Probabilistic Prediction of Student Affect from Hand Gestures

Abdul Rehman Abbasi*, Matthew N. Dailey*, Nitin V. Afzulpurkar*, and Takeaki Uno†
* Asian Institute of Technology, Bangkok, Thailand
Email: {abdulrehman.abbasi, mdailey, nitin}@ait.ac.th
† National Institute of Informatics, Tokyo, Japan
Email: uno@nii.ac.jp

***Abstract*— Affective information is vital for effective human-to-human communication. Likewise, human-to-computer communication could be potentiated by an "affective barometer" able to infer human affect using a machine vision system. For instance, during a classroom lecture, an affective barometer might provide useful feedback that a real or virtual instructor could use to improve pedagogical strategies. In this paper, we explore the feasibility of using students' unintentional hand gestures during a classroom lecture to predict their affective state. We propose a maximum a posteriori classifier based on a simple Bayesian network model. We then evaluate the classifier's ability to predict one of four affective states from five hand gestures observed in video recordings of a classroom lecture. Using four-fold cross validation, we find that the model's generalization accuracy is 100% over cases where the student reported an affective state, and 79.4% when we include cases where the student reported no affective state. The experiment demonstrates that there is a strong relationship between human affect and visually observable gestures. Future work will explore the applicability of these results in practical applications.**

***Index Terms*— Behavior recognition, Intelligent tutoring systems, Human-computer interaction, Probabilistic affect prediction, Unintentional hand gestures.**

I. Introduction

Imagine the following situation: a student is attending a lecture through videoconferencing. During the class, he starts rubbing his eyes. What might be the reason for this motion? Is he tired, or having a problem with his eyes? Suppose that later, in a question and answer session, he often scratches his head. Is he confused, having a hair problem, or suffering from a headache?

If the instructor in this scenario possessed an "affective barometer" capable of transducing student body movements into their likely affective states, she could use this information to adapt her teaching strategy in real time.

Affective barometers capable of interpreting unintentional cues such as hand movement, facial expressions, and body posture would be extremely useful across a variety of live classroom situations and intelligent tutoring applications.

Real-world applications that interpret affective cues are indeed beginning to emerge [1], [2]. However, the relationship between visible cues and affect is not only *uncertain*, but also *context-dependent*.

To solve the problems of uncertainty and context dependency in estimating affect from visual cues, we advocate a probabilistic approach. The idea is to first formulate a parametric statistical model expressing the *possible* cause and effect relationships between affect and action in a particular situational context such as a classroom. Then, given the form of the model, we estimate its parameters from a training set acquired through video recording and interviews with the recorded students. This approach solves the problem of uncertainty by providing an estimated probability distribution over possible affective states rather than committing to any one affective state. It also solves the problem of context-dependency by learning from observations recorded in the actual context rather than presupposing what a particular observation might mean in a particular situation.

In this paper, we describe a preliminary application of the method to a data set we previously acquired from a real-world student-instructor interaction [3]. The data set contains recorded activities of four students during a classroom lecture along with associated self-reported affective states as described in a post-experiment interview. Using this data set for training, we build a Bayesian network model able to estimate a posterior distribution over four affective states given a student's observed hand gestures. Using four-fold cross validation, we find that the model's generalization accuracy is 100% over cases where the student reported an affective state, and 79.4% when we include cases where the student reported no affective state. The experiment demonstrates the feasibility of our approach.

II. Related Work

Long ago, scientists realized that in many situations, visual clues can be more important than verbal communication in human-human interactions [4]. Gestures such as pointing to an object or waving hands to show denial often appear in human-human communication; these gestures are thought to be important for communicating intentions.

In human-machine interaction, gesture understanding is widely used to facilitate effective communication via sign language and in tasks such as robot control and graphical user interface control. Useful gestures include both whole body and partial body movements, especially hand movements [5].

Gestures often convey important affective information. Wallbott [6] finds that body movements help people cope with emotional situations. Meijer [7] claims that surprise, shame and fear are correlated with movements such as backward body motion. Similarly, Richmond and McCroskey [8] find that knees shake and hands tremble when someone experiences fear.

Other researchers have focused on *unintentional* communication of affect through gestures. Coulson [9] finds that a non-deliberate shoulder shrug can indicate uncertainty. Burgoon et al. [10] find that a contracted body can indicate fear. Pollick et al. [11] investigate the affective content of movements such as drinking and knocking. Atkinson [12] explores the effect of kinematics of body postures for perception of affect. Bernhardt and Robinson [13] find that it is possible to infer emotions such as sadness, anger, and happiness from motions such as knocking and walking.

Another central focus of research on the affective content of gestures is multimodal affect analysis, in which gestures are analyzed in combination with another communication channel. Ambady and Rosenthal [14] report findings from a human study indicating that body gesture information can help improve emotional facial expression recognition accuracy. Balomenos et al. [15] and Gunes et al. [16] reach similar conclusions from work combining machine recognition of facial expressions with hand gesture analysis.

Thus far, the research on predicting affect from gestures has primarily focused on *task-independent* inference. As such, the typical approach is to begin with a set of pre-defined emotional concepts then attempt to find signals that provide evidence for the presence or absence of those emotions. The premise is that the gesture is part of the physiological response to an affect-evoking stimulus, and that the response is more or less independent of the situational context. As already discussed, we advocate a very different approach: rather than presupposing any affective categories, we observe students behaving in context, identify the gestures that they make, then determine, through interviews, what they were feeling. Our previous study of a classroom situation [3] found that unintentional gestures such as head scratching and eye rubbing tend to co-occur with affective states such as recalling and weariness.

Although to our knowledge, this is the first attempt to tackle the specific problem of extracting affective information from students' gestures in a classroom situation, there has been a great deal of related research on the importance of understanding student affect in educational contexts.

Wentzel [17] reports that students using a tutoring system able to adapt to their level of frustration show more improvement than students using a system indifferent to their affective state. Conati's educational game system [18] predicts player emotions from their performance and uses that information to adapt its instructional style and content. Vicente and Pain's intelligent tutoring system [19] uses questionnaires to estimate student motivational state and uses that information to modify the teaching method. Dadgostar et al.'s intelligent tutoring system [20] is assisted by non-verbal information. The authors observe how childrens' gestures are related to their learning skills.

As previously discussed, we model the stochastic relationship between gesture and affect using probabilistic methods. Specifically, we use a Bayesian network in which presence or absence of a gesture is represented by an evidence node depending causally on a set of hidden parent nodes (see Fig. 1 and Fig. 4). We estimate the parameters of the model using training data collected in context and ground truth obtained through post-experiment interviews with students. Details are provided in Section IV.

A few researchers have proposed probabilistic methods for estimating affect in intelligent tutoring and classroom scenarios. For example, Hernández et al. [21] and Conati and Maclaren [22] propose probabilistic models for estimating students' affective state based on interactive questionnaires. McQuiggan et al. [23] predict student frustration and anxiety during interaction with intelligent tutoring systems. We propose to estimate affect based on the unintentional hand gestures students use during a classroom session. To our

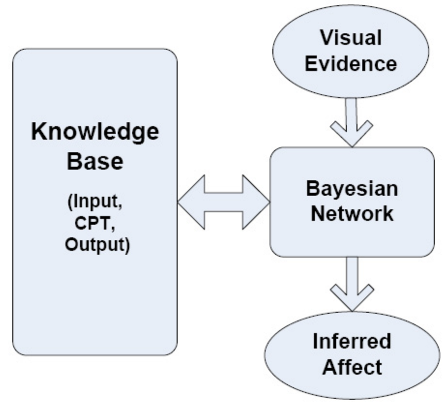

Fig. 1. Proposed system architecture.

(a) (b)

Fig. 2. Classroom experiment. (a) Recording setup. (b) Post-experiment interview with a student.

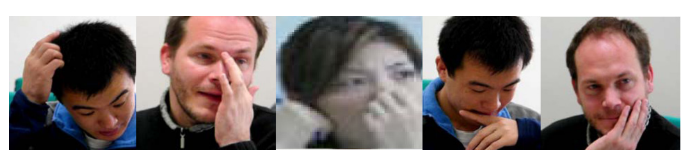

Fig. 3. Examples of gestures observed in the classroom experiment.

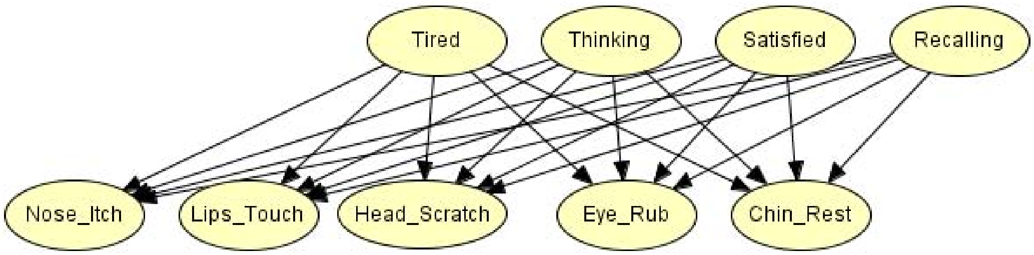

Fig. 4. Proposed model showing cause-affect relationship between affective states and observable gestures.

TABLE I

DISTRIBUTION OF SELF-REPORTED AFFECTIVE STATES FOR EACH STUDENT IN A REAL-WORLD CLASSROOM SCENARIO (IN PERCENT).

	Student				
Affective State	**A**	**B**	**C**	**D**	**Average**
Thinking	57.14	65.71	58.06	64.00	61.60
Recalling	9.54	0.00	0.00	12.00	4.5
Tired	4.76	11.42	6.45	0.00	6.25
Satisfied	14.28	8.57	3.22	4.00	7.12
Nothing	14.28	14.28	32.25	20.00	20.53

TABLE II

SELF-REPORTED AFFECT AND CO-OCCURRING HAND GESTURES. PERCENTAGES INDICATE THE PROPORTION OF THE TIME THAT THE GESTURE WAS ASSOCIATED WITH THE CORRESPONDING AFFECTIVE STATE.

Gesture	**Reported Affective States**
Head Scratch	Recalling (100%)
Chin Rest	Thinking (90%), Nothing (10%)
Eye Rub	Tired (81%), Nothing (19%)
Lip Touch	Thinking (88.75%), Nothing (11.25%)
Nose Itch	Satisfaction (77.5%), Nothing (22.5%)

knowledge, this is the first attempt to extract affective state from hand gestures such as head scratching and eye rubbing.

III. DATA ACQUISITION

Here we briefly summarize the data set we acquired in an earlier experiment [3]. We installed two video cameras in a classroom as shown in Fig. 2a, and in two sessions, we recorded students as they listened to a lecture. After recording, we manually isolated the gestures and facial expressions the students performed during the session. After segmenting the videos, in post-experiment interviews (see Fig. 2b), we asked each student what they were feeling at the time they performed each gesture. We recorded the students' free-form responses and clustered them using Geneva Affect Label Coding (GALC) [24], [25]. Details can be found in the original report [3].

In the original data, we observed 14 distinct gestures, but in this paper, we focus on 5 involving hand motion around the face: Head Scratch, Chin Rest, Eye Rub, Lip Touch, and Nose Itch. Some examples of the gestures we observed are shown in Fig. 3. The distribution over affective states for each student is shown in Table I, and the gestures with corresponding self-reported affective states, collapsed over all students, are shown in Table II.

To our knowledge, there has been relatively little study of unintentional hand gestures and their possible interpretations. The field would certainly benefit from a standard taxonomy and coding system similar to Ekman and Friesen's [26] Facial Action Coding System. However, the labels listed in Table II suffice for the purposes of the current experiment.

IV. PROPOSED MODEL

To model the stochastic relationship between affect and gesture, we propose the Bayesian network shown in Fig. 4. We represent each observable gesture (Head Scratch, Nose Itch, Lip Touch, Eye Rub, and Chin Rest) with a binary random variable depending probabilistically on the affective states (Tired, Thinking, Satisfied, Recalling). We also represent the affective states with binary random variables.

The model makes explicit the assumptions that the gestures are conditionally independent given the affective states and that the affective states are independent of each other. The first assumption (conditional independence of the evidence) is quite reasonable. The second assumption is clearly inaccurate: presumably some affective states are more likely to co-occur than others. However, without a substantial amount of co-occurrence data, we have little choice but to make this independence assumption.

For inference, we observe the values of the evidence variables (Nose Itch, Lip Touch, Eye Rub, and Chin Rest) then calculate the posterior distribution over the affective states. We denote the random variables for the gestures by $G_i, i = 1..n$, where n is the number of gestures in the model, in our case 5. We denote the observed value of gesture G_i by $g_i \in \{\text{true}, \text{false}\}$. We use the shorthand notation s_i to denote the event $S_i = \text{true}$ and g_j to denote the event $G_j = g_j$. With perfect detection of gestures and the independence assumptions just described, the posterior estimate of affective state S_i simply follows the naive Bayes model:

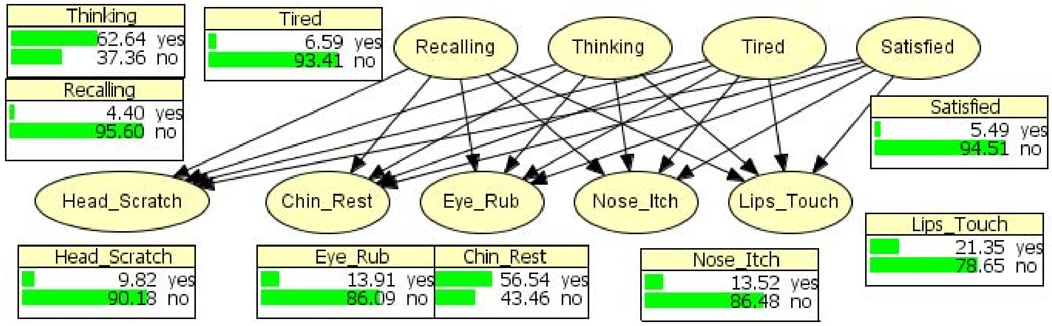

Fig. 5. Network trained on data from students B, C, and D. Each probability table shows the marginal prior distribution for one affective state or gesture over the training data.

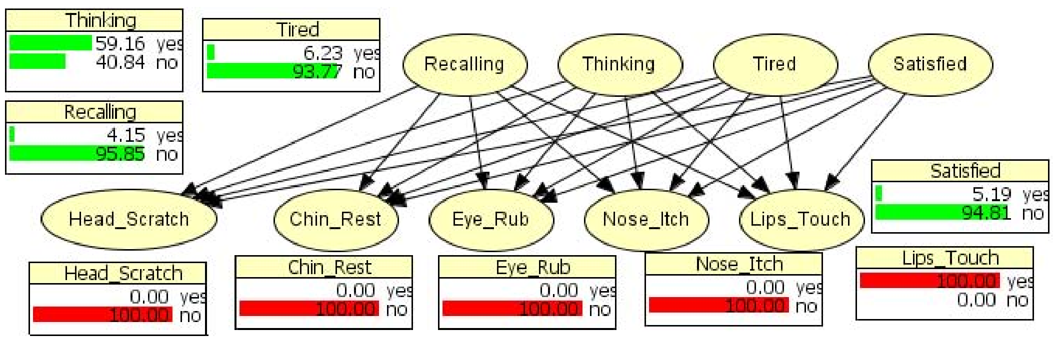

Fig. 6. Response of the network shown in Fig. 5 when the "Lip Touch" gesture is observed. The highest posterior probability is for affective state "Thinking," which agrees with student A's ground-truth self-reported affective state for this gesture.

$$
\begin{aligned}
P(s_i \mid g_1, \ldots, g_n) &\propto P(g_1, \ldots, g_n \mid s_i) P(s_i) \\
&= P(s_i) \prod_j P(g_j \mid s_i). \qquad (1)
\end{aligned}
$$

Training the model amounts to counting occurrences of each affective state i to obtain the priors $P(s_i)$ and counting co-occurrences of each gesture j with affective state i to obtain the class-conditional probabilities $P(g_j \mid s_i)$.

We use Hugin [27] to visualize our network. An example of trained network, showing the marginal prior distribution for each affective state and gesture variable, is shown in Fig. 5.

V. Model Evaluation

We tested the model using four-fold cross validation. That is, we performed four experiments in which one student's data was held out for testing and the other three students' data were used for training.

Since the human student interview protocol only requested a self-reported affective state at each point in time when we observed a gesture, we followed the same scheme when computing the priors $P(s_i)$ and class-conditional probabilities $P(g_j \mid s_i)$ for the model in Equation 1. There were 21 gestures for student A, 35 for student B, 31 for student C, and 25 for student D. Fig. 5 shows the marginal priors over gestures and affective states according to the network trained on students B, C, and D. The class-conditional probabilities, not shown, simply indicate the ratio of how many times gesture g_j co-occurred with affective state s_i compared to the total number of occurrences of affective state s_i.

For testing, at each point in time where a gesture occurred, we manually coded the gesture using a binary representation (e.g. when "Lip Touch" is observed, we code the evidence as $P(g_1) = 0, P(g_2) = 0, P(g_3) = 0, P(g_4) = 0, P(g_5) = 1$) then observe the inferred distribution over affective states according to the model in Equation 1. Response of network to one such input is shown in Fig. 6.

Table III shows the response of the network trained on student A, C, and D for all 35 gestures we observed for student B. Note that the forced-choice response of the network (obtained by finding the affective state with the highest estimated posterior probabililty) is always correct except for the "Nothing" state, which is not explicitly modeled by our network.

Table IV shows the test results averaged over all test folds and gestures as a confusion matrix. Table V shows the forced-choice version of the confusion matrix, wherein we select the affective state corresponding to the network's maximum estimated posterior probability and compare it to the student's reported ground truth for that gesture.

The confusion matrices show that when the ground truth is one of the four modeled affective states, the network's forced-choice classification accuracy is 100%, compared to a base rate of 77.5% that a trivial classifier always responding "Thinking" would obtain. This should not be entirely surprising given

TABLE III

POSTERIOR AFFECTIVE STATE ESTIMATES FOR NETWORK TRAINED ON STUDENTS A, C, AND D, IN RESPONSE TO THE GESTURES OBSERVED FOR STUDENT B. COLUMN "TRUTH" SHOWS THE THE CORRESPONDING GROUND-TRUTH SELF-REPORTED AFFECTIVE STATE.

	Network Response				
Gesture	**Thinking**	**Recalling**	**Tired**	**Satisfied**	**Truth**
Lip Touch	62.77	1.18	0.72	1.18	Thinking
Lip Touch	62.77	1.18	0.72	1.18	Thinking
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Lip Touch	62.77	1.18	0.72	1.18	Thinking
Lip Touch	62.77	1.18	0.72	1.18	Thinking
Eye Rub	13.45	5.82	62.13	5.82	Nothing
Nose Itch	7.78	3.37	2.05	61.41	Nothing
Nose Itch	7.78	3.37	2.05	61.41	Nothing
Eye Rub	13.45	5.82	62.13	5.82	Tired
Eye Rub	13.45	5.82	62.13	5.82	Nothing
Eye Rub	13.45	5.82	62.13	5.82	Nothing
Nose Itch	7.78	3.37	2.05	61.41	Satisfied
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Lip Touch	62.77	1.18	0.72	1.18	Thinking
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Eye Rub	13.45	5.82	62.13	5.82	Tired
Nose Itch	7.78	3.37	2.05	61.41	Satisfied
Eye Rub	13.45	5.82	62.13	5.82	Tired
Lip Touch	62.77	1.18	0.72	1.18	Thinking
Lip Touch	62.77	1.18	0.72	1.18	Thinking
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Eye Rub	13.45	5.82	62.13	5.82	Tired
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Chin Rest	75.22	0.47	0.28	0.47	Thinking
Nose Itch	7.78	3.37	2.05	61.41	Satisfied

TABLE IV

CONFUSION MATRIX OVER ALL FOUR TEST STUDENTS.

	Network Response			
Ground Truth	**Thinking**	**Recalling**	**Tired**	**Satisfied**
Thinking	77.87	0.84	0.96	1.07
Recalling	12.17	59.01	5.08	5.37
Tired	10.67	4.41	53.51	4.61
Satisfied	6.92	2.46	2.48	51.33
Nothing	49.89	9.83	11.41	12.12

TABLE V

FORCED-CHOICE CONFUSION MATRIX OVER ALL FOUR TEST STUDENTS.

	Network Response			
Ground Truth	**Thinking**	**Recalling**	**Tired**	**Satisfied**
Thinking	69	0	0	0
Recalling	0	5	0	0
Tired	0	0	7	0
Satisfied	0	0	0	8
Nothing	14	2	4	3

the tight relationship between gestures and affective states shown in Table II, but we should note that the results reported here are on the *test set*. This means that the students are consistent enough in making the gestures that learning from three students is always sufficient to correctly estimate the affective state of the fourth student.

Aggregated over the situations where our test students reported *no affective state*, however, we see that the network's response is diffuse. The total accuracy of the model is 79.4% correct overall, compared to a base rate of 61.6%. This is also unsurprising, as the network has no explicit concept of lack of affective state. But modeling lack of affect would not be as simple as adding a fifth affective state called "Nothing," because this state and the other affective states would clearly be strongly dependent on each other. The best way to model "Nothing" would be as a child of the four other affective states, as shown in Figure 7. Unfortunately, however, adding said "Nothing" node to the network in the current experiment, would have little effect on the network's forced choice behavior because it is infrequent. For example, referring to Table II, we see that in the training data, it happens to be the case that whenever a student was touching his or her lips, he or she reported an affective state of "Thinking" or "Nothing." The augmented network, of course, would learn this relationship, but when forced to make a choice without any additional information, it would always choose "Thinking" as the estimated affective state. Reliable classification of "Nothing" would clearly require additional contextual information, such as other aspects of the student's observable behavior, or some way of aggregating information over time.

VI. CONCLUSION

In this paper, we have demonstrated the feasibility of estimating students' affective state from visually observable gestures such as "Lip Touch" during a classroom lecture. We propose a simple Bayesian network incorporating observation of five hand gestures, and, using cross validation with a small training set, we find that a maximum a posteriori classifier based on the Bayes net obtains 100% accuracy over the cases where students report an affective state. When the gestures for which students reported no affective state are included, we obtain 79.4% accuracy.

Beyond the need for an explicit model of a "Nothing" state, which was already discussed, there are three main limitations

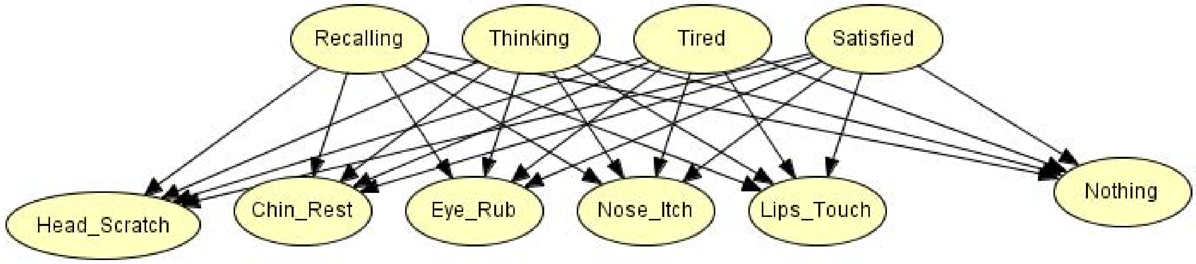

Fig. 7. One possible way to extend the model used in this paper account for the affective state "Nothing."

to our study. First, the experiment was small in scale; we only recorded four students in a single lecture. Second, we coded the gestures manually. Automatic gesture classification is a difficult problem in and of itself [28], so the uncertainty introduced by an automatic gesture recognition module would undoubtedly affect the model's performance in affective state classification. Finally, the statistical model we propose has no concept of building evidence for an affective state over time; it simply treats each observed gesture as a discrete event and attempts to classify it. In future work, we will focus on addressing these limitations and explore integrating the model into a pedagogical application.

Acknowledgements

We thank the students from the National Institute of Informatics, Tokyo, Japan who participated in our human study. We also thank the Higher Education Commission (HEC), Pakistan, for supporting this research. The work was partly supported by Thailand Research Fund through grant MRG4780209 to MND.

References

[1] R. Picard, *Affective Computing*. Cambridge, MA: MIT Press, 2000.
[2] S. Brave and C. Nass, "Emotion in HCI," *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, J. Jacko and A. Sears (Eds.)*, pp. 9–26, 2002.
[3] A. Abbasi, T. Uno, M. Dailey, and N. Afzulpurkar, "Towards Knowledge-Based Affective Interaction : Situational Interpretation of Affect," in *Proc. 2nd Intl. Conf. Affective Computing & Intelligent Interaction (ACII 2007), Lecture Notes in Computer Science*, vol. 4738, 2007, pp. 452–463.
[4] A. Mehrabian, "Communication without Words," *Psychology Today*, vol. 56, no. 4, pp. 53 – 56, 1968.
[5] S. Lee, "Automatic Gesture Recognition for Intelligent Human-Robot Interaction," in *Proc. 7th Intl. Conf. Automatic Face and Gesture Recognition (FGR06)*, 2006, pp. 645– 650.
[6] H. Wallbott, "Bodily Expression of Emotion," *European J. of Social Psychology*, vol. 28, no. 4, pp. 879–896, 1998.
[7] M. Meijer, "The Contribution of General Features of Body Movement to the Attribution of Emotions," *J. Nonverbal Behaviour*, vol. 13, no. 4, pp. 247–268, 1989.
[8] P. Richmond and C. McCroskey, *Nonverbal Behavior in Interpersonal Relationship*. West Virginia, USA: Allyn and Bacon, 2000.
[9] M. Coulson, "Attributing Emotion to Static Body Postures: Recognition Accuracy, Confusions, and Viewpoint Dependence," *J. Nonverbal Behavior*, vol. 28, no. 2, pp. 117–139, 1992.
[10] J. Burgoon, M. Jensen, T. Meservy, J. Kruse, and J. Nunamaker, "Augmenting Human Identification of Emotional States in Video," in *Proc. Intl. Conf. Intelligent Data Analysis*, 2005.
[11] F. Pollick, H. Paterson, A. Bruderlin, and A. Sanford, "Evidence for Distinct Contribution of Form and Motion Information to the Recognition from Body Gestures," *Cognition*, vol. 82, pp. 51–61, 2001.
[12] A. Atkinson, M. Tunstall, and W. Dittrich, "Evidence for Distinct Contribution of Form and Motion Information to the Recognition from Body Gestures," *Cognition*, pp. 59–72, 2007.
[13] D. Bernhardt and P. Robinson, "Detecting Affect from Non-Stylished Body Motions," in *Proc. 2nd Intl. Conf. Affective Computing & Intelligent Interaction (ACII 2007), Lecture Notes in Computer Science*, vol. 4738, 2007, pp. 59–70.
[14] N. Ambady and R. Rosenthal, "Thin Slices of Expressive Behavior as Predictors of Interpersonal Consequences : A Meta Analysis," *Psychological Bulletin*, vol. 111, no. 2, pp. 256–274, 1992.
[15] T. Balomenos, A. Raouzaiou, S. Ioannou, A. Drosopoulos, K. Karpouzis, and S. Kollias, "Emotion Analysis in Man-Machine Interaction Systems," in *Lecture Notes in Computer Science*, vol. 3361, 2004, pp. 318–328.
[16] H. Gunes and M. Piccardi, "A Bi-Modal Face and Body Gesture Database for Automatic Analysis of Human Nonverbal Affective Behavior," in *Proc. IEEE Intl. Conf. Pattern Recognition*, 2006, pp. 1148–1153.
[17] K. Wentzel, "Student Motivation in Middle School: The Role of Perceived Pedagogical Caring," *J. Educational Psychology*, vol. 89, no. 3, pp. 411–419, 1997.
[18] C. Conati, "Probabilistic Assessment of User's Emotions in Educational Games," *Applied Artificial Intelligence*, vol. 16, pp. 555–575, 2002.
[19] A. de Vicente and H. Pain, "Informing the Detection of the Students' Motivational State: An Empirical Study," in *Proc. 6th Intl. Conf. Intelligent Tutoring Systems*, 2002.
[20] F. Dadgostar, H. Ryu, A. Sarrafzadeh, and S. Overmyer, "Making Sense of Student Use of Nonverbal Cues for Intelligent Tutoring Systems," in *Proc. Intl. Conf. ACM SIGCHI*, vol. 122, 2005, pp. 1–4.
[21] Y. Hernández, J. Noguez, E. Sucar, and G. Arroyo-Figueroa, "A Probabilistic Model of Affective Behavior for Intelligent Tutoring Systems," in *Lecture Notes in Computer Science*, vol. 3789, 2005, pp. 1175–1184.
[22] C. Conati and H. Maclaren, "Evaluating a Probabilistic Model of Student Affect," *Intelligent Tutoring System*, pp. 55–66, 2004.
[23] S. McQuiggan, S. Lee, and J. Lester, "Early Prediction of Student Frustration," in *Proc. 2nd Intl. Conf. Affective Computing & Intelligent Interaction (ACII 2007), Lecture Notes in Computer Science*, vol. 4738, 2007, pp. 698–709.
[24] K. Scherer, "What are emotions? and how can they be measured?" *Social Science Information*, vol. 44, no. 4, pp. 695–729, 2005.
[25] ——, "Geneva affect label coder." [Online]. Available: http://www.unige.ch/fapse/emotion/resmaterial/GALC.xls
[26] P. Ekman and W. V. Friesen, *Facial Action Coding System (FACS): Manual*. Palo Alto, CA: Consulting Psychologists Press, 1978.
[27] Hugin Expert A/S, "Hugin expert: Advanced decision support using Bayesian networks and influence diagrams," 2004. [Online]. Available: http://www.hugin.com
[28] S. Mitra and T. Acharya, "Gesture Recognition: A Survey," *IEEE Trans. SMC - Applications and Reviews*, vol. 37, no. 3, pp. 311–324, 2007.