Bimodal HCI-related Affect Recognition

Zhihong Zeng, Jilin Tu, Ming Liu, Tong Zhang, Nicholas Rizzolo, Zhenqiu Zhang, Thomas S. Huang, Dan Roth and Stephen Levinson
Beckman Institute for Advanced Science and Technology
University of Illinois at Urbana-Champaign, USA
{zhzeng, jilintu, mingliu1, tzhang1, zzhang6, huang, sel}@ifp.uiuc.edu
{rizzolo, darn}@cs.uiuc.edu

ABSTRACT
Perhaps the most fundamental application of affective computing would be Human-Computer Interaction (HCI) in which the computer is able to detect and track the user's affective states, and make corresponding feedback. The human multi-sensor affect system defines the expectation of multimodal affect analyzer. In this paper, we present our efforts toward audio-visual HCI-related affect recognition. With HCI applications in mind, we take into account some special affective states which indicate users' cognitive/motivational states. Facing the fact that a facial expression is influenced by both an affective state and speech content, we apply a smoothing method to extract the information of the affective state from facial features. In our fusion stage, a voting method is applied to combine audio and visual modalities so that the final affect recognition accuracy is greatly improved. We test our bimodal affect recognition approach on 38 subjects with 11 HCI-related affect states. The extensive experimental results show that the average person-dependent affect recognition accuracy is almost 90% for our bimodal fusion.

Categories and Subject Descriptors: H.5.2 [Information Interfaces and Presentation]: User Interfaces---interaction styles, user-centered design; K.3.1 [Computers and Education]: Computer Uses in Education---collaborative learning

General Terms: Design, Performance

Keywords: affect recognition, emotion recognition, multimodal human-computer interaction, HCI, affective computing

1. INTRODUCTION
Affects play a critical role in rational decision-making, in perception, in human interaction, and in human intelligence. These facts inspired the very challenging research field "affective computing" which aims at enabling computers to express and recognize affect [11]. Perhaps the most fundamental applications of affective computing would be human-computer interaction in which the computer is able to detect and track users' affective states, and to initiate communications based on this knowledge, rather than to simply respond to users' commands.

The work in this paper is motivated by the ITR project [14] which is to contribute to the development of multimodal human-computer intelligent interaction environment. The ideas and tools resulting from the research are evaluated in an educational environment. This test bed concerns the learning of math and science for upper elementary and middle school children. It is an learning environment in which some children are energetically engaged and emotionally expressive while other children are timid and hesitant. The project will focus on using proactive computing to achieve two ends. The first one is helping children to participate in the activity who are timid, fearful or generally disinterested in science or construction activities, with the goal of helping them gain a vision of the benefits of these activities that will sustain a continued interest. The second one is encouraging children who already feel comfortable with the learning environment to explore further and to accept greater challenges. Thus, we will seek ways for the proactive computer to recognize the user's affective states (e.g. interest, boredom, frustration and confusion) and to apply the corresponding strategy (e.g. encouragement, transition/guidance, and confirmation). In this way, children will learn some basic skills, and their interest and creativity will be stimulated.

The psychological study [12] indicated that judging someone's affective states, people mainly rely on facial expressions and vocal intonations. Thus, affect recognition should inherently be the issue of multimodal analysis. From the perspective of engineering, even with improvements, the modalities of facial expression and prosody will undoubtedly be very fallible. The primary aim of our work in this paper is to combine cues from these modalities so that the affective state of a person can be inferred more accurately.

In this paper, we present our effort toward audio-visual HCI-related affect recognition. With HCI applications in mind, we take into account some special affective states which indicate user's cognitive/motivational states. Facing the fact that a facial expression is influenced by both an affective state and speech content, we apply a smoothing method to extract the information of the affective state from facial features. In our fusion stage, a voting method is applied to combine audio and visual modalities so that the final affect recognition accuracy is greatly improved. We test our bimodal affect recognition approach on 38 subjects with 11 affective states. The extensive experimental

ICMI'04, October 13–15, 2004, State College, Pennsylvania, USA.

results show that the average person-dependent affect recognition accuracy is almost 90% for our bimodal fusion.

2. RELATED WORK

Automatic affect analysis has attracted much interest from researchers in various research fields. Most current approaches to computer-based analysis of human affective states are uni-modal--information processed by the computer system is limited to either face images or the speech signals, and relatively little effort in multimodal affect analysis has been made. According to the latest overview of automatic affect recognition [1], only four reports [2-5] of bi-modal affect recognition are found. Thus, while the recent research and technology advances make multimodal analysis of human affective states tractable, progress toward this direction is only just beginning.

Compared with these four previous reports of bimodal affect recognition listed in [1], the progress in this paper includes:

1) 11 affective states are analyzed, especially including 4 HCI-related affective states (confusion, interest, boredom, and frustration). [2-5] only analyzed 5-6 basic emotions.

2) 38 subjects are tested. The numbers of subjects in [2-5] are at most five. Thus, the generality of their algorithm is not guaranteed.

3) A voting method is applied for bimodal fusion. [3-4] applied rule-based methods for combining two modalities. [2] applied the single-modal methods in a sequential manner for bimodal recognition. [5] used weighted intensity summation. It is not clear whether their rules or methods are suitable for more subjects.

4) Facing the fact that a facial expression is influenced by both an affective state and speech content, we apply a smoothing method to extract the information of the affective state from facial features. [3-5] did analyze this problem. [2] simply applied the single-modal methods in a sequential manner to avoid this problem.

3. DATABASE

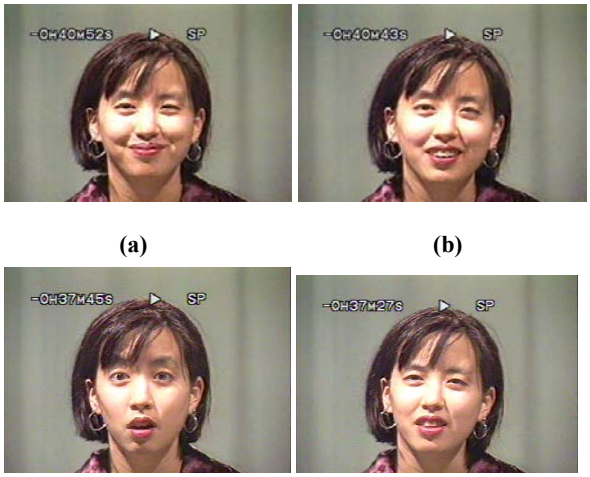
Most of the existed affect databases are for face-only affect recognition or prosody-only affect recognition. In these four previous bimodal affect recognition reports [2-5], the datasets used were so small that the generality of their methods are not guaranteed. In addition, their methods only detected 5-6 basic emotional states which are not directly related with human computer interaction. In the experiments of our ITR project, we noticed that facing the computer tutor subjects seldom expressed these basic emotions (i.e. happiness, sadness, fear, surprise, anger, disgust). Actually, detecting some special affects, including interest, boredom, confusion and frustration, is very important for the computer tutor to interact with users. We call these affects as HCI-related affects. These affects indicate the cognitive/motivational states of the subjects' learning. They provide information about whether the subject is engaged or whether the subject is having difficulties during the learning activities.

Therefore, a large-scale database was collected [15] in our group that is more related to the human-computer interaction. 11 affect categories were used which include 7 basic affects (i.e. happiness, sadness, fear, surprise, anger, disgust, and neutral), and 4 HCI-related affects (i.e. interest, boredom, confusion and frustration).

The subjects consist of mostly graduate and undergraduate students in various fields. Some staff and faculty members also volunteered to participate. Although the subjects displayed affect expression on request, minimal instruction was given to the subjects. In particular, no instruction on how to portray the affects was given. They were simply asked to display facial expressions and speak appropriate sentences. Each subject was required to pose pure facial expression without speech three times, followed by facial expression with speech three times, and then pure facial expression three more times.

In the dataset, we found that facing a camera, subjects look more natural to express affects with appropriate sentences than without them. In particular, pure facial expression is difficult to display subtle differences among the 4 HCI-related affects (confusion, interest, boredom and frustration). On the other hand, we also noticed, speaking reduces the discriminability of facial expressions among different affective states. That is because speaking influences facial expressions. Figure 1 illustrates the comparison between pure facial expressions without speech and facial expression with speech. It shows that compared with pure facial expression in (a) and (c), the facial expressions with speech in (b) and (d) are more natural, but they are more difficult to be distinguished from each other.

Data on 50 males and 50 females were recorded. By now, we only test our approach on 38 subjects (24 females and 14 males) of the 100 subjects. That is because our face tracker requires that the first frame of the video must be neutral facial expression with closed mouth. Only 38 videos meet this requirement. Other videos need a little edition before they are used. Figure 3 shows some examples of the videos used in our experiment.



(a) (b) (c) (d)

Figure 1. (a) Happy expression without speaking; (b) Happy expression with speaking; (c) Surprise expression without speaking; (d) Surprise expression with speaking

4. FACIAL FEATURE EXTRACTION

A tracking algorithm called Piecewise Bezier Volume Deformation (PBVD) tracking [16] is applied to extract facial features in our experiment.

Each facial expression is represented as a linear combination of 12 predefined Action Units (AU). These action units are listed in Table 1. A facial expression is represented as the formula (3.1)

$$V = BDP = B[D_1 D_2 ... D_{12}] \begin{bmatrix} p_1 \\ p_2 \\ \cdots \\ p_{12} \end{bmatrix} \quad (3.1)$$

where B is constructed by Bezier basis functions, $D_i (i = 1,...,12)$ corresponds to ith AU, and the $p_i (i = 1,...,12)$ corresponds to the magnitude of ith AU deformation. The overall motion of the head and face is represented as

$$R(V_0 + BDP) + T \quad (3.2)$$

where R is the 3D rotation matrix, T is the 3D translation matrix, and V_0 is the initial neutral face model.

The face tracker assumes that the expression of the initial frame is neutral with closed mouth where all magnitudes corresponding to the 12 AUs have the value of zero. During tracking procedure, all of the magnitudes corresponding to AUs are measured in relation to this initial frame. That means the tracking results are very sensitive to the initial frame.

After the initialization in which the face model is fitted to the face at the first image in the manual or automatic way [10], an optical flow method is applied to track these AU movements as facial features. The face tracker outputs 30 frames per second at the off-line condition. Figure 2 shows one example of tracking results.

Figure 2. One example of tracking result

In our dataset, we notice two aspects of facial expression with speaking: on one hand, speech signals provide the affect information which complements facial expression; on the other hand, speaking influences facial expression so that the discriminability of facial expressions among different affects decreases. The movements of facial features are related to both affective states and content of speech. Especially, the mouth features corresponding to AU1-AU6 are dependent more on speech content than on affective states. That causes decrease of the performance of face-only affect recognition. If we could obtain speech content and corresponding lip movement by audio-visual speech recognition, it is possible to separate the influence of speech on facial expressions. However, the study [8] shows that the accuracy of automated speech recognition, which is about 80% to 90% for neutrally spoken speech, tends to drop to 50% to 60% for emotional speech.

Table 1. 12 action units used in our face tracker

AU number	Description
1	Vertical movement of the center of upper lip
2	Vertical movement of the center of low lip
3	Horizontal movement of left mouth corner
4	Vertical movement of left mouth corner
5	Horizontal movement of right mouth corner
6	Vertical movement of right mouth corner
7	Vertical movement of right brow
8	Vertical movement of left brow
9	Lifting of right cheek
10	Lifting of left cheek
11	Blinking of right eye
12	Blinking of left eye

In this paper, we apply a smoothing method to reduce the influence of speech on facial expression to some extent, based on the assumption that the influence of speech on face features is temporary, and the influence of affect is relatively more persistent.

The smoothing features are calculated by the following formula

$$\bar{x}_{it} = \frac{1}{n} \sum_{j=t-n}^{j=t} x_{ij}$$

where x_{ij} is the ith original facial features at time j, $\bar{x}_{it}$ is the ith smoothing features at time t, and i=1,..,12 for 12 face features. In our experiment, we use 10 frames (i.e. n=10) as the length of the smoothing window.

We tested the method in our experiment. Compared with non-smoothing results, our smoothing method improves face-only average recognition accuracies by 18.55% for females and 15.44% for males.

5. PROSODIC FEATURE EXTRACTION

We use three kinds of prosody features for affect recognition: logarithm of energy, syllable rate, and two pitch candidates and corresponding scores. The log energy is computed by

$$E = \log(\sum_{i=1}^{N} x_i^2)$$

where N is the frame length. x_i is the i th signal in that frame. For pitch extraction, an autocorrelation based pitch detector is used to extract two candidates of pitch frequency. The mathematical definition of the autocorrelation function is shown as follows:

$$Xor_p = \sum_{i=1}^{N} x_{i+p} x_i \ .$$

where x_{i+p} is the $(i+p)$ th signal in that frame. The autocorrelation of periodic signal is also periodic. As the signal lag to the length of one period, the autocorrelation increases to the maximum, the first peak (excluding 0 tag) indicates the period of the signal. So the pitch is detected by

$$P_1 = \arg \max_{P_{MIN} \le p \le P_{MAX}} Xor_p$$

where P_{MIN} is the possible minimum pitch, P_{MAX} is the possible maximum pitch. In this paper, the search range for pitch is set to be 50~1000Hz. In addition to the pitch with the maximum autocorrelation score, the pitch P_2 with the second maximum of autocorrelation score is also chosen as a pitch candidate. Also, the autocorrelation scores are treated as features to detect whether the frame is a vowel. The syllable rate is computed by following formula

$$\frac{\#syllables}{duration}$$

where duration is the segment duration (0.5 second). In order to detect the numbers of syllables in the segments, a threshold-based speech detection method is used to detect the syllables in the signal. In detail, the frame is considered as speech if the following condition is satisfied:

$$E > EnThres(50.0) \wedge Xor_{p_1} > XorThres(0.5)$$

$$\wedge Xor_{p_2} > XorThres(0.5) \wedge |\frac{P_{i1}}{P_{i2}} - 2| < 0.2$$

where E is the log energy of one frame, Xor_{p_1} and Xor_{p_2} are the autocorrelation scores of the two pitch candidates, P_{i1} is the larger value of P_1 and P_2, P_{i2} is the smaller one. When this condition is satisfied, the frame is considered as speech, otherwise it's not speech. After speech detection, we can count the number of speech segments and compute the syllable rate as one dimension of prosody features. The prosody modality in our experiment can output 92 frames per second in real-time condition.

6. SINGLE MODAL AFFECT CLASSIFICATION

Affect recognition can be regarded as a multi-class classification problem. We applied SNoW (Sparse Network of Winnow) [6][9] to build two affect classifiers individually based on face-only and prosody-only features.

SNoW is a learning architecture framework that is specifically tailored for learning in the presence of a very large number of features. It has been successfully applied in several applications in natural language and visual processing domains.

In the SNoW architecture, each class label (i.e. target node) is represented as a linear function over the feature space. In the learning procedure, several update rules can be used, including winnow, perceptron and Naïve Bayes. In this paper, we choose Naïve Bayes as the update rule. In the case of Naïve Bayes, a feature's weight at a target node in the network is simply logarithm of the fraction of positive examples (positive with respect to the target) in which the feature is active. Predictions are done via winner-take-all policy.

6.1 Face-only Affect Classification

Face features extracted at 30Hz from the face tracker are the magnitudes of 12 AU which are measured in relation to this initial frame. They are floating-point numbers around zero. As the input of SNoW, each of them is discretized into 100 uniform bins. In other words, the original floating-point features are actually transformed to binary features in a higher dimensional space. In addition, conjunction features which are linear combinations of the binary features are used. In this way, the classes which are not linearly separable in the original space could become more linearly separable in the transformed higher dimensional space.

Finally, using SNoW, the face-only affect recognition accuracies in our experiment are 56.57% for males and 54.62% for females.

6.2 Prosody-only Affect Classification

The features extracted at 92Hz from prosody modality include six features, including one energy, two pitch candidates and corresponding autocorrelation scores, and one syllable rate.

Similarly, we transform these floating-point prosody features into binary features in a higher dimensional space for SNoW input. We chose 100 bins for discretizing one energy, and 1000 bins for discretizing two pitch candidates and one syllable rate. Because the autocorrelation scores could be viewed as confidence scores of pitch candidates, they are treated as feature strengths in the SNoW network corresponding to these two pitch candidates, and kept as floating-point numbers.

Finally, using SNoW, the prosody-only affect recognition accuracies in our experiment are 46.84% for males and 44.26% for females.

7. BIMODAL FUSION

Affect recognition accuracy based on a single modality is not good enough in our experiment. In this section, we discuss how to combine these two modalities to achieve better recognition performance.

The important problem in multimodal fusion which we have to face is asynchrony of different modalities. In our work, the face modality outputs one set of facial features about every 33 milliseconds, while prosody modality outputs one set of prosody features about every 10 milliseconds. This asynchrony makes multimodal fusion at the feature level difficult. Thus, we apply decision-level fusion which combines multiple single-modality

classifier outputs in a parallel architecture for final affective state classification. In this paper, we apply a voting method to combine the classification outputs from face and prosody modalities. In our experiment, every classification result in a one-second interval from the two modalities is treated as one vote. In every one-second interval, the system compute the vote number for every affective state, and the affective state with the largest number of votes will be regarded as the final classification result. We choose one second because this time is appropriate for affect expression judgment according to the study [7].

Voting in a one-second interval increases the classification accuracy of the affect recognizer. Logically, if the frame-level classification accuracy is above random, the possibility that the correct label wins in longer-level voting could increases. In addition, the more votes (i.e. frames) there are in a one-second interval, the more accurate the final estimate of affective states. In our experiment, although the frame-level recognition accuracies of prosody-only and face-only modalities are not good enough (only achieve about 50%), they are far larger than random (100/11=9.09%). Obviously, the chance that the correct label wins in a one-second interval is greatly increased. As a result, final affect recognition performance of bimodal classifier is tremendously improved. The average affect recognition accuracies of our bimodal fusion increase to 90.74% for males and 89.34% for females.

8. EXPERIMENTAL RESULTS

We test our person-dependent affect recognition algorithm on 38 subjects (24 females and 14 males) in the above-mentioned dataset. For every subject, the data is divided into two half parts. One half of every subject's frames are selected as training data, and the other half as testing data.

In order to understand the contributions to the final recognition improvement individually from smoothing, single modalities, bimodal fusion and different time levels, we did experiments whose results are illustrated in Chart 1. In this chart, blue and yellow bars present male and female results respectively. The vertical axis represents average affective recognition accuracy (percentage). And the chart includes:

1) The first column: face-only frame-level affect recognition without smoothing
2) The second column: face-only frame-level affect recognition with smoothing.
3) The third column: face-only affect recognition by voting at the one-second level.
4) The fourth column: prosody-only frame-level affect recognition.
5) The fifth column: prosody-only affect recognition by voting at the one-second level.
6) The sixth column: bimodal affect recognition by fusion at the one-second level

From the data, we summarize the following contributions to bimodal affect recognition:

1) The smoothing method increases face-only frame-level average accuracies by 18.55% for female and 15.44% for male
2) Voting at the one-second level improves face-only average accuracies by 19.14% for female and 19.13% for male
3) Voting at the one-second level improves prosody-only average accuracies by 28.34% for female and 27.78% for male
4) Fusion improves bimodal average accuracies by 16.66% for female and 15.58% for male, compared with single-modal one-second-level classifiers.

As a whole, our experimental results show that the average recognition accuracy is about 56% for our face-only frame-level classifier, about 45% for our prosody-only frame-level classifier, but almost 90% for our bimodal fusion at the one-second level. The test time is about 8.7 seconds for every one and every affect. Extensive experimental results prove the effectiveness of our approach.

The average confusion matrix for our bimodal classifier is presented in Table 2. The analysis of the confusion matrix shows that the neutral state is the state of highest recognition accuracy, and is also the state with which the other affective states are most confused. One reason behind this result is that every affective expression in our dataset started from the neutral state, and ended at the neutral state. Thus, the first and last few frames of each affect are very close to neutral states.

In addition, anger, boredom and happiness have high recognition accuracies while sadness, interest and frustation have low accuracies. Contrary to the neutral state, the frustrated state is the state of lowest recognition accuracy, and is also the state with which other affects are most unlikely to be confused.

9. CONCLUSION

With an automatic affect recognizer, the computer can respond appropriately to the user's affective state rather than simply responding to user commands. In this way, the computer would become more natural, persuasive, and friendly. Especially, the learning environment in our ITR project involves a host of associated affective responses. The proactive computing is to help the children learn some basic skill, and stimulate their interest and creativity by watching and responding to their affective states. The human multi-sensor affect system defines the expectation of multimodal affect analyzer.

In this paper, we introduce our effort toward multimodal affect recognition. Compared with previous studies listed in [1], our progresses include a smoothing method to reduce the detrimental influence of speech on face-only affect recognition, a voting method for bimodal fusion, and extensive testing on more subjects and more affective states.

Multimodal recognition of human affective states is an extremely challenging problem and is largely unexplored. In this paper, only person-dependent experiments were done. We are exploring person-independent affect recognition which is independent of variability in subjects' physiognomy, sex and age. In addition, our method to reduce the detrimental influence of speech on face-only affect recognition is not sophisticated. We believe, the performance of the face-only modality should have much space to be improved. Our algorithm was only tested on the database of requested emotional expression which is argued to differ from the authentic emotional expression [13]. Thus, our next steps also

include the effort to evaluate our affect recognition algorithm on some authentic affect database.

ACKNOWLEDGEMENT

We like to thank Dr. Lawrence Chen for collecting the valuable data in this paper for audio-visual affect recognition.

REFERENCES

[1] Pantic M., Rothkrantz, L.J.M., Toward *an affect-sensitive multimodal human-computer interaction*, Proceedings of the IEEE, Vol. 91, No. 9, Sept. 2003, 1370-1390

[2] Chen, L. and Huang, T. S., Emotional expressions in audiovisual human computer interaction, Int. Conf. on Multimedia & Expo 2000, 423-426

[3] Chen, L., Huang, T. S., Miyasato, T., and Nakatsu, R., *Multimodal human emotion/expression recognition*, Int. Conf. on Automatic Face & Gesture Recognition 1998, 396-401

[4] De Silva, L. C., and Ng, P. C., *Bimodal emotion recognition*, Int. Conf. on Automatic Face & Gesture Recognition 2000, 332-335

[5] Yoshitomi, Y., Kim, S., Kawano, T., and Kitazoe, T., Effect of sensor fusion for recognition of emotional states using voice, face image and thermal image of face, in Proc. ROMAN 2000, 178-183

[6] Roth, D., Learning to Resolve Natural Language Ambiguities: A Unified Approach. AAAI '98, July 1998

[7] Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., and Taylor, J.G., Emotion Recognition in Human-Computer Interaction, IEEE Signal Processing Magazine, January 2001, 32-80

[8] Steeneken, H.J.M. and Hansen, J.H.L., Speech Under Stress Conditions: Overview of the Effect on Speech Production and on System Performance, in Proc. ICASSP, vol. 4, 1999, 2079-2082

[9] Carlson, A.J., Cumby, C.M., Rizzolo, N.D., Rosen, J.L., and Roth, D., SNoW User Manual, UIUC Tech report UIUC-DCS-R-99-210

[10] Tu, J., Zhang, Z., Zeng, Z. and Huang, T.S., Face Localization via Hierarchical Condensation with Fisher Boosting Feature Selection, In Proc. Computer Vision and Pattern Recognition, 2004.

[11] Picard, R.W., Affective Computing, MIT Press, Cambridge, 1997.

[12] Mehrabian, A., Communication without words, Psychol. Today, vol.2, no.4, 53-56, 1968

[13] Sebe, N., Lew, M., Cohen, I., Sun, Y., Gevers, T., and Huang, T.S., Authentic Facial Expression Analysis, Int. Conf. on Automatic Face & Gesture Recognition 2004.

[14] itr.beckman.uiuc.edu

[15] Chen, L.S, Joint Processing of Audio-Visual Information for the Recognition of Emotional Expressions in Human-Computer Interaction, PhD thesis, UIUC, 2000

[16] Tao, H. and Huang, T.S., Connected Vibrations: A Model Analysis Approach to Non-rigid Motion Tracking, CVPR 1998, 735-740.

Figure 3. Some examples of the videos used in our experiment

Chart 1. Average Affect Classification Accuracies in our experiment

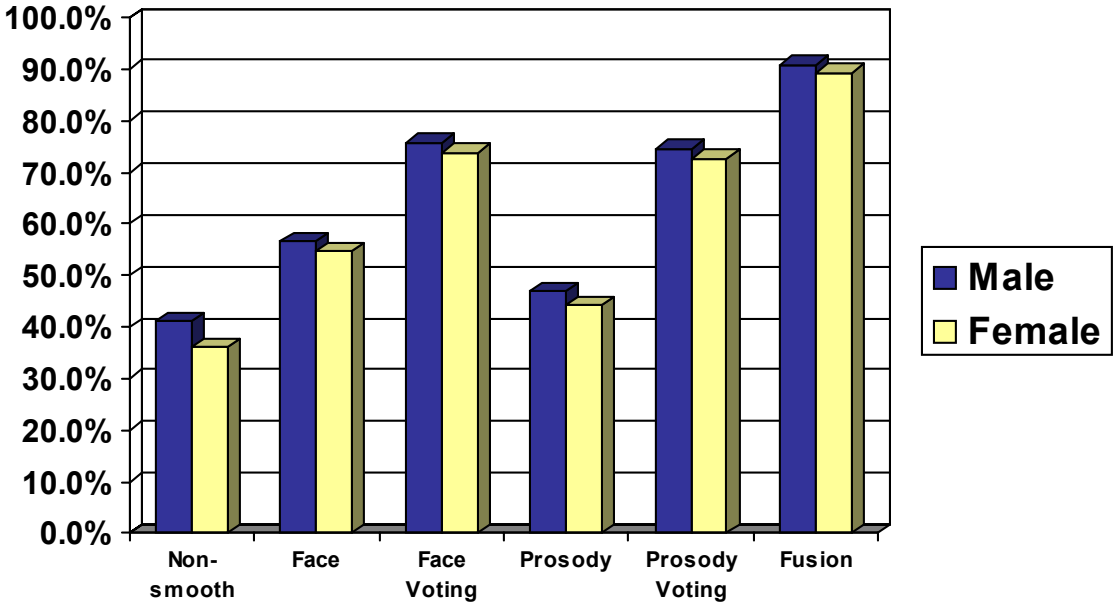


Table 2. Male-female Average Confusion Matrix for the Bimodal Classifier

Affect		Detected										
		neut	hap	surp	ang	disg	fear	sad	frust	puzz	inter	bore
	neut	98.66	0.27	0.27	0.00	0.00	0.00	0.54	0.00	0.00	0.27	0.00
	hap	4.07	89.15	0.68	0.68	0.68	0.68	1.36	0.00	0.00	1.02	1.69
	surp	4.70	0.31	88.40	1.57	0.00	1.57	1.88	0.00	0.94	0.00	0.63
	ang	1.16	0.23	0.47	94.65	1.40	0.23	1.16	0.00	0.47	0.23	0.00
Desired	disg	2.31	0.77	0.77	3.08	88.72	0.77	0.77	1.03	0.51	0.51	0.77
	fear	1.81	0.30	0.30	3.93	1.51	89.12	1.21	0.00	1.51	0.30	0.00
	sad	4.60	0.31	0.31	3.07	1.53	0.00	88.04	0.61	0.31	0.31	0.92
	frust	3.31	0.37	2.21	3.31	0.37	2.57	2.21	83.46	1.10	0.37	0.74
	puzz	1.88	0.31	0.63	2.82	1.25	0.94	0.31	0.63	88.71	1.25	1.25
	inter	2.14	0.36	1.42	3.20	1.78	1.07	1.42	0.36	1.07	85.05	2.14
	bore	2.32	0.99	0.66	1.32	0.33	0.66	2.32	0.66	0.99	0.33	89.40