Commentary: Contactless and Pose Invariant Biometric Identification Using Hand Surface

Student Name Anonymous
Student zID Anonymous

1. Introduction

Due to the fast advancements in IoT technologies, research efforts and applications towards using biometric traits for authentication has reached an all-time high over recent years. As a result, there already exist many established methods that utilize facial and limb features for authentication. In particular, the prescribed paper focuses on biometric authentication by extracting features from a person's hand/finger in a contact-free and constrained condition.

On such major limitation for currently existing methods is handling large pose variations in contact-free and unconstrainted conditions. Those methods usually utilize projective invariant features or homographic transformation which rely on multiple landmark point extraction. Such point extract is very inaccurate in pose variate conditions. Therefore, the author aims to address those drawbacks by proposing a measure that extracts pose-invariant features for greater accuracy and robustness. Moreover, a system is proposed basing on such measure that explores new options by using both 3D and 2D (via a 3D digitizer) features while not relying on multiple point extraction (only 1 point).

If the claim in the paper is true, it would address some of the major limitations and allow automated hand authentication systems to be more convenient to use. It will allow such systems to be used in a broader range of applications and improve the quality of life for many commercial users. Furthermore, it may also encourage more research effort towards methods that utilizes 3D geometric and pose invariant-variant features which in the long term may lead to better methods and faster advancements in the field of computer science.

2. Methods

The system proposed by the paper is segmented into 3 sections, pose-correction, feature extraction and dynamic fusion. We will comment on each section sequentially, starting with pose normalization. This section aims to localize the region of the hand and adjust its orientation among the x, y-axis

The section starts by applying Otsu's [1] method and then the distance transform to locate the center of the palm using local maximum points. ROI is then extracted from such center of both the 2&3D data. From a functional perspective, those methods have been known to be effective for binarization and foreground extraction for a long time, however, such a solution could be oversimplifying the problem. The author justifies the decision by claiming that the distance transform gives a good enough range of options but further improving ROI accuracy could greatly improve the accuracy of the detected angle. [2] proposes a method that uses the centroid method to extract the palm center with fast computation and possibly more accurate and stable performance. Moreover, applying Otsu alone will not suffice complex illumination conditions, hence [3] demonstrates a variation of the top-hat algorithm which normalizes illumination which may provide a better result for Otsu. (or a more general thresholding method can be used like GHT [4]).

Continuing with section 1, the method now adjusts the pose variation by fitting a plane in ROI via IRLS with a bi-square weighting function to eliminate outliers. The normal vector to such plane is then obtained and used to determine the angle to generate the rotation matrix. Matrix multiplication is then applied to adjust the pose of the acquired hand in 2&3D. Lastly, small holes in the rotated images are fixed via bicubic interpolation. Both the IRLS and the rotational matrix are well justified as they are proven to be mathematically stable if ROI is representative of the hand. However, its limitation is also quite clear, those irregular edges for 2D pose corrected images severely affects the geometrical accuracy. To combat such limitation, [5] has shown a promising method of orientation estimation using CNN, one may combine both methods or use a more advanced generative network for a better representation of the 2D hand.

Moving onto feature extraction, this section encodes the results from section 1 into bit feature vectors for comparison and authentication. Beginning with palmprint; for 2D hand, palmprints are encoded using a [6] competitive coding scheme. For 3D hand, palmprints are encoded using curvature via SurfaceCode [8]. Finally, both matching score is computed via the normalized Hamming distance. In terms of effectiveness, [6] uses multiple gabor filters to detect texture features of a skin surface at multiple different angles to generate an accurate representative mask; it also achieved an accuracy > 98%, beating the next best competitor [8]. For [7] , it has strong performance if the shape index does not change very rapidly, which is unlikely in this use-case. However, palmprint encoding methods using CNN [9] has shown some better result with > 99% validation score which may be a better encoding method.

Continuing with geometry features, for 2D hand, finger length, width, perimeter area and palm width are extracted from the binary image and concatenated to use as a feature vector. For 3D hand, 20 uniformly spaced cross-sectional finger segments are extracted and encoded using curvature at each point to form the feature vector [10]. In terms of effectiveness, considered 2D hand features are all relevant to geometric matching, but those quantitative features could bottleneck the accuracy as subtle differences will be ignored. For 3D hand, the normal vectors of fitted polynomials are computed for each sampled point and used as features ([10]). The accuracy of 2.6% FAR and FRR score from [10] further indicates that it is accurate and robust on small variations in hand posing and contour. Alternatively, using more complex features of hand biometrics may improve the accuracy of 2D hand matching, [11] has shown methods that use contour and section fitting to achieve a consistent result.

Lastly, we discuss the dynamic fusion section, its purpose is to use the weighted sum rule to combine the individual matching scores. The weighted sum rule is applied with individual scores of 2&3D palmprint and 3D hand geometry. The method also selectively drops the 3D hand geometric matching score when the orientation deviation angle is outside a certain threshold. This is to account for inaccurate geometric matching score due to a large deviation angle. Those high deviation angles will cause a loss of crucial information in the fingers during pose correction (PC) and skew the score. Although such a solution might do well for the quantitative result but from a theoretical perspective, the root cause of the inaccuracy is not solved but rather ignored. Perhaps more effort can be spent on the development of methods that preserve the critical information during PC. Newer Deep learning [9] or interpolation that may preserve finer details during rotation.

3. Results

For setup, 1140 right-hand images are acquired from 114 subjects(volunteers) using a 3D digitizer. For each subject, 5 different poses are collected: complete parallel hand position (1), anti-clockwise and clockwise rotation on x and y-axis (4). Note that the degree of rotation is not specifically instructed during collection. TABLE I displays the statistical result of the detected angles and we see that detection is effective as the axis corresponding to the deviation always has a relatively higher mean value with a somewhat constant standard deviation. Evaluation of the system includes 2 sets of experiments. 1st experiment targets the effect of PC and the 2nd experiment targets the dynamical fusion and feature extraction section. Both experiments utilize the leave-one-out validation among all collected samples.

Starting with the 1st experiment, the papers give a before and after view of PC by including the genuine-imposter score distributions with ROC curves for both 2&3D (palmprint and geometric features) along with TABLE II which gives a quantitative view of changes in EER (false positive for authentication). From the shown result, PC significantly decreases the overlap between the distributions and the ROC curves further confirms that the process improves the false acceptance rate for 2&3D palmprint and geometry. Quantitatively, TABLE II has also shown that all metrics in EER have reduced by at least 10%, hence we may conclude that PC is quite effective. However, the overall result for geometry features is not as capable as palmprint. It is likely that this is due to the loss of information during PC caused by oversimplifying the problem; Furthermore, the author can improve the evaluation by providing a more qualitative view e.g. the hand image corresponding to the distributions can be given to further show the capability of the system.

Moving onto the 2nd experiment, the paper gives a quantitative view via TABLE III along with a ROC curve representing the false acceptance rate of applying dynamic fusion (DF). From the given result, there is a > 40% improvement by applying DF, hence we can conclude that DF correctly handles the situation where erroneous geometrical data needs to be removed from the evaluation. The remaining errors could be caused by a violation of plane assumption, noise due to lighting condition etc. However, the paper has not given enough analysis in this experiment as almost no qualitative or visual analysis is provided. Further inclusion of genuine-imposter distribution or more ROC curves can better illustrate the effectiveness and flexibility of DF.

4. Conclusions

In general, the paper proposed an innovative system that attempts at the problem of contact-free and poses invariant biometric authentication. The paper has given an extensive explanation of the involved technically sufficient method as well as analysis on both the limitation and strength of the proposed method. The author has also provided many quantitative and qualitative result with visual aids and a realistic conclusion with future possibilities. Moreover, the paper explores a new area of research that has not seen much traction, yet which may encourage further effort towards the field of automated biometric authentication.

In terms of the method proposed by the paper, it addresses the initially stated problem by providing a well-defined and structured system that is built on both innovative and reliable classical methods with some promising result to back up its technical claims. However, there still exist many unsolved issues. e.g. the loss of information in geometrical pose adjustment in 2&3D; several areas of oversimplification such as illumination adjustment and the main rotational algorithm can be also improved. Despite those adversities, the paper still demonstrated its capability on actual data and has shown many areas of possible improvement and research.

In terms of future research, proposed methods in the paper majorly featured classical methods such as matrix rotation and IRLS, perhaps considering more modern ML methods such as CNN and deep learning [5] may address many limitations stated in the paper. Despite recent methods, older but more complex and capable methods can be also be considered to address oversimplified areas in this paper. Reducing Genuine-imposter overlap can maybe be done through particle filtering or Kalman filtering. Information loss due to large deviation can maybe be detected using Bayesian inference or more complex ML modelling.

References

[1] N. Otsu," A Threshold Selection Method from Gray-Level Histograms," in IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 62-66, Jan. 1979, doi: 10.1109/TSMC.1979.4310076.

[2] Lin, Sen & Xu, Tianyang & Yin, Xinyong. (2016). Region of interest extraction for palmprint and palm vein recognition. 538-542. 10.1109/CISP-BMEI.2016.7852769.

[3] Wang G, Wang Y, Li H, Chen X, Lu H, et al. (2014) Morphological Background Detection and Illumination Normalization of Text Image with Poor Lighting. PLoS ONE 9(11): e110991. doi:10. 1371/journal.pone.0110991

[4] Barron, J. 2020. A Generalization of Otsu's Method and Minimum Error Thresholding. *arXiv e-prints*, p.arXiv:2007.07350.

[5] P. Fischer, A. Dosovitskiy, & T. Brox (2015). Image Orientation Estimation with Convolutional Networks. In *German Conference on Pattern Recognition (GCPR)*. Springer.

[6] A. W. K. Kong and D. Zhang, "Competitive coding scheme for palmprint
verification," in Proc. IEEE Int. Conf. Pattern Recognit., Washington,
DC, 2004, pp. 1051–4651.

[7] C. Dorai and A. K. Jain, "COSMOS—A representation scheme for 3-D
free-form objects," IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no.
10, pp. 1115–1130, Oct. 1997.

[8] A. Kumar and H. C. Shen, "Palmprint identification using palmcodes," *Third International Conference on Image and Graphics (ICIG'04)*, Hong Kong, China, 2004, pp. 258-261, doi: 10.1109/ICIG.2004.110.

[9] Gong, Weiyong & Zhang, Xinman & Deng, Bohua & Xu, Xuebin. (2019). Palmprint Recognition Based on Convolutional Neural Network-Alexnet. 313-316. 10.15439/2019F248.

[10] V. Kanhangad, A. Kumar, and D. Zhang, "Combining 2-D and 3-D hand geometry features for biometric verification," in Proc. IEEE Workshop Biometrics, Miami, FL, Jun. 2009, pp. 39–44.

[11] Miroslav Bača; Petra Grd; Tomislav Fotak (2012). "4: Basic Principles and Trends in Hand Geometry and Hand Shape Biometrics" (PDF). *New Trends and Developments in Biometrics*. InTech. Retrieved 1 December 2013.