World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:6, No:5, 2012

Human Interactive E-learning Systems using Head Posture Images

Yucel Ugurlu

***Abstract*—**This paper explains a novel approach to human interactive e-learning systems using head posture images. Students' face and hair information are used to identify a human presence and estimate the gaze direction. We then define the human-computer interaction level and test the definition using ten students and seventy different posture images. The experimental results show that head posture images provide adequate information for increasing human-computer interaction in e-learning systems.

***Keywords*—**E-learning, image segmentation, human-presence, gaze-direction, human-computer interaction, LabVIEW

I. Introduction

TODAY, computers and information technologies play an important role in education through the use of e-learning environments and different computer-based systems. E-learning is one of the most popular approaches; it has enormous potential and is an important future trend in engineering education. E-learning is defined as learning and teaching online via network technologies. During the last decade, it has emerged as one of the most powerful approaches for addressing the growing need for education [1], [2].

Many traditional classes are currently augmented with virtual labs or e-learning environments. The biggest advantages of e-learning systems over traditional classes are that they are available on-demand, they provide rich content such as videos, graphics, and texts, and they offer a fully personalized structure.

However, e-learning differs from conventional education, in that e-learning separates teachers from students and students from other students. Therefore, the main challenges of these systems are the lack of face-to-face communication and of students' emotions. Therefore, modern e-learning should take human behavior and emotions into account so that the content can be changed adaptively or according to student needs. Several ongoing research studies are investigating human behavior and emotions for learning systems [3]–[5].

This paper is organized as follows. Section II presents an outline of an interactive e-learning system. Section III proposes a head posture analysis method and applies it to target frames; Section IV explains the evaluation process. The paper ends with some concluding remarks and implications for future work.

II. Human Interactive E-learning Systems

Human-computer interaction is a crucial component of e-learning systems. It influences the efficiency with which they are used and provides a means of communication between the user and the virtual environment [6]–[8].

Y. Ugurlu is with National Instruments Japan, Shibadaimon 1-9-9, Minato-ku, Tokyo 105-0012, Japan, and is also a research associate and lecturer at the Department of Integrated Information Technology at the Aoyama Gakuin University, Sagamihara, Japan (e-mail: yucel.ugurlu@ni.com).

Today, many e-learning systems have advanced graphical user interfaces, but the role of human-computer interaction has not developed and integrated yet. Efficient human-computer interaction is essential for the effective use of e-learning environments. There are three types of user interfaces: sensor-based, visual-based, and audio-based. In most e-learning systems, the human-computer interaction is primitive and limited to the mouse or keyboard, where it is used to progress to the next topic or for answering simple questions or quizzes. Visual-based human computer interaction is a particular focus of computer science research. It includes facial-expression analysis, body-movement tracking, gesture recognition, and gaze detection. Gaze detection is primarily an indirect form of interaction between user and machine that is used to measure the user's attention or intentions.

In recent years, we have developed an e-learning system that teaches graphical programming in the context of higher education [9]. The system is based on slides and programming videos; it requires only mouse and keyboard interactions. In this study, we propose a new approach for understanding and evaluating student interaction during the learning process, as shown in Fig. 1. Mouse clicks and keyboard entries provide limited information about user interactivity, making personal assessments difficult. In contrast, visual information provides continuous data about individuals that can be used to improve the learning content or to build smart virtual learning environments.

III. Head Posture Analysis

The detection of faces and facial features has received much attention in the context of face recognition [10]–[13]. However, detecting faces from a single image is a challenging task because of variations in the scale, location, orientation, and pose. The facial expression, features such as sunglasses, a mustache, or a beard, and the lighting conditions also significantly affect the appearance of the face.

We propose a robust approach that uses face and hair information to analyze human-head images. We analyzed a series of head posture images using the MIT-CBCL face recognition database [14].

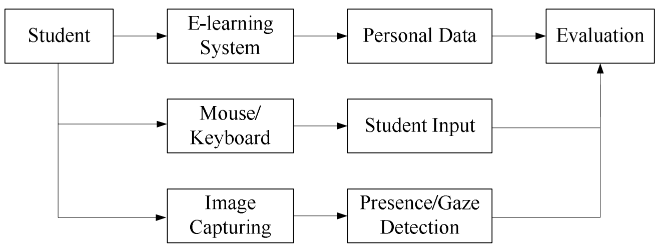


Fig. 1 Vision-augmented e-learning systems that use student-computer interaction

The images were of three women and seven men with different head postures, and the image sizes were fixed at 768 × 576 pixels. Fig. 2 shows seven different head views of one of the students. To analyze the images, we used the National Instruments (NI) Vision Assistant software and NI LabVIEW because of their simple interactive programming environments. The training and segmentation algorithm was first developed in NI Vision Assistant and then imported into the NI LabVIEW environment for some additional mathematical and logical calculations. As shown in Fig. 3, the students are required to select the training patterns for the image segmentation from the background, clothing, face, and hair regions. The front-view image is called the initial view. Since color features can handle a wide range of variations in a single image, we use a color classification algorithm. Fig. 4 shows the segmentation result for the initial image. We analyze the head posture images and use them to understand the student interaction during the e-learning process.

To simplify the process, we focus on the analysis of seven different head posture images per student.

A. Human-Presence Detection

Detecting a human presence in an e-learning system is relatively simple because the presence of a certain amount of skin and hair information suffices [15], [16]. Presence detection is an important step in checking the students' levels of interaction with their computers. The image segmentation based approach is fast and robust because it is independent of the movement of facial features such as the lips and eyes, and the position of head postures.

We use the face and hair pixel information to identify the level of a student's presence via

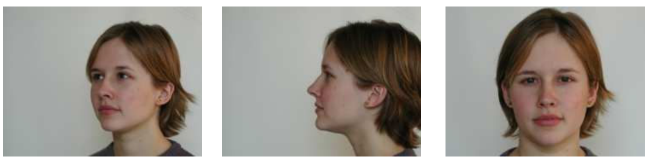
$$P_i = \frac{(f_i + h_i)}{(f_0 + h_0)} 100 \qquad (1)$$

where f_0 and h_0 are the numbers of face and hair pixels in the initial view, and f_i and h_i are the numbers of these pixels in a different head posture. For the initial view, $f_i = f_0$ and P_i equals 100.

These values change whenever a student moves his/her head to a different position or moves out of the range of the camera. For example, P_i becomes zero when the student is no longer present.

However, the presence value changes significantly for partial or near/distant views.

(a) (b) (c)



(d) (e) (f)

(g)

Fig. 2 Student's head posture images: a) initial image, b) slight right view, c) full right view, d) slight left view, e) full left view, f) near view, g) distant view.

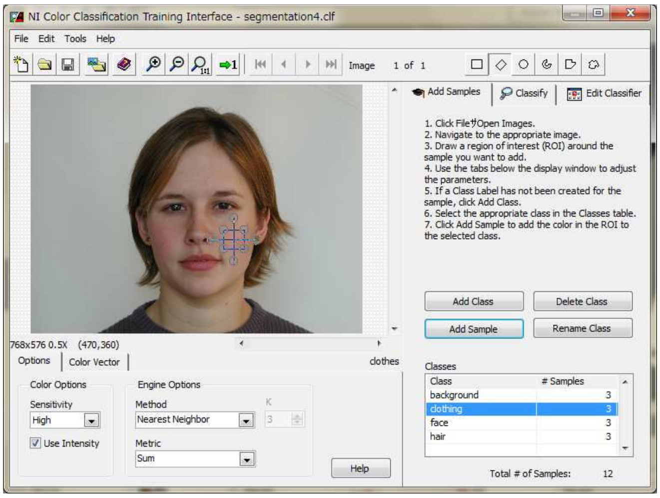


Fig. 3 Supervised training using NI Vision Assistant

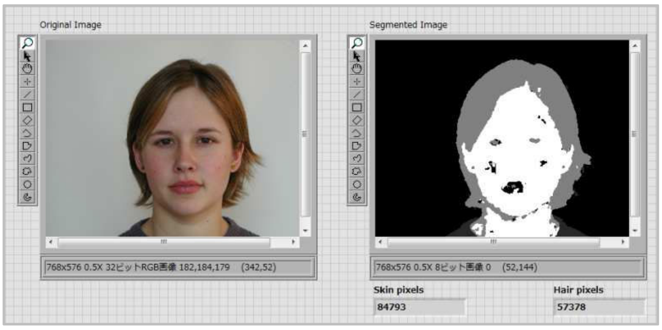


Fig. 4 Head posture image and corresponding segmentation

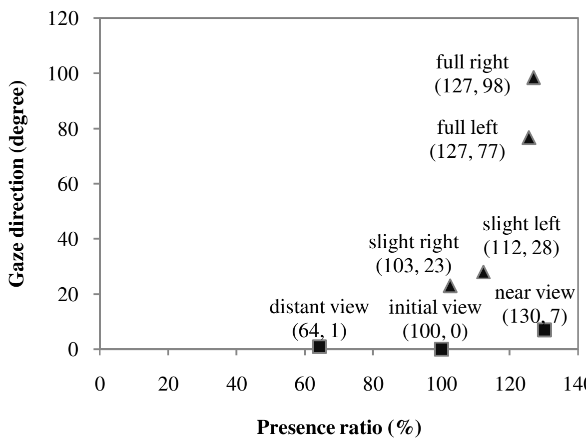


Fig. 5 Head posture analysis results

B. Gaze-Direction Estimation

Gaze-direction estimation is another criterion used to measure human-computer interaction [17]. When the students' gaze direction changes, the hair-to-face ratios of the segmented images also change. The following equation is used to estimate the gaze direction:

$$D_i = \left| 100 - \frac{h_i/(f_i+h_i)}{h_0/(f_0+h_0)} 100 \right| \frac{P_i}{P_0}\, m \qquad (2)$$

where m is the multiplier and

$$m = \frac{h_0}{I}\, n \qquad (3)$$

Here n is a constant, which is calculated experimentally and set to 10 for women and 20 for men, and I is the total number of pixels in the image.

The initial value of D_0 is zero; it indicates that the student is looking directly at the computer screen. If the student's face and hair ratio do not change, the D_i value remains close to zero. However, when the student looks in a different direction, this ratio changes significantly. Needless to say, more hair information is available in many situations, especially in the side or back views. There are clearly personal hair-visibility differences depending on gender and hairstyle, and in some cases there is no hair.

Fig. 5 shows the results for the previously introduced images of students in two-dimensional spaces. The rectangles indicate that the students' gaze direction is close to zero. Similarly, the triangles indicate that the presence ratio or gaze direction is larger than in the initial image. In case of the initial view, P_0 and D_0 are 100% and 0°, respectively.

However, when the head posture changes, the presence level and gaze direction are affected. In the initial view, distant view, and near view, the students appear to be properly interacting with the computer. However, a larger gaze direction usually indicates a human-computer interaction failure.

C. Human-Computer Interaction Check

A decision-making algorithm is used to identify the human computer interaction, based on the human-presence ratio and the gaze-direction estimation.

The following if case is used:

$$R_i = \begin{cases} True, & P_i \geq 40 \text{ and } D_i \leq 10 \\ False, & otherwise \end{cases} \qquad (4)$$

The results show that the presence ratio can be changed if students get closer to the camera or move beyond its range during the class. Therefore, a presence change of up to 40% is considered acceptable. However, the gaze-direction algorithm also has an error range in the case of head movements. Therefore, to be safe, a direction change of more than 10° is considered to indicate a human-computer interaction failure for the e-learning system.

In this decision-making process, *True* means that the student is sufficiently interacting with the computer and learning is in progress. *False* implies that no learning is happening because the student has weak computer interaction or has lost interest.

This information can be integrated into e-learning systems to improve the content or evaluate the performance and level of interest of individual students.

IV. Evaluation Results

As a final step, we applied our algorithm to ten students using the MIT-CBCL face recognition database, as shown in Fig. 6. Each person has a different hairstyle and color and wears different clothing. The initial images were used for training purposes, and the remainder of the sixty images were evaluated based on our algorithm, using (4).

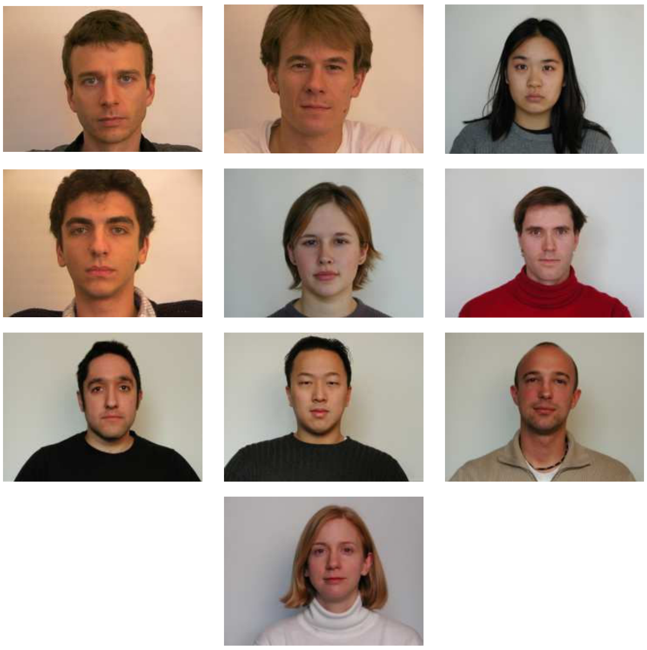


Fig. 6 Student head posture database

TABLE I
Student-Posture Image Evaluation results

	Distant view	Near view	Slight turn view	Full turn view
Presence ratio Gaze direction	64%	134%	28°	81°
Standard deviation	1.7%	7.3%	12.8°	16.8°

The results show that 92% of the head posture images were correctly identified; only 8% of the images were misclassified because of clothing and hair-color similarities. Our algorithm is therefore effective for most types of heads, hairstyles and genders. However, the failure rate may increase if the student does not have hair or is working in an environment with a complex background.

The human-presence detection algorithm was robust for both near and distant views for the ten students. The values calculated for the distant and near views were close to the average values; the standard deviation is less than 8%. Table I shows the average values of the ten subjects for different head postures.

The gaze-direction estimation was more sensitive to personal differences such as hairstyle, gender, or lighting conditions. These values fluctuated over the different head posture images, as shown in Table I. Despite this, the calculated average value for slight right/left views was 28° and that for full right/turn views was 81°. In addition, the average values of both the human-presence ratio and the gaze-direction estimation were close to the actual head positions despite the limited number of sample images.

V. Conclusion

We have proposed a new approach using head posture images for next-generation e-learning systems. Including human interaction data in e-learning systems will lead to new possibilities for building smart and adaptive learning environments. As a result, this information can also be used to improve the e-learning content or evaluate the student performance. However, the integration of the human-computer interaction data into the e-learning system is still an open topic and needs further investigation and experimental studies.

In future work, we plan to use video images to include sequential information to eliminate the fault detections. On the other hand, stereo cameras can also be used to increase the accuracy of the head posture analysis. The use of voice signals is another option for a better understanding of human-computer interaction. In addition, future work will consider integration and assessment methods of human-computer interaction.

Acknowledgment

We are grateful to the Massachusetts Institute of Technology and to the Center for Biological and Computational Learning for providing the facial image database.

References

[1] M. Rosenberg, *E-Learning: Strategies for Delivering Knowledge in the Digital Age*, McGraw Hill, Columbus, OH, 2001.
[2] S. Alexander, "E-learning developments and experiences," *Education + Training*, Vol. 43 No. 4/5, pp. 240-248, 2001.
[3] L. Shen, M. Wang, and R. Shen, "Affective e-learning: Using emotional data to improve learning in pervasive learning environment," *Educational Technology & Society*, Vol. 12, No. 2, pp. 176-189, 2009.
[4] H. Ekanayake, D.D. Karunarathna, and K.P. Hewagamage, "Cognitive architecture for affective e-learning," *Int. J. of the Computer, the Internet and Management*, Vol. 14, No. SP1, 2006.
[5] R.W. Picard, *Affective Computing*, MIT Press, Cambridge, Massachusetts, 1997.
[6] F. Karray, M. Alemzadeh, J.A. Saleh, and M.N. Arab, "Human-computer interaction: Overview on state of the art," *Int. J. on Smart Sensing and Intelligent Systems*, Vol. 1, No. 1, pp. 137-159, 2008.
[7] B. Shneiderman, "The future of interactive systems and the emergence of direct manipulation," *Behavior and Information Technology*, pp. 237-256, 1982.
[8] M. Eisenhauer, B. Hoffman, and D. Kretschmer, "State of the art human-computer interaction," *GigaMobile*/D2.7.1, 2002.
[9] Y. Ugurlu and H. Sakuta, "E-learning for graphical system design courses: A case study," *IEEE Int. Conf. on Technology Enhanced Education*, Kerala, India, 2012.
[10] S.Z. Li and A.K. Jain. *Handbook of Face Recognition,* Springer, 2004.
[11] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 15, pp. 1042-1052, 1993.
[12] P. Viola and M. Jones, "Robust real-time face detection," *Int. J. Computer Vision*, Vol. 57, No. 2, pp. 137-154, 2004.
[13] R. L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5, pp. 696-706, May 2002.
[14] B. Weyrauch, J. Huang, B. Heisele, and V. Blanz. "Component-based face recognition with 3D morphable models," *First IEEE Workshop on Face Processing in Video*, Washington, D.C., 2004.
[15] S.L. Phung, A. Bouzerdoum, and D. Chai, "Skin segmentation using color pixel classification: Analysis and comparison," *IEEE Trans. on Pattern Analysis and Machine Intelligence.,* Vol. 27, No.1, pp.148-154, 2005.
[16] M.J. Jones and J.M. Rehg, "Statistical color models with application to skin detection," *Proc. IEEE Conf. Computer Vision and Pattern Recognition.*, Fort Collins, pp. 274-280, 1999.
[17] L.E. Sibert and R.J.K. Jacob, "Evaluation of eye gaze interaction," *Proc. Human Factors in Computing Systems Conf*, Addison-Wesley/ACM Press, pp. 281-288, 2000.

Y. Ugurlu (M'09) was born in Turkey and received his B.S. and M.S. degrees in electronics engineering from Ankara University, Turkey, in 1993 and 1995, respectively. He won a Minister of Education (Monbusho) scholarship to continue his postgraduate study in Japan. He received his Ph.D. degree in information processing from the Tokyo Institute of Technology, Japan in 1999. He then joined Olympus Corporation's R&D Center to develop advanced ultra-high-vision natural vision systems. He is currently working at National Instruments Japan to promote educational and research programs in science and engineering. He has been a lecturer and research associate at Aoyama Gakuin University since 2010. His research interests are engineering education, learning environments and tools, e-learning, human computer interaction, machine vision, and image processing. Dr. Ugurlu is a member of IEEE, the Japanese Society for Engineering Education, and the Japanese Society of Applied Science.