REVIEW OF HUMAN-COMPUTER INTERACTION, RECOGNITION AND DETECTION SYSTEM

Mote Ganesh Mohanrao

Department of Computer Science and Engg ,S.T.B. College of Engg,Tuljapur,MS(India)

ABSTRACT

Human-Computer Interaction, Recognition and Detection System is the visual interaction between a human and a computer system, through various motions, posture, hand gestures, pointing gestures and other visual cues. The proposed system can be either a standalone system, or an additional system to make another (e.g. speech-based, gesture based) system more robust. The system is designed to be unnoticeable (marker-less) and fast enough (real-time) to provide a smooth user interaction. Such systems are useful as technically improved computerized systems which are more improved than traditionally keypad, mouse based systems towards more natural gestures like scroll wheels, multi-touch interfaces. Real-time analysis of human body pose or hand gestures enables to develop new, more natural interfaces.

The real-time nature of such systems shifts the focus away from complex trackers that rely on heavy iterative optimizations, as they require too much computation time for processing. The proposed work is of two methods. The difference between these two models namely model based and example based approaches, is that in example based approaches the current pose or gesture is detected from a database of preset poses or gestures. The benefit of such approach is that the system can be much faster, as compare to other.

Keywords *Design, interaction, interaction design, interaction design research, Pose/Gesture*

I. INTRODUCTION

The Human-Computer Interaction systems are the essential and emerging technology in the perspective of the computer vision trends now a days. The intelligent interaction of Human with the computer systems as on the user's side is constrained by the nature of human communication organs and abilities like motions, posture, hand gestures, pointing gestures and other visual cues; on the computer side, it is constrained only by input/output devices and methods that we can invent. The challenge before us is to design new devices and types of dialogues that better fit and exploit the communication-relevant characteristics of humans. The relationship between basic research and application development in this area ideally forms a circular chain: specific interface problems encountered in applications are generalized and then solved in basic research by inventing new interaction modes or techniques; and these general approaches can then be applied to the development of specific user interfaces.

Direct manipulation interfaces have enjoyed great success, particularly with new users, largely because they draw on analogies to existing human skills (pointing, grabbing, moving objects in space), rather than trained behaviors. As another example, research on eye movement-based computer input at NRL has tried to make use of natural eye movements (Jacob, 1990). Because eye movements are so different from conventional computer inputs, the overall approach in designing interaction techniques was, wherever possible, to obtain information from a user's natural eye movements while viewing the screen rather than requiring the user to make specific trained eye movements to actuate the system. This work began by studying the characteristics of natural eye movements and then attempted to recognize appropriate patterns in the raw data obtainable, turn them into tokens with higher-level meaning, and design interaction techniques for them around the known characteristics of eye movements.

According to Diaper (2005) the chronology of Human Computer Interaction starts in 1959 with Shakel's paper on "The ergonomics of a computer" which was the first time that these issues were ever addressed. This was followed by Licklider who produced what has come to be known as the seminal paper (1960)on "Man – Computer Symbiosis" which sees man and computer living together. Now it is emerging trends in many application areas and various papers and research work is going on.

II. LITERATURE REVIEW

2.1 Human Computer Interaction Research Threads

A blueprint for intellectual histories of Human-Computer Interaction was established by Ron Baecker in the opening chapters of the 1987 and 1995 editions of Readings in Human-Computer Interaction. It was followed in Richard Pew's chapter in the 2003 version of this handbook. Further insights and references can be found in Brian Shackel's (1997) account of European contributions, and specialized essays by Brad Myers (1998) on Human-Computer Interaction engineering history and Alan Blackwell (2006) on the history of metaphor in design. The Human-Computer Interaction threads are Human Factors and Ergonomics, Information systems, Computer human Interaction and the Information Fields.

In 1970 at Southborough University in England, Brian Shackel founded the Human Sciences and Advanced Technology (HUSAT) center, devoted to ergonomics research and emphasizing Human-Computer Interaction. Sid Smith and other human factors engineers worked on input and output issues, such as the representation of information on displays (e.g., Smith, Far-quhar & Thomas, 1965) and computer-generated speech (Smith & Goodwin, 1970). The Computer Systems Tech-nical Group (CSTG) of the Human Factors Society was formed in 1972 and became the largest technical group in the society.

The general Human Factors journal was joined in 1969 by the computer-focused International Journal of Man-Machine Studies (IJMMS). Companies acquired expensive business computers to address major organizational concerns. Even when the principal concern was simply to appear modern (Greenbaum, 1979), the desire to show benefits from a multi-million dollar investment could chain managers to a computer almost as tightly as the operator and data entry 'slaves.' In addition to being expected to make use of output, they might encounter resistance to system acceptance.

2.2 Human Computer Interaction Relating to Human Psychology

Human Computer Interaction relating to human psychology. Specific examples of research in the areas of icons and menus are then reviewed. The results of these experiments and the predictions they make about general human psychology and specific human interaction with computers is discussed. To understand the fundamentals of Human Computer Interaction, and how it relates to human psychology and physiology, we ask: "What can we learn about human perception and cognition from studying the way in which humans interact with computers?" To answer this question, this paper investigates two aspects of the computer interface icons and menus and reviews research that has been done on these interface elements in conjunction with how humans use them.

2.3 Design Research

In the Human Computer Interaction community and in the design practice community, the term design research is generally used to refer to the upfront research practitioners do to ground, inform, and inspire their product development process. However, in the design research community, including institutions such as the Design Research Society, the term design research implies an inquiry focused on producing a contribution of knowledge. This follows the convention of the design researchers, and intends the term design research to mean an intention to produce knowledge and not the work to more immediately inform the development of a commercial product.

2.4 A Survey of Human Computer Interaction Design in Science Fiction Movies

The motion picture industry is a major entertainment sector with a considerable impact on the general public mindset throughout all social classes. In return some producers or directors attempt to catch the spirit of the time, trying to pick up existing or emerging trends, which leads to interesting interactions between fictional worlds and the real one. Particularly science fiction movies due to their inherent theme are set in a world with advanced, fictional technology that is normally set in the future. Most of these movies expose their own unique vision of the future, with new technologies commonly being the most noticeable change in these hypothetical worlds. Besides visions regarding all kinds of technologic and scientific areas, human-computer interfaces are an important recurring component as they visualize otherwise abstract and possibly invisible technologies. In order to be able to draw comparisons to reality, it is often necessary to view the ideas and visions of a movie on a relatively abstract level. We observed an interesting development in the history of science fiction movies, beginning with films unrelated to technological tendencies. This phase was followed by movie makers taking current trends into account and by ideas of movies inspiring researchers, until finally Michael Schmitz et al. movie makers and Human Computer Interaction researchers are found to work together to create artificial yet authentic worlds [5].

III. SURVEY ON HUMAN COMPUTER INTERACTION

3.1 The Role of Design in Human computer Interaction

In the early days, the term "design" within the Human Computer Interaction community meant usability engineering: "It is the process of modeling users and systems and specifying system behavior such that it fitted the users' tasks, was efficient, easy to use and easy to learn." Over time, trained designers began working with software developers, bringing skills visual hierarchy, navigation, color, and typography they had developed designing printed artifacts. Jonas Löwgren labeled the process they brought to interaction design as "creative design" to distinguish it from the engineering approach [3]. In engineering design, developers created software to meet a specification, and in creative design, designers continually reframed the problem, constantly questioning the underlying assumptions during the design process.

Daniel Fallman's work casts Human computer Interaction as a design discipline [2]. He describes the research performed by engineers and behavioral scientists as "design-oriented research." Researchers engage in designing and making prototypes in order demonstrate a research contribution. In this case, the research community benefits from the processes of design and design thinking because they lead to better research prototypes. Christopher Alexander's work on Pattern Languages represents an example of how research performed by design researchers on design methods has had an impact on the Human computer Interaction community. His work asks design researchers to examine the context, system of forces, and solutions used to address repeated design problems in order to extract a set underlying "design patterns", thereby producing a "pattern language" [1]. The Human computer Interaction community has embraced this approach to address design of web sites [4]. The method turns the work of many designers addressing the same interaction problems into a discourse for the community, allowing interaction designers to more clearly observe the formation of conventions as the technology matures and is reinterpreted by users.

3.2 Current Research

In Human-Computer Intelligent Interaction, the three dominant clusters of research are currently in the areas of face analysis, body analysis, and complete systems which integrate multiple cues or modalities to give intuitive interaction capabilities to computers. Face and human body analysis includes modeling the face or body and tracking the motion and location. They are considered fundamental enabling technologies towards Human computer interaction. Understanding the dynamics of the human face is certainly one of the major research challenges. The potential of the following work is toward detecting and recognizing emotional states and improving computer to human communication. Valstar and Pantic[6] propose a hybrid technique where a Support Vector Machine (SVM) is used to segment the facial movements into temporal units. Then a Hidden Markov Model is used for classifying temporal actions. Chang et al. [7] propose a framework for using multiple cameras in pose and gaze estimation using a particle system approach.

Another major research challenge is human body analysis. The following research has the potential to endow computers to see where humans are, what their gestures and motions are, and track them over time. Oerlemans, et al.[8] propose a tracking system which uses the multidimensional maximum likelihood approach toward detecting and identifying moving objects in video. Their system has the additional novel aspect of allowing the user to interactively give feedback as to whether segmented objects are part of the background or foreground. Cooper and Bowden[9] address the problem of large lexicon sign language understanding. The method detects patterns of movement in two stages involving viseme classifiers and Markov chains. Angelopoulou et al.[8] are able to model and track non rigid objects using a growing neural gas network.

Jung, et al.[9] use a multi-cue method to recognize human gestures using AdaBoost and Fisher discriminant analysis. Chu and Nevatia[10] are able to track a 3D model of the human body from silhouettes using a particle filtering approach and infra-red cameras. In addition to novel face and human body analysis, current research is creating systems which are delving into the complete human-computer interaction, thereby endowing computers with new important functionality. State[11] had created a system where a virtual human can maintain exact eye contact with the human user thereby significantly improving the perception of immersion. Rajko and Qian[12] propose an automatic kinematic model building method for optical motion capture using a Markov random field framework. Vural, et al.[13] created a system for detecting Human-Computer Intelligent Interaction: A Survey 3 driver drowsiness based on facial action units and classified using AdaBoost and multinomial ridge regression. Thomee, et al.[14] used an artificial imagination approach to allow the computer to create example images to improve relevance feedback in image retrieval. Barreto, et al.[15] use physiological responses and pupil dilation to recognize stress using an SVM.

IV. RESEARCH GOALS

To carry out this research, we have to fulfill the following goals:

4.1 Segmentation

Segmentation will create a simpler user image in which targets are intermixed with other object/persons. Looking at the user image, pick a feature that seems likely to characterize the targets. The entire system depends on the segmentation of the user and/or his body parts. The better the segmentation, the easier the following tasks become. However, segmentation can never be perfect, hence the other components in the system will have to be robust to noise and reflections. The goal of the segmentation stages is to have a segmented shape of the user in each camera, a segmented outline of the hands, and an approximate location of face.

4.2 3D Reconstruction

Reconstruction captures the object/person of interest from different viewpoints Sample shape of the user taken from several cameras surrounding the user, 3D voxel reconstruction can provide us with a 3D voxel hull of the user, which is very useful for detecting the pose of the user.

4.3 Tracking

Difficulties in tracking object/persons can arise due to abrupt object/person motion, changing appearance patterns of both the object/person and the scene, non-rigid object/person structures, object/person to object/person and object/person to scene occlusions, and camera motion. Tracking is usually performed in the context of higher level applications that require the location or shape of the object/person. Tracking is then performed by first initializing the object/person position with its previous spatial position, then computing the shift vector which maximizes the likelihood of observing the object/person given the a priori object/person model. One of the most fundamental forms of tracking is tracking the position and orientation of the user from an overhead camera. The proposed tracking functional, which uses a support region around the contour, suppresses the noise and artifacts that generally occur during course of tracking.

4.4 Pose/Gesture Recognition

Although recent research has shown improved performance by embedding the color components, the effectiveness of color information in the RGB color space in terms of recognition performance depends on the type and angle of light source, often making recognition impossible. Poses, motions and gestures are detected using example based approaches, rather than full-body tracking or articulated hand tracking.

4.5 Classification

Computer vision involves the identification and classification of object/persons in an image, edge detection is an essential tool. In each case there is an intricate scheme for classification, based on overall shape (elliptical, circular, etc.), type and degree of irregularities (convex, rough or smooth outline, etc.), internal structures (holes, linear or curved features) that has been accumulated over many years of observation. Different classification algorithm can be implemented as per the object/person.

V. CONCLUSION AND FUTURE SCOPE

A Human-Computer Interaction systems trends are to be focus on natural movements of human which application of these results: it can lead to faster and more natural communication with interactive systems, enable better quality and efficiency in the operation of such systems, and improve the working conditions of their users by providing them with richer and more natural means of communication. We have highlighted major vision approaches for different modal human-computer interaction. We discussed techniques for large-scale body movement, gesture recognition, and gaze detection. We discussed facial expression recognition, emotion analysis from audio, user and task modeling, multimodal fusion, and a variety of emerging applications. There is no evidence that individuals in actual social interaction selectively attend to another person's face, body, gesture, or speech, or that the information conveyed by these channels is simply additive. The central mechanisms directing behavior cut across channels, so that, for example, certain aspects of face, body, and speech are more spontaneous and others are more closely monitored and controlled. It might well be that observers selectively attend not to a particular channel but to a particular type of information (e.g., cues to emotion, deception, or cognitive activity), which may be available within several channels. No investigator has yet explored this possibility or the possibility that different individuals may typically attend to different types of information. Human Computer Interaction becomes more complicated when the modern technologies will used to do the interaction. Human motions, pose and gestures are not unique so it is difficult when new action is to resolved.

REFERENCES

[1] C.Alexander, et al., A Pattern Language:Towns, Buildings, Construction. Oxford University Press, 1977.

[2] D.Fallman,"Design-Oriented Human-Computer Interaction. Proc". CHI 2003, ACM Press (2003), 225-232.

[3] J.Löwgren," Applying Design Methodology to Software Development. Proc. of DIS" 1995, ACM Press (1995), 87-95.

[4] T.Wolf, J.Rode, J.Sussman, W.Kellogg," Dispelling Design as the 'Black Art' of Chi. Proc. Of CHI" 2006, ACM Press (2006), 521-530.

[5] M.Schmitz and C.Endres,**"**A Survey of Human Computer Interaction Design in Science Fiction Movies ".

[6] M.Valstar, M. Pantic, "Combined Support Vector Machines and Hidden Markov Models for Modeling Facial Action Temporal Dynamics."In: Lew, M., Sebe, N., Huang,T.S., Bakker, E.M. (eds.) HCI 2007. LNCS, vol. 4796, pp. 118–127. Springer, Heidelberg (2007)

[7] C. Chang, C.Wu, C., Aghajan, ," Pose and Gaze Estimation in Multi-Camera Networks for Non-Restrictive HCI" In: Lew, M., Sebe, N., Huang, T.S., Bakker, E.M. (eds.) HCI 2007. LNCS, vol. 4796, pp. 128–137. Springer, Heidelberg (2007)

[8] Oerlemans, A., Thomee, B.: Interactive Feedback for Video Tracking Using Hybird Maximum Likelihood Similarity Measure. In: Lew, M., Sebe, N., Huang, T.S., Bakker, E.M. (eds.) HCI 2007. LNCS, vol. 4796, pp. 79–87. Springer, Heidelberg (2007)

[9] Cooper, H., Bowden, R.: Large Lexicon Detection of Sign Language. In: Lew, M., Sebe, N., Huang, T.S., Bakker, E.M. (eds.) HCI 2007. LNCS, vol. 4796, pp. 88–97. Springer, Heidelberg (2007)

[10] Jung, S., Guo, Y., Sawhney, H., Kumar, R.: Multiple Cue Integrated Action Detection. In: Lew, M., Sebe, N., Huang, T.S., Bakker, E.M. (eds.) HCI 2007. LNCS, vol. 4796, pp. 108–117. Springer, Heidelberg (2007)

[11] Chu, C.-W., Nevatia, R.: Real Time Body Pose Tracking in an Immersive Training Environment. In: Lew, M., Sebe, N., Huang, T.S., Bakker, E.M. (eds.) HCI 2007. LNCS, vol. 4796, pp. 146–156. Springer, Heidelberg (2007)

[12] State, A.: Exact Eye Contact with Virtual Humans. In: Lew, M., Sebe, N., Huang, T.S., Bakker, E.M. (eds.) HCI 2007. LNCS, vol. 4796, pp. 138–145. Springer, Heidelberg (2007)

[13] Vural, E., Cetin, M., Ercil, A., Littlewort, G., Bartlett, M., Movellan, J.: Drowsy Driver Detection Through Facial Movement Analysis. In: Lew, M., Sebe, N., Huang, T.S., Bakker, E.M. (eds.) HCI 2007. LNCS, vol. 4796, pp. 6–18. Springer, Heidelberg (2007)

[14] Thomee, B., Huiskes, M.J., Bakker, E., Lew, M.: An Artificial Imagination for Interactive Search. In: Lew, M., Sebe, N., Huang, T.S., Bakker, E.M. (eds.) HCI 2007. LNCS, vol. 4796, pp. 19–28. Springer, Heidelberg (2007)

[15] Barreto, A., Zhai, J., Adjouadi, M.: Non-intrusive Physiological Monitoring for Automated Stress Detection in Human-Computer Interaction. In: Lew, M., Sebe, N.,Huang, T.S., Bakker, E.M. (eds.) HCI 2007. LNCS, vol. 4796, pp. 29–38. Springer,Heidelberg (2007)