International Joint Conference on Biometrics (IJCB 2017)
October 1-4, 2017 Denver, USA

# Multi-Dim: A Multi-Dimensional Face Database Towards the Application of 3D Technology in Real-World Scenarios

Feng Liu, Jun Hu, Jianwei Sun, Yang Wang and Qijun Zhao
College of Computer Science, Sichuan University, Chengdu, China (http://www.scubrl.org/)

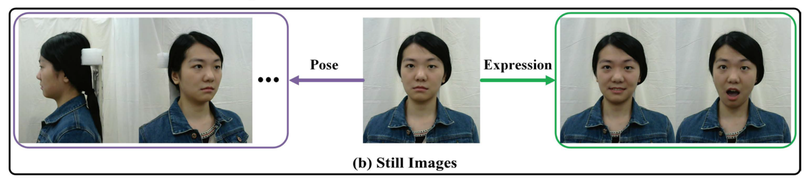
## 1. Introduction

- Despite the promising improvement achieved by 3D face technology, it is still hard to thoroughly evaluate the performance and effect of 3D face technology in real-world applications. This is due to the lack of **benchmark databases that contain both high precision full-view 3D faces and their 2D face images/videos under different conditions**.
  - 2D face image databases (e.g. FERET, CASIA, Multi-PIE, LFW, IJB-A) do not have ground truth 3D face data. Consequently, the evaluation only reports the face recognition accuracy, but can not give detailed assessments of 3D face alignment/reconstruction.
  - 3D face databases (e. g. Bosphorus, BU3DFE and BU4DFE) can be used to evaluate the 3D face alignment/reconstruction accuracy, but they usually do not have 2D face images/videos in the wild.
  - Hybird 2D/3D face databases (e.g. FRGC v2.0, UHDB11/31 and MICC) are limited in data diversity, e.g., lacking of surveillance video images, or lacking of acted still images with expression and pose variations.
- We present a multi-dimensional face database (namely **Multi-Dim**).
  - The Multi-Dim database incudes both 2D and 3D face data, both still images and surveillance videos, both acted and spontaneous data.
  - We believe that such a database will advance the 3D-based unconstrained 2D face recognition and related techniques towards applications in real-world.

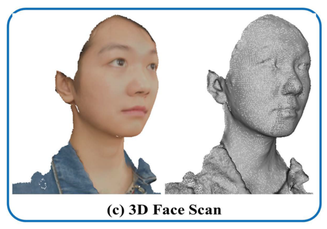
## 2. Multi-Dim Face Database

- Overview
  - 124 individuals (82 males, 42 females)
  - Each individual contributes **one** textured 3D face scan, **one** high definition neutral and frontal 2D face image, still face images of **seven** poses and **three** expressions, and with or without glasses, and **four** surveillance video clips.


(a) High Definition Photo

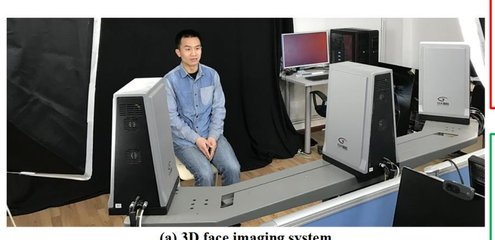


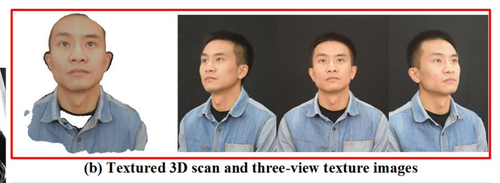
(b) Still Images

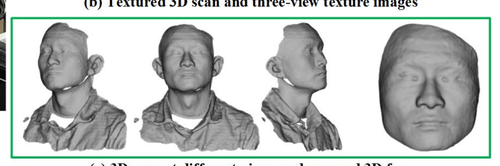

(c) 3D Face Scan

(d) Surveillance Video Clips

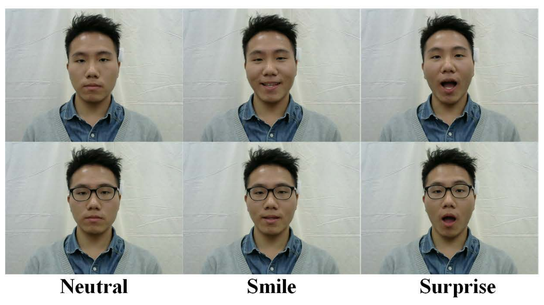
- 3D Face Scans [3]
  - Both raw and cropped 3D face scans.

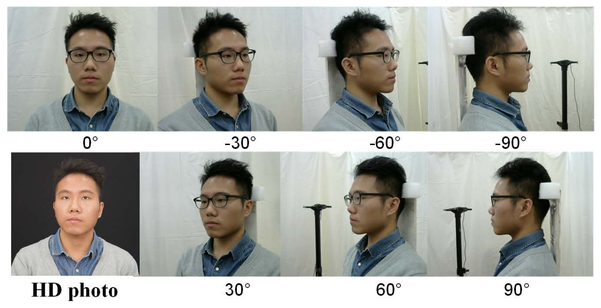

(a) 3D face imaging system


(b) Textured 3D scan and three-view texture images

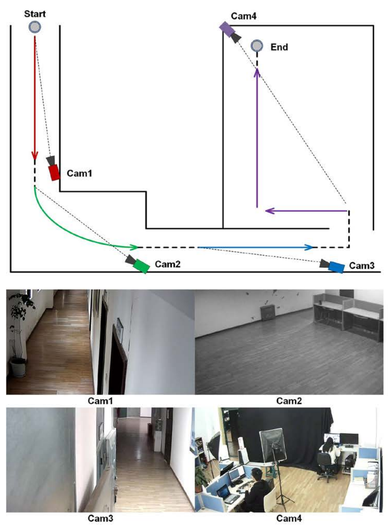

(c) 3D scan at different views and cropped 3D face

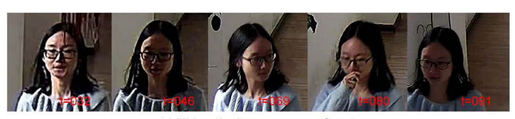
- Still Images
  - High definition photos, multi-view and multi-expression images




- Surveillance Videos
  - The subject is required to walk naturally through the regions covered by four cameras, resulting in images with varying qualities.





(a) Video clips by camera one (Cam1)

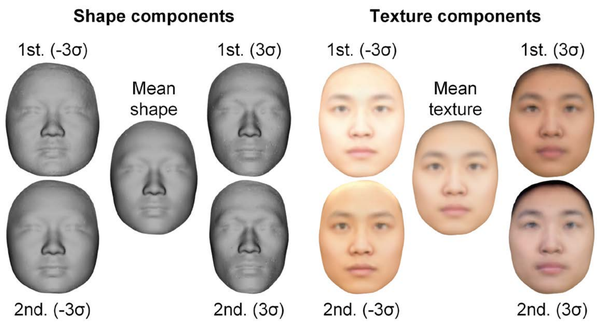

(b) Video clips by camera two (Cam2)

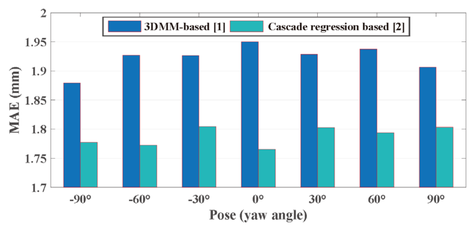

(c) Video clips by camera three (Cam3)


(d) Video clips by camera four (Cam4)

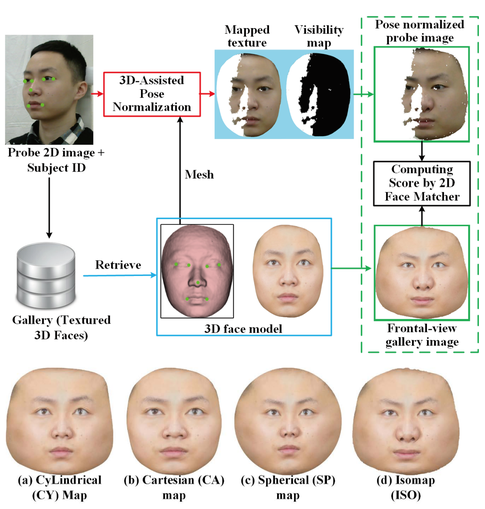
## 3. Evaluation Experiments

- Constructing 3DMM and 3D Face Reconstruction




- 3D-Assisted Pose Normalization for Face Verification

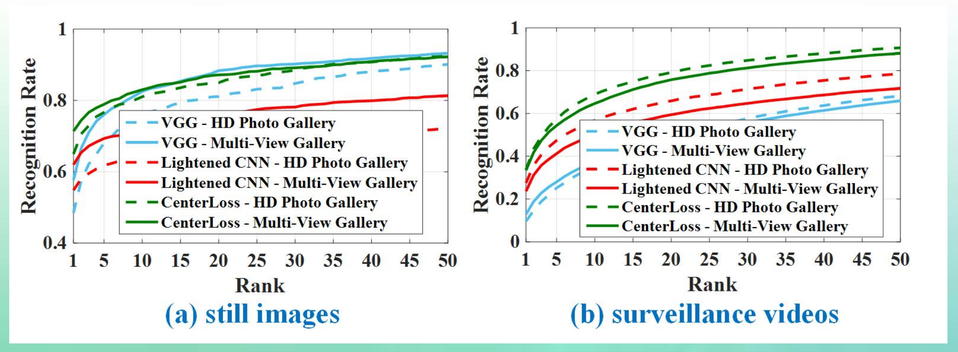




(a) CyLindrical (CY) Map (b) Cartesian (CA) map (c) Spherical (SP) map (d) Isomap (ISO)

| 2D Face Matcher | Method | TAR@FAR =10% | TAR@FAR =1% | EER | AUC |
|---|---|---|---|---|---|
| VGG | Baseline | 64.24% | 44.59% | 23.53% | 87.36% |
| | CY | 78.87% | 48.62% | 14.63% | 94.17% |
| | CA | 88.69% | 55.69% | 10.73% | 96.61% |
| | SP | 76.61% | 46.64% | 14.49% | 93.48% |
| | ISO | 83.75% | 53.71% | 12.91% | 95.25% |
| Lightened CNN | Baseline | 66.93% | 56.47% | 25.94% | 81.91% |
| | CY | 94.98% | 75.97% | 6.92% | 98.43% |
| | CA | 97.81% | 84.52% | 4.88% | 99.14% |
| | SP | 92.16% | 71.17% | 9.03% | 97.46% |
| | ISO | 93.00% | 74.42% | 8.05% | 97.79% |
| CenterLoss | Baseline | 77.17% | 66.71% | 17.74% | 92.30% |
| | CY | 94.06% | 78.02% | 7.49% | 98.04% |
| | CA | 93.50% | 79.01% | 7.70% | 98.18% |
| | SP | 93.29% | 74.42% | 8.24% | 97.83% |
| | ISO | 92.23% | 72.01% | 9.69% | 97.67% |

- 3D-Rendered Multi-view Gallery for Face Identification
  - .We generate virtual face images from the enrolled 3D faces at yaw view angle of 0°，±30°，±60°，±90°. In the baseline case, the gallery is composed by HD neutral and frontal face photos only. As for the probe images, we use the still images and surveillance video frames of varying expressions and occlusion.




## 4. Conclusions and Future Work

- **Multi-Dim** enables thorough and systematical evaluations of the performance of various 3D face technology as well as their effect in improving the performance of unconstrained 2D face recognition, particularly in real-world scenarios.
- **Multi-Dim** currently has a relatively small number of subjects, and is limited in the diversity of ethnicity and age. We are continuously acquiring the multi-dimensional face data of more subjects in more sessions. We anticipate the publication of new versions of Multi-Dim (containing 500 more subjects) in the next one to two years.

[1] Zhu, Xiangyu, et al. "High-fidelity pose and expression normalization for face recognition in the wild." *IEEE CVPR*. 2015.
[2] Liu, Feng, et al. "On 3D Face Reconstruction via Cascaded Regression in Shape Space." *Frontiers of Information Technology and Electronic Engineering*. 2017.
[3] Commercial 3D face imaging system. http://www.wisesoft.com.cn