

©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.

# Fusion of Summation Invariants in 3D Human Face Recognition

Wei-Yang Lin   Kin-Chung Wong   Nigel Boston   Yu Hen Hu  
University of Wisconsin-Madison  
Department of Electrical and Computer Engineering  
1415 Engineering Drive, Madison, WI 53706

{weiyanglin, kinchungwong}@wisc.edu, {boston, hu}@engr.wisc.edu \*

## Abstract

*A novel family of 2D and 3D geometrically invariant features, called summation invariants is proposed for the recognition of the 3D surface of human faces. Focusing on a rectangular region surrounding the nose of a 3D facial depth map, a subset of the so called semi-local summation invariant features is extracted. Then the similarity between a pair of 3D facial depth maps is computed to determine whether they belong to the same person. Out of many possible combinations of these set of features, we select, through careful experimentation, a subset of features that yields best combined performance. Tested with the 3D facial data from the on-going Face Recognition Grand Challenge v1.0 dataset, the proposed new features exhibit significant performance improvement over the baseline algorithm distributed with the dataset.*

## 1. Introduction

Human face recognition has received unprecedented interest in recent years [1, 24]. However, rigorous tests with real-world data such as FERET, FRVT have revealed many shortcomings of existing approaches [19, 18]. In short, for large scale, real world situations, current systems still cannot deliver the performance needed for practical applications.

A majority of current face recognition approaches make use of 2D frontal facial texture features. Nonetheless, facial texture features are sensitive to lighting, pose, distance, age (temporal) variations, and can easily be altered through simple make-up efforts. On the other hand, from 3D facial surfaces, one may exploit features that are invariant to appearance variations. For example, the facial surface around cheek bones or the nose would remain unchanged under varying lighting conditions, are less likely to change due to

aging, and are seldom covered with hairs. Hence, in this research, we focus on exploiting *invariant* features extracted from 3D facial surfaces.

Invariants for transformation groups play an important role in computer vision. Classical differential invariants such as curvature depend on derivatives that may be very sensitive to noise [5, 6]. Several approaches such as the semi-differential invariant introduced in [23, 15], integral invariant [8], and affine quasi-invariant arc-length [21] have been proposed. These invariants are defined on continuous functions. When applied to digitized object descriptions of contours or surfaces, numerical approximation will be needed and the results can be affected significantly by step size and other detailed settings.

Lin *et al.* [13] introduced a general method to generate invariants that are weighted summations of discrete data, as analogues to integral ones. Since these invariants are defined explicitly on discrete data, they do not require computationally intensive numerical integration to compute and will not be affected by the choice of step size. On the other hand, using weighted summation to compute the summation invariant will greatly reduce the impact of noise and hence promises a higher signal to noise ratio for the computed invariant features. Specifically, in [13], a semi-local summation invariant feature is proposed for two-dimensional (2D) closed contours. It delivers superior performance compared to those produced using integral invariants or wavelet invariant features.

In this paper, we exploit the feasibility of applying a summation invariant feature for 3D facial range images classification. Several key design issues are addressed: (a) To identify, among many possible variations of summation invariant features in the family, those features that yield the best performance. (b) To investigate the impacts of various preprocessing methods, including scaling and normalization on the proposed invariant features. (c) To explore proper feature reduction methods that will enhance the computation efficiency. (d) To explore fusion methods that will further enhance the performance of using a single summa-

---

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

tion invariant. In addition to analytical derivations, we have conducted extensive experiments using the Face Recognition Grand Challenge v1.0 dataset and the BEE (Biometric Experimentation Environment) package. To the best of our knowledge, the 3D face recognition algorithm developed by Kakadiaris *et al.* [11] report the best results on the FRGC v1.0 dataset. The proposed features yield a similar recognition rate to the one reported in [11].

The rest of this paper is organized as follows. Section 2 provides a brief review of previous work on 3D face recognition. Section 3 describes the summation invariant for the Euclidean transformation group. Both 2D and 3D cases are discussed. Section 4 illustrates our algorithm in detail. In Section 5, we present the experimental results and compare them with those of the FRGC 3D baseline algorithm. Finally, Section 6 summarizes our contributions and provides an overview of future directions.

## 2. Previous work in 3D face recognition

The majority of face recognition research focuses on using intensity images of the face. However, 3D images of the face have several advantages over the intensity-based features. In particular, 3D data provide a better representation for describing properties of the face in areas such as the cheeks, forehead, and chin, and are illumination invariant.

The early work of applying invariant functions on 3D face recognition was done over a decade ago. At that time, people began with the geometrical properties introduced in differential geometry, such as curvature. Cartoux *et al.* [3] proposed a face recognition algorithm based on Gaussian curvature of the face surface. Their approach yields a 100% verification rate on a small dataset (5 subjects and 18 range images). Lee and Milios [12] presented an algorithm for establishing a correspondence between features of two faces. Facial features are obtained by a segmentation of the range image based on the sign of the mean and Gaussian curvature at each point. Gordon [7] identifies the nose region, ridge and valley lines by using mean and Gaussian curvature. Tanaka *et al.* [22] also perform curvature-based segmentation and represent the face surface using an Extended Gaussian Image (EGI). Basically, these approaches use the invariant functions, *e.g.* Gaussian curvature is invariant under Euclidean transformations, to extract information from the face surface and then perform classification based on extracted information. The invariant functions they use all belong to the category called *differential invariants* which rely on derivative operations. The computation of differential invariants is not reliable because of the quantization error in discretized data. This fundamental issue limits their potential to achieve high recognition performance in a large dataset.

More recently, Medioni *et al.* [14] perform 3D face recognition using iterative closest point (ICP) matching.

Hesher *et al.* [10] perform principal component analysis (PCA) of range images. Similarly, Chang *et al.* [4] use a PCA-based method separately on 2D pictures and 3D range images. Matching scores from 2D and 3D are then combined for multimodal recognition. Further investigation into the 3D-PCA approach has been reported by Heseltine *et al.* [9]. Kakadiaris *et al.* [11] present a 3D face recognition approach that fits a deformable model to the probe face surfaces. They perform experiments on the FRGC v1.0 dataset and report a 97% verification rate measured at a 0.1% false accept rate (FAR). Note that none of these recent results utilizes any invariant feature from the face surface.

A 3D face recognition approach based on geometric invariants was introduced by Bronstein *et al.* [2]. The key idea of their algorithm is to approximate facial expressions as isometric transformations, *i.e.* length preserving transformations. The evaluation of their approach is limited and recognition performance is not reported.

## 3. Summation invariants

The method of moving frames [5, 6], originally introduced by Élie Cartan, is a powerful tool for constructing invariants under group actions. For expression-neutral face images, pose variations may be well modeled with Euclidean transformations (also known as rigid transformations). In this paper, we will use both profile curves and surface patches from 3D facial surfaces. In the following sections, we will develop novel invariants for the Euclidean groups acting on  $\mathbb{R}^2$  and  $\mathbb{R}^3$  respectively. This is similar to the work of Lin *et al.* [13], where summation invariants are derived for the affine transformation group acting on  $\mathbb{R}^2$ .

### 3.1. Euclidean summation invariants of curves

Given a curve  $(x[n], y[n])$  of length  $N$ , we use  $(\bar{x}[n], \bar{y}[n])$  to denote that it is transformed by a Euclidean transformation.

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

where  $a, b, \theta \in \mathbb{R}$ . We define the *potentials* of the curve as below.

**Definition 3.1** *The potential  $P_{i,j}$  of order  $k$  is given by*

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

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

One can construct a family of invariant functions  $\eta_{i,j}$  by applying a *moving frame* to potentials  $P_{i,j}$ . For example, we can find a moving frame by solving the following equations,

$$(\bar{x}[0], \bar{y}[0], \bar{y}[N-1]) = (0, 0, 0) \quad (3)$$

where  $\bar{x}$  and  $\bar{y}$  denote the x and y coordinate of a curve transformed by a moving frame. The equations above are called *normalization equations*. Note that we are free to specify normalization equations as long as they can be solved, *i.e.* a moving frame can be found. Let  $\bar{P}_{i,j}$  be the potentials transformed by a moving frame,

$$\bar{P}_{i,j} = \sum_{n=0}^{N-1} \bar{x}^i[n] \bar{y}^j[n] \quad (4)$$

where  $\bar{x}$  and  $\bar{y}$  denote the x and y coordinates transformed by a moving frame. The  $\bar{P}_{i,j}$  are invariant functions under Euclidean transformation acting on  $\mathbb{R}^2$ , *i.e.*  $\eta_{i,j} = \bar{P}_{i,j}$ . The first and second order invariant functions,  $i + j = 1$  or  $2$ , have been explicitly derived as shown below:

$$\eta_{1,0} = P_{1,0}(x_1 - x_0) + P_{0,1}(y_1 - y_0) + Nx_0(x_0 - x_1) + Ny_0(y_0 - y_1) \quad (5)$$

$$\eta_{0,1} = P_{1,0}(y_1 - y_0) + P_{0,1}(x_0 - x_1) + N(x_1y_0 - x_0y_1) \quad (6)$$

$$\begin{aligned} \eta_{2,0} = & -2P_{1,0}(x_0 - x_1)(x_0^2 - x_0x_1 + y_0^2 - y_0y_1) \\ & -2P_{0,1}(y_0 - y_1)(y_0^2 - y_0y_1 + x_0^2 - x_0x_1) \\ & + P_{2,0}(x_0 - x_1)^2 + P_{0,2}(y_0 - y_1)^2 \\ & + 2P_{1,1}(x_0 - x_1)(y_0 - y_1) \\ & + N(x_0(x_0 - x_1) + y_0(y_0 - y_1))^2 \end{aligned} \quad (7)$$

$$\begin{aligned} \eta_{1,1} = & P_{1,1}((x_0 - x_1)^2 - (y_0 - y_1)^2) \\ & + P_{1,0}(y_0^3 + 2x_0x_1y_1 - 2y_1x_0^2 \\ & + x_0^2y_0 - 2y_0^2y_1 + y_0y_1^2 - x_1^2y_0) \\ & - P_{0,1}(x_0^3 + 2y_0y_1x_1 - 2x_1y_0^2 \\ & + y_0^2x_0 - 2x_0^2x_1 + x_0x_1^2 - y_1^2x_0) \\ & + (P_{0,2} - P_{2,0})(x_0 - x_1)(y_0 - y_1) \\ & + N(x_1y_0 - x_0y_1) \\ & (x_0(x_1 - x_0) + y_0(y_1 - y_0)) \end{aligned} \quad (8)$$

$$\begin{aligned} \eta_{0,2} = & 2(x_1y_0 - x_0y_1)(P_{1,0}(y_1 - y_0) - P_{0,1}(x_1 - x_0)) \\ & + P_{2,0}(y_0 - y_1)^2 + P_{0,2}(x_0 - x_1)^2 \\ & - 2P_{1,1}(x_0 - x_1)(y_0 - y_1) \\ & + N(x_0y_1 - x_1y_0)^2 \end{aligned} \quad (9)$$

where  $x_0 = x[0]$ ,  $x_1 = x[N-1]$ ,  $y_0 = y[0]$ ,  $y_1 = y[N-1]$ .

### 3.2. Euclidean summation invariants of surfaces

In this section, we consider a Euclidean transformation acting on  $\mathbb{R}^3$ . Given a parameterized surface  $(x[m, n], y[m, n], z[m, n])$  with  $m = 0, \dots, M-1$  and  $n = 0, \dots, N-1$ , we first give the definition of the corresponding *potentials*.

**Definition 3.2** The *potential*  $Q_{i,j,k}$  of order  $\ell$  is given by

$$Q_{i,j,k} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x^i[m, n] \cdot y^j[m, n] \cdot z^k[m, n] \quad (10)$$

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

Then, the summation invariants  $\kappa_{i,j,k}$  of surfaces under Euclidean transformation can be expressed in a compact form

$$\kappa_{i,j,k} = \bar{Q}_{i,j,k} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \bar{x}^i \cdot \bar{y}^j \cdot \bar{z}^k \quad (11)$$

where  $\bar{x}, \bar{y}$  and  $\bar{z}$  denote the x, y and z coordinates transformed by a moving frame. The remaining task is simply to find a moving frame. In the previous section, we find a moving frame by solving normalization equations. Here, solving normalization equations is not trivial since the Euclidean transformation group acting on  $\mathbb{R}^3$  has more degrees of freedom than the one acting on  $\mathbb{R}^2$ . Instead of solving normalization equations, a moving frame can be expressed in terms of rotation matrices  $\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3$  and translation vector  $\mathbf{T}$ .

$$\begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} = \mathbf{R}_3 \mathbf{R}_2 \mathbf{R}_1 \left( \begin{bmatrix} x \\ y \\ z \end{bmatrix} - \mathbf{T} \right) \quad (12)$$

We choose a moving frame such that

$$\begin{aligned} (\bar{x}[0, 0], \bar{y}[0, 0], \bar{z}[0, 0], \bar{y}[M-1, 0], \bar{z}[M-1, 0], \\ \bar{z}[0, N-1]) = (0, 0, 0, 0, 0, 0) \end{aligned} \quad (13)$$

Hence, we have

$$\mathbf{T} = \begin{bmatrix} x(u_0, v_0) \\ y(u_0, v_0) \\ z(u_0, v_0) \end{bmatrix} \quad (14)$$

so that  $(\bar{x}[0, 0], \bar{y}[0, 0], \bar{z}[0, 0]) = (0, 0, 0)$ . Then, we need to fix the other coordinates in the Eq. (13). Let

$$\mathbf{A} = \begin{bmatrix} x[M-1, 0] \\ y[M-1, 0] \\ z[M-1, 0] \end{bmatrix} \quad (15)$$

and express  $\mathbf{A}$  in spherical coordinates  $(R_a, \theta_a, \phi_a)$ , where  $\theta$  and  $\phi$  are angular displacements measured from the positive z-axis and positive x-axis respectively and  $R$  is the distance from the origin to the point. By setting

$$\mathbf{R}_1 = \begin{bmatrix} \cos \phi_a & \sin \phi_a & 0 \\ -\sin \phi_a & \cos \phi_a & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (16)$$

$$\mathbf{R}_2 = \begin{bmatrix} \sin \theta_a & 0 & \cos \theta_a \\ 0 & 1 & 0 \\ -\cos \theta_a & 0 & \sin \theta_a \end{bmatrix} \quad (17)$$

, the rotation matrices  $\mathbf{R}_1$  and  $\mathbf{R}_2$  will rotate  $\mathbf{A}$  so that it is lying on the x-axis, *i.e.*  $(\bar{y}[M-1, 0], \bar{z}[M-1, 0]) = (0, 0)$ . Now, there is one degree of freedom left to be fixed. Let

$$\mathbf{B} = \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} = \mathbf{R}_2 \mathbf{R}_1 \begin{bmatrix} x[0, N-1] \\ y[0, N-1] \\ z[0, N-1] \end{bmatrix} - \mathbf{T} \quad (18)$$

and

$$\mathbf{C} = \begin{bmatrix} B_y \\ B_z \\ B_x \end{bmatrix} \quad (19)$$

Again, the vector  $\mathbf{C}$  can be expressed in spherical coordinates  $(R_c, \theta_c, \phi_c)$ . By setting

$$\mathbf{R}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi_c & \sin \phi_c \\ 0 & -\sin \phi_c & \cos \phi_c \end{bmatrix} \quad (20)$$

, the rotation matrices  $\mathbf{R}_1$ ,  $\mathbf{R}_2$ ,  $\mathbf{R}_3$  and translation  $\mathbf{T}$  form a moving frame. We can construct summation invariants  $\kappa_{i,j,k}$  under Euclidean transformation by applying the moving frame to the potential  $Q_{i,j,k}$ .

### 3.3. Semi-local summation invariants

The summation invariant can be defined over a local region of any 2D curve or 3D surface. In order to enhance the discriminating power of this feature, Lin *et al.* [13] have proposed a *semi-local* summation invariant feature that evaluates an invariant function for each point of the curve over a local window surrounding that pixel. As such, a curve consisting of  $N$  pixels will generate a feature vector of the same length. Similarly, a surface consisting of  $M \times N$  points will also generate a feature vector of the same size.

## 4. Application to face recognition

### 4.1. 3D face dataset and BEE

We use the 3D facial images distributed with the Face Recognition Grand Challenge (FRGC) dataset [17] to conduct face recognition experiments. FRGC is sponsored by the US National Institute of Standard and Technology (NIST) and other government agencies. Its testing data contains comprehensive 3D range images of human faces. It also provides an XML-based framework, called the Biometric Experimentation Environment (BEE) to document and describe computation experiments. All experiments in this paper are conducted using BEE.

The 3D data provided by FRGC v1.0 contains 275 subjects (1 to 8 range scans per subject) and a total of 943 range scans. Each range scan has a resolution of  $640 \times 480$  pixels. In addition to the 3D range map, it also comes with a 2D texture image of the same subject. However, in this paper, we use only the 3D range data in all experiments.

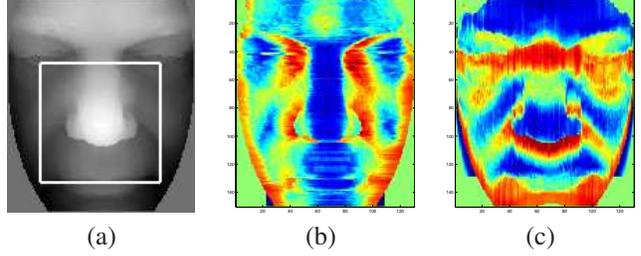


Figure 1. (a) Normalized range image and the selected  $81 \times 81$  region; (b) The computed result of  $\eta_{0,1}$  using a horizontal window; and (c) The computed result of  $\eta_{0,1}$  using a vertical window.

### 4.2. Procedures

In the official FRGC contest, four sets of experiments are defined. Among them, experiment 3 concerns the 3D face recognition task and hence is the focus of this paper. Our experiment procedures closely follow those defined for the baseline algorithm provided by FRGC, except for the following modifications:

1. 3D only experiment: In order to focus only on 3D face recognition performance, we do not use 2D texture images in our experiments.
2. Arc-length resampling : For 2D curve invariants  $\eta_{i,j}$ , range data are resampled uniformly with respect to arc-length. Specifically, for each row on the range data, we first compute its arc-length and resample 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.
3. Semi-local summation invariants: While the baseline algorithm uses 3D range data directly, we extract semi-local summation invariants from a normalized range image and use the results as invariant features. At each pixel, a semi-local summation invariant is computed from a local window. We compute curve invariants  $\eta_{i,j}$  from both a horizontal window and a vertical window. The results are shown in Figure 1. The length of the local window is chosen to be  $L = 21$ . Similarly, we also compute the  $\kappa_{i,j,k}$  from a local window surrounding each pixel. The window size for surface invariants is  $17 \times 17$ .
4. Specify a region of interest: Instead of using the entire 3D range image as the baseline algorithm does in BEE, we crop invariant features from an  $81 \times 81$  rectangular region centered at the nose tip (see Figure 1).
5. Alignment refinement: The location of the selected  $81 \times 81$  region is refined by minimizing the SSD (sum

of squared differences) with the averaged invariant features. We compute the averaged invariant features during the training stage.

6. Dimension reduction : 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 containing 183 range images. During the classification, the similarity metric is Mahalanobis cosine. Note that in our experiments, the training set and the similarity metric are the same as those in baseline algorithm. Please refer to [4] for more details.
7. Decision fusion: There are many ways of combining different features to achieve a better recognition performance. We use sum rule to combine the similarity metrics produced by different invariant features.

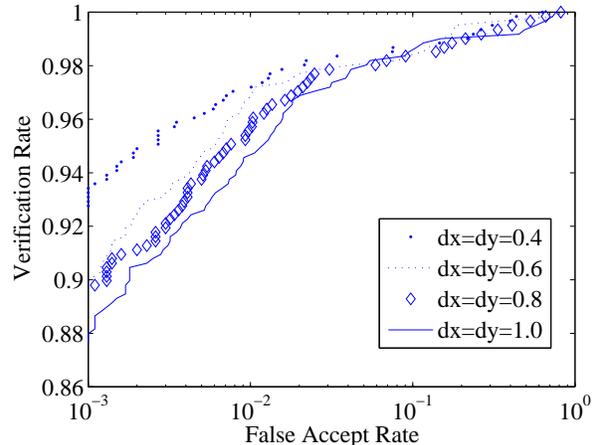
The code is implemented in non-optimized C language. The experiments are conducted under a Linux operating system with 3.40GHz XEON processor and 2GB memory. Given the normalized range images, it takes about 5 minutes to compute  $\eta_{i,j}$  and accomplish the following analysis (e.g. computing similarity metrics and performing ROC analysis). For computing  $\kappa_{i,j,k}$  and the following analysis, it takes 8 minutes. For the currently available datasets (e.g. FRGC v2.0 dataset has 50,000 images), computation cost will be one of the major concerns in designing a face recognition algorithm. We think summation invariant is an efficient and feasible solution to operate on the huge datasets like FRGC v2.0.

## 5. Experimental results

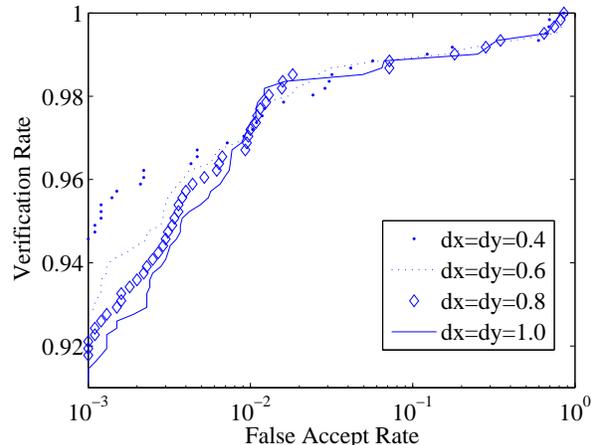
We have conducted a series of experiments to assess the performance of the proposed algorithm. The first part is to examine how the recognition performance is affected by the undesirable scaling in 3D data. The second part is to evaluate the discriminative capability of an individual summation invariant. Decision fusion is consider the third part. We use a simple fusion strategy to combine results at the metric level. Finally, we compare with the 3D face recognition algorithms which also operate on the FRGC v1.0 dataset.

### 5.1. Effects of difference scaling of x and y coordinates

In this experiment, we explore the issue of scaling on range data by using the summation invariant  $\eta_{0,1}$  without decision fusion. Recall that the summation invariants developed in section 3 are for the Euclidean groups acting on  $\mathbb{R}^2$  or  $\mathbb{R}^3$ . While this feature is invariant to Euclidean transformations, it will be affected by improper scaling of coordinates. In the FRGC 3D baseline algorithm,  $(x, y)$  coordinates in the range data are discarded and the  $z$  coordinate



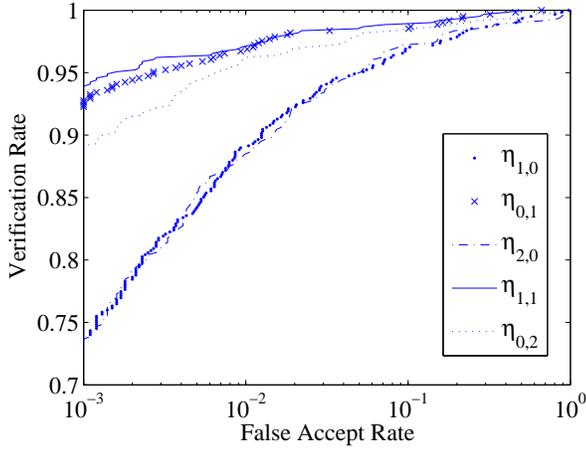
(a)



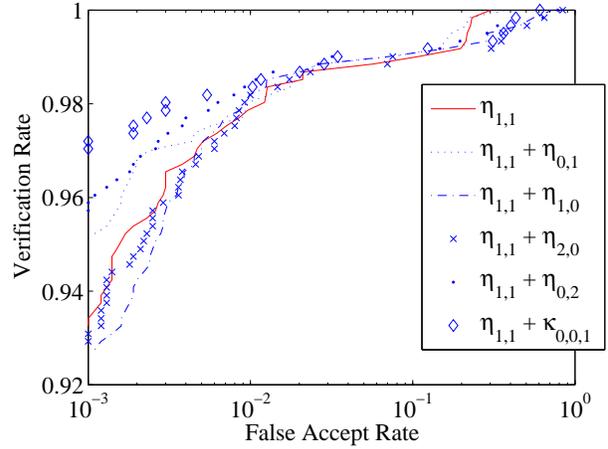
(b)

Figure 2. ROC performance for using different unit difference on the x and y coordinates. The results indicate that geometrical distortion will cause performance degradation of summation invariants. (a) ROC curves of  $\eta_{0,1}$  and (b) ROC curves of  $\kappa_{0,0,1}$ .

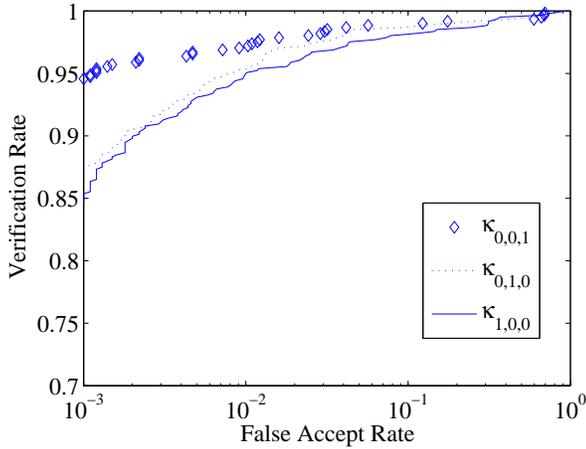
is multiplied by a scaling factor. In our method, however, we need to retain these pieces of information. We experimented with different values of increments,  $dx = dy \in \{0.4, 0.6, 0.8, 1.0\}$ , and a fixed scaling factor 12.5 on z values. In Figure 2, we observe a significant impact of these changes on both  $\eta_{0,1}$  and  $\kappa_{0,0,1}$ . In both cases, the verification rates decrease when  $dx$  and  $dy$  are smaller than 0.4. Hence, we assume that  $dx = dy = 0.4$  and a fixed scaling factor 12.5 on z values can reasonably restore the original geometry of the face surface and use these parameters in the following experiments. In general, better recognition performance is expected if the original x, y, z information is preserved.



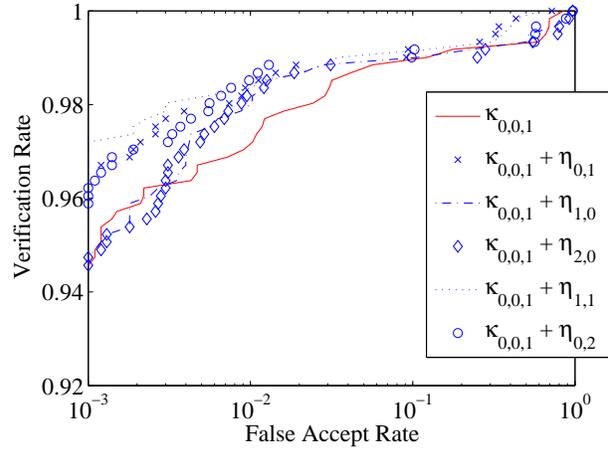
(a)



(a)



(b)



(b)

Figure 3. ROC performance obtained using different summation invariants. (a)  $\eta_{i,j}$  and (b)  $\kappa_{i,j,k}$

Figure 4. ROC performance obtained by combining two summation invariants. (a) Fusion of  $\eta_{1,1}$  and others. (b) fusion of  $\kappa_{0,0,1}$  and others.

## 5.2. Effects of different summation invariants

Summation invariants are functions which are unaffected under the actions of a particular transformation group. In other words, they do not guarantee any discriminative power for object recognition purpose. Hence, it is important to know how they perform individually in a face recognition application. In this experiment, we do not use decision fusion so that the contribution of each summation invariant can be identified. From the ROC curves in Fig 3, it is apparent that not all semi-local summation invariants are created equal in terms of discriminating power. Also, summation invariant  $\eta_{1,1}$  achieves the same level of recognition performance as those of summation invariants  $\kappa_{i,j,k}$  while their computation costs are quite different. The  $\eta_{i,j}$  and the  $\kappa_{i,j,k}$  are  $O(L)$  and  $O(L^2)$  operations respectively, where  $L$  is the size of the local window.

## 5.3. Effects of combining summation invariants

This experiment is to investigate the value of using more than one summation invariant. The hypothesis for this experiment is that the combined decision of these summation invariants will further improve the face recognition accuracy. We only choose summation invariants which have higher recognition rate to perform fusion because there are too many possible combinations to be shown. Figure 4 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 not always yield a higher verification rate than a single summation invariant. If we only focus on lower false accept rate, e.g. 0.1%, the improvement provided by fusion is significant and stable.

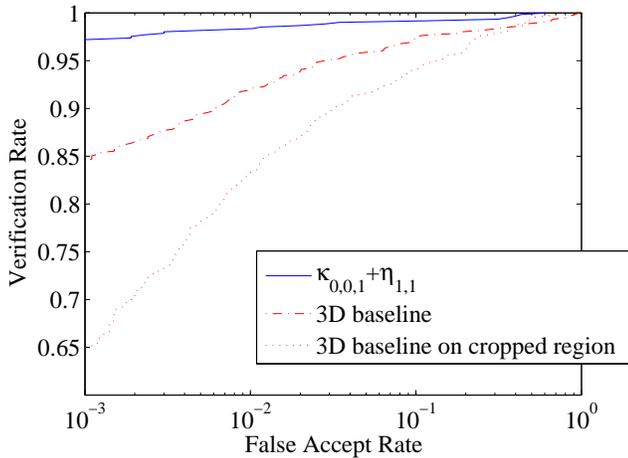


Figure 5. Comparison with FRGC 3D baseline algorithm. We apply the FRGC 3D baseline algorithm on the normalized depth map and the cropped region shown in Fig 1. Their corresponding ROC curves are shown by the solid line and the dash line respectively.

#### 5.4. 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 5, 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 verification rate as one can see in Fig 5. The proposed algorithm achieves the verification rate of 97.2% at the false accept rate of 0.1% and the EER (Equal Error Rate) of 1.43%. The results clearly indicate that summation invariants offer statistically significant better recognition performance than the range data itself. Except for the FRGC 3D baseline, some recent results on the FRGC v1.0 dataset are briefly discussed below. Kakadiaris *et al.* [11] present an approach based on the Annotated Face Model (AFM) and report about 97% verification rate at 0.1% false accept rate. Pan *et al.* [16] propose a novel mapping of the range image, called *mapped relative depth image*, and use it as feature vector to perform classification. Their method achieves the EER of 2.83% on the FRGC v1.0 dataset. Russ *et al.* [20] develop a 3D face recognition approach based on Hausdorff distance metric. Their experimental validation are conducted on part of FRGC v1.0 dataset, using only a single probe per person rather than all available probes. Performance is reported as the verification rate of 93.5% at

	Proposed algorithm	FRGC baseline
ROC I	83.18%	56.22%
ROC II	82.08%	49.55%
ROC III	80.82%	42.78%

Table 1. Verification rates on FRGC v2.0 dataset (FAR = 0.1%).

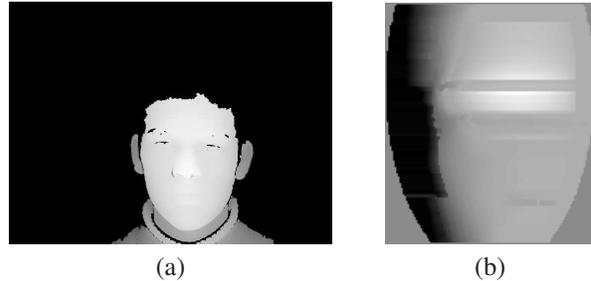


Figure 6. Failure example of the FRGC 3D normalization process: (a) Original range image and (b) Normalized range image.

the false accept rate of 0.1%. So, the proposed method has comparable performance with the state-of-the-art algorithm for 3D face recognition.

#### 5.5. Preliminary results on FRGC v2.0 dataset

We also verify the performance of our algorithm using the FRGC v2.0 dataset. Here, the selected region is shifted upward about 20 pixels to avoid using the mouth region. This will reduce the effect of facial expressions which are introduced in the FRGC v2.0 dataset. By comparing the verification rate at 0.1% FAR, we observe that our method also yields significant improvement over the baseline (see Table 1). Please refer to [17] for the specifications of ROC I, II and III.

## 6. 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 apply PCA on the resulting representation. To the best of our knowledge, Kakadiaris *et al.* [11] report the best previous results in FRGC v1.0 dataset. Our algorithm can yield the same level of recognition performance. Furthermore, such good performance can be achieved using only the nose portion of the whole face. In general, the results support the following conclusions:

1. Geometric features extracted by using summation invariants provide useful information for recognition purposes.
2. The combination of two or more summation invariants generally improves performance over using a single

summation invariant.

Also, we should note that the results reported in this paper are obtained using the normalized range images provided by the FRGC 3D baseline. In theory, we can compute summation invariants from unnormalized range images and obtain the same results. Note that summation invariants are designed to be unaffected by rotations and translations at the very beginning. Currently, the normalization is the most time-consuming part in the FRGC 3D baseline. The normalization alone takes about 20 minutes and the rest take about 5 minutes. Thus, an important part of future research along this direction is the development of a more efficient 3D face recognition technology in which normalization efforts are reduced by using summation invariants.

Some recognition errors reported by our algorithm are due to the failure of the normalization step (Figure 6). The baseline normalization algorithm works for most of the range images but not in some special cases. Note that the general quality of 3D range images is not as good as 2D pictures. In some regions like eyes, cheeks and eyebrows, current 3D sensing devices, such as the Minolta vivid 900, tend to either produce wrong depth values or have missing data. These issues would limit the 3D verification rate in general. Algorithms for reliably recovering data in such areas is an area where more research is needed.

## References

- [1] K. W. Bowyer, K. Chang, and P. Flynn. A survey of approaches to three-dimensional face recognition. In *Proc. of Intl. Conf. on Pattern Recognition*, volume 1, pages 358–61, 2004.
- [2] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Expression-invariant 3d face recognition. In *Proc. of the 4th Intl. Conf. on Audio- and Video-Based Biometric Person Authentication*, pages 62–9, 2003.
- [3] 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.
- [4] K. I. Chang, K. W. Bowyer, and P. J. Flynn. Multimodal 2d and 3d biometrics for face recognition. In *2003 IEEE Intl. Workshop on Analysis and Modeling of Faces and Gestures*, pages 187–94, 2003.
- [5] M. Fels and P. J. Olver. Moving coframes: I. a practical algorithm. *Acta Applicandae Mathematicae*, 51(2):161 – 213, 1998.
- [6] M. Fels and P. J. Olver. Moving coframes: II. regularization and theoretical foundations. *Acta Applicandae Mathematicae*, 55(2):127 – 208, 1999.
- [7] G. G. Gordon. Face recognition based on depth and curvature features. In *Proc. IEEE Conf. on CVPR*, pages 808–10, 1992.
- [8] C. E. Hann and M. S. Hickman. Projective curvature and integral invariants. *Acta Applicandae Mathematicae*, 74(2):177–193, 2002.
- [9] T. Heseltine, N. Pears, and J. Austin. Three-dimensional face recognition: An eigensurface approach. In *Proc. of the Intl. Conf. on Image Processing*, volume 5, pages 1421–1424, 2004.
- [10] C. Heshner, A. Srivastava, and G. Erlebacher. A novel technique for face recognition using range imaging. In *Proc. of the Seventh Intl. Symposium on Signal Processing and its Applications*, volume vol.2, pages 201–4, 2003.
- [11] I. A. Kakadiaris, G. Passalis, T. Theoharis, G. Toderici, I. Konstantinidis, and N. Murduza. Multimodal face recognition: Combination of geometry with physiological information. In *Proc. IEEE Conf. on CVPR*, volume 2, pages 1022–1029, 2005.
- [12] J. C. Lee and E. Miliotis. Matching range images of human faces. In *Proc. of Third Intl. Conf. on Computer Vision*, pages 722–6, 1990.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] G. Pan, S. Han, Z. Wu, and Y. Wang. 3d face recognition using mapped depth images. In *IEEE Workshop on Face Recognition Grand Challenge Experiments*, June 2005.
- [17] 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.
- [18] P. J. Phillips, P. Grother, R. Micheals, D. M. Blackburn, E. Tabassi, and M. Bone. Face recognition vendor test 2002. In *Intl. Workshop on Analysis and Modeling of Faces and Gestures*, page 44, 2003.
- [19] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss. The feret evaluation methodology for face-recognition algorithms. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(10):1090–104, 2000.
- [20] T. Russ, M. Koch, and C. Little. A 2d range hausdorff approach for 3d face recognition. In *IEEE Workshop on Face Recognition Grand Challenge Experiments*, June 2005.
- [21] 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.
- [22] 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.
- [23] L. Van Gool, P. Kempenaers, and A. Oosterlinck. Recognition and semi-differential invariants. In *Proc. IEEE Conf. on CVPR*, pages 454 – 60, 1991.
- [24] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys*, 35(4):399–458, 2003.