# Face Recognition Using String Grammar Nearest Neighbor Technique

P. Kasemsumran[1], S. Auephanwiriyakul[1,2], and N. Theera-Umpon[2,3]
[1]Computer Engineering Department, Faculty of Engineering, Chiang Mai University, Chiang Mai, Thailand
[2]Biomedical Engineering Center, Chiang Mai University, Chiang Mai, Thailand
[3]Electrical Engineering Department, Faculty of Engineering, Chiang Mai University, Chiang Mai, Thailand
Email: payungsak2517@gmail.com, {sansanee, nipon}@ieee.org

***Abstract*—Face recognition has become one of the important biometrics in many applications. However, there is a problem of collecting more than one image per person in the training data set, the so-called "one sample per person problem". Hence in this paper, we develop a face recognition system with a string grammar nearest neighbor (sgNN) to cope with the problem. We implement our system in three data sets, i.e., ORL, MIT-CBCL, and Georgia Tech databases. The recognition rates of the test data set from three databases are 88.25%, 87.50%, and 70.71%, respectively.**

***Index Terms*—face recognition, one sample per person, string grammar, nearest neighbor, Levenshtein distance**

## I. Introduction

Face recognition is one of the important biometric techniques that are used in many applications including law enforcement, surveillance, security, etc. It can be done even though the subjects are not cooperated, e.g., identify a person from a surveillance camera. In the case of uncooperative subjects, many challenging problem occurs such as a variation of light, pose, face expression, and so on. However, there have been many research groups working in the face recognition area with some of these challenging problems [1]-[12]. Although the recognition results could reach 95%, those algorithms needed more than one image per person in the training process. It has been shown in [13] that the accuracy dropped with the decreasing of number of samples per person in the training process. Also, since some real world applications, it is not easy to collect more than one image per person. This "one sample per person problem" [13] is another challenging problem in face recognition that recently has become an interest of many research groups [14]-[25]. The accuracies from these research works were not quite high (around 70% to 86%). However, there were some other works on "one sample per person problem" that yielded high correct classification [13], [26]-[29]. Those works usually had some pre-processing, e.g., crop image according to eye coordinate or background removal.

In this paper, we develop a face recognition system with "one sample per person problem" using string grammar nearest neighbor without cropping image or background removal.

This paper is organized as follows. The next section describes the proposed system in details. The experimental results are illustrated in Section III. The conclusions are drawn in Section IV.

## II. System Description

The face recognition system is shown in Fig. 1. We need to preprocess a face image so that we can generate a string for each image because the String Grammar Nearest Neighbor (sgNN) [30], [31] is a syntactic relative of the Nearest Neighbor method. To make the recognition simple, we resize each image to 200×200. Since the face images are in color, in the pre-process step, we firstly convert them into gray-scale images using luminance (Y) component [32]. Then we compute the average (*Ave_f*) of all the images in the training data set. Please be noted that there is only one image per person in the training data set. Then we compute the difference between each image ($Ori\_f_i$) (the original image of the $i^{th}$ person) and the average, i.e.,

$$Dif\_f_i = Ori\_f_i - Ave\_f \quad (1)$$

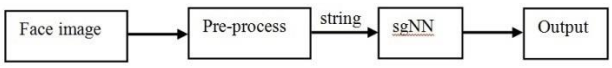



Figure 1. Face recognition system

Fig. 2 shows an example of the original image, an average of the training data set from ORL database [33] and its difference with respect to the average.

(a) (b) (c)

Figure 2. An example of (a) original image, (b) average image, and (c) the difference between (a) and (b)

To reduce the effect of the variation of illumination, we apply the self quotient normalization [34]. Firstly, the

Manuscript received July 10, 2014, revised June 5, 2015.


doi: 10.18178/joig.3.1.6-10

original image is blurred by convolving with the 200×200 Gaussian kernel with $\sigma$=1. This blurred image of $i^{th}$ person is called $Blur\_f_i$. Then, the difference of the $i^{th}$ person is divided pixel-wise by the blurred image of the same person. The final image of the $i^{th}$ person is called $Fi\_f_i$. Fig. 3 shows an example of this process.


$Dif\_f_i$ $Blur\_f_i$ $Fi\_f_i$

Figure 3. An example of self quotient normalization process

Now, we are ready to create a string for each image by first dividing $Fi\_f_i$ into nonoverlapped subimages, each has a size of 5×5. Hence, we have 1600 subimages. Then for each subimage, we implement the Histogram of Gradients (HoG) method [35]-[37] with 8 bins as shown in Fig. 4. The orientation in each bin is shown Table I. The orientation of each pixel $(x,y)$ in the $r^{th}$ subimage is computed by

$$\theta_r(x,y) = 360 - \tan^{-1}\left(\frac{Fi\_f_{ir}(x,y+1) - Fi\_f_{ir}(x,y-1)}{Fi\_f_{ir}(x+1,y) - Fi\_f_{ir}(x-1,y)}\right) \quad (2)$$

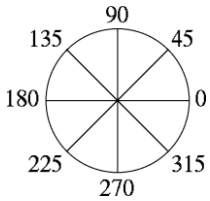



Figure 4. Eight bins in HoG

TABLE I. BIN ORIENTATION

| Bin No. | Orientation |
|---|---|
| 1 | $0<=\theta_r(x,y)<45$ |
| 2 | $45<=\theta_r(x,y)<90$ |
| 3 | $90<=\theta_r(x,y)<135$ |
| 4 | $135<=\theta_r(x,y)<180$ |
| 5 | $180<=\theta_r(x,y)<225$ |
| 6 | $225<=\theta_r(x,y)<270$ |
| 7 | $270<=\theta_r(x,y)<315$ |
| 8 | $315<=\theta_r(x,y)<360$ |

Then the bin with the maximum frequency will be a representative of that subimage. For simplicity, we represent each bin number by a character in the string representing the image, e.g., the bin number 5 has the maximum frequency of the $15^{th}$ subimage, the character of this subimage is also 5. This process is repeated for all subimages. Then the string of the image is achieved. Fig. 5 shows the process of computing a string.

In the recognition process, we implement the string grammar nearest neighbor (sgNN) [30], [31] which is a counterpart of the nearest neighbor [38]. Identifying a string of image $i$ ($st_i$) to the closest string of image $j$ ($st_j$) is as following

$$st_i \text{ is } j^{th} \text{ person if } d(st_i, st_j) = \min_{1\le k\le TN}\left(d(st_i, st_k)\right) \quad (3)$$

where $TN$ is the number of persons in the training data set. The distance between string $i$ and string $k$, $d(st_i, st_k)$, is Levenshtein distance [30].

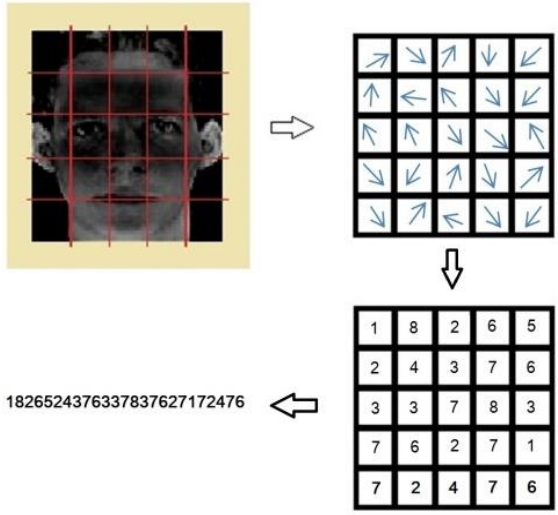
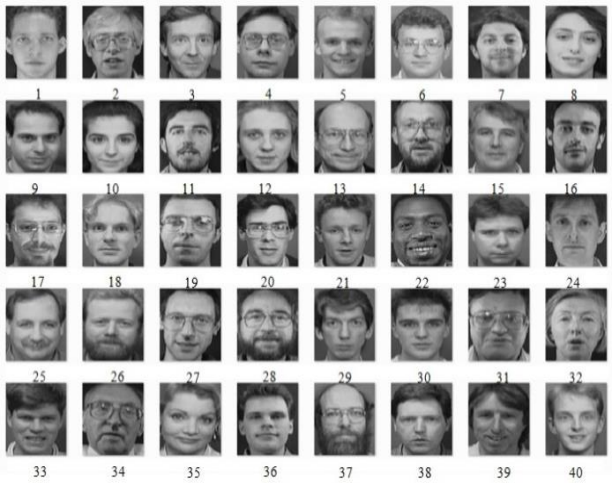



Figure 5. String created process

## III. EXPERIMENTAL RESULTS

We implement the system on three standard data sets, i.e., ORL or AT&T database [10], [33], MIT-CBCL database [39] (Credit is hereby given to the Massachusetts Institute of Technology and to the Center for Biological and Computational Learning for providing the database of facial images), and Georgia Tech face database [40]. In the ORL database, there are 40 persons with 10 different images (variation of facial expression, poses, illumination, rotation and scale) per person in the data set. There are 10 individual and each has 200 images with different rotations and illuminations in the MIT-CBCL database. In the Georgia Tech face data set, there are 50 persons with 15 face images per person. These images are varied in size, facial expression, illumination, and rotation. We utilize the same setting for each data set, i.e., we manually select one frontal face position of each subject from a separate database to be our training data set. Hence in each training data set, there is only one image for each subject. Fig. 6 shows the train images of ORL, MIT-CBCL, and Georgia Tech data sets. The recognition rate of ORL database is 88.25%, while that of the MIT-CBCL database is 87.50%, and that of the Georgia Tech data base is 70.71%. An example of a successful recognition from the ORL data base is shown in Fig. 7. The recognition rates from all 3 data sets are very promising, however, the system sometimes recognizes wrong subjects. An example of misrecognition is shown in Fig. 8. This is because the face in the testing face image has similar structure with the identified person when the testing subject has a variation of face detail or pose. Since our method is utilizing only the structure of the image, not the color, there might be an advantage if the hair color is added into our system also.



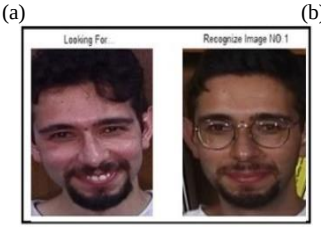
(a)

(b)

(c)

Figure 6. Images in (a) ORL training data set, (b) MIT-CBCL training data set, and (c) Georgia tech training data set

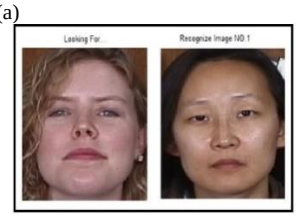

(a) (b)




(c)

Figure 7. Example of successful recognition from (a) ORL, (b) MIT-CBCL, and (c) Georgia tech databases.


(a) (b)




(c)

Figure 8. Example of wrong recognition from (a) ORL data set, (b) MIT-CBCL data set, and (c) Georgia Tech data set.

## IV. Conclusions

In this paper, we develop a face recognition system for "one sample per person problem" using the string grammar nearest neighbor (sgNN). The system provides 88.25%, 87.50%, and 70.71% recognition rates for the ORL, MIT-CBCL, and Georgia Tech databases. Although we have satisfied results, we still need an improvement in the preprocessing, e.g., the string generating process. Also, the hair color is needed to be added in the recognition process. Another problem is that the Levenshtein distance [30], [31] is based on the transformation of the strings that may not be appropriate for computing the distance since there may be two strings that are close from the view of this distance, but actually they are far apart in the normal sense.

## Acknowledgement

The authors would like to thank Pol.Lt.Col Apichart Hattasin from The Children and Women Protection Unit & Transnational Crime Unit, Crime Investigation Division of Police Region 5, Royal Thai Police who is supported by the Destiny Rescue Foundation for valuing and foreseeing the upcoming of computing technology to serve the public.

## References

[1] C. Kotropoulos, A. Tefas, and I. Pitas, "Frontal face authentication using morphological elastic graph matching," *IEEE Trans. on Image Processing*, vol. 9, no. 4, pp. 555-560, Apr. 2000.

[2] L. Chen, H. Man, and A. V. Nefian. (2005). Face recognition based on multi-class mapping of fisher scores. *Pattern Recognition*. [Online]. 38(6). pp. 799-811. Available: http://www.ElsevierComputerScience.com

[3] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210-227, Feb. 2009.

[4] D. Cai, X. He, Y. Hu, J. Han, and T. Huang, "Learning a spatially smooth subspace for face recognition," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1-7.

[5] S.-W. Lee, S.-H. Moon, and S.-W. Lee. (2007). Face recognition under arbitrary illumination using illuminated exemplars. *Pattern Recognition*. [Online]. 40(5). pp. 1605-1620. Available: http://www.sciencedirect.com

[6] D. L. Swets and J. Weng, "Using discriminant eigenfeatures for image retrieval," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 18, no. 8, pp. 831-836, 1996.

[7] W. Zhao, R. Chellappa, and P. J. Phillips, "Subspace linear discriminant analysis for face recognition," Technical Report CAR-TR-914, Center for Automation Research, University of Maryland, Apr. 1999.

[8] J. Wang, J. You, Q. Li, and Y. Xu, "orthogonal discriminant vector for face recognition across pose," *Pattern Recognition*, vol. 45, pp. 4069-4079, 2012.

[9] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," *IEEE Trans. Neural Networks*, vol. 8, no. 1, pp. 98-113, Jan. 1997.

[10] S. J. Wang, H. L. Chen, W. J. Yan, Y. H. Chen, and X. Fu, "Face recognition and micro-expression recognition based on discriminant tensor subspace analysis plus extreme learning machine," *Neural Process Letter*, vol. 39, pp. 25-43, 2014.

[11] X. Zhou, W. Jiang, Y. Tian, and Y. Shi, "Kernel subclass convex hull sample selection method for SVM on face recognition," *Neurocomputing*, vol. 73, pp. 2234-2246, 2010.

[12] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognitive Neurosci.*, vol. 3, no. 1, pp. 71-86, 1991.

[13] X. Tan, S. Chen, Z. H. Zhou, and F. Zhang, "Face recognition from a single image per person," *Pattern Recognition*, vol. 39, pp. 1725-1745, 2006.

[14] S. C. Chen, J. Liu, and Z.-H. Zhou, "Making FLDA applicable to face recognition with one sample per person," *Pattern Recognition*, vol. 37, no. 7, pp. 1553-1555, 2004.

[15] F. Samaria and A. Harter, "Parameterisation of a stochastic model for human face identification," in *Proc. Second IEEE Workshop on Applications of Computer Vision*, Dec. 1994, pp. 138-142.

[16] D. Beymer and T. Poggio, "Face recognition from one example view," *Science*, vol. 272, 1996.

[17] A. M. Martinez, "Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 6, pp. 748-763, 2002.

[18] J. Yang, D. Zhang, A. F. Frangi, and J. Yang, "Two-Dimensional PCA: A new approach to appearance-based face representation and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 131-137, 2004.

[19] J. Wu and Z.-H. Zhou, "Face recognition with one training image per person," *Pattern Recognition Lett.*, vol. 23, no. 14, pp. 1711-1719, 2002.

[20] S. C. Chen, D. Q. Zhang, and Z.-H. Zhou, "Enhanced (PC) 2A for face recognition with one training image per person," *Pattern Recognition Lett.*, vol. 25, no. 10, pp. 1173-1181, 2004.

[21] D. Q. Zhang, S. C. Chen, and Z.-H. Zhou, "A new face recognition method based on SVD perturbation for single example image per person," *Appl. Math. Comp.*, vol. 163, no. 2, pp. 895-907, 2005.

[22] S. C. Chen, J. Liu, and Z.-H. Zhou, "Making FLDA applicable to face recognition with one sample per person," *Pattern Recognition*, vol. 37, no. 7, pp. 1553-1555, 2004.

[23] J. Huang, P. C. Yuen, W. S. Chen, and J. H. Lai, "Component-Based LDA method for face recognition with one training sample," in *Proc. AMFG*, 2003, pp. 120-126.

[24] J. Wang, K. N. Plataniotis, and A. N. Venetsanopoulos, "Selecting discriminant eigenfaces for face recognition," *Pattern Recognition Lett.*, vol. 26, no. 10, pp. 1470-1482, 2005.

[25] K.-M. Lam and H. Yan, "An analytic-to-holistic approach for face recognition based on a single frontal view," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 7, pp. 673-686, 1998.

[26] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," *Computer Vision, Lecture Notes in Computer Science*, vol. 3021, pp. 469-481, 2004.

[27] L. Wiskott, R. Fellous, N. Kruger, and C. V. Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 775-779, Jul. 1997.

[28] H.-S. Le and H. Li, "Recognizing frontal face images using hidden Markov models with one training image per person," in *Proc. 17th International Conference on Pattern Recognition (ICPR04)*, 2004, vol. 1, pp. 318-321.

[29] S. G. Shan, W. Gao, and D. Zhao, "Face identification based on facespecific subspace," *Int. J. Image Syst. Technol.*, vol. 13, no. 1, pp. 23-32, 2003.

[30] J. C. Bezdek, J. Keller, R. Krishnapuram, and N. R. Pal, *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*, USA: Kluwer Academic Publishers, 1999.

[31] K. S. Fu, *Syntactic Pattern Recognition and Applications*, Englewood Cliffs, NJ: Prentice Hall, 1982, pp. 155-182.

[32] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed., New Jersey: Pearson Education, Inc., 2008.

[33] ORL Face Database, AT&T Laboratories Cambridge, 1992-1994.

[34] P. J. Phillips, *et al.*, "The good, the bad, and the ugly face challenge problem," *Image and Vision Computing*, vol. 30, pp. 177-185, 2012.

[35] V. Chandrasekhar, G. Takacs, D. Chen, S. Tsai, R. Grzeszczuk, and B. Girod, "CHoG: Compressed histogram of gradients a low bit-rate feature descriptor," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 2504-2511.

[36] A. Psyllos, C. N. Anagnostopoulos, and E. Kayafas, "Vehicle model recognition from frontal view image measurements," *Computer Standards & Interfaces*, vol. 33, pp. 142-151, 2011.

[37] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, pp. 886-893.

[38] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Trans on Inform. Theory*, vol. 13, pp. 21-27, 1967.

[39] B. Weyrauch, J. Huang, B. Heisele, and V. Blanz, "Component-Based face recognition with 3D morphable models," in *Proc. First IEEE Workshop on Face Processing in Video*, Washington, DC, 2004.

[40] Georgia Tech Face Database. [Online]. Available: ftp://ftp.ee.gatech.edu/pub/users/hayes/facedb/.


**Payungsak Kasemsumran** received the B.S. degree in computer science from Chiang Mai Rajabhat University, Chiang Mai, in 1996, M.S. degree in computer science from Chiang Mai University, Chiang Mai, in 2008. He is currently a Ph.D. candidate in computer engineering at Chiang Mai University, Chiang Mai, Thailand. In 1989, he completed a certificate of mathematic outstanding at Mathematic Olympic. His research interests include expert system, image and signal processing, computer vision, pattern recognition, fuzzy set theory.

**Sansanee Auephanwiriyakul** (S'98–M01) received the B.Eng. (Hons.) degree in electrical engineering from the Chiang Mai University, Thailand (1993), and the M.S. degree in electrical and computer engineering and Ph.D. degree in computer engineering and computer science, both from the University of Missouri-Columbia in 1996, and 2000, respectively. She had worked as a Postdoctoral Fellow at the Computational Intelligence Laboratory, University of Missouri, Columbia, on clustering techniques for humanitarian land mine detection. She is currently an Assistant Professor in the Department of Computer Engineering, Chiang Mai University, Thailand. Her research interests include neural networks, fuzzy set theory, linguistic algorithms, pattern recognition, image processing, computer vision and medical image processing. Dr. Auephanwiriyakul is a senior member of the Institute of Electrical and Electronics Engineers (IEEE). She is an Associate Editor of the IEEE Transactions on Fuzzy System, and is an Editorial Board of the Engineering Journal, Chiang Mai. She was a Guest Editor of the Journal of Advanced Computational Intelligence and Intelligent Informatics in 2004. She was a Technical Program Chair, Organizing Committee in several major conferences including the IEEE International Conference on Fuzzy Systems.

**Nipon Theera-Umpon** (S'98-M'00-SM'06) received his B.Eng. (Hons.) degree from Chiang Mai University, M.S. degree from University of Southern California, and Ph.D. degree from the University of Missouri-Columbia, all in electrical engineering. He has been with the Department of Electrical Engineering, Chiang Mai University since 1993. He has served as editor, reviewer, general chair, technical chair and committee member for several journals and conferences. He was associate dean of Engineering and chairman for graduate study in electrical engineering. He is serving as the director of Biomedical Engineering Center and the chairman for graduate study in biomedical engineering, Chiang Mai University. He is a member of Thai Robotics Society, Biomedical Engineering Society of Thailand, Council of Engineers in Thailand. He has served as Vice President of the Thai Engineering in Medicine and Biology Society. Dr. Theera-Umpon is a senior member of the IEEE and IEEE-IES Technical Committee on Human Factors. His research interests include Pattern Recognition, Digital Image Processing, Neural Networks, Fuzzy Sets and Systems, Medical Signal and Image Processing.