

# 3D FACE ANALYSIS FOR DISTINCT FEATURES USING STATISTICAL RANDOMIZATION

Chunxiao Zhou<sup>1,2</sup>, Yuxiao Hu<sup>1</sup>, Yun Fu<sup>1</sup>, Huixia Wang<sup>5</sup>, Thomas S. Huang<sup>1</sup>, Yongmei Michelle Wang<sup>2,3,4</sup>

Dept. of <sup>1</sup>Electrical & Computer Engineering, <sup>2</sup>Statistics, <sup>3</sup>Psychology, <sup>4</sup>Bioengineering  
University of Illinois at Urbana-Champaign, IL 61801, USA  
Dept. of <sup>5</sup>Statistics, North Carolina State University, NC 27695, USA

## ABSTRACT

It is a fascinating yet challenging problem to accurately and efficiently localize regionally distinct features between face groups in multi-dimensional signal processing and analysis. Given a data with unknown distribution and small sample size, we propose a new statistical analysis framework using hybrid randomization (i.e., permutation) tests to improve the system's efficiency in identifying distinct features. The proposed method fits the nonparametric distribution of the test statistic with Pearson distribution series. We bypass the tedious online randomization via calculating the first four moments of the permutation distribution. This can reduce the computational complexity from  $O(n!)$  to  $O(n^2)$  over traditional methods for the modified Hotelling's  $T^2$  test statistics. Experiments on simulated data and 3D face analysis demonstrate the efficiency, accuracy and robustness of the proposed approach.

**Index Terms**— 3D face analysis, feature selection, randomization test

## 1. INTRODUCTION

Machine learning methods have been widely used to identify important features for statistical pattern classification, with broad applications in biometrics including human face, gender, age and ethnic recognition. Recently, researchers have been working on 3D data or other multi-dimension/multi-modality signal analysis as opposed to traditional static image-based recognition [1]. This demands accurate and efficient algorithms to detect distinct features between data groups since the redundant information carried in the original high dimensional data may greatly degrade the performance of machine learning algorithms. Discriminate analysis [2] or analysis of variance (ANOVA) selects feature vectors that are consistent within a class but have a large separation from other classes of objects. These methods work well when the size of the dataset is large so that the distributions of features can be approximated by Gaussian. However, multi-dimensional data are often carried out with small sample size, as 3D face

data collection requires expensive acquisition devices such as laser scanner and structure light based reconstruction system. Tedious post processing is also required to build the vertices correspondence that is critical for recognition or synthesis tasks [1, 3, 4].

In this paper, we focus on the 3D face group analysis for two classes, though our method can be extended to multi-class cases. At each corresponding position on the face surfaces, we test whether there is a significant mean difference between location vectors (3D coordinates of the face vertexes) of two groups. If a hypothesis test leads to a  $p$ -value smaller than the pre-chosen  $\alpha$ -level, we reject the null hypothesis and conclude that a significant shape difference exists at this face location.

For face surface data with small sample size and little information about the distribution of location vectors, nonparametric approaches such as randomization tests, i.e., permutation tests, provide a flexible and robust alternative to Gaussian-theory inference. Randomization tests only require exchangeability but no parametric distributional assumptions, and they have been successfully used in biomedical image analysis [5, 6, 7, 8]. The null hypothesis can be accepted or rejected based on the permutation distribution of the test statistic. Permutation distribution is created by repeatedly permuting the group labels of samples and calculating the test statistic for each permutation. The  $p$ -value is then the proportion of permutations giving a value of the test statistic as or more extreme than the observed value. There are three major approaches to construct the permutation distribution [9]. First, exact permutation enumerates all possible arrangements. Due to the factorial increase in the number of permutations with the number of samples, the computational cost is the main disadvantage of the exact permutation. The second approach, random permutation, provides an approximate permutation distribution based on random sampling from all possible permutations. This technique has the problem of replication and causes more type I errors. The third method, permutation distribution approximation uses the analytical moments of the exact permutation distribution under the null hypothesis. This approach has two main limitations. For some special scenarios, the moments of the exact permutation distribution do not actually exist. For the

scenarios that these moments exist, they are usually difficult to obtain.

In this work, we propose a general theoretical method to derive moments of permutation distribution for any linear test statistics on multivariate data. Here, the term “linear test statistic” refers to a linear function of test statistic coefficients, instead of that of data. Given the first four moments, the permutation distribution can be well fitted by Pearson distribution series. The  $p$ -values can then be estimated from the fitted Pearson distribution without any real permutation. The main advantages of the proposed approach are: 1) It achieves computation efficiency by only calculating the analytically derived first four moments of exact permutation distribution without any real permutation. Given the sample size  $n_1 = n_2 = 21$ , and the number of surface locations  $m = 8000$ ,  $m \times (n_1 + n_2) / n_1 / n_2! \approx 4.4 \times 10^{15}$  permutations are needed for an exact permutation test, or  $m \times 20000 = 1.6 \times 10^8$  for random permutations, which is eliminated in our method. 2) Our hybrid permutation method achieves good robustness and high accuracy in addition to its efficiency.

## 2. HYBRID PERMUTATION

### 2.1. Pearson Distribution Series

The Pearson distribution series (Pearson I ~ VII) is a family of probability distributions that are more general than the normal distribution [7]. It covers all distributions in the  $(\beta_1, \beta_2)$  plane including normal, beta, gamma, log-normal and etc., where distribution shape parameters  $\beta_1, \beta_2$  are the square of standardized skewness and kurtosis measurements, respectively. Given the first four moments, the Pearson distribution series can be utilized to approximate the permutation distribution of the test statistic without conducting real permutation.

### 2.2. Theoretical Derivation of Moments

In order to approximate the permutation distribution with Pearson distribution, the moments of the exact permutation distribution need to be computed. Let  $X = [x_1 \dots x_n]^T$  be an univariate data of  $n$  observations,  $T(X, \pi) = C^T P X = \sum_i c_{\pi(i)} x_i$  be the linear test statistic after permutation,  $\pi(i)$  be the permutation operation, and  $C = [c_1 \dots c_n]^T$  be an  $n \times 1$  test statistic coefficients vector. The permutation matrix  $P$  is a matrix that has exactly one entry 1 in each row/column and 0's elsewhere, and it is obtained by permuting the rows of an identity matrix according to some permutation of the numbers 1 to  $n$ . The  $r$ -th conditional moment of  $T(X, \pi)$  can be derived as:

$$\begin{aligned} E(T^r(X, \pi) | X) &= E((C^T P X)^r) = \frac{1}{n!} \sum_{\pi} \left( \sum_i c_{\pi(i)} x_i \right)^r = \frac{1}{n!} \sum_{\pi} \sum_{i_1, \dots, i_r} \left( \prod_{k=1}^r c_{\pi(i_k)} x_{i_k} \right) \\ &= \frac{1}{n!} \sum_{\pi} \sum_{i_1, \dots, i_r} \left( \prod_{k=1}^r x_{i_k} \prod_{k=1}^r c_{\pi(i_k)} \right) = \sum_{i_1, \dots, i_r} \left( \prod_{k=1}^r x_{i_k} \frac{1}{n!} \sum_{\pi} \prod_{k=1}^r c_{\pi(i_k)} \right). \end{aligned} \quad (1)$$

To compute  $\frac{1}{n!} \sum_{\pi} \prod_{k=1}^r c_{\pi(i_k)}$ , it is natural to partition the

index space  $U = \{1 2 \dots n\}^r$  into  $\bigcup_{(\lambda_1, \lambda_2, \dots, \lambda_q) \in L} U^{(\lambda_1, \lambda_2, \dots, \lambda_q)}$ , where

$$L = \{(\lambda_1, \lambda_2, \dots, \lambda_q) : \lambda_1, \lambda_2, \dots, \lambda_q \in \mathbb{Z}^+; \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_q; \lambda_1 + \lambda_2 + \dots + \lambda_q = r\},$$

$U^{(\lambda_1, \lambda_2, \dots, \lambda_q)}$  means that all  $r$  indices are permuted into  $q$  different numbers. Each number corresponds to  $\lambda_i$  indices. When  $r=3$ ,  $U = U^{(1,1,1)} \cup U^{(1,2)} \cup U^{(3)}$ , where

$U^{(1,1,1)}$  is the set of  $\{i_1 \neq i_2 \text{ and } i_1 \neq i_3 \text{ and } i_2 \neq i_3\}$  with  $q=3$  and  $\lambda_1 = \lambda_2 = \lambda_3 = 1$ ,  $U^{(1,2)}$  is the set of

$\{i_1 = i_2 \neq i_3 \text{ or } i_1 = i_3 \neq i_2 \text{ or } i_2 = i_3 \neq i_1\}$  with  $q=2$  and  $\lambda_1 = 1, \lambda_2 = 2$ , and  $U^{(3)}$  is the set of  $\{i_1 = i_2 = i_3\}$  with  $q=1$

and  $\lambda_1 = 3$ . Since permutation is equally related to all  $r$

indices,  $\frac{1}{n!} \sum_{\pi} \prod_{k=1}^r c_{\pi(i_k)}$  is invariant in each category, we

define it as moment coefficient  $a(\lambda_1, \lambda_2, \dots, \lambda_q)$ , if

$(i_1 i_2 \dots i_r) \in U^{(\lambda_1, \lambda_2, \dots, \lambda_q)}$ . Thus,

$$E(T^r(X, \pi) | X) = \sum_{(i_1, \dots, i_r) \in U^{(\lambda_1, \lambda_2, \dots, \lambda_q)}} a(\lambda_1, \lambda_2, \dots, \lambda_q) \sum_{i_1, \dots, i_r} \left( \prod_{k=1}^r x_{i_k} \right). \quad (2)$$

The proposed approach can be easily extended to multivariate case. Let the linear test statistic

$T(X, \pi) = \sum_{i,j} e_{\pi(i,j)} x_{i,j}$ , where  $X = (x_{i,j})$  is an  $n \times m$  matrix

with  $n$  observations on  $m$  variables,  $E = (e_{i,j})$  is an  $n \times m$  test

statistic coefficient matrix and  $\pi(i,j)$  is a general permutation which permutes the row index  $i$  to from 1 to  $n$ , and column index  $j$  from 1 to  $m$ . Following the similar way as in Eq. (1), we obtain the conditional moments for the multivariate data as

$$\begin{aligned} E(T^r(X, \pi) | X) &= \frac{1}{(nm)!} \sum_{\pi} \left( \sum_{i,j} e_{\pi(i,j)} x_{i,j} \right)^r = \frac{1}{(nm)!} \sum_{\pi} \sum_{i_1, \dots, i_r} \left( \prod_{k=1}^r e_{\pi(i_k, j_k)} x_{i_k, j_k} \right) \\ &= \frac{1}{(nm)!} \sum_{\pi} \sum_{i_1, \dots, i_r} \left( \prod_{k=1}^r x_{i_k, j_k} \prod_{k=1}^r e_{\pi(i_k, j_k)} \right) = \sum_{i_1, \dots, i_r} \left( \prod_{k=1}^r x_{i_k, j_k} \frac{1}{(nm)!} \sum_{\pi} \prod_{k=1}^r e_{\pi(i_k, j_k)} \right). \end{aligned}$$

We can partition the index space  $U = \{1 2 \dots n\}^r \times \{1 2 \dots m\}^r$  into

$\bigcup_{(\lambda_1, \lambda_2, \dots, \lambda_{q_1}; \beta_1, \beta_2, \dots, \beta_{q_2}) \in L} U^{(\lambda_1, \lambda_2, \dots, \lambda_{q_1}; \beta_1, \beta_2, \dots, \beta_{q_2})}$ , where

$$L = \{(\lambda_1, \lambda_2, \dots, \lambda_{q_1}; \beta_1, \beta_2, \dots, \beta_{q_2}) : \lambda_1, \lambda_2, \dots, \lambda_{q_1} \in \mathbb{Z}^+, \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{q_1}, \lambda_1 + \lambda_2 + \dots + \lambda_{q_1} = r; \beta_1, \beta_2, \dots, \beta_{q_2} \in \mathbb{Z}^+, \beta_1 \leq \beta_2 \leq \dots \leq \beta_{q_2}, \beta_1 + \beta_2 + \dots + \beta_{q_2} = r\}.$$

$U^{(\lambda_1, \lambda_2, \dots, \lambda_{q_1}; \beta_1, \beta_2, \dots, \beta_{q_2})}$  means that all  $r$  row indices are permuted into  $q_1$  different numbers and all  $r$  column indices are permuted into  $q_2$  different numbers. Thus the  $r$ -th conditional moment for multivariate case is:

$$E(T^r(X, \pi) | X) = \sum_{(i_1, \dots, i_r, j_1, \dots, j_r) \in U^{(\lambda_1, \lambda_2, \dots, \lambda_{q_1}; \beta_1, \beta_2, \dots, \beta_{q_2})}} a(\lambda_1, \lambda_2, \dots, \lambda_{q_1}; \beta_1, \beta_2, \dots, \beta_{q_2}) \sum_{i_1, \dots, i_r, j_1, \dots, j_r} \left( \prod_{k=1}^r x_{i_k, j_k} \right).$$

All the  $a$ 's can also be calculated by computer simulation

without analytical derivation. Note that the computation cost of data summation terms for the  $r$ -th moment in each index subspace can be reduced from  $O(n^{2r})$  to  $O(n^2)$ .

### 2.3. Moments for Modified Hotelling's $T^2$

Assume  $X_A$  is an  $n_1 \times m$  and  $X_B$  is an  $n_2 \times m$  matrices from groups A and B, respectively, where  $n_1$  and  $n_2$  are the number of observations in two groups, and  $m$  is the dimension of variables. Under multivariate normal assumption, Hotelling's  $T^2$  is an optimal multivariate test statistic for testing the mean difference between two groups. The test statistic is defined as:

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} (\bar{X}_A - \bar{X}_B)^T S_{pooled}^{-1} (\bar{X}_A - \bar{X}_B).$$

where  $S_{pooled} = \frac{(n_1 - 1)S_A + (n_2 - 1)S_B}{n_1 + n_2 - 2}$  is the pooled

variance-covariance matrix,  $\bar{X}_A$  and  $\bar{X}_B$  are sample mean vectors of, and  $S_A$  and  $S_B$  are the sample variance-covariance matrices of groups A and B, respectively.

Since it is difficult to avoid real permutations with the pooled covariance matrix involved in Hotelling's  $T^2$ , we replace  $S_{pooled}$  with the sample variance-covariance matrix  $S$  over all the subjects of the two groups. This leads to a modified  $T^*$  as below

$$T^* = (\bar{X}_A - \bar{X}_B)^T S^{-1} (\bar{X}_A - \bar{X}_B).$$

Since it has been proved in Wald and Wolfowitz (1944) that the modified  $T^*$  is a monotonic function of Hotelling's  $T^2$ , these two statistics result in the same permutation  $p$ -value.

For the modified Hotelling's  $T^*$  test statistic in 3D case, let

$$D = \begin{bmatrix} x_1 & \dots & x_{n_1} & x_{n_1+1} & \dots & x_{n_1+n_2} \\ y_1 & \dots & y_{n_1} & y_{n_1+1} & \dots & y_{n_1+n_2} \\ z_1 & \dots & z_{n_1} & z_{n_1+1} & \dots & z_{n_1+n_2} \end{bmatrix}^T$$

denote the location vector matrix over all subjects in both groups, then  $T^* = (C^T P D) S^{-1} (C^T P D)^T = (C^T P V) (C^T P V)^T$ , where  $V = D S^{-1/2}$  is used for simplification of derivation. Then

$$E(T^* | W, \pi) = E((C^T P V) (C^T P V)^T) = E((C^T P V V^T P^T C)^T) = E((C^T P W P^T C)^T).$$

Since  $W = V V^T$  is a square matrix, there is no difference between permutation of row and column indices.

$$E(T^* | W, \pi) = \sum_{(i_1, \dots, i_r, j_1, \dots, j_r) \in U^{(1,1,1)}} a_{(i_1, i_2, \dots, i_r)} \sum_{i_1, \dots, i_r, j_1, \dots, j_r} \prod_{k=1}^r w_{i_k, j_k}.$$

For the first moment, we get  $U = U^{(1,1)} \cup U^{(2)}$ ,

$$a_{(1,1)} = -1 / (n_1 n_2 (n_1 + n_2 - 1)), \quad a_{(2)} = 1 / (n_1 n_2),$$

$$E(T^* | W, \pi) = a_{(1,1)} \sum_{(i_1, i_2) \in U^{(1,1)}} w_{i_1, i_2} + a_{(2)} \sum_{(i_1, i_2) \in U^{(2)}} w_{i_1, i_2} = a_{(1,1)} \sum_{i_1 \neq i_2} y_{i_1, i_2} + a_{(2)} \sum_{i_1} w_{i_1, i_1}.$$

For the second moment, we get

$$U = U^{(1,1,1,1)} \cup U^{(1,1,2)} \cup U^{(1,3)} \cup U^{(2,2)} \cup U^{(4)},$$

$$a_{(1,1,1,1)} = \binom{n_1}{4} \frac{1}{n_1^4} - \binom{n_1}{3} \binom{n_2}{1} \frac{1}{n_1^3 n_2} + \binom{n_1}{2} \binom{n_2}{2} \frac{1}{n_1^2 n_2^2} - \binom{n_1}{1} \binom{n_2}{3} \frac{1}{n_1 n_2^3} + \binom{n_2}{4} \frac{1}{n_2^4} \binom{n_1 + n_2}{4},$$

$$\begin{aligned} a_{(1,1,2)} &= \binom{n_1}{3} \frac{1}{n_1^4} - \frac{2}{3} \binom{n_1}{2} \binom{n_2}{1} \frac{1}{n_1^2 n_2} + \frac{1}{3} \binom{n_1}{1} \binom{n_2}{2} \frac{1}{n_1 n_2^2} \\ &\quad + \frac{1}{3} \binom{n_1}{1} \binom{n_2}{2} \frac{1}{n_1^2 n_2^2} - \frac{2}{3} \binom{n_1}{1} \binom{n_2}{1} \frac{1}{n_1 n_2^3} + \binom{n_2}{3} \frac{1}{n_2^4} \binom{n_1 + n_2}{3}, \\ a_{(1,3)} &= \binom{n_1}{2} \frac{1}{n_1^4} - \frac{1}{2} \binom{n_1}{1} \binom{n_2}{1} \frac{1}{n_1^3 n_2} - \frac{1}{2} \binom{n_1}{1} \binom{n_2}{1} \frac{1}{n_1 n_2^3} + \binom{n_2}{2} \frac{1}{n_2^4} \binom{n_1 + n_2}{2}, \\ a_{(2,2)} &= \binom{n_1}{2} \frac{1}{n_1^4} + \binom{n_1}{1} \binom{n_2}{1} \frac{1}{n_1^2 n_2^2} + \binom{n_2}{2} \frac{1}{n_2^4} \binom{n_1 + n_2}{2}, \\ a_{(4)} &= \binom{n_1}{1} \frac{1}{n_1^4} + \binom{n_2}{1} \frac{1}{n_2^4} \binom{n_1 + n_2}{1}, \\ E(T^{*2} | W, \pi) &= a_{(1,1,1,1)} \sum_{(i_1, i_2, i_3, i_4) \in U^{(1,1,1,1)}} w_{i_1, i_3} w_{i_2, i_4} + a_{(1,1,2)} \sum_{(i_1, i_2, i_3, i_4) \in U^{(1,1,2)}} w_{i_1, i_3} w_{i_2, i_4} + \\ &\quad + a_{(1,3)} \sum_{(i_1, i_2, i_3, i_4) \in U^{(1,3)}} w_{i_1, i_3} w_{i_2, i_4} + a_{(2,2)} \sum_{(i_1, i_2, i_3, i_4) \in U^{(2,2)}} w_{i_1, i_3} w_{i_2, i_4} + a_{(4)} \sum_{(i_1, i_2, i_3, i_4) \in U^{(4)}} w_{i_1, i_3} w_{i_2, i_4}. \end{aligned}$$

The third and fourth moments can be obtained in a similar way [6].

### 3. EXPERIMENT

To evaluate the accuracy and efficiency of our hybrid permutation tests, we consider six simulated cases in the first experiment for testing the difference between two groups A and B. For group A,  $n_1$  observations are generated independently from Normal(0,1) in Cases 1-2, from Gamma(3,3) in Cases 3-4, and from Beta(0.8, 0.8) in Cases 5-6. For group B,  $n_2$  independent observations are generated from Normal(1, 0.5) in Cases 1-2, from Gamma(3,2) in Cases 3-4, and from Beta(0.1, 0.1) in Cases 5-6. The design is balanced in Cases 1, 3 and 5 with  $n_1 = n_2 = 10$ , and unbalanced in Cases 2, 4 and 6 with  $n_1 = 6, n_2 = 18$ .

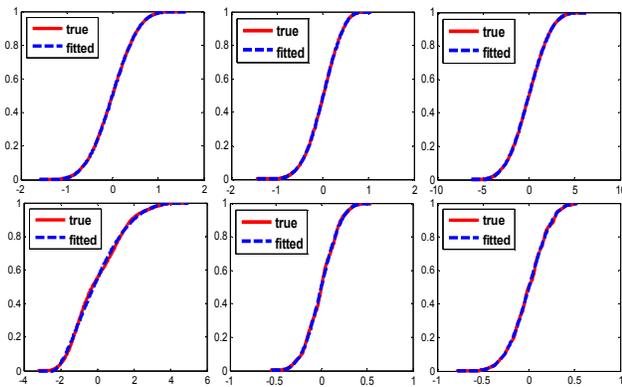
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
t_HP	0.01130	0.01133	0.0123	0.01378	0.01723	0.00181
t_RP	1.15841	1.14383	1.13693	1.12501	1.12620	1.13843
t_EP	4.43899	4.27950	1.29839	1.32407	4.13205	4.29486
p_HP	0.04991	0.13144	0.00101	0.02497	0.09083	0.08052
p_RP	0.04955	0.12695	0.0012	0.0242	0.0889	0.0818
p_EP	0.04980	0.13016	0.00101	0.02508	0.09250	0.08033

**Table 1.** Comparison of computation costs and  $p$ -values of three permutation methods: hybrid permutation (HP), random permutation (RP) and exact permutation (EP). The t\_HP, t\_RP and t\_EP denote the computation time (in seconds), and p\_HP, p\_RP and p\_EP are the  $p$ -values of three permutation methods.

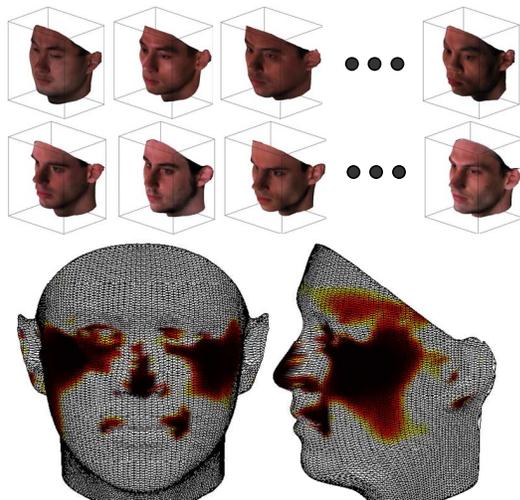
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
Mean_ABias_HP	8.79e-5	8.97e-6	9.54e-5	2.16e-4	6.79e-4	4.53e-4
Mean_ABias_RP	2.82e-4	6.64e-5	2.14e-4	1.30e-3	2.78e-4	4.59e-4
VAR_ABias_HP	5.99e-8	1.34e-7	2.10e-6	3.66e-7	9.55e-7	7.97e-6
VAR_ABias_RP	1.98e-6	1.42e-7	1.41e-6	5.34e-6	1.05e-5	1.00e-5

**Table 2.** Robustness and accuracy comparison of hybrid permutation and random permutation across 10 simulations, considering the  $p$ -values of exact permutation as gold standard. Mean\_ABias\_HP and VAR\_ABias\_HP are the mean and variance of the absolute biases of  $p$ -values of hybrid permutation; Mean\_ABias\_RP and VAR\_ABias\_RP are the mean and variance of the absolute biases of  $p$ -values of random permutation, respectively.

Table 1 and Fig 2 illustrate the high accuracy of our hybrid permutation technique. Furthermore, comparing with exact permutation or random 20,000 permutations, the hybrid permutation tests reduce more than 99% computation cost, and this efficiency gain increases with sample size. Table 1 shows the computation time and  $p$ -values of three permutation methods from one simulation. In order to demonstrate the robustness of our method, we repeated the simulation for 10 times in each case, and calculate the mean and variance of the absolute biases of  $p$ -values of both hybrid permutation and random permutation, treating the  $p$ -values of exact permutation as gold standard. In most cases, hybrid permutation is less biased and more stable than random permutation (Table 2), which demonstrates the robustness and accuracy of our method.



**Figure 2.** Cumulative distribution functions (CDF) of exact permutation (in red) and Pearson distribution fitting (in blue) results. Left Column: for Case 1 (top) and Case 2 (bottom); Middle Column: for Case 3 (top) and Case 4 (bottom); Right Column: for Case 5 (top) and Case 6 (bottom).



**Figure 3.** Facial differences between Asian male and white male. Locations in red on the 3D surface denote significant face shape differences (significance level  $\alpha=0.01$  with false discovery rate control).

In experiment 2 on 3D face analysis, we choose 10 Asian males and 10 white males out of the USF face database [10] to calculate their differences. Results from our algorithm in Fig. 3 show that significant differences occur at eye edge, nose, lip corners and cheeks. They are consistent with anthropology findings [11] and suggest the discriminant surface regions for ethnic group recognition.

## 4. CONCLUSION

We propose and develop a statistical 3D face analysis method by using novel hybrid randomization tests where the permutation distributions are approximated through Pearson distributions. The proposed method reduces computation cost considerably without loss of accuracy. General and analytical formulations for the moments of permutation distribution are derived for multivariate test statistics. The proposed hybrid strategy takes advantage of nonparametric permutation tests and parametric Pearson distribution approximation to achieve both accuracy/flexibility and efficiency. Note that the theoretical derivations described in this work are general and can be applied to any linear test statistics on multivariate data, not limited on the test statistics demonstrated in this paper. For future work, we plan to apply our hybrid permutation tests to multidimensional data for feature selection and discriminate analysis such as face gender and facial expression recognition.

## 5. REFERENCES

- [1] Hu, Y., Zhang, Z., Xu, X., Fu, Y., and Huang, T. S., "Building large scale 3D face database for face analysis," *International Workshop on MCAM'07*, 343-350, 2007.
- [2] Lachenbruch, P., and Goldstein, M., "Discriminant analysis," *Biometrics* 35, 69-85, 1979.
- [3] Fu, Y., and Zheng, N., "M-Face: An appearance-based photorealistic model for multiple facial attributes rendering," *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, 16, 830-842, 2006.
- [4] Blanz, V., and Vetter, T., "A morphable model for the synthesis of 3D faces," in Proc. *SIGGRAPH'99*, pp. 187-194, 1999.
- [5] Zhou, C., Park, D. C., Styner, M. and Wang, Y. M., "ROI constrained statistical surface morphometry," *IEEE ISBI 2007*, 1212-1215, Washington D.C, 2007.
- [6] Zhou, C., and Wang, Y. M., "Hybrid permutation test with application to surface shape analysis," *Statistica Sinica*, in press, 2008.
- [7] Nichols, T.E., and Holmes, A.P., "Nonparametric permutation tests for functional neuroimaging: A primer with examples," *Human Brain Mapping*, 15, 1-25, 2001.
- [8] Pantazis, D., Leahy, R.M., Nichols, T.E., and Styner M., "Statistical surface-based morphometry using a non-parametric approach," *IEEE ISBI*, 1283-1286, 2004.
- [9] Good, P., *Permutation, Parametric and Bootstrap Tests of Hypotheses*, 3rd ed., Springer, New York, 2005.
- [10] USF DARPA Human ID 3D face database, Courtesy of Prof. Sarkar, Sudeep, University of South Florida, Tampa, FL.
- [11] McNeill, D., and McNeill, D., *The Face*, Back Bay Press, 1998.