INVESTIGATING THE FEASIBILITY OF IMAGE-BASED NOSE BIOMETRICS

Niv Zehngut, Felix Juefei-Xu, Rishabh Bardia, Dipan K. Pal, Chandrasekhar Bhagavatula, M. Savvides

Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA

ABSTRACT

The search for new biometrics is never ending. In this work, we investigate the use of image based nasal features as a biometric. In many real-world recognition scenarios, partial occlusions on the face leave the nose region visible (*e.g.* sunglasses). Face recognition systems often fail or perform poorly in such settings. Furthermore, the nose region naturally contain more invariance to expression than features extracted from other parts of the face. In this study, we extract discriminative nasal features using Kernel Class-Dependence Feature Analysis (KCFA) based on Optimal Trade-off Synthetic Discriminant Function (OTSDF) filters. We evaluate this technique on the FRGC ver2.0 database and the AR Face database, training and testing exclusively on nasal features and have compared the results to the full face recognition using KCFA features. We find that the between-subject discriminability in nasal features is comparable to that found in facial features. This shows that nose biometrics have a potential to support and boost biometric identification, that has largely been under utilized. Moreover, our extracted KCFA nose features have significantly outperformed the PittPatt face matcher which works with the original JPEG images on the AR facial occlusion database. This shows that nose biometrics can be used as a stand-alone biometric trait when the subjects are under occlusions.

***Index Terms*—** Nose Biometrics

1. INTRODUCTION

Biometric studies over the years have explored many different characteristics that help distinguish one person from another. Some previously explored biometrics include fingerprints, faces [1], iris, hand geometry, DNA, palm print, voice, and gait [2]. However, out of these a few have emerged more practically useful than others. A few characteristics that a good biometric has is discriminability, robustness, and ease of extraction. Although biometric systems based on iris and fingerprint recognition have proved themselves to be very discriminative and robust, special acquisition systems required in both cases raise costs and reduces their implementation

Fig. 1. Example images from the Internet and the AR Face database showing individuals wearing sunglasses and with non-neutral expressions. In such cases, nose biometric recognition can be a good alternate for the full face. In addition, the nose region is more stable under various facial expressions.

flexibility. Image based face recognition has proved to be a cheaper and more flexible biometric to acquire. However, sophisticated algorithms are needed to deal with the large number of nuisance parameters that are introduced such as pose, illumination, expressions, and age variations. Nonetheless, a lot of effort has led to significant improvements in unconstrained face recognition capability [3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. However, these algorithms typically assume access to the entire face. This assumption is violated in many real-world scenarios as illustrated in Figure 1. In such cases, techniques which rely on component based recognition achieve more success. Image based nose biometrics has not yet been studied extensively in the biometrics community. Song *et al.* [13] studied extracting the quantity and positions of the pores on the nose skin as a biometric while they were able to achieve good results. The effort needed to get images in which the pores could be extracted made it somewhat infeasible. Image based nose biometrics intrinsically have some advantages over standard full face recognition such as expression invariance and higher likelihood of availability.

In this work, we present a preliminary feasibility study of image based nose biometrics. The nose is arguably the part of the face which contains the most depth information. However, in this study we refrain from using that depth information and explore how much of the identity information is
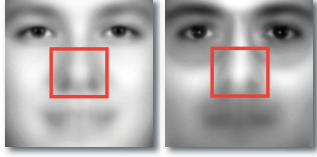
Fig. 2. Mean faces from FRGC (left) and AR (right) database with the orange box indicating the cropped nose region.

available to linear methods for exploitation. This is a harder setting since depth information is also a biometric available which we choose not to utilize. Thus, we work with frontal images in this study. Nonetheless, we find in our experiments that frontal images of noses does contain a lot of subject based identity information. There are some other works that deal with component-based face recognition [14, 15, 16, 17, 18]. Tome *et al.* [16] take one identification scheme and compare different distances and regions. It answers the question 'For this specific scheme which regions one should consider'. But we dive into the specific examination of one region, and answer the question 'How good is identification based only on this region', and 'how does it perform in different methods'. Bonnen *et al.* [14] requires per-component alignment using Procrustes analysis which minimizes the MSE between two ordered sets of coordinates, after the landmarks are obtained by ASM. Such a procedure is computationally expensive. In that sense, our contribution is still very unique.

2. DATA AND PREPROCESSING

2.1. Large-scale Face Database

The NIST's Face Recognition Grand Challenge (FRGC) ver2.0 database [19] is the largest face database that is publicly available. It has three components: the generic **training**, the **target**, and the **probe** set. The training set contains 12,776 facial images from 222 subjects, with each subject under both controlled and uncontrolled environment. The target set contains 16,028 images from 466 subjects, under controlled environment. The probe set contains 8,014 images from the same 466 subjects as in the target set, but under uncontrolled environment. There are 153 overlapping subjects in training set and target/probe set. We follow the FRGC matching protocols.

2.2. Occluded Face Database

The AR Face database [20] is one of the most widely used face databases with occlusions. It contains 3,288 color images from 135 subjects (76 male subjects + 59 female subjects). Typical occlusions include scarves and sunglasses. The database also captures expression variations and lighting changes. This database will be used for our Experiment I where we deal with faces with occlusions and study whether

nose biometrics has an advantage over full face biometrics under such scenarios. The bottom row in Figure 1 shows some samples images with occlusions from the AR Face database.

2.3. Nose Region Extraction

For both the FRGC and AR database, we register the faces based on the eye locations and crop out the square face images. Figure 2 shows the mean face images from the two databases. We then determine the nose region which is shown in the orange box in Figure 2 for cropping out nose images for subsequent training/testing experiments. The nose region is about 15% of the entire face region.

3. FEATURE EXTRACTION AND CLASSIFICATION

3.1. Correlation Filter Based Classification

Correlation filters have a successful history in biometric pattern recognition [21]. One of the most successful correlation filter based approaches is Kernel Class-Dependence Feature Analysis (KCFA) [21]. The KCFA framework is built on Optimal Tradeoff Synthetic Discriminant Function (OTSDF) filters [22]. We use the same approach to show the feasibility of recognition of the nose regions.

3.1.1. OTSDF Filters

Correlation filters are classifiers which simultaneously localizes and recognizes an object in an image. This is accomplished by correlating the filter with the image and inspecting the resulting correlation surface for some desired pattern or property. OTSDF filters are formulated as a tradeoff between Minimum Average Correlation Energy (MACE) filters [23] and another type of filter, the Minimum Variance Synthetic Discriminant Function (MVSDF) filter [24]. MVSDF filters aim to create consistent output correlation surfaces in the prescence of some known noise. By trading off between the peak sharpness of MACE filters and the noise tolerance of MVSDF filters, OTSDF filters tend to outperform both. MACE filters try to achieve as sharp peaks as possible in order to reduce the chances of error. This is accomplished by minimizing the average energy over all the correlation surfaces using the training data. By taking advantage of the fact that correlation in the space domain can be accomplished by an element wise multiplication and conjugation in the frequency domain as well as Parseval's theorem, which states that energy is proportional in these two domains, the average correlation energy (ACE) can be formulated as

$$ACE = \mathbf{h}^+ \left[\frac{1}{d \cdot N} \sum_{i=1}^{N} \mathbf{X}_i \mathbf{X}_i^* \right] \mathbf{h} \qquad (1)$$

Here, $\mathbf{h}$ represents the Fourier domain MACE filter we are trying to find, d is the dimensionality of your training data

o, in our case, the number of pixels in the image, N is the number of training samples, and $\mathbf{X}_i$ is a dxd diagonal matrix conataining the ith training sample in the Fourier domain along the diagonal. We often rewrite this as

$$ACE = \mathbf{h}^+\mathbf{D}\mathbf{h} \tag{2}$$

where $\mathbf{D} = \frac{1}{d*N} \sum_{i=1}^{N} \mathbf{X}_i\mathbf{X}_i^*$. Since just minimizing the average correlation energy will not lead to sharp peaks, we also add in a constraint at the origin of correlation surface. It can be shown that the value at the origin of the correlation surface can be expressed in the frequency domain as

$$\mathbf{X}^+\mathbf{h} = d \cdot \mathbf{u} \tag{3}$$

where $\mathbf{X}$ in this equation is the matrix containing the Fourier domain representations of all the training samples along the columns of the matrix and $\mathbf{u}$ is a vector of desired peak values. Usually this is set to a vector of ones. By minimizing Eq. 2 subject to the constraints in Eq. 3, it can be shown that

$$\mathbf{h} = \mathbf{D}^{-1}\mathbf{X}(\mathbf{X}^+\mathbf{D}^{-1}\mathbf{X})^{-1}\mathbf{u} \tag{4}$$

MACE filters tend not to perform well under noise degradation however. MVSDF filters follow a similar derivation and take a similar form of

$$\mathbf{h} = \mathbf{C}^{-1}\mathbf{X}(\mathbf{X}^+\mathbf{C}^{-1}\mathbf{X})^{-1}\mathbf{u} \tag{5}$$

where $\mathbf{C}$ is the power spectrum of the noise along the diagonal. OTSDF filters end up taking the form of

$$\mathbf{h} = \mathbf{T}^{-1}\mathbf{X}(\mathbf{X}^+\mathbf{T}^{-1}\mathbf{X})^{-1}\mathbf{u} \tag{6}$$

where $\mathbf{T} = (\alpha\mathbf{C} + \sqrt{1 - \alpha^2}\mathbf{D})^{-1}$ and In most cases, we assume Gaussian white noise which means $\mathbf{C} = \mathbf{I}$.

3.1.2. KCFA Framework

Redundant Class-Dependence Feature Analysis (CFA) is a technique built upon these correlation filters originally used for face matching on the FRGC dataset [25]. At its heart, CFA uses these correlation filters as sets of linear feature extractors in order to generate new feature vectors for classification. Since the filters are designed to be discriminative, they turn out to work well in this regard. However, this technique has been improved upon by incorporating kernel techniques into it. All the feature extraction ends up being done through inner products which can then be replaced with a kernelized version leading to the KCFA framework [26, 27, 28]. As with other kernel techniques, this allows us to extract higher dimensional features without needing to store those dimensions leading to nonlinear feature extractors in the original dimensionality.

In our experiments we train our KCFA framework on the FRGC generic training set following the FRGC training protocol. One kernel correlation filter was trained for each of the 222 subjects in the generic training set using a Gaussian RBF kernel. This model was used in experiments involving the FRGC database and the AR Face database.

4. EXPERIMENTS

In this section, we detail our major experiments: (1) a comparison between nose and face biometrics under large-scale setting, and (2) a comparison between nose and face biometrics for occluded faces. In each of the following subsections, the experimental setup will be detailed followed by the experimental results and discussions.

4.1. Experiment I: Large-scale Nose Biometrics vs. Face Biometrics

In this experiment, we carry out the large-scale verification experiments by following the FRGC matching protocols. The FRGC Experiment 1 protocol matches the entire FRGC target set to itself, involving over 256 million face match comparisons. Facial images in the target set are taken under controlled environment with good lighting conditions. A more challenging matching protocol is the FRGC Experiment 4 protocol, where the entire unconstrained probe set is matched against the entire constrained target set, which involves over 128 million face match comparisons. We compare the nose crop with the full face following the same protocols. The training is carried out solely on the training set using the KCFA method [21] on discrete transform encoded local binary patterns (DT-LBP) features whose effectiveness has been previously observed [29, 30, 31, 32, 33, 34].

The experimental results are consolidated in Table 1 where the verification rates (VR) at 0.1% and 1% false accept rate (FAR), as well as the equal error rates (EER) are reported. The ROC curves are shown in Figure 3. In this experiment, the nose crop is only about 15% of the size of the full face, and yet it is doing a very good job as compared to the full face. The performance of the two is even closer to each other when it is under the constrained matching protocol.

4.2. Experiment II: Nose Biometrics vs. Face Biometrics for Occluded Faces

In this experiment, we work with facial occlusion database: the AR Face database. As a benchmark, we first run PittPatt (SDK 5.2.2) on the original JPEG images and obtain the matching scores. The protocol we follow in this experiment is straight forward: the entire AR database is matched against itself, resulting in a similarity matrix of size $3,288 \times 3,288$, ideally resulting in a block-diagonal matrix.

Next, we extract the KCFA features for the nose crop of every image in the AR database. It is worth noting that the training is done using the generic training set from the FRGC database as discussed in our Experiment I, and once the subspace is learned, we project all 3,288 images from the AR database onto the subspace and obtain the final KCFA features. Not a single AR image is observed during training.

We have shown the experimental results in Table 2, where VR at 0.1% and 1% FAR and EER are reported. The ROC
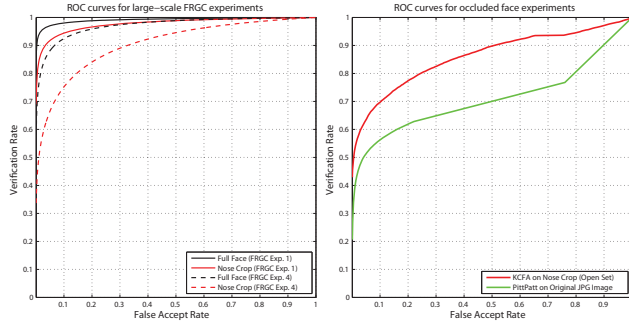
Fig. 3. (L) ROC curves for Experiment I: large-scale FRGC experiments. The solid lines are FRGC Experiment 1, and the dashed lines are FRGC Experiment 4. (R) ROC curves for Experiment II: occluded face matching experiment.

Table 1. VR at 0.1% and 1% FAR, and EER for Exp. I.

	VR at 0.1% FAR	VR at 1% FAR	EER
Exp.1 Full Face	0.8638	0.9381	0.035
Exp.1 Nose Crop	0.7024	0.8414	0.070
Exp.4 Full Face	0.5929	0.7623	0.081
Exp.4 Nose Crop	0.3452	0.4981	0.174

curves are shown in Figure 3. We can see that our nose biometrics with KCFA feature can achieve a high 43.5% verification rate at 0.1% FAR on such a challenging occlusion database, which significantly outperforms the PittPatt SDK 5.2.2 face matcher. With real-world occlusions, we have shown that the nose biometrics is significantly more feasible and robust than traditional full face biometrics.

This protocol is more challenging than the one that allows same subjects being observed and trained on during the training stage. We would expect further performance improvement if trained on the same database.

5. CONCLUSIONS

In this work we present, to the best of our knowledge, the first feasibility study which examines image based nose biometrics. Nose images have some advantages over full face images such as more tolerance to expressions and availability under certain facial occlusions such as sunglasses. Our experiments show that image based nasal features contain enough discriminative information which can be capitalized on by nonlinear techniques, such as advanced kernel correlation filters. We

Table 2. VR at 0.1% and 1% FAR, and EER for Exp. II.

	VR at 0.1% FAR	VR at 1% FAR	EER
PittPatt on Original JPG Img	0.2221	0.3934	0.334
KCFA on Nose Crop (Open Set)	0.4350	0.5297	0.211

find that nasal features have enough biometric information to perform competitively to full face recognition with all methods that we explore. The efficacy of our extracted nose biometric feature is benchmarked against the full face biometric under a large scale verification setting by following the FRGC Experiment 1 and Experiment 4 matching protocols. Moreover, when tested on the AR facial occlusion database, our KCFA features extracted from the nose region significantly outperforms one of the best commercial face matchers which utilizes the whole JPEG image.

Overall, nose biometrics prove themselves to be worthy of further investigation. Since most occluded faces in the real-world can be argued to have either the periocular region or the nose accessible, powerful combinations of periocular and nose biometrics would be interesting to explore. Such models might be flexible yet powerful enough to handle highly occluded faces robustly in unconstrained scenarios. Another interesting thrust would be to develop models which capitalize on the biometric information encoded in the 3-D structure of the nose.

6. REFERENCES

[1] F. Juefei-Xu, D. K. Pal, K. Singh, and M. Savvides, "A Preliminary Investigation on the Sensitivity of COTS Face Recognition Systems to Forensic Analyst-style Face Processing for Occlusions," in *CVPRW*, June 2015.

[2] F. Juefei-Xu, C. Bhagavatula, A. Jaech, U. Prasad, and M. Savvides, "Gait-ID on the Move: Pace Independent Human Identification Using Cell Phone Accelerometer Dynamics," in *BTAS*, Sept 2012, pp. 8–15.

[3] T. Kanade, "Picture processing system by computer complex and recognition of human faces," in *doctoral dissertation, Kyoto University*. November 1973.

[4] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE TPAMI*, vol. 28, no. 12, pp. 2037–2041, Dec 2006.

[5] F. Juefei-Xu and M. Savvides, "Single Face Image Super-Resolution via Solo Dictionary Learning," in *ICIP*, Sept 2015.

[6] F. Juefei-Xu and M. Savvides, "Encoding and Decoding Local Binary Patterns for Harsh Face Illumination Normalization," in *ICIP*, Sept 2015.

[7] F. Juefei-Xu and M. Savvides, "Pareto-optimal Discriminant Analysis," in *ICIP*, Sept 2015.

[8] F. Juefei-Xu, D. K. Pal, and M. Savvides, "NIR-VIS Heterogeneous Face Recognition via Cross-Spectral Joint Dictionary Learning and Reconstruction," in *CVPRW*, June 2015.

[9] K. Seshadri, F. Juefei-Xu, D. K. Pal, and M. Savvides, "Driver Cell Phone Usage Detection on Strategic Highway Research Program (SHRP2) Face View Videos," in *CVPRW*, June 2015.

[10] F. Juefei-Xu and M. Savvides, "Facial Ethnic Appearance Synthesis," in *ECCVW*, 2015, pp. 825–840.

[11] F. Juefei-Xu, Dipan K. Pal, and M. Savvides, "Hallucinating the Full Face from the Periocular Region via Dimensionally Weighted K-SVD," in *CVPRW*, June 2014.

[12] F. Juefei-Xu and M. Savvides, "An Image Statistics Approach towards Efficient and Robust Refinement for Landmarks on Facial Boundary," in *BTAS*, Sept 2013.

[13] S. Song, K. Ohnuma, Z. Liu, L. Mei, A. Kawada, and T. Monma, "Novel biometrics based on nose pore recognition," *Optical Engineering*, vol. 48, no. 5, May 2009.

[14] K. Bonnen, B. Klare, and A. K. Jain, "Component-based representation in automated face recognition," *IEEE TIFS*, vol. 8, no. 1, pp. 239253, 2013.

[15] B. Heisele, T. Serre, and T. Poggio, "A component-based framework for face detection and identification," *IJCV*, vol. 74, no. 2, pp. 167181, 2007.

[16] P. Tome, J. Fierrez, R. Vera-Rodriguez, and D. Ramos, "Identification using face regions: Application and assessment in forensic scenarios," *Forensic Science International*, vol. 233, pp. 75–83, 2013.

[17] F. Li, H. Wechsler, and M. Tistarelli, "Robust fusion using boosting and transduction for component-based face recognition," in *Proceedings of IEEE 10th International Conference on Control, Automation, Robotics and Vision, ICARCV*, 2008, p. 434439.

[18] O. Ocegueda, S. K. Shah, and I. A. Kakadiaris, "Which parts of the face give out your identity?," in *CVPR*, 2011, p. 641648.

[19] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, Jin Chang, K. Hoffman, J. Marques, Jaesik Min, and W. Worek, "Overview of the face recognition grand challenge," in *CVPR*, June 2005, vol. 1, pp. 947–954.

[20] A.M. Martinez and R. Benavente, "The AR Face Database," *CVC Technical Report No.24*, June 1998.

[21] B. V. K Vijaya Kumar, M. Savvides, and C. Xie, "Correlation pattern recognition for face recognition," *Proc. of the IEEE*, vol. 94, no. 11, pp. 1963–1976, Nov 2006.

[22] J. Figue and Ph. Réfrégier, "Optimality of trade-off filters," *Appl. Opt.*, vol. 32, no. 11, pp. 1933–1935, Apr 1993.

[23] Abhijit Mahalanobis, B. V. K. Vijaya Kumar, and David Casasent, "Minimum average correlation energy filters," *Appl. Opt.*, vol. 26, no. 17, pp. 3633–3640, Sep 1987.

[24] B. V. K. Vijaya Kumar, "Minimum-variance synthetic discriminant functions," *J. Opt. Soc. Am. A*, vol. 3, no. 10, pp. 1579–1584, Oct 1986.

[25] C. Xie, M. Savvides, and B. V. K Vijaya Kumar, "Redundant Class-Dependence Feature Analysis Based on Correlation Filters Using FRGC2.0 Data," in *CVPRW*, June 2005, pp. 153–153.

[26] F. Juefei-Xu and M. Savvides, "Subspace Based Discrete Transform Encoded Local Binary Patterns Representations for Robust Periocular Matching on NIST's Face Recognition Grand Challenge," *TIP*, vol. 23, no. 8, pp. 3490–3505, aug 2014.

[27] F. Juefei-Xu and M. Savvides, "An Augmented Linear Discriminant Analysis Approach for Identifying Identical Twins with the Aid of Facial Asymmetry Features," in *CVPRW*, June 2013, pp. 56–63.

[28] F. Juefei-Xu and M. Savvides, "Can Your Eyebrows Tell Me Who You Are?," in *ICSPCS*, Dec 2011, pp. 1–8.

[29] F. Juefei-Xu and M. Savvides, "Unconstrained Periocular Biometric Acquisition and Recognition Using COTS PTZ Camera for Uncooperative and Non-cooperative Subjects," in *WACV*, Jan 2012, pp. 201–208.

[30] F. Juefei-Xu, Khoa Luu, M. Savvides, T.D. Bui, and C.Y. Suen, "Investigating Age Invariant Face Recognition Based on Periocular Biometrics," in *IJCB*, Oct 2011, pp. 1–7.

[31] F. Juefei-Xu, M. Cha, J. L. Heyman, S. Venugopalan, R. Abiantun, and M. Savvides, "Robust Local Binary Pattern Feature Sets for Periocular Biometric Identification," in *BTAS*, sep 2010, pp. 1–8.

[32] F. Juefei-Xu and M. Savvides, "Image Matching Using Subspace-Based Discrete Transform Encoded Local Binary Patterns," Sept. 2013, US Patent 2014/0212044.

[33] F. Juefei-Xu and M. Savvides, "Weight-Optimal Local Binary Patterns," in *ECCVW*, 2015, pp. 148–159.

[34] F. Juefei-Xu, M. Cha, M. Savvides, S. Bedros, and J. Trojanova, "Robust Periocular Biometric Recognition Using Multi-level Fusion of Various Local Feature Extraction Techniques," in *DSP*, 2011.