Motion Recognition and Students’ Achievement

Wen-Fu Pan  
Department of Educational Administration and Management  
National Dong Hwa University, Hualien, Taiwan  
s1210@gms.ndhu.edu.tw

Anton Subarno  
Doctoral Student in Education, Department of Educational Administration and Management  
National Dong Hwa University, Hualien, Taiwan  
Department of Office Administration Education  
Sebelas Maret University, Surakarta, Indonesia  
antonsubarno@fkip.uns.ac.id

Mei-Ying Chien  
Department of Educational Administration and Management  
National Dong Hwa University, Hualien, Taiwan  
myc101@gms.ndhu.edu.tw

Ching-Dar Lin  
Institute of Education  
Tzu Chi University, Hualien, Taiwan  
abcd1812@mail.tcu.edu.tw

Abstract—Human motion has multifarious meanings that can be recognized using a facial detection machine. This article aims to explore body motion recognition to explain the relationship between students’ motions and their achievement, as well as teachers’ responses to students’ motions, and especially to negative ones. Students’ motions can be identified according to three categories; facial expression, hand gestures, and body position and movement. Facial expression covers four categories, namely, contempt, fear, happiness, and sadness. Contempt is used to express conflicted feelings, fear to express unpleasantness, happiness to express satisfaction, and sadness to express that the environment is uncomfortable. Hand gestures can likewise be grouped into four categories: conversational gestures, controlling gestures, manipulative gestures, and communicative gestures. Conversational gestures refer to communicative gestures. Controlling gestures refer to vision-based interface communications, like the ones popular in current technology. Manipulative gestures refer to ones used in human interaction with virtual objects. Communicative gestures relate to human interaction, and therefore involve the field of psychology. Body position and movement also can be classified into four categories, namely: leaning forward, leaning backward, correct posture, and physical relocation. Leaning forward happens when a user is working with a high level of concentration. Leaning backward occurs when a user has been highly concentrated on work for several hours, and needs a break or change. Correct posture is the sign of an enjoyable working position which involves sitting in a free and relaxed manner. Movement refers to a change to the student’s sitting location, reflecting some inadequacy of the learning environment.

Teachers can anticipate changes of students’ emotions by good learning design, teaching metacognitive skills, self-regulated performance, exploratory talks, mastery approach/avoidance, using hybrid learning environments, and controlling
space within classrooms. Teachers’ responses to students’ motions will be explored in this article.

**Keywords**—motion recognition, facial expression, hand gestures, body position

## I. INTRODUCTION

Human emotion can be recognized through body movements and dynamical gestures [1]. In the academic context, emotions have an effect on students’ learning and achievement that is mediated by attention, self-regulation, cognitive resources, and motivation [2]. Humans’ bodies will naturally react to convenient or inconvenient environments, just as they will to interesting or uninteresting situations. A fundamental of humans’ body motions will change in appearance in order to respond an action based on its emotion [3].

Teachers, in their role as mediators in the classroom [4], should pay careful attention to student’s motions. Students’ emotions can be detected by looking at their visible expressions. [5] showed that posture prediction can be computed in a few seconds based on sub-zone volume, zone differentiation, collision avoidance, and reduced target set. The computation based on sub-zone volume focuses on the approach of zone volume and ignore the avatar volume, zone differentiation emphasizes the sphere of constraints that relate to the zone avoidance, collision avoidance concerns to the calculation of zone volume, and reduced target set is the probability of reaching a target. During the teaching-learning process, astute teachers will become aware of both the positive and negative emotions of students. Teachers have to change their teaching methods or take other remedial action to stimulate positive student emotions if they perceive physical cues from their students indicating negative emotions. This article aims to explore body motion recognition, the relationship between motions and students’ achievement, and teachers’ responses to students’ motions.

## II. LITERATUR REVIEW

### A. Recognition of Students’ Motions

Visualization is the easiest way to present data so as to make it perceptible and engage human sensory systems [6]. Data is any sort of information, ranging from news, stories, and maps to other resources, and is generated every day. Images are meant to be representations of data. [7] state that the Phase Congruency-based model of images is a reliable tool for recognizing emotions. [8] explored facial expression recognition using deep conspicuous neural networks based on seven basic facial expressions; anger, contempt, disgust, fear, happiness, sadness, and surprise. Only four facial expressions; contempt, fear, happiness and sadness, were recognized well. [1] analyzed 240 items of gesture data from the Human-Machine Interaction Network on Emotion (HUMAINE), in which human emotional states were classified into eight types, namely; anger, despair, interest, pleasure, sadness, irritation, joy and pride. Four of these eight emotions, anger, joy, pleasure, and sadness, were further investigated by [1]. Negative emotions (anger and sadness) were ambiguous, as were positive emotions (joy and pleasure), but the movements generated by positive emotions were observed to be more expansive than those associated with negative emotions.

Technologies that can detect human gestures are widely available. [9] recognized hand gestures through an artificial neural network (ANN), template matching, hidden Markov models (HMMs), and dynamic time warping (DTW). Another type of equipment popular for detecting the actions of the human body is a Kinect sensor. Kinect is widely-known as the set of tools the Xbox 360 game console (released in 2010) uses to allow players to control games using gestures and spoken commands [10]. Kinect sensors work using an RGB camera, depth sensors and a multi-array microphone to capture movement in three dimensions. The depth sensors use an infrared laser projector combined with a monochrome CMOS sensor.
The accuracy of gesture detection is beyond dispute. [11] stated that feature-based faces can be detected with 98% accuracy using local orientation histograms (LOH). The same assumption is also stated by [12], who found that the stochastic context-free grammar (SFCG) system can detect 10 gestures recorded in video frames (swipe right, swipe left, swipe up, swipe down, horizontal wave, vertical wave, circle, point, palm up and fist) with over 94% accuracy. [13] stated that full body movement can be recognized using Rivers-Gesture Description Language (R-GDL), such that all human movements can be detected with between 91% and 100% accuracy. Similar results, in the form of a 91% detection accuracy rate, were obtained using Microsoft Kinect by [14].

B. Motions on Student's Achievement

The human body has different ways of responding to convenient and inconvenient environments, using its posture to express emotions. [15] stated that “usually, the posture we subconsciously adopt to match certain moods is temporary, but in some cases, it persists if the emotional state is habitual. Consider the posture of a person who is grieving, or the muscle tone of a person who is angry” (p. 38). [16] stated that slight head turns and other head poses can express emotions. In addition, hand gestures can be recognized as human expressions. Hands give signals related to the body's movement, reduce reliance on sound in noisy environments, and can be scaled and broken down into meaningful movements. [9] stated that static hand gesture can be classified into four categories: conversational gestures, controlling gestures, manipulative gestures, and communicative gestures. Conversational gestures refer to communicative gestures. These are useful in noisy environments, reducing the reliance on sound, and are also suitable for communication with and between disabled persons. Controlling gestures, in contrast, refer to vision-based interfaces that are so popular in current technologies. Some technologies use gesture controlled applications to control machines or remove virtual objects, while other gestures involve navigating. Manipulative gestures relate to human interaction using virtual objects. Communicative gestures relate to human interaction, and therefore involve the field of psychology. [17] analyzed primitive dynamic motion based on four different styles: neutral, happy, angry and sad, and showed that primitive motions corresponded to four actions; raising the arm, lowering the arm, knocking, and retracting. After removing the individual motion bias, the recognition of motion was recognized with a significant level of accuracy. The position and meaningful movement of hand gestures is important for their categorization. [18] developed a measurement system and showed that static hand gesture might not actually be a reliably consistent way of accurately recognizing emotions. Therefore, the choice of a given application, hand motion trajectory (hand shape and position features), and choice of recognizers should all be considered.

Human-level expression recognition is achievable with machine learning technology in real life [19]. [20] stated that the angles of joints relative to the human body’s position can be detected using a skeleton analysis. For instance, these postures can be observed this way: leaning forward, leaning backward, and correct posture. Sitting in front of a computer for long hours without interruption and inappropriate workstation environments make for poor sitting posture. This can be interpreted as the human body changing its sitting position, indicating the environment is inconvenient. Leaning forward to express that one is working with a high level of concentration, leaning backward to show that one has been highly concentrated on work for several hours, and correct posture to indicate one is in an enjoyable working position. Of course, this analysis has limitations. For instance, camera position is constrained, and a clear workstation environment is required for accurate observation, but this is still fertile ground for further investigation. Spontaneous emotion has high impact on human-computer interaction, education, psychology, and psychiatry [21].
[22] argued the role of emotions in the learning environment is not so much about how to let students have a feel-good experience, but rather they represent the psycho-socio-emotional glue that leads student to forge their way into new areas that reflect, and indeed allow them to practice, their skills and capabilities. It makes sense that students will either move forward or regress as a reaction to the learning process. This assumption relates to the work of [23], who found a statistically significant relationship between the students’ locations in the classroom and their attendance and educational achievements. Students sitting at the front of the class tended to be more motivated, and interacted with teachers more than their classmates did. [24] found no effect of seat location on exam scores in a Biology class, but found that the opposite was true for a Physics class, in which seat location had an effect [25]. This situation has been found to also happen in higher education settings with older students. [24] found that students who sat in the front of the lecture hall had higher GPAs than ones who sat at the back. [26] stated that the seating locations of higher education students had a significant impact on the students’ performance, but this effect was less pronounced in teamwork class courses and classes where teachers employed good teaching practices.

III. RESULTS AND DISCUSSION

Teachers, when facing inconvenient learning environments, have to take urgent action to stimulate students’ emotions. A boredom or negative emotion in relation to academics has a significant effect on student performance. Such effects, and indeed the emotions underlying them, can be reduced using a motivational regulation skill that is called metacognitive skill [27]; [2]. [28] stated that well-designed learning materials and round face-like shapes with warm colors induced positive emotions in situations related to the learning process. [29] argued that self-regulated performance has a positive effect on strategic knowledge. Students given a faded manuscript were capable of finding out the answers to question by themselves. [30] stated that exploratory talks improve group reasoning, and social reasoning, in turn, can improve individual reasoning capability. Teachers can teach by using exploratory talks to lead students to improve their skills as they transfer knowledge to their peer group. [31] state that mastery approach/avoidance can predict achievement goals and academic performance, with these effects being mediated by emotions, pride, and hope. Emotion can also be stimulated using technology. [32] argue that synchronous hybrid learning environments are correlated to students’ control, values, emotions, and perceived success in terms of achieving goals and technology use. Another tack teachers can employ involves changing sitting positions or giving enough space to each student to assist them to feel free to express opinions. [33] explained how space within classrooms stimulates students’ imagination, and pointed to the importance of creating convenient learning environments by controlling light, color, sound, and micro-climate. Visual environments also stimulate students’ creativity. Based on the analysis above, the relationships between motion recognition and teacher response can be conceptualized as depicted in Figure1.

Figure 1. Levels of human body response

There are three areas, or zones, related to human body responses. Zone A relates to facial expression, zone B relates to hand...
gestures, and zone C relates to body position and movements. Zone A indicates the lowest level of response, as indicated by expression, to the environment. Students express their feelings, like contempt, fear, happiness, sadness or other expressions indicating their mood and feeling. Rotating the head indicates thinking process and nodding is used to express agreeing and understanding [16]. Other head motions are used to express feelings, too. Sometimes someone changes the sound of his or her voice to express a feeling [15] and an inconsistent tone can show unstable thinking or confusion.

Zona B refers to hand gestures, and comprises what might be called middle level responses to the environment. For example, students will raise their hands to ask questions or give answers. Swipe left and Swipe down were he first recognized detections, followed by Swipe Right, Circle, Point and Fist [12]. The four basic categories of hand gestures are conversational gestures, controlling gestures, manipulative gestures, and communicative gestures [9]. Conversational gestures can be used when reacting to noisy environments, to ask someone to perform an action, to reduce the requirement for sound, and for other conversational purposes in daily communication. Conversational gestures are suitable for disabled communication, to explain ideas or to ask someone to perform an action [17]. Controlling gestures have become especially popular in the last decade as technology has developed. Gesture-based applications are now used to control mobile phones, open virtual objects, and perform other navigational functions. Manipulative gestures have been used in practice in teaching virtual classes, virtual libraries, for medical procedures and to handle patients, and for other applications related to virtual objects. Communicative gestures relate to human cultural interaction, and therefore involve the study of psychology and practical communications. The most familiar actions, when it comes to communication gestures, have been found to be raising the arm, lowering the arm, knocking, and retracting [17]. Zone C represents a high-level response to the environment in the form of an action or body movement. Students will change their sitting position, indicating their feeling toward the environment. For instance, leaning forward expresses a high level of motivation toward the learning process, while leaning backward shows that they need a break or it is time to change to other material, while a correct posture position indicates they feel comfortable to study. In this way, spontaneous emotion can describe the education process [21]. The role of emotions in the learning environment is not so much about how to make students feel good during the experience, but emotions are more properly seen as indicator behaviors [22]. Hence, students will move forward in the classroom to get a clearer explanation of the material or they move back to little bit to avoid taking part in the learning process. This condition cannot be assumed to hold for all classes and all subjects, however, because it will not necessarily happen in situations where good teaching practice is implemented in class or in teamwork class courses [26].

Teachers can anticipate overcoming the inconvenient learning environment by preparing, and specifically by devoting their attention to seven steps, namely: designing learning well, teaching metacognitive skills, self-regulated performance, exploratory talks, mastery approach/avoidance, hybrid learning environments, and space within classrooms. An interesting learning environment can be prepared using good materials, assorted games, and colorful presentation, that, taken together, induce positive emotions [28]. Metacognitive skill teaching should be provided in relation to motivational regulation, aimed at identifying and implementing effective strategies by which to create positive emotional states in students and facilitate their academic performance [27]. Self-regulated performance guides students to find detailed materials and answers to practical questions. Clear instruction is important, a prerequisite to supporting students’ performance and something that promotes their learning independently [29].
Exploratory talk is effective to stimulate peer-learning and to help students understand materials in a deep way that will be transferred to classmates, or to give feedback [29]; [30]. Mastery approach/avoidance relates to goal orientation and goal avoidance [31]. Students will be engaged in their learning if they know the final goal and avoid goal disturbance. This goal should be clearly explained at the beginning of learning to pique students’ interest and attract their motivation. Hybrid learning environments relate to the synchronization of learning environments and technology use. Technology can be used to describe goals and to help students to reach those goals. Space within the classrooms refers to the spatial layout and setting of the classroom, and how comfortable it is to learn there. Students’ imaginations are stimulated when they have enough space to play, and can partake in games or other activities with classmates [33]. The room should be supported by light, color, sound, and a micro-climate that enables students to explore all their capabilities.

V. CONCLUSIONS

Students’ motions have several meanings that can be recognized using a facial detection machine. The accuracy of gesture detection is relatively high. Prior research results show that in excess of 91% of human gestures and motions can be recognized using technology. Students’ motions can be identified as falling into three categories; facial expression, hand gestures, and body position and movement. The low-level response of environment expression can be indicated through head area movements. Students express their feelings likes contempt, fear, happiness, and sadness. They also express agreeing and understand changes in the tones of sounds to express feelings. The middle level responses to the environment are using hand gestures. Students will use their hands to respond to a question or give an answer or provide other direction in the course of communication. The four basic categories of hand gestures are conversational gestures, controlling gestures, manipulative gestures, and communicative gestures. The high-level response relates to an action or body movement. Students will physically change their sitting position to indicate their feeling. Body position and movement can be classified into four categories, namely; leaning forward, leaning backward, correct posture, and movement. Teachers can anticipate the change of students’ motions using well-designed learning, metacognitive skills, self-regulated performance, exploratory talks, mastery approach/avoidance, hybrid learning environments, and space within classrooms. An interesting learning environment can be prepared before the start of teaching, including design and learning material selection, and can be continued during the teaching process using hybrid learning environments, technology, and short games.

ACKNOWLEDGMENT

The deepest gratitude goes to the Ministry of Science and Technology of Taiwan for its funding (MOST 105-2511-S-259-007), and the Review Panel for its valuable advice.

REFERENCES

[1]. G. Castellano, S.D. Villalba, and A. Camurri, Recognising human emotions from body movement and gesture dynamics. In Proceeding Second International Conference, ACII 2007 (Lisbon, Portugal, September 12-14, 2007) LNCS 4738, 71–82, 2007. doi 10.1007/978-3-540-74889-2_64

[2]. R. Pekrun, T. Goetz, W. Titz, and R.P. Perry, Academic emotions in students’ self-regulated learning and achievement: a program of qualitative and quantitative research. Educational Psychologist. 37, 2, Jun 2010, 91-105. http://dx.doi.org/10.1207/S15326985EP3702_4
[3]. D. Bernhardt, and P. Robinson, Detecting emotions from connected action sequences. In Proceeding First International Visual Informatics Conference, IVIC 2009 (Kuala Lumpur, Malaysia, November 11-13, 2009) 738-747, 2009. DOI: 10.1007/978-3-642-05036-7_1.

[4]. A.M. Hamamorad, Teacher as mediator in the efl classroom: a role to promote students’ level of interaction, activeness, and learning. International Journal of English Language Teaching. 4,1 (Jan 2016), 64-70.

[5]. T. Marler, S. Beck1, U. Verma1, R. Johnson1, V. Roemig1, and B. Dariush, A digital human model for performance-based design. In Proceedings 5th International Conference, DHM 2014 Held as Part of HCI International 2014 (Heraklion, Crete, Greece, June 22–27, 2014). LNCS 8529, 35-42. DOI: 10.1007/978-3-319-07725-3_4

[6]. A. Wexelblat, Virtual reality: Applications and explorations, Academic Press: New York, 1993.

[7]. S. Shojaeilangari, W. Y. Yau, J. Li, and E. K. Teoh, Feature extraction through binary pattern of phase congruency for facial expression recognition. In Proceeding 12th International Conference on Control Automation Robotics & Vision (ICARCV), 166-170, 2012. http://dx.doi.org/10.1109/ICARCV.2012.6485152.

[8]. J.P. Can’ario, and L. Oliveira, Recognition of facial expressions based on deep conspicuous net. In Proceeding 20th Iberoamerican Congress, CIARP 2015 (Montevideo, Uruguay, November 9-12 2015) LNCS 9423, 255-262, 2015. DOI: 10.1007/978-3-319-25751-8_31

[9]. H. S. Badi, and S. Hussein, Hand posture and gesture recognition technology. Neural Computing & Applications, 25 (Apr 2014), 871-878. doi:10.1007/s00521-014-1574-4

[10]. Q. Luo, Study on three dimensions body reconstruction and measurement by using kinect. In Proceedings 5th International Conference, DHM 2014 Held as Part of HCI International 2014 (Heraklion, Crete, Greece, June 22–27, 2014). LNCS 8529, 35-42. DOI: 10.1007/978-3-319-07725-3_4

[11]. S. Majed, H. Arof, and Z. Hashmi, Orientation features-based face detection by using local orientation histogram framework. In Proceeding First International Visual Informatics Conference, IVIC 2009 (Kuala Lumpur, Malaysia, November 11-13, 2009), 2009. 738-747. DOI: 10.1007/978-3-642-05036-7_13

[12]. M. Gavrilascu, Recognizing human gestures in videos by modeling the mutual context of body position and hands movement. Multimedia Systems. 22, 2 (Mar 2016), 1–13. DOI 10.1007/s00530-016-0504-y

[13]. T. Hachaja, and M. R. Ogilab, Full body movements recognition – unsupervised learning approach with heuristic R-GDL method. Digital Signal Processing. 46 (Jul 2015), 239–252. DOI: http://dx.doi.org/10.1016/j.dsp.2015.07.004

[14]. M. Gowing, A. Ahmadi, F. Destelle, D. S. Monaghan, N. E. O’Connor, and K. Moran, Kinect vs. Low-cost inertial sensing for gesture recognition. In Proceedings Part I 20th Anniversary International Conference, MMM 2014 (Dublin, Ireland, January 6-10, 2014). LNCS 8325, 484–495. DOI 10.1007/978-3-319-04114-8_41

[15]. J. Johnson, Postural assessment. International Therapist. 99, (Jan 2012), 36-38.

[16]. S. Shojaeilangari1, W. Yau, and E. Teoh, Pose-invariant descriptor for facial emotion recognition. Machine Vision and Applications. 27, 5 (Jul 2016), 1063–1070. DOI 10.1007/s00138-016-0794-2
[17]. D. Bernhardt, Posture, gesture and motion quality: A multilateral approach to affect recognition from human body motion. In Proceedings of the Doctoral Consortium at the Second International Conference on Affective Computing and Intelligent Interaction, ACII’07 (Lisbon, Portugal, September 12-14, 2007), 49–56. DOI: 10.1007/978-3-540-74889-2

[18]. S. Bilal, R. Akmeliawati, A. A. Shafie, and M.J.E. Salami, Hidden Markov model for human to computer interaction: a study on human hand gesture recognition. Artif Intell Rev, 40, 2013, 495–516. DOI: 10.1007/s10462-011-9292-0

[19]. J. Whitehill, G. Littlewort, I. Fasel, M. Bartlett, and J. Movellan, Towards practical smile detection. IEEE Transactions on Pattern Analysis and Machine Intelligence. 31, 11 (Nov 2009), 2106-2111.

[20]. K. Wongpatikaseree, H. Kanai, and Y. Tan, Context-aware posture analysis in a workstation-oriented office environment. In Proceedings 5th International Conference, DHM 2014 Held as Part of HCI International 2014 (Heraklion, Crete, Greece, June 22–27, 2014). LNCS 8529, 148-159. DOI: 10.1007/978-3-319-07725-3_14

[21]. Z. Zeng, Y. Fu, G.I. Roisman, Z. Wen, Y. Hu, and T.S. Huang, Spontaneous emotional facial expression detection. Journal of Multimedia. 1, 5 (Aug 2006), 1-8.

[22]. G. Gläser-Zikuda, I. Stuchlíková, and J. Janík, Emotional aspects of learning and teaching: Reviewing the field – discussing the issues. Orbis Scholae. 7, 2 (2013), 7–22.

[23]. K. Zomorodian, M. Parva, I. Ahrari, S. Tavana, C. Hemyari, K. Pakshir, P. Jafari, and A. Sahraian, The effect of seating preferences of the medical students on educational. Med Educ Online. 17, 2012, 10448 - http://dx.doi.org/10.3402/meo.v17i0.10448

[24]. S. Kalinowski, and M.L. Taper, The effect of seat location on exam grades and student perceptions in an introductory biology class. Journal of College Science Teaching. (Feb 2007), 54-57.

[25]. K.K. Perkins, and C.E. Wieman, The surprising impact of seat location on student performance. Phys. Teach. 43, 30 (Dec 2005). DOI: http://dx.doi.org/10.1119/1.1845987

[26]. M.D. Meeks, T.L. Knotts, K.D. James, F. Williams, J.A. Vassar, and A.O. Wren, The impact of seating location and seating type on student performance. Educ. Sci. 3, (Oct 2013), 375-386. DOI:10.3390/educsci3040375

[27]. I. Fritea, and R. Fritea, Can motivational regulation counteract the effects of boredom on academic achievement? Procedia - Social and Behavioral Sciences. 78, (2013), 135 – 139. doi: 10.1016/j.sbspro.2013.04.266

[28]. J. L. Plass, S. Heidig, E.O. Hayward, B.D. Homer, and E. Um, Emotional design in multimedia learning: Effects of shape and color on affect and learning. Learning and Instruction. 29, (Feb 2014), 128-140. DOI: http://dx.doi.org/10.1016/j.learninstruc.2013.02.006

[29]. C. Wecker, and F. Fischer, From guided to self-regulated performance of domain-general skills: The role of peer monitoring during the fading of instructional scripts. Learning and Instruction. 21, (May 2011), 746-756. DOI:10.1016/j.learninstruc.2011.05.001

[30]. R. Wegerif, N. Mercer, and L. Dawes, From social interaction to individual reasoning: An empirical investigation of a possible sociocultural model of cognitive development. Learning and Instruction. 9, (1999), 493-516.
[31]. D.W. Putwain, P. Sander, and D. Larkin, Using the 2x2 framework of achievement goals to predict achievement emotions and academic performance. Learning and Individual Differences. 25, (Jan 2013) 80–84. DOI:http://dx.doi.org/10.1016/j.lindif.2013.01.006

[32]. N.T. Butz, R.H. Stupnisky, and R. Pekrun, Students’ emotions for achievement and technology use in synchronous hybrid graduate programmes: a control-value approach. Research in Learning Technology. 23, (Mar 2015), 26097.DOI: http://dx.doi.org/10.3402/rlt.v23.26097

[33]. D. Davies, D. Jindal-Snape, C. Collier, R. Digby, P. Hay, and A. Howe, Creative learning environments in education—A systematic literature review. Thinking Skills and Creativity. 8, (April 2013), 80-91. DOI: http://dx.doi.org/10.1016/j.tsc.2012.07.004