

©2006 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

FACE RECOGNITION USING 3D SUMMATION INVARIANT FEATURES

Wei-Yang Lin, Kin-Chung wong, Yu Hen Hu, and Nigel Boston *

University of Wisconsin-Madison
Department of Electrical and Computer Engineering
1415 Engineering Drive, Madison WI, 53706 USA

weiyanglin@wisc.edu

ABSTRACT

In this paper, we developed a family of 2D and 3D invariant features with applications to 3D human faces recognition. The main contributions of this paper are: (a) systematically deriving a family of novel features, called *summation invariant* that are invariant to Euclidean transformation in both 2D and 3D; (b) developing an effective method to apply summation invariant to the 3D face recognition problem. Tested with the 3D data from the Face Recognition Grand Challenge v1.0 dataset, the proposed new features exhibit achieves a performance that rivals the best 3D face recognition algorithms reported so far.

1. INTRODUCTION

Human face recognition has received unprecedented interest in recent years [11]. However, a majority of current face recognition systems are designed for 2D facial images. There are much fewer works on 3D face recognition. Among those reported 3D face recognition algorithms, there are three main approaches:

The first approach [6], [4] regards a 3D facial range image as a gray-level image and apply standard 2D algorithm such as *eigen-face* to perform the classification task. The second approach [4] is to apply a deformable 3D facial surface model. The third approach [1], [10] is to extract features from a 3D facial surface that are invariant to pose variations. These approaches focus on the classical differential invariants such as Gaussian curvature.

A fundamental weakness of differential invariants is that the high-order derivative operations are sensitive to noisy data. Hann and Hickman proposed a new family of *integral invariant* [3] that are based on integration rather than derivative operators. However, since their formulation is defined on continuous functions, when applied to sampled data, such as digital images, it suffers from excessive numerical errors. Several related approaches such as the semi-differential invariants [7], and affine quasi-invariant arc-length [9] are also defined

on continuous functions and hence share the same numerical concerns as the integral invariants.

Recently, we proposed a new family of invariants, called *summation invariant* [5] that is based on summation of discretized sample data. As such, these invariants are likely to be more robust to noise and numerical errors. Based on this preliminary result, in this paper, we developed a new family of summation invariants in both \mathbb{R}^2 and \mathbb{R}^3 with respect to the Euclidean transformation (rotation and translation). Moreover, we developed an efficient approach to extract these 2D and 3D summation invariants from a given 3D range image of human face and demonstrated superior performance. We have conducted extensive experiments using the Face Recognition Grand Challenge v1.0 dataset and the BEE (Biometric Experimentation Environment) package. Our method yields the best performance that has been reported so far [4].

The rest of this paper is organized as follows. Section 2 describes the summation invariant for the Euclidean transformation group. Both 2D and 3D cases are discussed. The application to 3D face recognition is presented in section 3, with experimental results summarized in section 5.

2. SUMMATION INVARIANTS

We employ the method of moving frames [2] by Élie Cartan, to systematically construct mathematical invariants under group actions. For 3D face recognition, we seek features that are invariant to pose variations which can be modeled by Euclidean geometrical transformations (rotation and translation). Our approach is similar to that of our early work [5] for the affine transformation group acting on \mathbb{R}^2 .

2.1. Euclidean summation invariants of curves

Denote $(x[n], y[n])$ and $(\bar{x}[n], \bar{y}[n])$ $0 \leq n \leq N - 1$ respectively to be the corresponding point on a curve in \mathbb{R}^2 before and after an Euclidean transformation

$$\begin{bmatrix} \bar{x}[n] \\ \bar{y}[n] \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x[n] \\ y[n] \end{bmatrix} + \begin{bmatrix} a \\ b \end{bmatrix} \quad (1)$$

*The authors have been partially supported by the National Science Foundation under Grant No. CCF-0434355.

where $a, b, \theta \in \mathbb{R}$ are parameters. Define the i, j^{th} potentials of this curve as

$$P_{i,j} = \sum_{n=0}^{N-1} x^i[n] \cdot y^j[n] \quad (2)$$

where i, j are nonnegative integers, $i+j = k$ and $k \neq 0$. Similarly, we denote $\bar{P}_{i,j}$ to be the potentials after Euclidean transformation.

To systematically derive a family of invariants with respect to the Euclidean transform group, our approach is as follows:

1. Construct a *prolonged potential jet space* $(\bar{x}_0, \bar{y}_0, \bar{x}_1, \bar{y}_1, \bar{P}^{i,j})$ where $\bar{x}_0 = \bar{x}[0]$, $\bar{y}_0 = \bar{y}[0]$, $\bar{x}_1 = \bar{x}[N-1]$, $\bar{y}_1 = \bar{y}[N-1]$.
2. Solve for parameters (called *moving frame* a, b , and θ of the Euclidean transform from a set of a *normalization equations*:

$$(\bar{x}_0, \bar{y}_0, \bar{y}_1) = (0, 0, 0) \quad (3)$$

3. Express the analytical expression of the solution in terms of x_0, y_0 , and potentials. Substitute these solutions into a transformed potential $\bar{P}_{i,j} = \sum_{n=0}^{N-1} \bar{x}^i[n] \bar{y}^j[n]$. The result is the desired invariant $\eta_{i,j}$

Note that we are free to specify normalization equations as long as they admit explicit analytical solution, *i.e.* a moving frame can be found. We have derived all five first and second order invariant functions, $i+j = 1$ or 2 , $\eta_{1,0}, \eta_{0,1}, \eta_{2,0}, \eta_{1,1}$, and $\eta_{0,2}$. Due to space limitation, we only list the first two here:

$$\eta_{1,0} = (x_1 - x_0)P_{1,0} + (y_1 - y_0)P_{0,1} + Nx_0(x_0 - x_1) + Ny_0(y_0 - y_1) \quad (4)$$

$$\eta_{0,1} = (y_1 - y_0)P_{1,0} + (x_0 - x_1)P_{0,1} + N(x_1y_0 - x_0y_1) \quad (5)$$

2.2. Euclidean summation invariants of surfaces

Above procedure can readily be generalized to a surface in \mathbb{R}^3 . Albeit, some complications must be overcome. Given a parameterized surface $(x[m, n], y[m, n], z[m, n])$ with $m = 0, \dots, M-1$ and $n = 0, \dots, N-1$, its corresponding *potentials* is defined by

$$Q_{i,j,k} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x^i[m, n] \cdot y^j[m, n] \cdot z^k[m, n] \quad (6)$$

where i, j, k are nonnegative integers, $i+j+k = \ell$ and $\ell \neq 0$.

In 3D, the Euclidean transformation is parameterized by 3 elementary rotation matrices $\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3$, and a 3D translation vector \mathbf{T} . This increases the number of parameters to

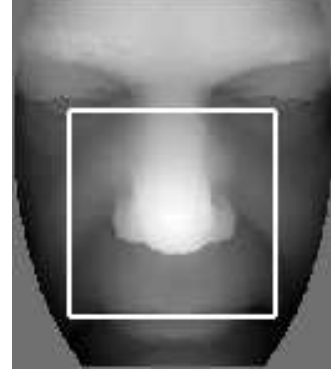


Fig. 1. Normalized range image and an 81×81 region centered at the nose tip.

six. These parameters are to be solved from a set of six normalization equations:

$$\begin{aligned} (\bar{x}[0, 0], \bar{y}[0, 0], \bar{z}[0, 0], \bar{y}[M-1, 0], \bar{z}[M-1, 0], \\ \bar{z}[0, N-1]) = (0, 0, 0, 0, 0, 0) \end{aligned} \quad (7)$$

The first three equations yields $\mathbf{T} = [x[0, 0], y[0, 0], z[0, 0]]^T$.

In order to satisfy the next two equations, denote $\mathbf{A} = [x[M-1, 0], y[M-1, 0], z[M-1, 0]]^T$ and compute its polar coordinates (R_a, θ_a, ϕ_a) . If two of the elementary matrices are defined on θ_a , and ϕ_a respectively, then it is guaranteed that $\bar{y}[M-1, 0] = 0$, and $\bar{z}[M-1, 0] = 0$. Thus, the fourth and fifth normalized equations are solved.

Finally, to solve the last equation, denote $\mathbf{B} = \mathbf{R}_2 \mathbf{R}_1 \cdot \left(\begin{bmatrix} x[0, N-1] & y[0, N-1] & z[0, N-1] \end{bmatrix}^T - \mathbf{T} \right)$. Again, by representing \mathbf{B} in polar coordinates, (R_c, θ_c, ϕ_c) , one may determine \mathbf{R}_3 using ϕ_c . By substituting these solutions into the transformed potentials $\bar{Q}_{i,j,k}$, we obtain desired invariants $\kappa_{i,j,k}$.

3. APPLICATION TO FACE RECOGNITION

We use the 3D facial images distributed with the Face Recognition Grand Challenge (FRGC) dataset [8] to conduct face recognition experiments. The 3D data contains 275 subjects (1 to 8 range scans per subject) and a total of 943 range scans. Each range scan has a resolution of 640×480 pixels.

Our experiment procedures closely follow those defined for the baseline algorithm provided by FRGC. We extract invariant features from an $N \times N$ ($N = 81$) rectangular region centered at the nose tip of the 3D facial map. An example is shown in Figure 1. The positions of the nose tips are manually selected and are provided by the FRGC dataset.

For 2D invariants $\eta_{i,j}$, facial range data are resampled uniformly with respect to arc-length. Specifically, for each row on a range data, we first compute its arc-length and resample

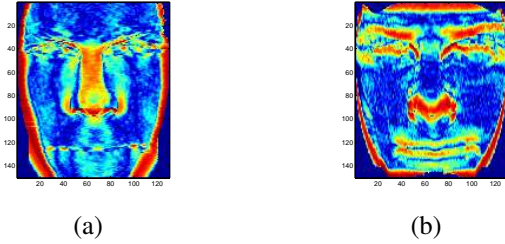


Fig. 2. For each range image, we compute summation invariants at each pixel: (a) Summation invariant is computed from a $1 \times L$ horizontal window and (b) from a $L \times 1$ vertical window.

it uniformly with respect to arc-length. Then, we perform the same resampling on each column. For surface invariants $\kappa_{i,j}$, we do not perform resampling on the normalized range data.

We extract a semi-local summation invariant from each row and each column of the 81×81 rectangular region and use the results as invariant features. For curve invariants $\eta_{i,j}$, we compute their semi-local summation invariants from a local window. At each pixel, semi-local summation invariants are computed both vertically and horizontally. The length of local window is chosen to be $L = 21$. Similarly, we also compute semi-local summation invariants $\kappa_{i,j,k}$ from a local window surrounding each pixel. The window size for surface invariants is 17×17 .

In order to reduce the size of the feature vectors, we use principal component analysis (PCA) to compute their subspace projections. The PCA basis is computed using the training set (183 range images) specified by the experiment 3 of the FRGC v1.0. The classification is accomplished by using subspace projections. The Mahalanobis cosine distance metric is used in these experiments.

The code is implemented in non-optimized C language. The experiments are conducted under Linux operation system with 3.40GHz XEON processor and 2GB memory.

4. EXPERIMENTAL RESULTS

We have conducted a series of experiments to assess the performance of the proposed algorithm. Due to space limitation, we only report the final results.

4.1. Effects of combining summation invariants

The purpose of this experiment is to investigate how to best combine several different summation invariant to achieve highest performance. Only summation invariants that give higher performance individually are combined. Figure 3 shows the ROC curves of combining two summation invariants. The fusion strategy is simply adding the similarity scores from two different summation invariants. We observe that fusion does

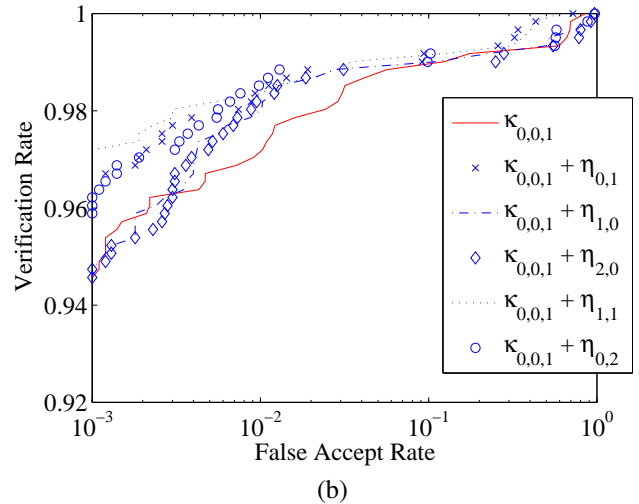


Fig. 3. ROC performance obtained by combining $\kappa_{0,0,1}$ and others.

not always yields higher recognition rate than that of a single summation invariant. However, at low false alarm rate, *eg.* 0.1%, the improvement provided by fusion two or more summation invariant is quite significant.

4.2. Comparison with other 3D face recognition algorithms

In this section, we conduct three experiments to evaluate the performance of our algorithm and the FRGC baseline algorithm. In the first, we simply run the FRGC baseline algorithm on 3D data alone. In the second, we still run the FRGC baseline algorithm on 3D data but using only the cropped region rather than the whole normalized range data. In Fig 4, the second experiment shows a lower recognition rate than the first one. This is reasonable because the second experiment uses less data to perform recognition. In the third experiment, we compute $\eta_{1,1}$ and $\kappa_{0,0,1}$ from the cropped region. Our algorithm yields the highest recognition rate as one can see in Fig 4 (At 0.1% false alarm rate, the recognition rate is 97.2%). The results clearly indicate that summation invariants offer statistically significant better recognition performance than the range data itself. Except for the FRGC 3D baseline, Kakadiaris *et al.*[4] also apply their algorithm to the range images from FRGC v1.0 dataset and report about 97.0% recognition rate at 0.1% false alarm rate. So, the proposed method has the same recognition performance as the the best results ever published.

5. DISCUSSION AND CONCLUSION

The value of summation invariants in the context of 3D face recognition is evaluated in this paper. We extract geometric features of facial surfaces using summation invariants and

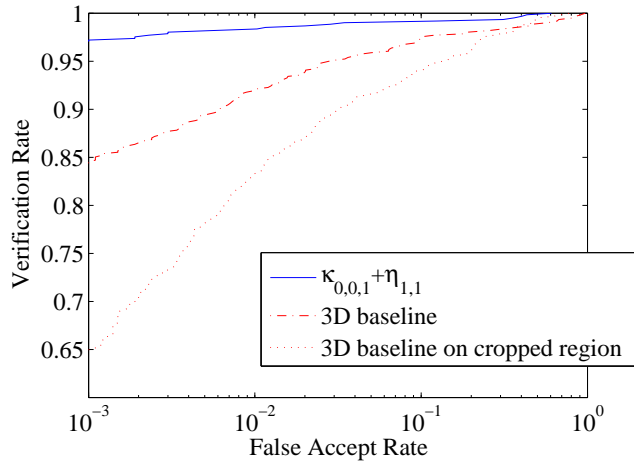


Fig. 4. Comparison with FRGC 3D baseline algorithm. We apply the FRGC 3D baseline algorithm on the normalized depth map and the cropped region shown in Fig 1. Their corresponding ROC curves are shown by the solid line and the dash line respectively.

apply PCA on the resulting representation. Our results compares favorably to those reported by Kakadiaris *et al.*[4]. We are currently working on FRGC version 2.0 dataset, and explore additional 2D and 3D feature extraction method to improve the 3D face recognition performance.

6. REFERENCES

- [1] J. Y. Cartoux, J. T. Lapreste, and M. Richetin. Face authentication or recognition by profile extraction from range images. In *Proc. of Workshop on Interpretation of 3D Scenes*, pages 194–9, 1989.
- [2] M. Fels and P. J. Olver. Moving coframes: I. a practical algorithm. *Acta Applicandae Mathematicae*, 51(2):161–213, 1998.
- [3] C. E. Hann and M. S. Hickman. Projective curvature and integral invariants. *Acta Applicandae Mathematicae*, 74(2):177–193, 2002.
- [4] I. A. Kakadiaris, G. Passalis, T. Theoharis, G. Toderici, I. Konstantinidis, and N. Murtuza. Multimodal face recognition: Combination of geometry with physiological information. In *Proc. IEEE Conf. on CVPR*, volume 2, pages 1022–1029, 2005.
- [5] W. Y. Lin, N. Boston, and Y. H. Hu. Summation invariant and its application to shape recognition. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, volume V, pages 205–208, 2005.
- [6] G. Medioni and R. Waupotitsch. Face modeling and recognition in 3-d. In *2003 IEEE Intl. Workshop on*

Analysis and Modeling of Faces and Gestures, pages 232–3, 2003.

- [7] T. Moons, E. J. Pauwels, L. J. V. Gool, and A. Oosterlinck. Foundations of semi-differential invariants. *Intl. Journal of Computer Vision*, 14(1):25–47, 1995.
- [8] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the face recognition grand challenge. In *Proc. IEEE Conf. on CVPR*, volume 1, pages 947–54, 2005.
- [9] J. Sato and R. Cipolla. Affine integral invariants and matching of curves. In *Proc. of 13th Intl. Conf. on Pattern Recognition*, volume 1, pages 915–19, Vienna, Austria, 1996.
- [10] H. T. Tanaka, M. Ikeda, and H. Chiaki. Curvature-based face surface recognition using spherical correlation. principal directions for curved object recognition. In *Proc. of Third IEEE Intl. Conf. on Automatic Face and Gesture Recognition*, pages 372–7, 1998.
- [11] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys*, 35(4):399–458, 2003.