Development of Optical Computer Recognition (OCR) for Monitoring Stress and Emotions in Space

Nicholas Michael[1], Fei Yang[1], Peng Yang[1], Lin Zhong[1], Dimitris Metaxas[1], David F. Dinges[2]

[1]Center for Computational Biomedicine Imaging and Modeling Center (CBIM), Rutgers University, New Brunswick, NJ, USA

[2]Unit for Experimental Psychiatry, Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, PA, USA

Introduction

While in space, astronauts are required to perform mission-critical tasks on very expensive equipment at a high level of functional capability. Stressors can compromise their ability to do so, thus it is very important to have a system that can unobtrusively and objectively detect neurobehavioral problems involving elevated levels of behavioral stress and negative emotions. Computerized approaches involving inexpensive cameras offer an unobtrusive way to detect distress and to monitor observable emotions of astronauts during critical operations in space, by tracking and analyzing facial expressions and body gestures in video streams. Such systems can have applications beyond space flight, e.g., surveillance, law enforcement and human computer interaction.

Active Shape Model (ASM) Face Tracking

We have developed a framework that is capable [1] of real-time tracking of facial landmarks (e.g., eyes, eyebrows, nose, mouth), using a group of deformable statistical models of facial shape variation and local texture distribution. Our model tolerates partial occlusions as well as out of plane rotations. It can automatically recover from lost track and reinitialize.

- Facial shapes represented by (x,y) coordinates of 79 landmarks.

$$\mathbf{x} = (x_1, y_1, \ldots, x_n, y_n)^T$$

- Statistical model of permissible facial shape variation and deformation, learned by application of Principal Components Analysis

$$\mathbf{x} \approx \mathbf{x}_{mean} + \mathbf{Pb}$$

- Learn non-linear shape manifold as separate ASM models for each major pose, dynamically switching models as head.
- Use mixture of experts model to predict 3D head pose (pitch, yaw, tilt) given the tracked landmarks

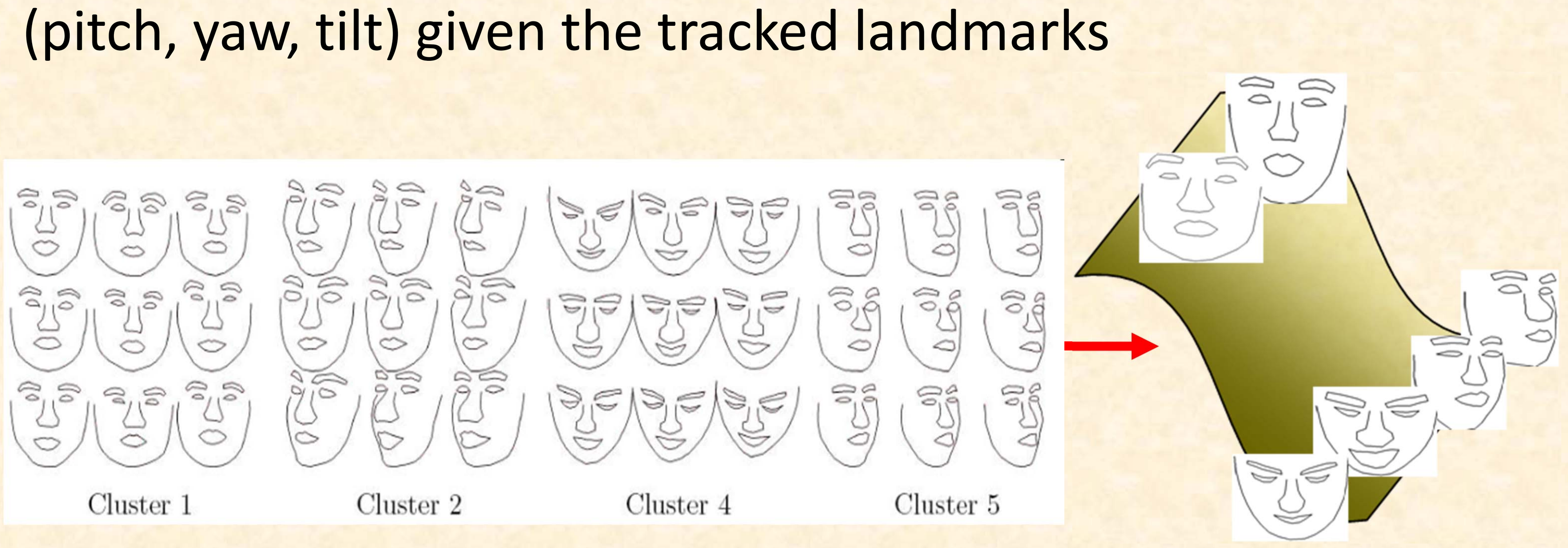

Figure 1: (Left) Labeled Face Shapes clustered by pose (Right) Learn Non-Linear Shape Manifold as ASM Shape Clusters

SUPPORT: NSBRI through NASA NCC 9-58

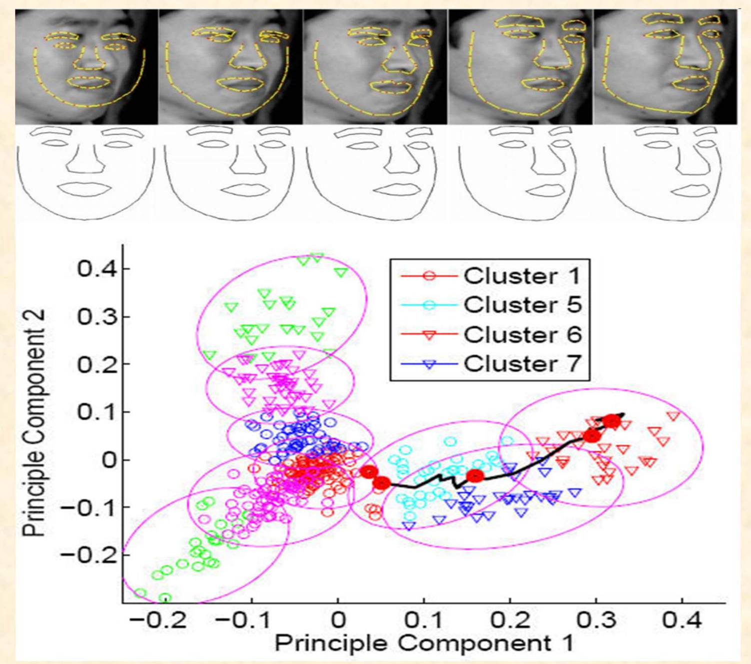

Figure 2: The overlapping between clusters ensures smooth deformation of the shape during cluster switching while tracking full profile rotations.

Figure 3: Sample processed frame showing the skin blobs detected and tracked by the skin blob tracker (blobs marked as ellipsoids).

Skin Blob Tracking

Our skin blob tracker [2] is capable of detecting and tracking regions of skin color in real time (i.e., head and hands). It is automatically initialized with a generic skin color model (which is learned offline), dynamically learning the specific color distribution of the current subject for online adaptive tracking. Detected blobs are filtered online, both in terms of motion and shape, using Eigen-space analysis and temporal dynamical models to prune false detections (see Figure 3 for an illustration).

Recognizing Facial Expressions

In order to train classifiers for facial expression recognition we need training images depicting facial expressions of the 6 universal emotions (i.e., sadness, anger, disgust, fear, surprise, happiness) as well as neutral facial expressions. Images are cropped based on eye and nose positions, for image normalization and dimensionality reduction. We then compute Haar features and encode dynamic information of neighboring frames. For training the classifier for a particular expression, we first form a set of positive instances and a set of negative instances of that class. We then use Adaboost learning to select discriminative features, hence training the classifiers [3]. All 7 classifiers are evaluated on each test frame and the max score is used as its label.

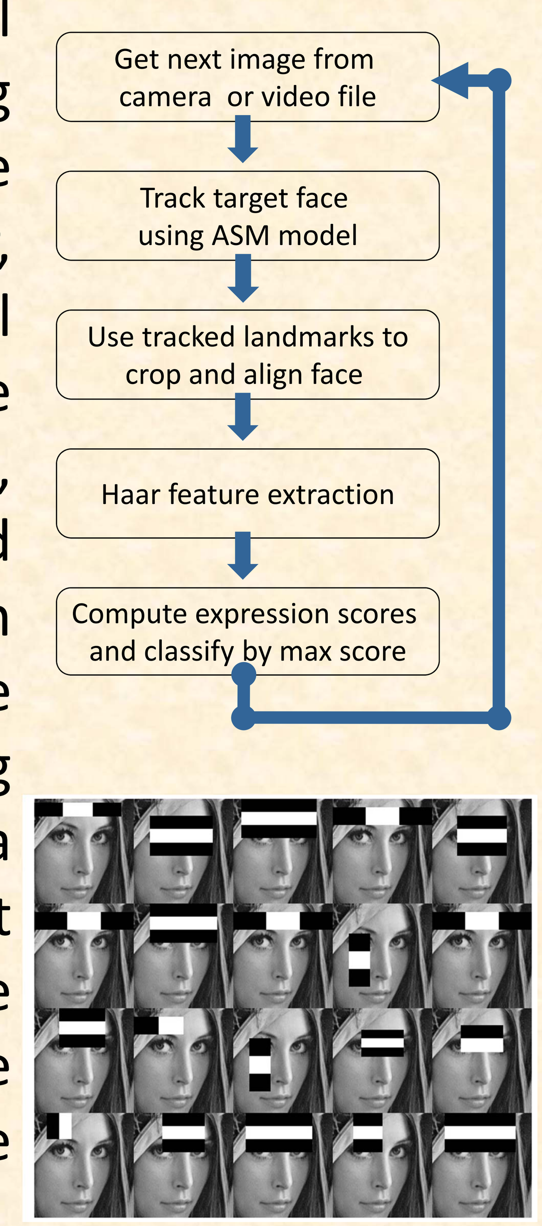

Figure 4: (Top) Expression recognition flowchart, (Bottom) Illustration of extracted Haar features.

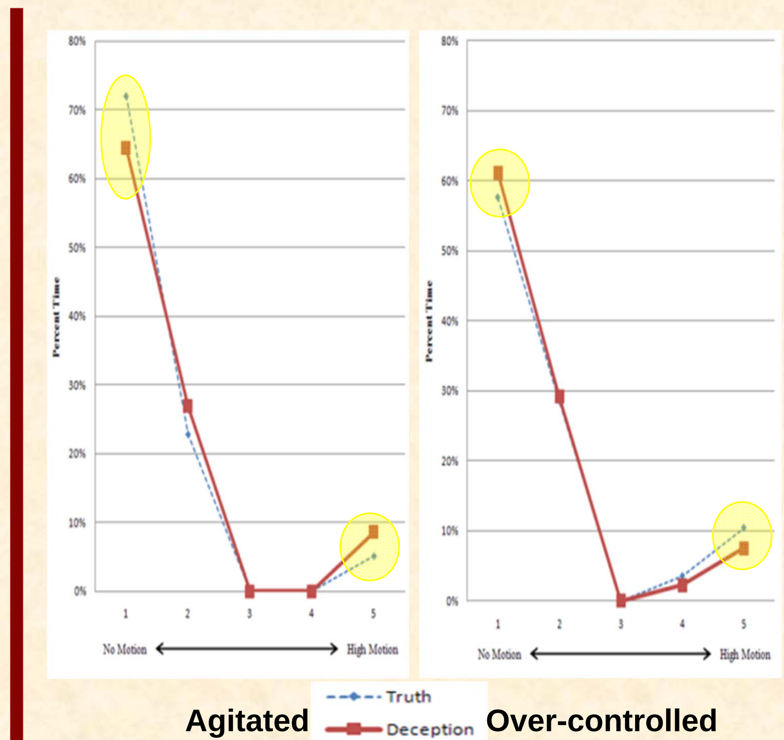

Figure 5: (Above) Graphical representation of motion profiles of two different subjects exhibiting over-controlled and agitated behaviors [2]

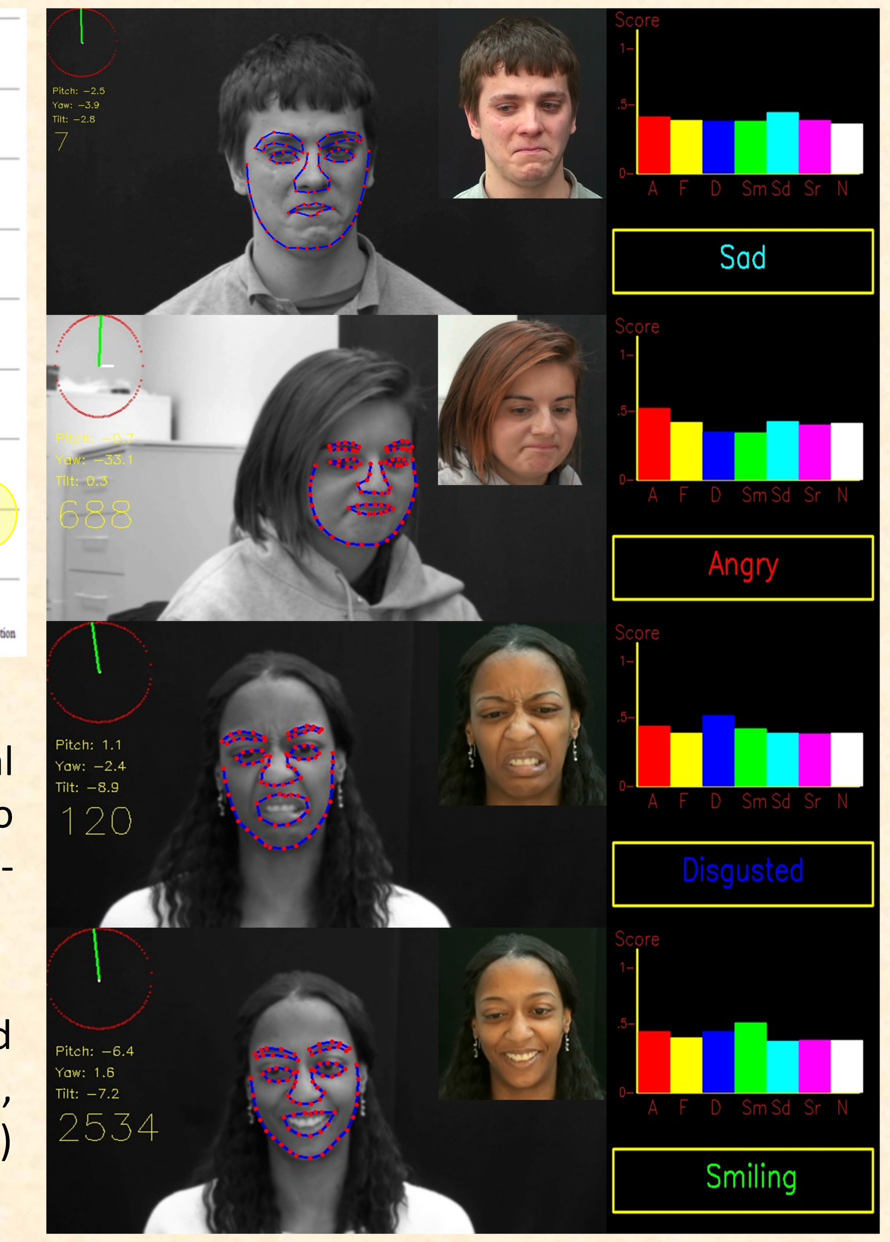

Figure 6: (Right) Sample processed frames showing tracked landmarks, estimated head pose (top left corner) and predicted facial expression scores.

Experimental Results

We recorded 147 subjects in 12-question interviews following a mock-crime, tracking their facial expressions and body gestures (via ASM and skin blob tracking) to recognize deception by detecting *subject-specific* stressed and relaxed behaviors. Leave One Out Cross Validation using Nearest Neighbor classification discriminated truthful and deceptive responses with an average accuracy of 81.6%.

We used 352 training sequences (96 subjects, 3-6 samples per expression each) from the CMU database [4] to train 7 expression classifiers (table shows accuracy of 5-fold CV). We also used 240 neutral expression from 10 subjects (see sample predictions in Figure 6). Evaluation is shown in the poster by Moreta et. al.

Expression	Recognition Rate
Angry	87%
Disgust	61%
Fear	72%
Happiness	94%
Sadness	59%
Surprise	90%
Mean	77%

Conclusions

While validation of the OCR output is underway, preliminary results suggest that OCR technology will serve as an on-line detection system to help identify risks to performance posed by changes in mood and emotion during spaceflight.

References

1. A. Kanaujia et. al. Tracking Facial Features Using Mixture of Point Distribution Models. ICVGIP 2006.
2. N. Michael et. al. Motion Profiles for Deception Detection Using Visual Cues, ECCV 2010.
3. P. Yang et. al. Boosting Encoded Dynamic Features for Facial Expression Recognition. Pattern Recognition Letter, 2009.
4. James Lien et. al. Detection, tracking, and classification of action units in facial expression, J. Robotics Autonomous Syst. 1999.