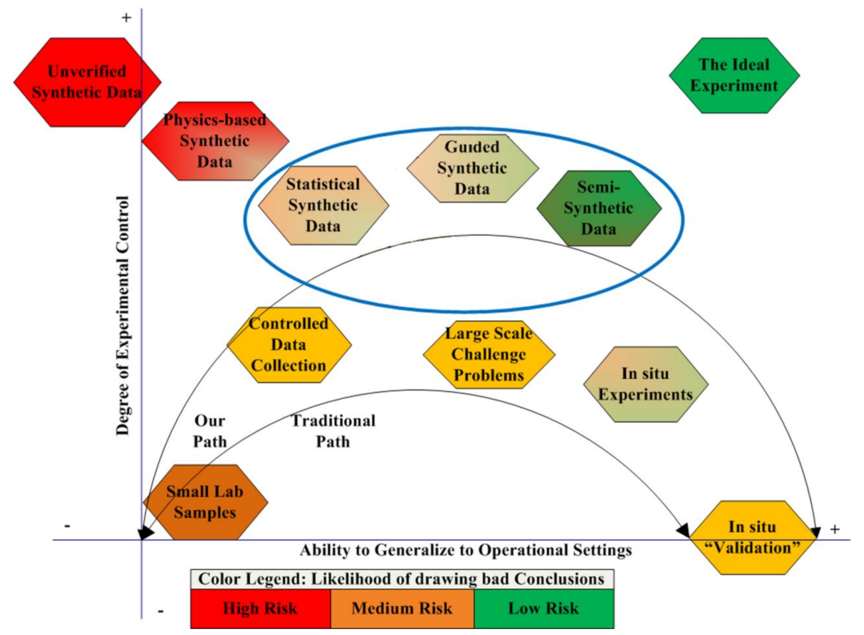
# A Taxonomy of Face-models for System Evaluation

{viyer,skirkbride,bparks,wscheirer,tboult}@vast.uccs.edu
Vijay N. Iyer, Shane. R. Kirkbride, Brian C. Parks, Walter J. Scheirer and Terrance. E. Boult

## Motivation and Data Types




Generating statistically significant size datasets for face matching system evaluation is both a laborious and expensive process.

There is a gap in datasets that allow for evaluation of system issues including:

- Long distance recognition
- Blur caused by atmospherics
- Various weather conditions
- End to end systems evaluation

Our contributions:

- Define a taxonomy of face-models for controlled experimentations
- Show how Synthetic addresses gaps in system evaluation
- Show a process for generating and validating synthetic models
- Use these models in long distance face recognition system evaluation

### Synthetic Data Types

**Unverified** – Have no underlying physical or statistical basis

**Physics -Based** – Based on structure and materials combined with the properties formally modeled in physics.

**Statistical** – Use statistics from real data/experiments to estimate/learn model parameters. Generally have measurements of accuracy

**Guided Synthetic** – Individual models based on individual people. No attempt to capture properties of large groups, a unique model per person. For faces, guided models are composed of 3D structure models and skin textures, capturing many artifacts not easily parameterized. Can be combined with physics-based rendering to generate samples under different conditions.

**Semi–Synethetic** – Use measured data such as 2D images or 3D facial scans. These are not truly synthetic as they are re-rendering's of real measured data.

Semi and Guided Synthetic data provide higher operational relevance while maintaining a high degree of control.

## Experimental Setup

### Example Models

Original Pie

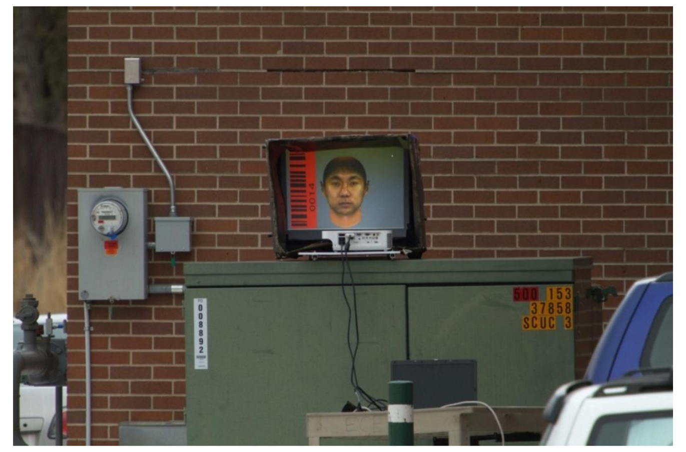
Semi-Synthetic

FaceGen
http://www.facegen.com

Animetrics
http://www.animetrics.com/products/Forensica.php

Guided-Synthetic Models

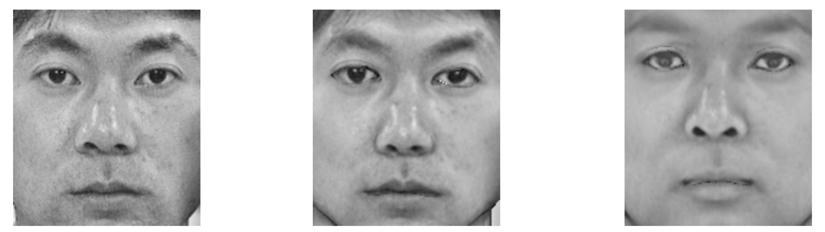
- Models generated using the well known CMU PIE [18] dataset. Each of the 68 subjects of PIE were modeled using a right profile and frontal image from the lights subset.
- Two modeling programs were used, Facegen and Animetrics. Both programs create OBJ files and textures
- Models are re-rendered using custom display software built with OpenGL, GLUT and DevIL libraries



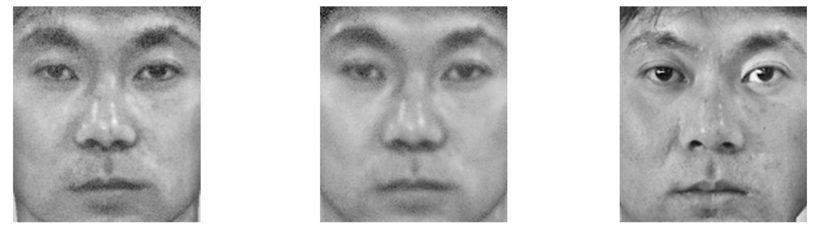
- Custom Display Box housing a BENQ SP820 high powered projector rated at 4000 ANSI Lumens
- Canon EOS 7D withd a Sigma 800mm F5.6 EX APO DG HSM lens a 2x adapter imaging the display from 214 meters

### Normalized Example Captures

Real PIE 1 | Animetrics | FaceGen



81M inside | 214M outside | Real PIE 2



- Pre-cropped images were used for the commercial core
- Ground truth eye points + geometric/lighting normalization pre processing before running through the implementation of the V1 recognition algorithm found in [1].
- Geo normalization highlights how the feature region of the models looks very similar to that of the real person.

## Results and Conclusions

Each test consisted of using 3 approximately frontal gallery images NOT used to make the 3D model used as the probe, best score over 3 images determined score.

Even though the PIE-3D-20100224A–D sets were imaged on the same day, the V1 core scored differently on each highlighting the synthetic data's ability to help evaluate data capture methods and effects of varying atmospherics. The ISO setting varied which effects the shutter speed, with higher ISO generally yielding less blur.

| Dataset | Range(m) | Iso | V1 | Comm. |
|---|---|---|---|---|
| Original PIE Images | N/A | N/A | 100 | 100 |
| FaceGen ScreenShots | N/A | N/A | 47.76 | - |
| Animetrics Screenshots | N/A | N/A | 100 | 100 |
| PIE-3D-20100210B | 81m | 500 | 100 | - |
| PIE-3D-20100224A | 214m | 125 | 58.82 | 100 |
| PIE-3D-20100224B | 214m | 125 | 45.59 | 100 |
| PIE-3D-20100224C | 214m | 250 | 81.82 | 100 |
| PIE-3D-20100224D | 214m | 400 | 79.1 | 100 |

- The same (100 percent) recognition rate on screenshots as original images validate the Anmetrics guided synthetic models and fails FaceGen Models.
- 100% recognition means dataset is too small/easy; exapanding pose and models underway.
- Expanded the photohead methodology into 3D
- Developed a robust modeling system allowing for multiple configurations of a single real life data set.
- Gabor+SVM based V1[15] significantly more impacted by atmospheric blur than the commercial algorithm

**Key References:**
**[6 of 21] R. Bevridge, D. Bolme, M Teixeira, and B. Draper. The CSU Face Identification Evaluation System Users Guide: Version 5.0. Technical report, CSU 2003**
**[8 of 21] T. Boult and W. Scheirer. Long range facial image acquisition and quality. In M. Tisarelli, S. Li, and R. Chellappa.**
**[15 of 21]** N. Pinto, J. J. DiCarlo, and D. D. Cox. How far can you get with a modern face recognition test set using only simple features? In IEEE CVPR, 2009.
**[18 of 21] T. Sim, S. Baker, and M. Bsat. The CMU Pose, Illumination and Expression (PIE) Database. In Proceedings of the IEEE F&G, May 2002.**