Pairwise Macropixel Comparison Can Work at Least as Well as Advanced Holistic Algorithms for Face Recognition

Liang Chen
lchen@ieee.org;chen.liang.97@gmail.com
University of Northern British Columbia
3333 University Way, Prince George, BC, Canada V2N 4Z9

Abstract

It is well known that, due to illumination effects and the registration/alignment problem, it does not make sense to compare the "values" of two single-pixels for face recognition. But does that mean that the comparison of two "big" pixels makes no sense either? This paper shows that, by taking a few pixels together as one "big" pixel, called macropixel, and measuring the similarity of macropixels by simple Euclidean distance, a method that counts best matched macropixels indeed works very well for face recognition – experiments show that it is not only much better than traditional holistic algorithms, but is also at least comparable with recently developed ones, if not better.

The superiority of the extremely naive macropixel counting approach over well-established ones stimulates us to rethink: Does the seemingly dedicated process of our brains for face pattern recognition involve dimensionality reduction? Has the current advance of computer vision research touched the underlying problem in face recognition?

1 Introduction

Face recognition is a special pattern recognition problem and has been a hot topic in the last thirty years, of which numerous approaches have been developed and published, although we have to confess that we understand very little of the real process our brains use in performing such a task. Holistic subspace approaches (aka holistic approaches, subspace approaches in face recognition community) – the dimensionality techniques in face recognition research – are most popular approaches so far for face recognition, although popularity does not necessarily directly mean superiority; algorithms including Eigenface approach (aka Principal Component Analysis, or PCA approach)[17], Linear Discriminant Analysis (LDA)[12], Fisherface approach[1], Independent Component Analysis (IDA)[10], Gabor feature approach, Spectral Regression Discriminant Analysis (SRDA)[4], Spatially Smooth Subspace Learning (SSSL)[3], Random Subspace Method (SRM) [20], and so on, have continued to be developed and tried. Being encouraged by many promising experiments, new holistic subspace approaches for face recognition have been developed and published in almost every, if not every, computer vision conference in last few years; although the pioneering work of Wang and Tang [19] has shown, using three most representative subspace approaches as examples, that many holistic subspace approaches can indeed be unified under the same framework.

BMVC 2010 doi:10.5244/C.24.5

All these holistic approaches are developed based on the general understanding that (1) due to illumination effects etc., comparing the “values" of two pixels in two face images is unlikely to be meaningful; (2) because there is a registration/alignment problem, we cannot precisely know which pixel is the “corresponding" pixel of another; (3) dimensionality reduction may be able to extract informative or discriminative facts from raw facial image arrays which are of high dimensional, just as in many other pattern recognition problems.

Encouraged by the success of unifying subspace approaches [19], we have the following arguments regarding the general understandings above: (1) We all understand that comparing the “values" of two single-pixels really does not make sense; but does that imply that the comparison of two “big" pixels makes no sense either? (2) The negative effects caused by the registration/alignment problem, if some kind of pixel wise matching is to be used, can be efficiently reduced by “shifting" groups of pixels. (3) Although we do not know how our brains perform face recognition; evidences from many sources, including the studies on prosopagnosia, seem to show that face recognition is a dedicated process of our brains different from other pattern recognition tasks [23]. Then why should we believe that dimensionality reduction, an idea rooted in general pattern recognition, should work for face recognition?

In this paper, we will show that: If we take a few pixels together as one “big" pixel, called a macropixel, then, a method that counts matched “big" pixels, or macropixels, should work well; at least not worse than most of recently developed holistic algorithms, if not better.

Indeed, since a macropixel is still very small, consisting of only a few pixels, the simple Euclidean distance is used to decide how two macropixels are different. In our entire algorithm, no dimensionality reduction is required – actually as it is shown in our experiments in Section 3 that 4×4 is a typical size of a macropixel that works quite well, it is unlikely to be meaningful and there is not much space left to perform any dimensionality reduction inside such a small macropixel[1]; and counting in our approach is just adding up matched pixels, where no dimensionality reduction can fit in. Therefore, our work also suggests that, it is possible that dimensionality reduction technique is unnecessary for face recognition.

The remainder of this paper is organized as follows: Section 2 will provide our detailed approach; the experiments on known datasets together with the comparisons with many holistic algorithms, including newly developed ones, will be shown in Section 3. We will briefly show other possible settings of our macropixel based approach in Section 4. Conclusions and future work will be presented in the last section.

Before going to details, we note here:

1. This paper is intended to demonstrate that the full space macropixel counting algorithm works well such that it is at least comparable with recently developed holistic/subspace algorithms. Although there are many ways to further improve its performances, including weighting macropixels and using distance metrics other than Euclidean, it is not our focus in this paper to include many different settings to explore such further improvements.

2. Recently local patch/component based approaches (e.g., [5, 7, 9, 15, 16]) are hot in face recognition.[2] *Since this paper (as indicated by its title) only intents to show that dimensionality reduction in face recognition, represented by subspace approaches, does not do better than the simple full space macropixel counting approach. We do not claim that our full space*

[1] It is worth noting that, in most of patch/component based approaches in face recognition, a “component" usually consists of lots of pixels, e.g. the area of an eye or nose, so that dimensionality reduction are used there.

[2] Although they may use some of dimensionality reduction techniques in the patches; it is generally understood that they do not belong to the category of dimensionality reduction approaches.

macropixel counting approach is better than local patch/component approaches[3]*; therefore, we will not compare our results with local patch/component approaches. Note again that the macropixels are indeed very small so that the macropixel counting approach does not belong to the category of local patch/component based approaches.(See also footnote 1.)*
3. Knowing that the performance of our full space macropixel counting approach might be somewhat surprising to many researchers, the program codes of our approach are included in "Supplementary Material", and also available at `http://web.unbc.ca/~chenl/DataCode.html`.

2 Macropixel Counting Approach

We assume that all the face images are pre-cropped, scaled and rotated to a size of $m \times n$ pixels, such that the positions of the eyes are aligned. We align all face images into a common coordinate system for the convenience of description.

We shall now take a face image as a union of "macropixels".

Definition (*Macropixel*)**:***Partition an image into windows of size $K \times K$ (We set K to be 4 in most of our experiments, although we will also discuss other choices in Section 4.), leaving s pixels (We let $s = 2$ in most of our experiments, although we will also discuss other choices in Section 4.), called* **shift range**, *in each side of the image for shifting purpose. Each macropixel consists of all the $K \times K$ pixels within a window.*

The effect of illumination is always a problem in face recognition; one or several "pre-processing" strategies, such as vector normalization and histogram equalization and etc., for reducing the effects of illuminations are usually applied to face recognition *before* a holistic algorithm takes place. Note that all such "pre-processing" strategies should have no problem to be applied in any images or any block of images (as long as the size of the block is greater than one pixel). Following the way that holistic algorithms are introduced in most of the publications, we concentrate on introducing the macropixel counting approach and do not discuss pre-processing strategies in this paper. At the time when we compare our algorithm on datasets with other approaches, we use exactly same pre-processing algorithms to reduce illumination effects as those approaches do. When a holistic algorithm is applied to measure the similarity of two images, these two images should first be pre-processed to reduce illumination effects; when two macropixels are to be compared, they should also be pre-processed first. A macropixel of size $K \times K$ can be represented by a K^2 dimensional vector so that Euclidean distances between macropixels can be calculated.

Definition *(Distance)***:** *The distance of two macropixels is defined as the Euclidean distance between these two macropixels after being pre-processed for reducing illumination effects.*

As we mentioned in Section 1, the algorithm we introduce here is based on macropixel comparison – We'll compare "corresponding macropixels". Due to the registration / alignment problem, we shall not restrict ourselves to macropixels with exact same coordinates.

Definition *(Corresponding Macropixel)***:** *For a macropixel M of an image A, assume that the coordinates of its center are (x_o, y_o). To find the corresponding macropixel in image B, we first calculate the Euclidean distances between X and the macropixels with center coordinates $(x_o + s_1, y_o + s_2)$, $-s \le s_1, s_2 \le s$; the one with the shortest distance to M is called the* **corresponding macropixel** *in image B. If there is a tie, then we arbitrary choose one.*

[3] It is worth noting here: Ours is way simpler than any of them; although it seems that most of people do not care about the simplicity/complexity of algorithms.

Note that, above definition indicates that we **shift** a window up to s pixels in each direction in finding a corresponding macropixel. This is why we leave s pixels in each side of the images.

Our algorithm for face identification is as follows:

Macropixel Counting Algorithm for Face Recognition

Input: face image X to be identified.

1. Set a counter for each identity in the gallery.[4]
2. Partition X into blocks of size $K \times K$, leaving s pixels in each side for shifting purpose; a block of $K \times K$ pixels is taken to be one macropixel. (We assume both $m-2s$ and $n-2s$ are dividable by K, where $m \times n$ is the image size.)
3. For each macropixel P in X,
   (a) Find the corresponding macropixel of each image in the gallery and its distance to P.
   (b) Among all macropixels found in (a), find the one with shortest distance to P;
   (c) Assuming the macropixel found in (b) belongs to identity T, we increase its counter by 1. [5]
4. Check all the counters, the one with greatest value is returned as the identity for X. [6]

3 Experimental Results

We assume all image faces are standardized, i.e., cropped, scaled and rotated to same size such that the positions of the eyes are aligned. To be focused on the macropixel counting approach, we do not discuss the algorithms for locating the eyes, which can be found in [18].

To avoid any possible bias, including scaling approach and eye locating approach used for "standardization" above mentioned[7], we decide to use the UIUC versions of Yale, PIE and ORL face sets at http://www.cs.uiuc.edu/homes/dengcai2/Data/data.html[8], with the following reasons:

A. Faces are already standardized according to eye locations (or also nose tips? We don't know and it isn't important here.) – so that when we compare the performances of identification algorithms, we do not need to worry if the standardization approaches are the same.

B. There are two versions, 32×32 pixels and 64×64 pixels, available for each set, so that we can see if our approach works for different sizes.

C. The above UIUC site provides many holistic algorithms, including newly developed ones, in source codes, they also provided best results of many algorithms (5 to 10 algorithms) for 32×32 versions of above data sets.

D. These datasets have been used to evaluate the performances of face recognition algorithms in many recent papers (e.g.,[4, 6, 11, 13, 14, 21, 24]) and many of the algorithms

[4] There may have more than one image per identity in the gallery.

[5] If there's a tie among several identities,we increase each counter by 1.

[6] If there is a tie among q identities for making the final decision, we can randomly choose one of them as the returned ID. We shall only count this as $1/q$ when we evaluate the system performance.

[7] The authors want to claim that, we believe that many have the similar experiences, many times we have difficulties in implementing the face recognition algorithms of others to reach the same performances. Part of the reasons is that it is difficult to standardize the faces with exactly the same way. For example, eye location data might be different, scaling approaches might also be different. Therefore, it is clear that it should be more fair to use published data that have already been standardized in order to compare performances of face recognition approaches.

[8] Also available at http://www.zjucadcg.cn/dengcai/Data/FaceData.html. These two sites also contain the descriptions of the diversities of the datasets.

mentioned in C have been published recently in leading journals and conferences (e.g.,R-LDA [22], LPP[8], OLPP [2], SRDA[4], S-LDA and S-LPP [3]) – therefore, the results on these datasets should be important and sufficient to demonstrate the advantages of our approach. These algorithms to be compared with ours should be able to represent state-of-the-art algorithms.

Following above UIUC website, for each k, denoted by kTrain, k images per individual with labels are used as training images, and also as gallery images, the rest are used for testing. The best results (average error percentages) over 50 random splits of 5 to 10 algorithms for the 32x32 versions of following $k$trains are available in above website: $k = 2,3,4,5,6,7,8$ for ORL and Yale, $k = 5,10,20,30,70,90,110,130$ for PIE. The random 50 splits of each $k$train are also available there.

The common pre-processing strategy for reducing illumination effects for all above reported results is: normalize each image vector to unit. This pre-processing strategy will also be used in our experiments in each macropixel for reducing illumination effects.[9]

3.1 Data sets of 32×32 **pixel images** As it is mentioned early, UIUC website has published many experiments on 32 by 32 version of these data sets; and indeed these algorithms on these data sets has published in many recent top ranked journals and conferences. We conduct the macropixel counting approach on these datasets. Here we let the macropixel size be 4×4 and shift range be 2 pixels. Note that, $32 - 2 \times 2$ (image length/width minus side spaces) is dividable by 4.

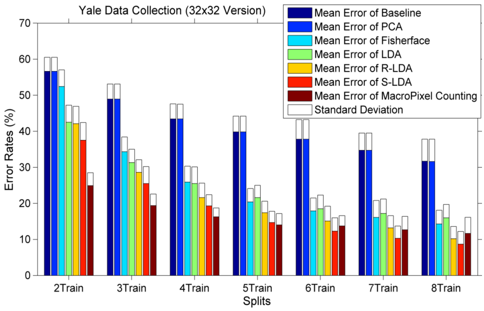
The average errors and standard deviations of our macropixel counting algorithm, together with the results of all those available in above UIUC site[10] are shown in Figures 1(a)-1(c). The results of following approaches (Note that all these approaches except Baseline algorithms are holistic algorithms.) can be found in above UIUC site: LDA(Linear Discriminant analysis); R-LDA(regularized linear discriminant analysis [22]), S-LDA (Spatially Smooth Subspace Learning based Linear Discriminant Analysis [3]), PCA(Principal Component Analysis), Fisherface Approach, LPP(Locality Preserving Projections [8]), R-LPP (Regularized Locality Preserving Projection), S-LPP(Spatially Smooth Subspace Learning based Locality Preserving Projection [3]), OLPP(Orthogonal Locality Preserving Projections [2]), and Baseline (recognition is performed by comparing Euclidean distances between image vectors), although we note that the numbers/sets of algorithms they used for Yale, ORL and PIE are not the same; the best results of all the approaches[11] used by UIUC site owners for each dataset are already shown in Figures 1(a)-1(c).

It is easy to see that, although in a few cases (3 out of 7 cases in Yale experiments, 1 out of 7 cases in ORL experiments an 4 out of 9 in PIE experiments), our approaches are not as good as the best results of the best holistic approaches with best parameters; most of the time our results are indeed the best results. Also, we should note that there are a few parameters for adjustments in many holistic algorithms; the results of all these holistic approaches shown in these figures are already the best ones (obtained by “trial and error", we believe); the results of our macropixel counting approach shown in these figures are obtained

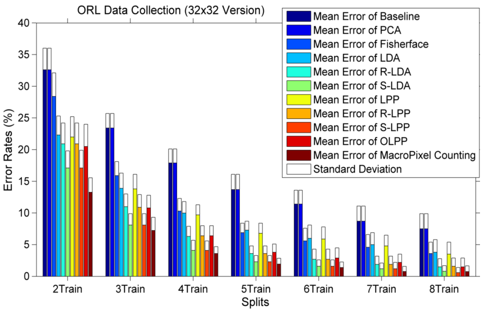
[9]The authors of UIUC site provided a linear scaling approach, i.e. dividing the pixel values of each image by 256, as another pre-processing method for ORL set only; but it is not an approach for reducing illumination effects. The results we report in this paper use only the illumination effect reducing strategy, i.e., normalizing into unit, used in all above three datasets.

[10]Which are actually better than the results in their formal paper publications in many cases, since they adjusted parameters after formal paper publications.

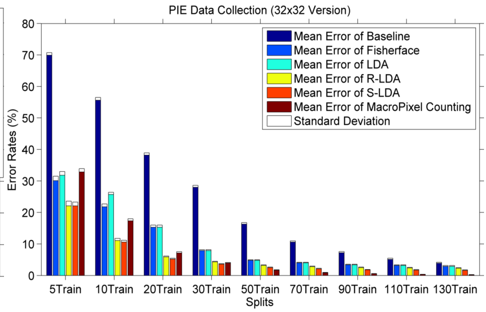
[11]The UIUC website provided results of both NN (nearest neighbour) and NC(nearest center) versions for each algorithm, we show their best ones.



(a) Yale face set (32 by 32 image version)



(b) ORL face set (32 by 32 image version)



(c) PIE face set (32 by 32 image version)

Figure 1: Average Error Recognition Rates and Standard Deviations of A Few Known Holistic Algorithms, Baseline Algorithm and Macropixel Counting Algorithm for 32 by 32 versions of Yale, ORL and PIE

without using any "trial and error" – the parameters (macropixel size and shift range) used are consistent for all these datasets and all these different splits. [12]

As we will see in section 4 and section 5, indeed we may be able to get even better results by adjusting the size of macropixels and shift ranges.

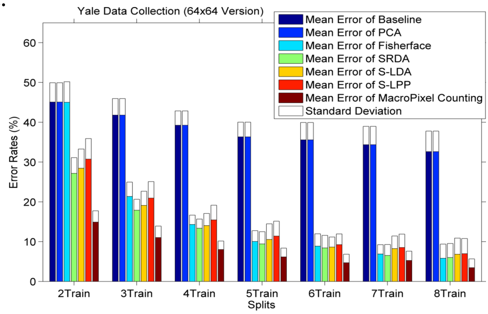
3.2 Data sets of 64×64 **pixel images** The 64×64 versions of these datasets are available in UIUC website, but no experimental results are shown there. We conduct experiments of 6 algorithms for these versions of data sets: Baseline (recognition is performed by comparing Euclidean distances between image vectors), Fisherface, S-LDA(Spatially Smooth Subspace Learning based Linear Discriminant Analysis [3]), PCA(Principal Component Analysis), S-LPP(Spatially Smooth Subspace Learning based Locality Preserving Projection [3]), and SRDA(Spectral Regression Discriminant Analysis [4]). We use UIUC codes directly, except the "Baseline" approach. During the experiments, we follow the suggested values of UIUC implementation for parameter selections if there are suggested ones, otherwise use the default values.[13] The nearest neighbour strategy is applied to find the best matches.

We then conduct experiments using our macropixel counting approach, with exactly the

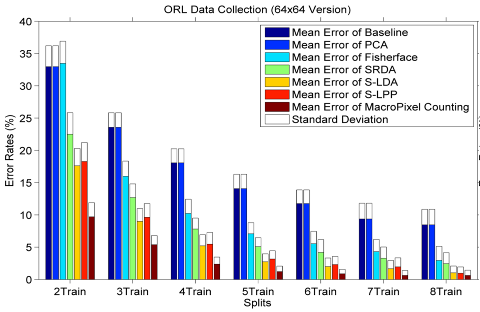
[12] Our intention of this paper (as stated in Section 1) is only to compare our macropixel counting approach with holistic subspace approaches; we do not rule out the possibilities that a complicated local patch approach with proper dataset-dependent empirical parameters might be able to outperform our macropixel counting approach.

[13] As an example, we let the number of reduced dimensions be the suggested value $\min(M,S)-1$ for PCA approach, where M is the total number of total images and $S(=64\times 64)$ is the number of pixels in each image, as it has been known that, when the number of reduced dimensions is $\min(M,S)-1$, the PCA approach always reach best or almost best accuracy.

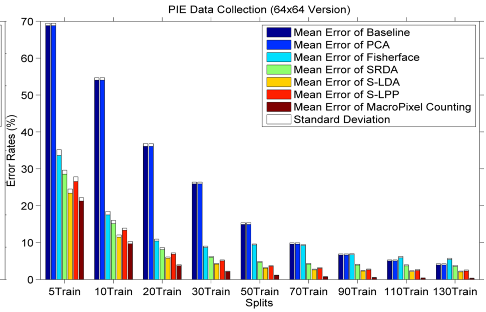
same settings in subsection 3.1, i.e., the macropixel size is 4 by 4 and shift range is 2. Again, note that $64 - 2 \times 2$ (image length/width minus side spaces) is dividable by 4. The results are shown in Figures 2(a)-2(c). This time, the advantage of macropixel counting approach is much more significant.



(a) Yale face set (64 by 64 image version)



(b) ORL face set (64 by 64 image version)



(c) PIE face set (64 by 64 image version)

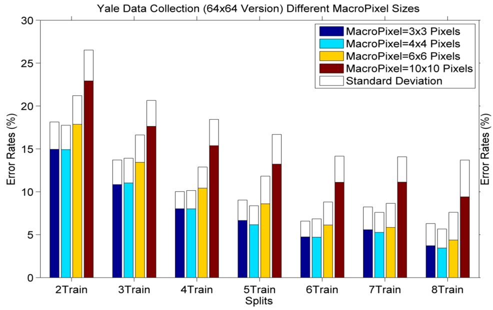
Figure 2: Average Error Recognition Rates and Standard Deviations of 5 Known Holistic Algorithms, Baseline Algorithm and Macropixel Counting Algorithm for 64 by 64 versions of Yale, ORL and PIE

4 Discussion

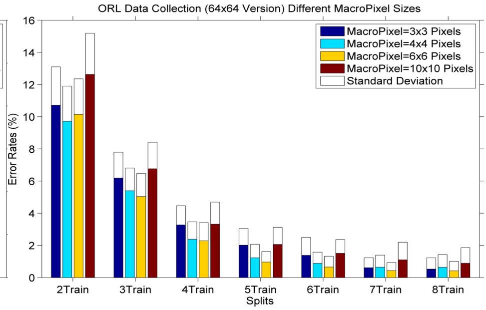
4.1 Possible Further Improvements It is of course possible to define macropixel with different sizes and allow different shift ranges, and it is possible to achieve better results by “trial and error" on the macropixel sizes and shift ranges, although our experiments in Section 3 seem to have shown that the macropixel size of 4×4 and shift range of 2 work quite well to demonstrate that the macropixel counting approach does work as well as, or better than most of holistic algorithms, including newly developed ones.

Our experiments on different macropixel sizes with fixed shift range of 2 are shown in Figures 3(a)-3(c). We also have experiments using different shift ranges (For the experiments with shift range 1 or 3, we remove one pixel in each side of all images so that the image sizes minus the side spaces (i.e., shift range in each side) are divisible by the macropixel size of 4 by 4.) with the fixed macropixel size of 4×4 pixels as shown in Figures 3(d)-3(f). These results indicate that, further improvement by “trial and error" on macropixel sizes and shift ranges is possible – but dataset-dependent.

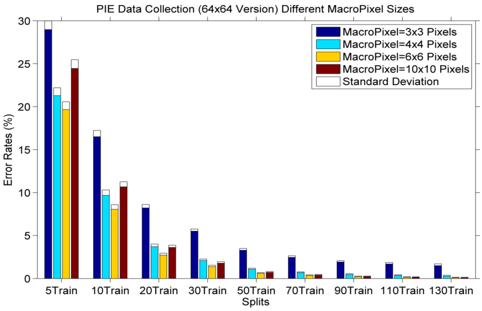
4.2 Summing the distances People may argue that, we here make decision by counting the most close macropixels of each identity. Why don’t we try to use other strategies, such as



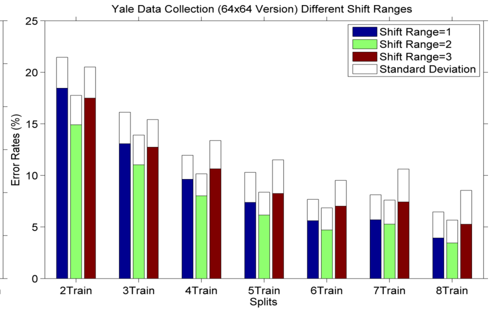
(a) Experiments for Yale face set (64 by 64 image version), with Different MacroPixel Sizes



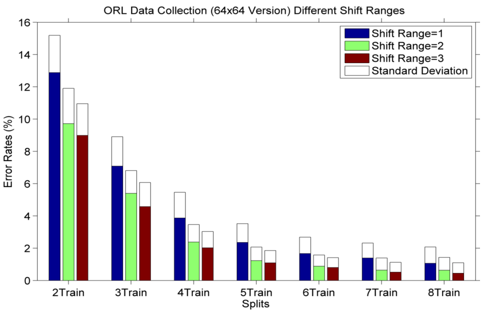
(b) Experiments for ORL face set (64 by 64 image version), with Different MacroPixel Sizes



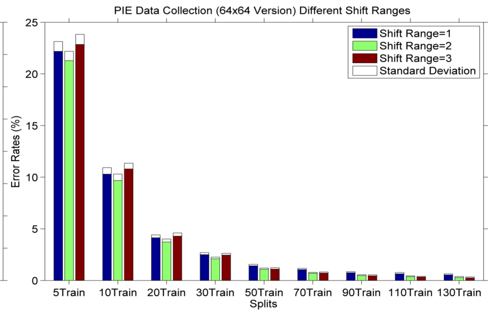
(c) Experiments for PIE face set (64 by 64 image version), with Different MacroPixel Sizes



(d) Experiments for Yale face set (64×64 image version), with Different Shift Ranges



(e) Experiments for ORL face set (64×64 image version), with Different Shift Ranges



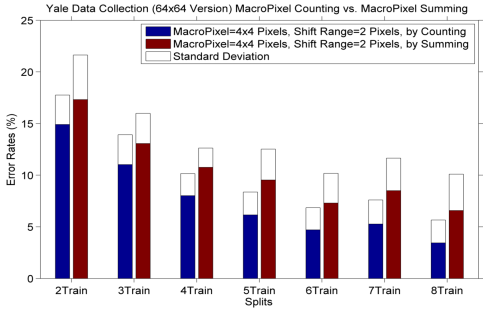
(f) Experiments for PIE face set (64×64 image version), with Different Shift Ranges

Figure 3: Average Error Recognition Rates and Standard Deviations of Macropixel Counting Algorithm with Different MacroPixel Sizes or Different Shift Ranges

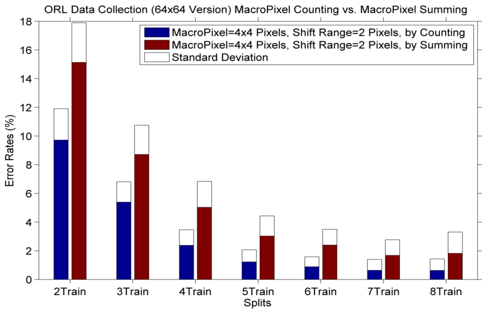
summing up the Euclidean distances? We have done some experiments of which the results are shown in Figures 4(a)-4(c). It seems that, such "summing-up" may still work, but it is unlikely to have significant improvements over the "counting" strategy.

5 Conclusions and Future Work

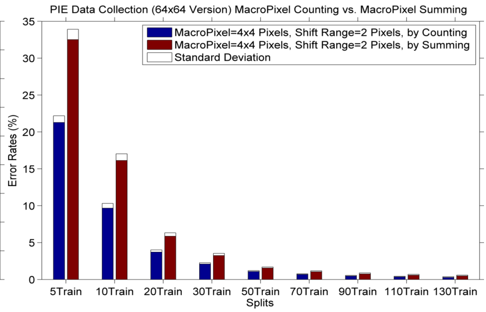
Experiments show that macropixel counting approach can work, at least, as well as recently developed holistic approaches, if not better. Note that, even for holistic approaches, although



(a) Yale face set (64 by 64 image version)



(b) ORL face set (64 by 64 image version)



(c) PIE face set (64 by 64 image version)

Figure 4: Average Error Recognition Rates and Standard Deviations of Macropixel Counting Algorithm and Macropixel based Algorithm with Summing-up Strategy

it is the general understanding that most recently developed approaches have better performances than very old ones, the order of precedence among newly developed algorithms seems to be example dependent. Indeed many times one approach is better than the other in some cases but worse in some other cases.

The success of the macropixel counting approach using small macropixels (i.e., macropixels of size 4×4) also encourages people to reconsider some accepted hypotheses, e.g., the hypothesis that comparison among small areas of face images, which are often near uniform in appearance to our naked eyes, is unlikely to be useful.

The fact that the macropixel counting approach, which has an excellent performance, does not involve any dimensionality reduction seems to suggest (implicitly) that the dimensionality reduction is probably not a necessary process for face recognition.

The superiority of this extremely naïve approach over well established and highly popular ones also stimulates us to rethink: Has the current advance of computer vision research touched the underlying problem in face recognition?

It is clear that, there are different shift ranges and different macropixel sizes (as shown in Subsections 4.1 & 4.2), and there are many different ways of counting and many different distance metrics in summing up the distances of macropixels we can choose. In this paper, we only intend to *point* out that the pairwise macropixel comparison based full space approach *can* work well, i.e., it is at least comparable with most of holistic algorithms, if not better. It is not our intention to show the up-boundary of this approach; therefore, we feel it is not necessary to try many different ways of counting and different ways of summing-up in this paper. In practice, however, we may be able to go through “trial and error" to get a better result – although “trial and error" will be example/dataset-dependent.

The Macropixel approach involves the computations of the distances of all corresponding Macropixels, therefore it requires lots of computing resources. However, it should be noted that the distance computation of each Macropixel pair is indeed very simple in comparing with any holistic/global approach due to the small sizes of Macropixels so that computation costs can be reduced if it is cleverly programed. Taking advantage of the complete independence of the computations of the distances of different Macropixel pairs, parallel computation can be implemented if it is necessary.[14]

Acknowledgements: This research is supported by a Discovery Grant of NSERC, Canada. The author wishes to thank the anonymous reviewers for their helpful comments.

References

[1] Peter N. Belhumeur, Joao P. Hespanha, and David J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Analysis & Machine Ingelligence*, 19(7):711–720, 1997.

[2] D. Cai, X. He, J. Han, and H. Zhang. Orthogonal laplacianfaces for face recognition. *IEEE Trans. Image Processing*, 15(11):3608–3614, Nov. 2006.

[3] D. Cai, X. He, Y. Hu, J. Han, and T. Huang. Learning a spatially smooth subspace for face recognition. In *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, June 2007.

[4] D. Cai, X. He, and J. Han. SRDA: An efficient algorithm for large scale discriminant analysis. *IEEE Trans. Knowledge and Data Engineering*, 20(1):1–12, January 2008.

[5] H. Ekenel and R. Stiefelhagen. Analysis of local appearance-based face recognition: Effects of feature selection and feature normalization. In *Proc. of Computer Vision & Pattern Recognition Workshop*, pages 34–34, June 2006.

[6] Y. Fu and S.J.D. Prince. Investigating the spatial support of signal and noise in face recognition. In *ICCV Subspace Workshop*, 2009.

[7] R. Guttumukkal and V. K. Asari. An improved face recognition technique based on modular pca approach. *Pattern Recgnition Letters*, 24(4):429–436, 2004.

[8] X. He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang. Face recognition using laplacianfaces. *IEEE Trans. Pattern Analysis & Machine Ingelligence*, 27(3):328–340, March 2005.

[9] B. Heisele and T. Koshizen. Components for face recognition. In *Proc. of the Sixth IEEE Int. Conf. on Automatic Face and Gesture Recognition*, pages 153–58, 2004.

[10] A. Hyvärinen, J. Karhunen, and E.Oja. *Independent Component Analysis*. John Wiley & Sons, New York, 2001.

[14] Considering the extreme simplicity for computing the distance of each Macropixel pair, a single thread is to the possible parallel computing structure as a single neuron is to a nervous system.

[11] R. Kumar, A. Banerjee, and B.C. Vemuri. Volterrafaces: Discriminant analysis using volterra kernels. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 150–155, Los Alamitos, CA, USA, 2009. IEEE Computer Society. ISBN 978-1-4244-3992-8.

[12] A.M. Martinez and A.C. Kak. Pca versus lda. *IEEE Trans. Pattern Analysis & Machine Intelligence*, 23(2):228–233, 2001.

[13] D.-S. Pham and S. Venkatesh. Robust learning of discriminative projection for multi-category classification on the stiefel manifold. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1 –7, June 2008.

[14] H. Shan and G. W. Cottrell. Looking around the backyard helps to recognize faces and digits. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, Los Alamitos, CA, USA, 2008.

[15] Y. Su, S. Shan, X. Chen, and W. Gao. Hierarchical ensemble of global and local classifiers for face recognition. *IEEE Trans. Image Process*, 18(8):1885–96, 2009.

[16] K. Tan and S. Chen. Adaptively weighted sub-pattern pca for face recognition. *Neurocomputing*, 64:505–511, 2005.

[17] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

[18] P. Wang, M.B. Green, Q. Ji, and J. Wayman. Automatic eye detection and its validation. In *Proc. of 2005 IEEE Conf. on Computer Vision and Pattern Recognition -Workshops*, volume 3, page 164, June 20-26 2005.

[19] X. Wang and X. Tang. A unified framework for subspace face recognition. *IEEE Trans. Pattern Analysis & Machine Intelligence*, 26(9):1222–1228, 2004.

[20] X. Wang and X. Tang. Random sampling for subspace face recognition. *Int. Journal of Computer Vision*, 70(1):91–104, Oct. 2006.

[21] Z. Wang and Z. Miao. Scale invariant face recognition using probabilistic similarity measure. In *International Conference of Pattern Recognition*, 2008.

[22] J. Ye, T. Xiong, Qi Li, R. Janardan, J. Bi, V. Cherkassky, and C. Kambhamettu. Efficient model selection for regularized linear discriminant analysis. In *Proc. of the 15th ACM international conference on Information and knowledge management*, pages 532–539, Arlington,Virginia, 2006.

[23] W. Zhao and R. Chellappa. A guided tour of face processing. In Wenyi Zhao and Rama Chellappa, editors, *Face Processing: Advanced Modeling and Methods*, pages 3–54. Academic Press, 2006.

[24] W. Zuo, H. Zhang, D. Xhang, and K. Wang. Post-processed LDA for face and palmprint recognition: What is the rationale. *Signal Processing*, 90(8):2344-2352, Aug. 2010.