ANALYZING ASYMMETRY BIOMETRIC IN THE FREQUENCY DOMAIN FOR FACE RECOGNITION

*Sinjini Mitra**

Department of Statistics
Carnegie Mellon University
Pittsburgh, PA 15213

Marios Savvides

Electrical and Computer Engineering Department
Carnegie Mellon University
Pittsburgh, PA 15213

ABSTRACT

The present paper introduces a novel set of facial biometrics based on quantified facial asymmetry measures in the frequency domain. In particular, we show that these biometrics work well for images showing expression variations. A comparison of the recognition rates with those obtained from spatial domain asymmetry measures based on raw intensity values suggests that the frequency domain representation is more robust to intra-personal distortions and indeed provides an efficient approach for performing classification or recognition. The role of asymmetry of the different regions (e.g., eyes, mouth, nose) of the face is investigated to determine which regions provide the maximum discrimination among individuals in the presence of different expressions for better classification results in such a scenario.

1. INTRODUCTION

Human faces have two kinds of asymmetry - intrinsic and extrinsic. The former is caused by growth, injury and age-related changes, while the latter is affected by viewing orientation and lighting direction. We are however interested in intrinsic asymmetry which is directly related to the individual face structure while extrinsic asymmetry can be controlled to a large extent. Psychologists have long been interested in the relationship between facial asymmetry and attractiveness and its role in identification. [1] observed that the more asymmetric a face, the less attractive it is. Furthermore, the less attractive a face is, the more recognizable it is ([2]). All these studies indicate the potential significance of asymmetry in human face-identification problems.

A commonly accepted notion in computer vision is that human faces are bilaterally symmetric ([3]) and [4] reported no differences whatsoever in recognition rates while using only the right and left halves of the face. However, a well-known fact is that manifesting expressions cause a considerable amount of facial asymmetry, they being more intense on the left side of the face ([5]). Indeed [6] found differences in recognition rates for the two halves of the face under a given facial expression. Craniofacial research generally describes that asymmetries exist mostly in the middle and lower third of the face ([7]).

Despite extensive studies on facial asymmetry, its use in human identification started in the computer vision community only in 2001 with the seminal work by Liu ([8]), who for the first time showed that certain facial asymmetry measures are efficient human identification tools under expression variations. This was followed by more in-depth studies ([9], [10]) which further investigated the role and locations of different types of asymmetry measures both for human as well as expression classifications. But no work has been done on developing asymmetry measures in the frequency domain as per our knowledge, and their use as a biometric for face identification.

The paper is organized as follows. Section 2 describes the dataset used. Section 3 introduces the new asymmetry measures in the frequency domain and Section 4 contains the classification results and a feature set analysis. Finally, a discussion appears in Section 5.

2. DATA

The dataset used is a part of the "Cohn-Kanade AU-coded Facial Expression Database" ([11]), consisting of images of 55 individuals expressing three different kinds of emotions - joy, anger and disgust. Each person was asked to express one emotion at a time by starting with a neutral expression and gradually evolving into its peak form. The data consists of video clips of people showing an emotion, each clip being broken down into several frames. The raw images are normalized using an affine transformation, the details are included in [9]. Some normalized images from our database are shown in Figure 1. This database is the only known one for studying facial asymmetry in the presence of severe expression variations. We chose this subset as our initial test-bed but hope to extend to larger databases in future.

*The research is funded in part by the Army Research Office contract DAAD19-02-1-3-0389 to CyLab.

Fig. 1. Sample images from our database. Courtesy [9]

We use a total of 495 frames, which include 3 frames from each emotion for each subject ($55 \times 3 \times 3$). These are chosen from the most neutral (the beginning frame), the most peak (the final frame) and a middle frame in the entire expression sequence.

3. THE FREQUENCY DOMAIN

Many signal processing applications in computer engineering involve the frequency-domain representation of signals. The frequency spectrum consists of two components, the *magnitude* and *phase*. In 2D images particularly, the phase component captures more of the image intelligibility than magnitude and hence is very significant for performing image reconstruction ([12]). [13] showed that correlation filters built in the frequency domain can be used for efficient face-based recognition. Recently, the significance of phase has also been used in biometric authentication. [14] proposed correlation filters based only on the phase component of an image, which performed as well as the original filters. Later [15] demonstrated that performing PCA in the frequency domain by eliminating the magnitude spectrum and retaining only the phase not only outperformed spatial domain PCA, but also have attractive features such as illumination tolerance, can handle partial occlusions. All these point out the benefits of considering classification features in the frequency domain for potentially improved results.

Symmetry properties of the Fourier transform are often very useful ([16]). Any sequence $x(n)$ can be expressed as a sum of a *symmetric* part $x_e(n)$ and an *asymmetric* part $x_o(n)$. Specifically,

$$x(n) \; = \; x_e(n) + x_o(n),$$

where $x_e(n) = \frac{1}{2}(x(n) + x(-n))$ and $x_o(n) = \frac{1}{2}(x(n) - x(-n))$. When a Fourier transform is performed on a real sequence $x(n)$, the even part ($x_e(n)$) transforms to the real part of the Fourier transform and the odd part ($x_o(n)$) transforms to its imaginary part (Fourier transform of any sequence is generally complex-valued). The Fourier transform of a real and even sequence is thus real; that of a real and odd sequence is purely imaginary. Now, since phase is defined as $\theta = \tan^{-1}\left(\frac{I}{R}\right)$, it will be zero in case the imaginary component is zero. In other words, a symmetric sequence gives rise to zero-phase frequency spectrum. These observations therefore imply that the imaginary component of the Fourier transform (1D Fourier transform slices of the face) can be considered as a measure of facial asymmetry in the frequency domain, and also establish a nice relationship between facial asymmetry and the phase component of the frequency domain. Given the role played by both phase and asymmetry in face-based recognition, this presents an opportunity to exploit this correspondence for the development of more refined classification tools.

3.1. The Asymmetry Biometric

Following the notion presented in the earlier section, an obvious choice for the asymmetry measures in the frequency domain seems to be some function of the imaginary part of the Fourier transform. One such metric can be defined as the *energy* of these imaginary components. For a sequence $a_k + ib_k, \; k = 1, \ldots, N$, the energy of the imaginary part of the entire sequence is given by

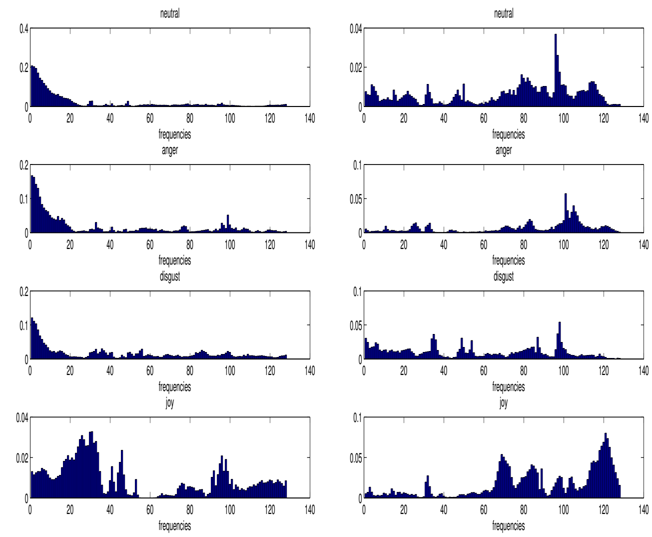
$$e_N = \sum_{k=1}^{N} b_k^2.$$

The lower the value of e_N, the less the amount of asymmetry (and hence more symmetry) and vice versa.

We considered each row of an image as a sequence and computed the energy of the imaginary components of each row as its measure of asymmetry. Figure 2 shows how this measure of asymmetry varies among the different expressions of the different people. For instance, for person 1, joy produces the greatest degree of asymmetry, and neutral expression the lowest, whereas, for person 2, joy and neutral expressions show maximum asymmetry followed by anger and disgust. On comparing the two people, we observe that overall, person 1 has more facial asymmetry than person 2 for the three emotions, but less asymmetry for the neutral expression. These are only exploratory analysis but they give a preliminary idea that these measures may be helpful in recognizing people in the presence of expression variations. This hence constitutes a work parallel to that of [9], using a frequency domain representation instead.

4. RESULTS

For the images in our database, the identification features are computed as follows. A Fourier transform is performed on each row of a face image, and the energy of the imaginary components for each row computed. The entire face image is then divided into blocks of two rows and the energies over each block of two rows averaged. Thus, the total



(a) Person 1 (b) Person 2

Fig. 2. Asymmetry of the different facial features for the four expressions of two persons. The horizontal axis shows different frequencies at which the features are computed.

number of features we have for each image is half of that of the total number of rows in that image. Since our images are of dimension 128×128, this means that our feature vectors are each of length 64. The averaging over rows is done in order to smooth out noise in the image which can possibly create artificial asymmetry artifacts and give misleading results. Averaging over more rows, on the other hand, can lead to over-smoothing and a loss of relevant asymmetry information. So, we selected blocks of two rows as the optimal after some experimentation.

We tried a few different classification methods, and the best results were obtained with the individual PCA approach (or, IPCA for short) [14]. For each person p, subspaces W_p are computed and each test image is projected onto each individual subspace using $y_p = W_p^T(x - m_p)$. The image is then reconstructed as $x_p = W_p y_p + m_p$ and the reconstruction error computed $||e_p||^2 = ||x - x_p||^2$. The final classification chooses the subspace with the smallest $||e_p||^2$.

The training was done on the neutral frames of the 3 emotions of joy, anger and disgust from all the 55 individuals in the dataset and testing on the peak frame of the 3 emotions from all the people. We compare the results with those reported in [9], which uses a simplistic measure of facial asymmetry in the spatial domain called *D-face*. The results show that our proposed frequency domain measures are more effective for expression-invariant human identification, resulting in a 11.22% absolute improvement and a 63.8% relative improvement over D-face.

We also studied the discriminative power of these asymmetry measures to determine which parts of the face actually contributes to the identification process. Ideally, those features which contribute to inter-class differences should have large variation between subjects and small variation within the same subject. Hence, a measure of discrimination can be provided by a variance ratio type quantity, in particular, we use what is known as an *Augmented Variance Ratio* or AVR, which was also used by ([9]). AVR compares within class and between class variances and at the same time penalizes features whose class means are too close to one another. For a feature F with values S_F in a data set with C total classes, AVR is calculated as

Asymmetry features	Misclassification rates
Spatial D-face	17.58%
Frequency-based	6.36%

Table 1. Misclassification rates using asymmetry measures.

$$AVR(S_F) = \frac{Var(S_F)}{\frac{1}{C}\sum_{k=1}^{C} \frac{Var_k(S_F)}{min_{j \neq k}(|mean_k(S_F) - mean_j(S_F)|)}},$$

where $mean_i(S_F)$ is the mean of the subset of values from feature F belonging to class i. AVR thus imposes a penalty on features which may have small intra-class variance but close inter-class mean values. The higher the AVR value of a feature, the more discriminative it is for classification. For our problem, the individual subjects form the classes ($C = 55$).

Figure 4(a) shows the AVR values for all the features, from which it an be observed that the forehead region has much higher values than the rest of the face. Looking at this more carefully, we find that a very few of the subjects in the the database have some artificial asymmetry in that region arising from either falling hair or edge artifacts introduced in the normalization procedure. Two such images are shown in Figure 3. This is highly undesirable and causes spurious results by masking the actual asymmetry of that facial region. We thus removed the top 3 features and the new AVR plot appears in Figure 4(b). It is clearly evident from this

(a) Hair (b) Edge artifacts

Fig. 3. Images with artificial asymmetry in the forehead.

that the features around the nose bridge contain the most discriminative information pertaining to recognition of individuals based on facial asymmetry under different expressions, followed by those around the chin. Moreover, this is consistent with results in [9], which noted that nose bridge is the most discriminating facial region for similar recognition tasks based on D-face. This establishes a firm basis for the confidence that these features are the most relevant and useful, and may yield fewer misclassifications that any other features.

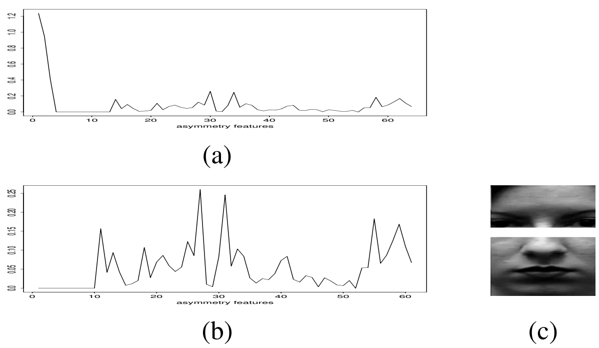


Fig. 4. AVR values for our frequency-based asymmetry measures for (a) all features, (b) all features except the top three. (c) The white strip on the image denotes the regions corresponding to the two peaks in (b). The features $0 - 64$ represent the regions from the forehead to the chin of a face.

5. DISCUSSION

We have thus shown in this paper that facial asymmetry measures in the frequency domain offer a promising potential as an useful biometric in practice, especially, in the presence of expression variations in face images. An error rate as low as 6.36% is very impressive and desirable indeed given that the test images are very different from the training ones. This in turn is very important for recognition routines in practice, for example, in biometric identification applications since surveillance photos captured at airports and other public places is expected to be quite diverse with respect to the expressions of an individual's face. Hence any algorithm that can deal with such variations is supposed to be attractive to users. Thus far, we have only used very simple classification methods and despite this, we have obtained impressive results over previous assymetry measures. Moreover, the fact that these measures vary considerably with the different expressions can be exploited to perform expression classifications as well, and a comparison of the features for these apparently conflicting classification goals (expression-invariant human identification and expression classification) may be interesting, as done in [10]. The next direction of research will consist of evaluating the frequency domain measures for identification in the presence of illumination variations in face images. The advantage of working in the frequency domain as opposed to that in the spatial or the image domain is that it is much easier to adjust for these common distortions that occur frequently in images.

6. REFERENCES

[1] Thornhill, R. and Gangstad, S. W., "Facial attractiveness," *Transactions in Cognitive Sciences*, vol. 3, no. 12, pp. 452–460, 1999.

[2] Troje, N. F. and Buelthoff, H. H., "How is bilateral symmetry of human faces used for recognition of novel views?," *Vision Research*, vol. 38, no. 1, pp. 79–89, 1998.

[3] Seitz, S.M. and Dyer, C.R., "View morphing," *SIGGRAPH*, pp. 21–30, 1996.

[4] Gutta, S., Philomin, V., and Trajkovic, M., "An investigation into the use of partial-faces for face recognition," in *International Conference on Automatic Face and Gesture Recognition*, Washington D.C., 2002, pp. 33–38.

[5] Borod, J.D., Koff, E., Yecker, S., Santschi, C., and Schmidt, J.M., "Facial asymmetry during emotional expression: gender, valence and measurement technique," *Psychophysiology*, vol. 36, no. 11, pp. 1209–1215, 1998.

[6] Martinez, A.M., "Recognizing imprecisely localized, partially occluded and expression variant faces from a single sample per class," *PAMI*, vol. 24, no. 6, pp. 748–763, 2002.

[7] Vig, P.S. and Hewitt, A.B., "Asymmetry of the human facial skeleton," *The Angle Orthodontist*, vol. 45, pp. 125–129, 1975.

[8] Liu, Y., Schmidt, K., Cohn, J., and Weaver, R.L., "Human facial asymmetry for expression-invariant facial identification," in *Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, 2002.

[9] Liu, Y., Schmidt, K., Cohn, J., and Mitra, S., "Facial asymmetry quantification for expression-invariant human identification," *CVIU*, vol. 91, no. 1/2, pp. 138–159, 2003.

[10] Mitra, S. and Liu, Y., "Local facial asymmetry for expression classification," in *Proceedings of CVPR*, 2004.

[11] Kanade, T., Cohn, J.F., and Tian, Y.L., "Comprehensive database for facial expression analysis," in *4th IEEE International Conference on Automatic Face and Gesture Recognition*, 2000, pp. 46–53.

[12] Hayes, M.H., "The reconstruction of a multidimensional sequence from the phase or magnitude of its fourier transform," *ASSP*, vol. 30, no. 2, pp. 140–154, 1982.

[13] Savvides, M., Vijaya Kumar, B.V.K., and Khosla, P., "Face verification using correlation filters," in *3rd IEEE Automatic Identification Advanced Technologies*, Tarrytown, NY, 2002, pp. 56–61.

[14] Savvides, M. and Kumar, B.V.K., "Eigenphases vs.eigenfaces," in *IEEE International Conference in Pattern Recognition*, 2004.

[15] Savvides, M., Kumar, B.V.K., and Khosla, P.K., "Corefaces - robust shift invariant pca based correlation filter for illumination tolerant fac4e recognition," in *Proceedings of CVPR*, 2004.

[16] Oppenheim, A.V. and Schafer, R.W., *Discrete-time Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1989.