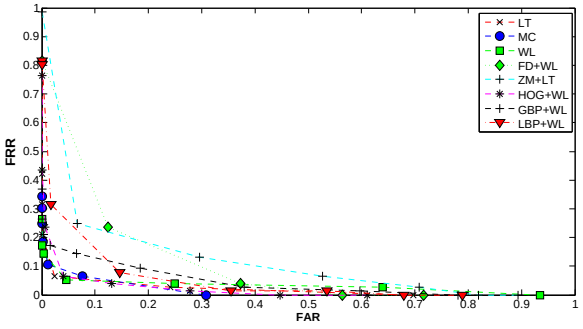
Figure 5: Change of EER values under rotation, translation and noise using Earth Mover's Distance. (a) Rotation. (b) Translation. (c) Salt and Pepper Noise. (d) Gaussian Noise.



(a)



(b)



(c)

Figure 6: FAR vs. FRR graphs for the best of each method (e.g., the best FD among FD+LT, FD+MC and FD+WL) compared with LT+mismatch, MC+mismatch and WL+mismatch (a) Using χ^2 distance. (b) Using Euclidean distance. (c) Using Earth Mover's distance.

4 Conclusion

Using a publicly available database, we have compared several feature extraction methods for finger-vein recognition. Among the feature extraction methods we have considered, Histogram of Oriented Gradients (HOG), Fourier Descriptors (FD), Zernike Moments (ZM), Local Binary Patterns (LBP) and Global Binary Patterns (GBP) that have not been applied to finger-vein recognition before. We show that Wide

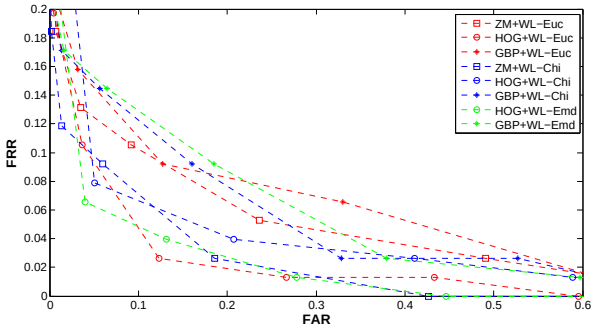


Figure 7: Best of Earth mover's, χ^2 and Euclidean distances.

Line Detector method outperforms others in terms of accuracy, however, it is affected by rotation and noise. Methods like HOG and GBP and ZM provide comparable performance where HOG and GBP have better running time and ZM is invariant to rotation and translation with slowest running time in comparison to HOG and GBP. Our results indicate that in the case of no noise, no rotation and no translation, WL is the best method. In a scenario involving rotation, ZM+WL, and in a scenario involving noise, GBP+WL can be used. If time is a critical issue, then MC can be chosen since it is the next best method in terms of EER values.

ACKNOWLEDGMENTS

We would like to express our thanks to the MLA Group of Shandong University for the SDUMLA-HMT Database. This project is partially funded by the Turkish National Science and Research Organization (TUBITAK) project no. 7110393.

REFERENCES

Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. *Int. Conf. on Computer Vision and Pattern Recognition (CVPR)*.

Gonzalez, R. C. and Woods, R. E. (2001). *Digital Image Processing*. Addison-Wesley Longman Publishing Co., Inc., 2nd edition.

Huang, B., Dai, Y., Li, R., Tang, D., and Li, W. (2010). Finger-vein authentication based on wide line detector and pattern normalization. In *20th International Conference on, Pattern Recognition (ICPR)*, pages 1269–1272. IEEE.

Khotanzad, A. and Hong, Y. H. (1990). Invariant image recognition by zernike moments. *IEEE Transac-*

tions on, Pattern Analysis and Machine Intelligence, 12(5):489–497.

Lee, E., Jung, H., and Kim, D. (2011). New finger biometric method using near infrared imaging. *Sensors*, 11(3):2319–2333.

Mahri, N., Suandi, S., and Rosdi, B. (2010). Finger vein recognition algorithm using phase only correlation. In *International Workshop on Emerging Techniques and Challenges for Hand-Based Biometrics (ETCHB)*, pages 1–6.

Miura, N., Nagasaka, A., and Miyatake, T. (2004). Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. *Machine Vision and Applications*, 15(4):194–203.

Miura, N., Nagasaka, A., and Miyatake, T. (2007). Extraction of finger-vein patterns using maximum curvature points in image profiles. *IEICE Transactions on Information and Systems*, 90(8):1185–1194.

Ojala, T., Pietikainen, M., and Harwood, D. (1994). Performance evaluation of texture measures with classification based on kullback discrimination of distributions. *12th IAPR Int. Conf. on Pattern Recognition*, 1:582–585.

Proakis, J. G. and Manolakis, D. K. (2006). *Digital Signal Processing (4th Edition)*. Prentice Hall.

Qin, H., Qin, L., and Yu, C. (2011). Region growth–based feature extraction method for finger-vein recognition. *Optical Engineering*, 50(5):057208–057208.

Sivri, E. (2013). Global binary patterns: A novel shape descriptor. *IAPR Conference on Machine Vision and Applications*.

Teague, M. R. (1980). Image analysis via the general theory of moments. *Journal of the Optical Society of America (1917-1983)*, 70:920–930.

Teh, C. H. and Chin, R. T. (1988). On image analysis by the methods of moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4):496–513.

Xueyan, L., Shuxu, G., Fengli, G., and Ye, L. (2007). Vein pattern recognitions by moment invariants. *1st Int. Conf. on Bioinformatics and Biomedical Engineering*.

Yanagawa, T., Aoki, S., and Ohyama, T. (2007). Human finger vein images are diverse and its patterns are useful for personal identification. *Kyushu University MHF Preprint Series: Kyushu, Japan*, pages 1–7.

Yang, G., Xi, X., and Yin, Y. (2012). Finger vein recognition based on a personalized best bit map. *Sensors*, 12(2):1738–1757.

Yang, J. and Li, X. (2010). Efficient finger vein localization and recognition. In *20th International Conference on Pattern Recognition (ICPR)*, pages 1148–1151. IEEE.

Yang, J., Shi, Y., and Yang, J. (2011). Personal identification based on finger-vein features. *Computers in Human Behavior*, 27(5):1565–1570.

Yang, J., Yang, J., and Shi, Y. (2009). Finger-vein segmentation based on multi-channel even-symmetric gabor filters. *IEEE Int. Conf. on Intelligent Computing and Intelligent Systems*, 4:500–503.

Zhang, D. and Lu, G. (2002). A comparative study of fourier descriptors for shape representation and retrieval. In *Proceedings of the 5th Asian Conference on Computer Vision (ACCV)*, pages 646–651.