Genetic-Based Selection and Weighting for LBP, oLBP, and Eigenface Feature Extraction

Tamirat Abegaz[#], Gerry Dozier[#], Kelvin Bryant[#], Joshua Adams[#], Brandon Baker[#], Joseph Shelton[#], Karl Ricanek^, Damon L. Woodard*

#North Carolina A&T State University

^The University of North Carolina at Wilmington

**Clemson University*

Abstract— **In this paper, we have investigated the use of genetic-based feature selection (GEFeS), genetic-based feature weighting (GEFeW) on feature sets obtained by Eigenface and LBP. Our results indicate that GEFeS and GEFeW enhances the overall performance of both the Eigenface and LBP-based techniques. Compared to Eigenface hybrid, our result shows that both LBP and oLBP hybrids perform better in terms of accuracy. In addition, the results show that GEFeS reduces the number of features needed by approximately 50% while obtaining a significant improvement in accuracy.**

Keywords— Local Binary Pattern (LBP), Eigenface, Steady State Genetic Algorithm (SSGA), Overlapping Patches, Feature Selection.

I. Introduction

Feature Selection is a computational technique that attempts to identify a subset of features that are most relevant to a particular task (such as biometric identification) [1]. The ideal feature selection technique removes those features that are less discriminative and keeps those features that have high discriminatory power. A number of feature selection techniques have been developed and can be classified as: Enumeration Algorithms (EAs), Sequential Search Algorithms (SSAs), and Genetic Algorithms (GAs). EAs guarantee the optimal subset of features by evaluating all possible subsets of the features. This works well for a very small sized feature sets, however, it is computationally infeasible when the size of the feature set is large [2].

SSAs attempts to divide a feature set, *U,* into two subsets of features, *X,* and *Y,* where *X* denotes the selected features and *Y* denotes the remaining ones. Based on user specified criteria, SSAs select the least significant features from the subset *X* and moves those features into *Y* while selecting the most significant features from *Y* and moving them into *X.* While SSAs are suitable for small and medium size problems, they are too computationally expensive to use on large problems [2].

GAs attempt to find an optimal (or near optimal) subset of features for a specific problem [3, 4, 5, 6, 7, 8, 9, 10]. First, a number of individuals or candidate Feature Subsets (FSs) are generated to form an initial population. Each FS is then evaluated and assigned a fitness obtained from the evaluation function specific to the problem at hand. Parents are then selected based on fitness. New FSs are produced from the selected parents by the processes of reproduction. Survivors are selected from the previous generation and combined with the offspring to form the next generation. This process continues for user specified number of cycles.

This work is an extension of the research performed by Abegaz et. al [10]. In their work, Abegaz et al. used Genetic and Evolutionary Feature Selection (GEFeS), GEFeS+ (which is a co-evolutionary version of GEFeS) , and Genetic and Evolutionary Feature Weighting (GEFeW), Eigenface algorithm. In their work, Abegaz et. al. reported that Eigen-GEFeS, Eigen-GEFeS+, and Egen-GEFeW enhanced the overall performance of the Eigenface method while reducing the number of features needed. Comparing Eigen-GEFeS, Eigen-GEFeS+, and Eigen-GEFeW, they reported that Eigen-GEFeW performed best in terms of accuracy even though it used a significantly larger number of features as compared to either Eigen-GEFeS or Eigen-GEFeS+. In this paper, we extend the work of Abegaz et. al compare GEFeS, GEFeS+, and GEFeW hybrids using Eigenface, LBP, and overlapped LBP (oLBP).

Our work is partly motivated by the research of Gentile et. al [11, 12]. Gentile et. al proposed a hierarchical two-stage process to reduce the number of feature checks required for an iris-based biometric recognition system. The claimed that a shorter representation of the iris template by pre-aligning the probe to each gallery sample and generate a shortlist of match candidates. Our target is a similar system for Face Recognition (FR) based on short length biometric templates that are able to achieve higher recognition accuracies.

The remainder of this paper is as follows. Section II explains the feature extraction techniques used as input for the GEFeS, GEFeS+, and GEFeW. Section III provides an overview of GEFeS, GEFeS+, and GEFeW. Section IV presents our experiment, and in Section V we present our results. Finally, our conclusions and future work are presented in Section VI.

II. Feature Extraction using Eigenface, LBP, and oLBP

In a typical biometric system, the task of sample acquisition and feature extraction are always performed [13]. Sample acquisition is the gathering of biometric traits such as fingerprints, iris scan, periocular images, or facial images.

From the acquired sample, feature extraction is performed to create a feature vector to be used for comparison. In the case of a facial biometric sample, Eigenface and LBP are commonly used feature extractors. For a typical feature extractor, the pre-enrolled images (and their associated feature vectors) are stored in a database commonly referred to as gallery [13], while newly acquired images (and their feature vectors) are called probes [13].

For Eigenface based feature extraction [14], each image in the training dataset was converted into a single vector. This conversion is necessary because one needs a square matrix (transformation matrix or covariance matrix) to compute the Eigenvectors (Eigenfaces) and the Eigenvalues . The gallery images have been used to construct a face space spanned by the Eigenfaces. Each image is then projected into the face space spanned by the Eigenfaces. 560 discriminatory feature weights were extracted for each image and stored for the feature selection experiments.

For LBP based feature extraction [15, 16], an image is first divided into several patches (blocks) from which local binary patterns are extracted to produce histograms from every non-border pixels. The histogram obtained from each patch is concatenated to construct the global feature histogram that represents both the micro-patterns and their spatial location. In other words, the histograms contain description of the images on three different levels of localities. The first one indicates that the labels for histograms contain information about the pattern on a pixel level. Second, the summation of the labels obtained in the patch level to produce the information on a regional level. Finally, the histograms at the regional level are concatenated to produce the global descriptor of the image.

The standard LBP uses those labels which have at most one 0-1 and one 1-0 transitions when viewed as a circular bit string. Such labels are known as uniform patterns [17] For uniform pattern LBP, every patch (block) consists of $P(P-1)+3$ bins where $\mathrm{P}(\mathrm{P}-1)$represents the bins for the patterns with two transitions [18]. The remaining three bins represents the bins for the patterns with 0 transitions (all zeros (00000000) and all ones (11111111), and for all non-uniform patterns (bin that represents more than two transitions) [18]. The total number of histogram is computed using the formula, $\mathrm{B}(\mathrm{P}(\mathrm{P}-1)+3)$, where B represents the number of blocks and P represents the of sampling points. For our research, we use P =8, and B =36 to obtain a feature vector of 2124.

oLBP based feature extraction [18] is a variant of LBP that attempts to include the internal border pixels that are left out during the process of logical portioning on the standard LBP feature extraction method. This is done by logically overlapping the patches horizontally, vertically, and both horizontally and vertically with a one pixel overlap. This provides information to determine whether including the middle border pixels have impact on the recognition rate of the LBP based face recognition algorithm.

III. GEFeS, GEFeS+, and GEFeW

GEFeS, GEFeS+, and GEFeW were designed for selecting and/or weighting the most discriminatory features for recognition. GEFeS, GEFeS+, and GEFeW are instances of a Seady State GA(SSGA) with in eXplanatory Toolset for the Optimization Of Launch and Space Systems (X-TOOLSS) [19]. In order to describe GEFeS, consider the following feature vector.

[30 16 21 151 117 27 113 44 28 54 78 240 77 124 189 . . .]

Figure 1: Sample feature vector

Furthermore, consider the vector shown in Figure 2 as a candidate real-coded feature mask.

[0.5 0.11 0.7 0.2 0.9 0 0.15 0.6 0.2 0.44 0.7 0 0.3 0.6 0.2 . . .]

Figure 2: Real-Coded Feature Mask

For GEFeS a masking threshold value of 0.5 is used to create a binary coded candidate feature mask which will be used as condition for masking features. If the random real number generated is less than the threshold (0.5 in this case), then the value corresponding to the real generated number is set to 0 in the candidate feature mask vector or 1 otherwise. The candidate feature mask is used to mask out a feature set extracted for a given biometric modality. Figure 3 shows the candidate binary coded feature mask matrix obtained from the random real numbers generated in Figure 2. The masking threshold value is applied on the real numbers to obtain the binary representation

[1 0 1 0 1 0 0 1 0 0 1 0 0 1 0 . . .]

Figure 3: Binary coded candidate feature mask

When Comparing the candidate feature mask with the feature matrix, if a position corresponding to the feature matrix value in the candidate feature mask is 0 then that feature value will be masked out from being considered in the distance computation. Figure 4 shows the result of the features in Figure 1 when feature masking (Figure 3) is applied to a feature vector.

[30 0 21 0 117 0 0 44 0 0 78 0 0 124 0 . . .]

Figure 4: The Resulting feature vector after feature masking

GEFeS + is a co-evolutionary version of GEFeS where that instead of using the static threshold value of 0.5, we evolve a threshold value between 0 and 1. So each random number generated using a uniform distribution has a masking threshold value that determines whether the feature corresponding to features is masked out or not.

For GEFeW, the real-coded candidate feature mask is used to weight features within the feature matrix. The real-coded candidate feature mask value is multiplied by each feature value to provide a weighted feature. If the number generated is 0 (or approximately equal to 0) the feature value is 0, which basically means that the feature is masked.

As given in Equation 1, the fitness returned by the evaluation function is the number of recognition errors encountered after applying the feature mask multiplied by 10 plus the percentage of features used. The selection of the parent is based on smaller fitness values because the optimization goal is to reduce the number of recognition errors (i.e. increasing the accuracy) while reducing the number of features.

$$fitness = (number\ of\ errors) * 10 + \%Features\ Used \quad (1)$$

IV. Experiments

The dataset used in this research is a subset of the Face Recognition Grand Challenge (FRGC) dataset [20]. In our dataset, 280 subjects were used, with each subject having a total of 3 associated images with it. Out of 840 images, 280 were used as probe and 560 images were selected for training images. The images had passed the pre-processing stages such as eye rotation alignment, histogram equalization, masking resizing (each with 225 by 195), and conversion of the images into greyscale.

For the GEFeS, GEFeS+, and GEFeW, the inputs used were the features extracted using Eigenface, LBP, and oLBP feature extraction methods. These methods were used on a subset of the FRGC dataset. This subset was selected because it contains a variety of imaging conditions such as different ethnic origins, frontal images that were neutral, and frontal images that had facial expressions.

The objective of this experiment is to compare the impact of applying GEFeS, GEFeS+, GEFeW on the Eigenface, LBP, and oLBP based feature extraction methods.

V. Results

For our experiment, nine GEFeS, GEFeS+, GEFeW instances were used. These instances all have a population size of 20, Gaussian mutation rate of 1 and mutation range of 0.2. The Mutation rate value of 1 implies that all children (100%) must undergo mutation. The mutation range provides a window from the current value (obtained value after recombination) that the new value will be mutated. Furthermore, they were each run a total of 30 times with a maximum of 1000 function evaluations. GEFeS, GEFeS+, and GEFeW were designed for selecting and/or weighting the most discriminatory features for recognition. Our results are shown in Tables I.

In Table I, the columns represent the method used, the percentage of the average features, the average accuracy, and the best accuracy obtained. The percentage of the average accuracy is computed using the results obtained from the 30 runs. The best accuracy is selected from the run that resulted in the smallest number of errors.

ANOVA and t-Tests were used to divide the GEFeS, GEFeS+, GEFeW instances and the baseline algorithms into equivalence classes. As shown in Table 1, comparing the baseline algorithms, the Eigenface method performs best. The results show that when using 100 percent of the features, the maximum accuracy obtained for the baseline LBP was 70.36%. While the $\text{Baseline}_{\text{LBPBest}}$ performs slightly better than the baseline $\text{Baseline}_{\text{LBP},}$ it still uses the entire feature set As can be seen in Table 1, applying GEFeS on the feature set extracted by the standard LBP significantly improves accuracy from a 70.36% to 96.62%. This result shows that GEFeS is actually masking out those features which are less relevant for recognition. This improvement in accuracy comes also with a reduction in the number of features used for recognition.

TABLE I
EXPERIMENTAL RESULTS OF THE LBP BASELINE, OLBP AND THE EIGENFACE METHODS

Methods	Number of Features Used	% Accuracy	Best Accuracy
$\text{Baseline}_{\text{LBP}}$	2124	70.36	70.36
$\text{Baseline}_{\text{oLBPbest}}$	2124	70.71	70.71
$\text{Baseline}_{\text{Eigenface}}$	560	87.14	87.14
Eigen-GEFeS	291.2	86.67	87.85
LBP-GEFeS	1022.1	96.62	97.14
oLBP-GEFeS	1018.46	96.43	96.79
Eigen-GEFeS+	476	88.48	88.92
LBP-GEFeS+	463.24	96.52	97.14
oLBP-GEFeS+	446.89	96.50	97.14
Eigen-GEFeW	492.8	91.42	92.5
LBP-GEFeW	1865.29	95.33	95.71
oLBP-GEFeW	1865.08	95.33	96.07

Compared to GEFeS and GEFeS+, all of the results show that GEFeW used a larger number of features. Using a larger number of features brings a better result in the case Eigen-GEFeW as compared to Eigen-GEFeS, and Eigen-GEFeS+. Surprisingly, in the case of LBP-GEFeW and oLBP-GEFeW the result is the opposite. Utilizing a significantly larger number of features actually decreases the accuracy for both LBP-GEFeW and oLBP-GEFeW as compared to their corresponding methods.

LBP-GEFeS, LBP-GEFeS+, oLBP-GEFeS, and oLBP-GEFeS+ fall in the best equivalence class with respect to accuracy. This means that there is no statistical difference among them. All performed well in terms of reducing the number of features needed and in producing a significant improvement in accuracy from their corresponding baseline methods.

Figure 1 shows the Cumulative Match Characteristic (CMC) curve for the $\text{Baseline}_{\text{LBP}}$, $\text{Baseline}_{\text{oLBPbest,}}$ $\text{Baseline}_{\text{Eigenface}}$ and for the methods that fall in the first equivalent class. As can be seen from the Figure 1, LBP-GEFeS, LBP-GEFeS+, oLBP-GEFeS, and oLBP-GEFeS+ obtain approximately 97.5%

accuracy at rank 10. However, both $Baseline_{Eigenface}$ and Eigen-GEFeS performed well (approximately 96%) at rank 10. $Baseline_{LBP}$ performed relatively poorly in terms of accuracy.

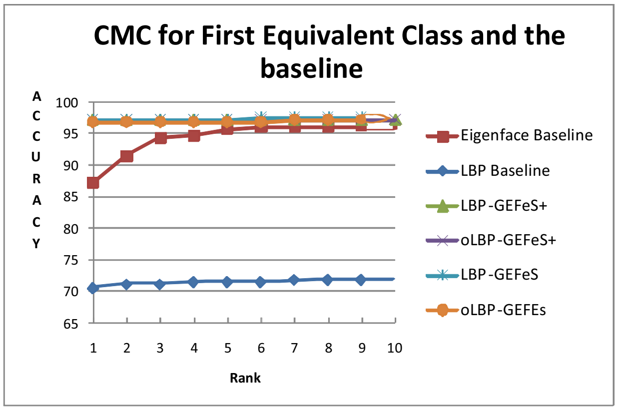

Figure 5: Comparisons of CMC results for baseline and the best performing algorithms

VI. Conclusion and Future Work

Our results using GEFeS, GEFeS+, and GEFeW suggests that hybrid GAs for feature selection/weighting enhances the overall performance of the Eigenface, LBP, and oLBP methods while reducing the number of features needed. When comparing the baseline accuracy, the Eigenface method performed far better than both LBP and oLBP. However, the hybrid GAs result show that both LBP and oLBP hybrids performed much better than the Eigenface hybrid method.

Our future work will be devoted towards the investigation of GEFeS, GEFeS+, and GEFeW based on other forms of Genetic and Evolutionary Computation[21, 22, 23, 24]

Acknowledgment

This research was funded by the Office of the Director of National Intelligence (ODNI), Center for Academic Excellence (CAE) for the multi-university Center for Advanced Studies in Identity Sciences (CASIS) and by the National Science Foundation (NSF) Science & Technology Center: Bio/computational Evolution in Action CONsortium (BEACON). The authors would like to thank the ODNI and the NSF for their support of this research

References

[1] Ajay Kumar "ENCYCLOPEDIA OF BIOMETRICS" 2009, Part 6, 597-602, DOI: 10.1007/978-0-387-73003-5_157

[2] M. Dash and H. Liu, Bartlett, Javier R. Movellan, and Terrence J. Sejnowski,, *"Feature Selection for Classification" Genetic Algorithms for Feature Selection",* Intelligent Data Analysis, vol. 1, no. 3, pp. 131-156, 1997.

[3] J. Adams, D. L. Woodard, G. Dozier, P. Miller, K. Bryant, and G. Glenn. *Genetic-based type II feature extraction for periocular biometric recognition: Less is more*. In Proc. Int. Conf. on Pattern Recognition, 2010. to appear.

[4] *Huang C. L. and Wang C. J. "*GA-based feature selection and parameters optimization for support vector machines*",. C.-L. Huang,* C.-J. Wang / Expert Systems with Applications. Vol. 31(2), 2006, pp231–240.

[5] Adams, J., Woodard, D. L., Dozier, G., Miller, P., Glenn, G., Bryant, K. "*GEFE: Genetic & Evolutionary Feature Extraction for Periocular-Based Biometric Recognition*," Proceedings 2010 ACM Southeast Conference, April 15-17, 2010, Oxford, MS.

[6] Dozier, G., Adams, J., Woodard, D. L., Miller, P., Bryant, K. "A Comparison of Two Genetic and Evolutionary Feature Selection Strategies for Periocular-Based Biometric Recognition via X-TOOLSS,", Proceedings of the 2010 International Conference on Genetic and Evolutionary Methods (GEM'10: July 12-15, 2010, Las Vegas, USA).

[7] Simpson, L. , Dozier, G., Adams, J., Woodard, D. L., Dozier, G., Miller, P., Glenn, G., Bryant, K.. "*GEC-Based Type-II Feature Extraction for Periocular Recognition via X-TOOLSS*," Proceedings 2010 Congress on Evolutionary Computation, July 18-23, Barcelona, Spain.

[8] Dozier, G., Bell, D., Barnes, L., and Bryant, K. (2009). "Refining Iris Templates via Weighted Bit Consistency", Proceedings of the 2009 Midwest Artificial Intelligence & Cognitive Science (MAICS) Conference, Fort Wayne, April 18-19, 2009.

[9] Dozier, G., Adams, J., Woodard, D. L., Miller, P., Bryant, K. "A Comparison of Two Genetic and Evolutionary Feature Selection Strategies for Periocular-Based Biometric Recognition via X-TOOLSS", (to appear in) The Proceedings of the 2010 International Conference on Genetic and Evolutionary Methods (GEM'10: July 12-15, 2010, Las Vegas, USA).

[10] Tamirat Abegaz, Gerry Dozier, Kelvin Bryant Joshua Adams, Khary Popplewell, Joseph Shelton, ,Karl Ricanek, Damon L. Woodard" "*Hybrid GAs for Eigen-Based Facial Recognition"*, accepted for IEEE Symposium Series in Computational Intelligence 2011 (SSCI 2011)

[11] J.E Gentile, N. Ratha, and J. Connell, "SLIC: Short-length iris codes," In Proc. IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems, 2009. BTAS '09, 28-30 Sept. 2009, pp.1-5.

[12] J.E. Gentile, N. Ratha, and J. Connell, "An efficient, two-stage iris recognition system", In *Proc. 3rd International Conference on Biometrics: Theory, Applications, and Systems (BTAS),* 2009.

[13] Peter T. Higgins, *"*Introduction to Biometrics*",* The Proceeding of Biometrics consortium conference 2006, Baltimore"*, MD, USA, Sept. 2006.*

[14] M. Turk and A. Pentland, "Eigenfaces for recognition", *Journal of Cognitive Neuroscience*, Vol. 13, No. 1, pp. 71-86, 1991.

[15] Caifeng Shan and Tommaso Gritti, " Learning Discriminative LBP-Histogram Bins for Facial Expression Recognition", Proc. of 15th EUSIPCO, Poznan, Poland, September 2007.

[16] Goldberg, Toimo Ahonen, Abdenour Hadid, and Matti Pietik¨ainen " Learning Face Expression Recognition", http://www.ee.oulu.fi/mvg/, visited on sept 10, 2120.

[17] J. Zhao, H. Wang, H. Ren, and S. C. Kee," LBP discriminant alalysi for face verification," in Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol 3, pp. 167-172, June 2005.

[18] Tamirat Abegaz, "*GENETIC AND EVOLUTIONARY FEATURE SELECTION AND WEIGHTING FOR FACE RECOGNITION*", thesis submitted to North Carolina A&T State University

[19] *M. L. Tinker, G. Dozier, and A. Garrett, "The exploratory toolset for the optimization of launch and space systems (x-toolss)," http://xtoolss.msfc.nasa.gov/, 2010.*

[20] P. Jonathon Phillips1, Patrick J. Flynn2, Todd Scruggs3, Kevin W. Bowyer2 Jin Chang2, Kevin Hoffman3, Joe Marques4, Jaesik Min2, William Worek3," Overview of the Face Recognition Grand Challenge", IEEE Conference on Computer Vision and Pattern Recognition, 2005.

[21] Danial Ashlock. "Evolutionary Computation for Modeling and Optimization.", Springer, 2005.

[22] Dozier, G., Homaifar, A., Tunstel, E., and Battle, D. (2001). "*An Introduction to Evolutionary Computation*" (Chapter 17), Intelligent Control Systems Using Soft Computing Methodologies, A. Zilouchian & M. Jamshidi (Eds.), pp. 365-380, CRC press.

[23] D. Guillamet, & J. Vitri`a, "*Evaluation of distance metrics for recognition based on non-negative matrix factorization*", Pattern Recognition Letters, 24(9-10), 2003, 1599-1605.

[24] Fogel, D. Evolutionary Computation: Toward a New Philosophy of Machine Intelligence . IEEE Press, 2nd Edition., 2000.