Human-Computer Interaction: Preamble and Future in 2020.

Ashish Palsingh

J. G. Institute of Computer Applications
Gujarat University, Ahmedabad, India.
ranvirsingh1069@rediffmail.com

Ekta Palsingh

Shree M. M. Patel Institute of Sciences and Research.
Kadi Sarva Vishwavidyalaya, Kadi, India.
ekta_5balva@rediffmail.com

Abstract-The growth in Human-Computer Interaction (HCI) field has not only been in quality of interaction, it has also experienced different branching in its history. The intention behind writing this paper is to provide an overview on the subject of Human-Computer Interaction. The main topics which are covered in the paper include history, Unimodal of HCI Systems, and the future of HCI. The paper also includes the references.

Keywords-Human-Computer Interaction; History; Unimodal HCI; Future.

I. INTRODUCTION

Human-computer interaction, also known as Man-Machine Interaction, is a concept that emerged side by side with computers. If machines are not using by men then they are worthless. The method is traveled a long way by which human has been interacting with computers. The journey still continues and new designs of technologies and systems appear more and the research in this area has been growing very fast in the last some decades. HCI (human-computer interaction) is the study of how people interact with computers and to what extent computers are or are not developed for successful interaction with human beings. Software engineering focuses on the production of software application solutions, whereas HCI focuses on discovering methods and techniques that support people. HCI has expanded rapidly and steadily for three decades, attracting professionals from many other disciplines and incorporating diverse concepts and approaches.

II. ERA OF HUMAN-COMPUTER INTERACTION

The first Electronic Numerical Integrator and Computer (ENIAC) were developed in 1943 which was the world's first computer. Then, in 1944 the Mark-I paper was developed which was used to read the paper tape. In the year 1960 to 1980 the use of computer was dramatically increased in which the use of input devices such as data tablets (1964) – display processors capable of real-time manipulation of images (1968) was in main focus and so the requirement to think on Human-Computer Interaction was also increased. In 1962, A SRI Report on "A Conceptual Framework for Augmenting Human Intellect" was developed by Douglas Engelbart. The first mouse was developed in 1964. In 1974, Ted Nelson wrote a book on "Computer Lib/Dream Machines" which was very popular at that time; the book included what computers can do for people instead of business! The first commercial personal computer designed for "business professionals" was developed in 1981. The features of that computer includes: familiar user's conceptual model (simulated desktop), promoted recognizing/pointing rather than remembering/typing, property sheets to specify appearance of objects, what you see is what you get (WYSIWYG), small set of generic commands that could be used throughout the system, high degree of consistency and simplicity, mode less interaction, limited amount of user tailor ability. In early 1983, Apple introduced Lisa, a commercial machine, but because of its cost and its product positioning and its inadequate application base, it was the commercial failure. Then in January of 1984, Apple introduced Macintosh at a price of approximately $2,500. In about a year, its success was clear. The main reasons behind Mac' success was: it did not need to trial blaze, it was second generation of Lisa, so Apple had the opportunity to learn from the experience and eliminated many bugs, the product had the excellent graphics.

III. UNIMODAL HCI SYSTEMS

Modality means each of the various independent single channels. A system that is based on only one modality is known as unimodal. And on the contradictory, those which are based on many modality is multimodal. In this paper, the detail concept of unimodal is given and we will include the detail of multimodal in the next paper. They can be divided into three categories based on the nature of different modalities:

A. Vision-Based

The most extensive area in HCI research is apparently the visual based human computer interaction. Following are the some of the main research areas:

1) Facial Expression Analysis:

People usually recognize the face easily and without many efforts, but then, it has remained a difficult problem in the area of computer visual, where to yield useful technological solutions, it took some 20 years of research.

As a biometric technology, there are number of delectable properties for automated face recognition that are submerging research with practical techniques. Every medium of input which is used for face recognition brought robustness to certain conditions, for example, infra-red face imaging is practically invariant to lighting conditions while 3-dimensional data in theory is invariant to head pose. Because of the large quantity of legacy data and the cheapness of photographic capture equipment, picturing in the visible light spectrum, however, will remain the leading domain for research and application of face recognition

2) Body Movement Tracking (Large-scale):

Motion capture is the concept of registering the movement of people or object. It is used in various fields like sports, medical applications, military, entertainment, and for validation of computer vision and robotics.

3) Gesture Recognition:

Identification of gesture makes it possible for humans to interact effortlessly and to connect with the machine deprived any mechanical devices. By the use of theory of gesture recognition, it is feasible to move the cursor as accordance with our point of a finger at the computer. It is same like the conventional input devices such as mouse, keyboards and even touch-screens redundant.

It cut down the impact on hardware on a system and also boosts the scope of physical world items usable for control beyond traditional digital items like mice and keyboards. Such implements can empower a new wide range of hardware which does not require monitors. This idea may lead to the creation of holographic display. The term gesture recognition has been used to refer more narrowly to non-text-input handwriting symbols, such as inking on a graphics tablet, multi-touch gestures, and mouse gesture recognition.

The capacity to record a person's gestures and determine what movement they may be accomplishing can be carry out through different tools. However, there is a huge number of research has been carried out in image based gesture and in video based gesture recognition, there is some differentiation within the tools and environments used between implementations.

- *Wired gloves:* Those gloves which can provide a computer input by the help of position and movement of the hands by the use of magnetic or inertial tracking devices are the wired gloves. Moreover, few gloves can discover finger bending with a the accuracy of a high degree about 5-10 degrees. DataGlove was the first merchandising available hand-tracking glove-type device. This device is useful in detecting hand position, movement and finger bending. The main part of the wired glove is it uses fiber optic cables which runs down the back of the hand and which is used to detect the gesture of a hand. It creates light pulses and registers the pulses when the fingers are bent, light leaks through small cracks and the loss, giving an approximation of the hand pose.

- *Depth-aware cameras:* Using particular cameras such as structured light, one can bring out a depth map of a short range of what is being seen by the help of camera, and use this data to approximate a 3-dimensional illustration of what is being seen. These can be effective for detection of hand gestures due to their short range capabilities.

- *Stereo cameras:* A 3-dimensional representation can be approximated by the output of the cameras, using two cameras whose associations to each other are known. To get the cameras' relations, one can use a arranging reference such as a lexian-stripe or infrared emitters. In combination with direct motion measurement, 6-dimensional gestures can directly be detected.

- *Controller-based gestures:* These types of controllers act as an enlargement of the body so that when actions are performed, some of their gestures can be conveniently conquer by software. One such example is mouse gesture; where the movement of the mouse is correspond to a symbol being drawn by a person's hand, as is the Wii Remote, which can study changes in acceleration over time to represent gestures. Devices such as the LG Electronics Magic Wand, the Loop and the Scoop use Hillcrest Labs' Freespace technology, which uses MEMS accelerometers, gyroscopes and other sensors to translate gestures into cursor movement.

- *Single camera:* A single standard 2-dimensional camera can be used for gesture recognition where the environment would not be agreeable for other forms of image-based recognition. Previously, it was

thought that single camera may not be as impressive as stereo or depth aware cameras, but few companies are demanding this ideology. Software-based gesture recognition technology using a standard 2-dimensional camera that can encounter booming hand gestures, hand signs, as well as track hands or fingertip at high certainty has already been fixed in Lenovo's Yoga ultra books, Pantech's Vega LTE smart phones, Hisense's Smart TV models, among other devices.

4) Gaze Detection (Eyes Movement Tracking):

Gaze detection is the process of electronically discovering the point of a person's gaze, or following and registering the action of the point of gaze. Different technologies exist for achieving this task; few methods involve connections to the eye, while others depend on images of the eye taken without any physical contact.

B. Audio-Based HCI

The audio based interaction is a important area of Human-Computer Interaction. This field accord with information captured by various audio signals. While the nature of signals may not be as variable as visual signals but the collected information from audio signals can be more trust able, and is some cases unique providers of information. Research areas in this section can be segmented to the following categories:

1) Speech Recognition:

SRT means Speech recognition technology, other name is ASR (automated speech recognition), CSR (continuous speech recognition) and VR (voice recognition), indicates that the computer software systems that convert the spoken word to text. This technology is becoming more and more popular in the Medicare field, as it is being marketed to institutions and physicians as a way to boost productivity and decrease the costs.

Many voice recognition systems are available. The most capable can identify thousands of words. Although, they generally need an lengthy training session during which the computer system becomes accustomed to a appropriate voice and accent. Those systems are known as *speaker dependent.*

Front-End SRT.

Front-end Speech Recognition Techniques includes such consumer applications as Via Voice by IBM, etc. Users of front-end SRT edict into a microphone and those spoken words are converted to text in any type of word processing application. In order for front-end SRT to be as accurate as possible, a user must immediately correct the errors made by the software so the program will learn the nuances of the user's speech patterns.

Back-End SRT

The category of Speech Recognition Techniques used by many institutions and clinics is back-end SRT. By the help this method, the certain speech-to-text conversion takes place after the speaker has dictated, rather than concurrently. The dictation is recorded in digital form at the time of dictation, and then the digital voice files are processed by a powerful computer running SRT software and converted to a draft text document. A human speech recognition editor must then listen to the voice file while proofreading the draft document because even the most sophisticated SRT applications are not nearly accurate enough to eliminate the need for human review.

2) Speaker Recognition:

The process in which a user's claimed individuality can be finding is by using the technique of Speaker Recognition. This technique is one of the most popular and beneficial biometric identification techniques in the world especially in the areas in which security is a major concern. It can be used for authentication, surveillance, forensic speaker recognition and a number of related activities.

This technique can be categorized into identification and verification. The process of determining which registered speaker provides a given utterance is known as speaker identification. On the other hand, speaker verification, is the process of accepting or rejecting the identity claim of a speaker. Speaker verification systems are: Time and Attendance Systems, Access Control Systems, Banking/Broking, Biometric Login to telephone aided shopping systems, Information and Reservation Services, Security control for confidential information, Forensic purposes.

3) Human-Made Noise/Sign Detections (Gasp, Sigh, Laugh, Cry, etc.):

A Human-Made Noise/Sign Detection is the process which process and analyzes the voice and find the emotions on the basis of different fluctuation in the voice like open/close ratio of the vocal chords, and the quality of the voice. Sadness for example influences the voice quality so that creaky voice may be produced. In this case the speech is of low pitch and low intensity. Other than the tone and pitch of speech data, typical human auditory signs such as sigh, gasp, and etc helped emotion analysis for designing more intelligent HCI system.

4) Musical Interaction:

Music interaction refers to "Music and Human Interaction". Music Interaction encompasses the design, refinement, evaluation, analysis and use of interactive systems that involve computer technology for any kind of musical activity. Music interaction has serious implications for music, musicians, educators, learners and those seeking deeper involvement in music. But music interaction is also a valuable source of challenges, new ideas, and new techniques for HCI.

C. Sensor-Based HCI

This section is a combination of variety of areas with a wide range of applications. The commonality of these different areas is that at least one physical sensor is used between user and machine to provide the interaction. These sensors as shown below can be very primitive or very sophisticated.

1) Pen-Based Interaction:

Pen based interaction is an appropriate input device for direct interaction. Using a stylus in blend with an electronic display almost looks like the mundane pen-and-paper situation. Drawings, writing and commands can be pen-produced directly on the display tablet. The intentions of the users do not need to be mediated by a command language nor by a sequence of actions for selecting icons, positions or keys. Pen-based interaction indeed offers advantages that other input devices cannot provide

2) Mouse & Keyboard:

The most common interaction equipments are mouse and keyboards. A **computer mouse** is an input device that is most often used with a personal **computer**. Moving a **mouse** along a flat surface can move the on-screen cursor to different items on the screen. Items can be moved or selected by pressing the **mouse** buttons (called clicking).

A light source, generally a light-emitting diode, and a light indicator, such as an image sensor, to detect movement relative to a surface is use by optical computer mouse. On the other hand a mechanical, traditional mouse uses moving parts to perform the same function.

A form of computer input device is a projection keyboard whereby the image of a virtual keyboard is visualized onto a surface: when a user touches key which is there on the surface covered by an image, the device records the corresponding keystroke.

3) Joysticks:

Joysticks are conventionally utilized to provide positioning information in a two dimensional system. For example, joysticks are generally used to position items on the screen of a video game or to manipulate a machining tool about a two dimensional work surface.

The hand controller contains of six linear force sensors connected between the base and the handle so as to support the handle and measure any force applied against it. The linear force sensors used in the hand controller are not typical and thus must be custom-manufactured. Also, The hand controller requires a relatively complex coordinate transformation mechanism to convert the sensor measurements to useable directional values.

4) Motion Tracking Sensors:

Motion capture is the process of registering the motion and the actions of items or people. Motion capture is being used highly to produce films which attempt to approximate the look of live-action cinema, with nearly photo realistic digital character models. The first movie made primarily with motion capture was Sinbad: Beyond the Veil of Mists. Another main application is video games; it uses motion capture to animate athletes, martial artists, and other in-game characters.

5) Haptic Sensors:

A tactile sensor that measure forces exerted by the user on the interface is called Haptic devices. This haptic technology has made it possible to examine how the human sense of touch works by allowing the formulation of cautiously controlled haptic virtual objects.

6) Pressure Sensors:

The degree of force of a sensor is generally of gases or liquids is called pressure sensor. Pressure means to register the amount of force required to stop a fluid from developing, and is generally stated in terms of force per unit area. A pressure sensor typically acts as a transducer; it produces a signal as a function of the pressure imposed.

7) Taste/Smell Sensors:

Recent research says that the ability to discriminate various tastes of beverages or foods has stimulated by the development of several kinds of sensing systems. Maximum of these sensors includes the idea of global sensing.

Smell is the other "chemical" sense. Unlike taste, there are many olfactory receptors; each has its particular molecular feature. These odor molecules have various features and, therefore, it excites specific receptors more or less strongly. This combination signals from other receptors makes up what we notice as the molecule's smell.

IV. LIMITATION OF UNIMODAL HCI SYSTEMS

- Unimodal Human-Computer Interaction is not a natural way of human interaction.
- It is generally designed for the average user.
- It has failed to cater the requirement of different categories of people.
- Unimodal Human-Computer Interaction is difficult to use by disable, illiterate and untrained people.
- It cannot provide universal interface.
- It is more error prone.

V. CONCLUSION

The broad area like Human-Computer Interaction cannot be concluded as the expansion in this field will never stop. But, the conclusion of this paper is research on Human-Computer Interaction will go ahead and ahead and probably in 2020, we can imagine some of the new tools like:

1. The Reactable: a multitouch interface for playing music. Person can at the same time interact with it by moving and rotating physical objects on its surface.
2. Animated Textiles developed by Studio subTela at the Hexagram Institute, Canada. These two jackets after synchronization when the wearers hold hands, and the message scrolls from the back of one person to the other.
3. The Rovio robotic webcam is wirelessly connected to the Internet. It wanders around the home providing an audio and video link to keep an eye on family or pets when you are out. Etc.

The bottom line is that computer technologies are not impartial; they are loaded down with human, cultural and social values. The more we rely upon technologies to carry out our everyday activities, the more we will need to trust them to do so.

REFERENCES

[1] Fakhreddine Karray, Milad Alemzadeh, Jamil Abou Saleh and Mo Nours Arab, "Human-Computer Interaction: Overview on State of the Art." INTERNATIONAL JOURNAL ON SMART SENSING AND INTELLIGENT SYSTEMS 1.March (2008): 137-159.
[2] "Allegory of the Five Senses". The Walters Art Museum.
[3] "Human-Computer Interaction " By Alan Dix, Janet Finlay, Gregory D. Abowd, Russell Beale.
[4] Art Graesser, "Unimodal and Multimodal Human Perceptionof Naturalistic Non-Basic Affective Statesduring Human-Computer Interactions", IEEE Transactions on Affective Computing, vol.4, no. 4, pp. 452-465, Oct.-Dec. 2013, doi:10.1109/T-AFFC.2013.19
[5] "Three Faces of Human–Computer Interaction" By Jonathan Grudin, Microsoft Research, IEEE Annals of the History of Computing, OCT-DEC, 2005. CA: University Science, 1989.

AUTHOR'S PROFILE

Mr. Ashish Palsingh has completed his M.Sc. in Computer Application & Information Technology from Hemchandracharya North Gujarat University, Patan with First-Distinction in 2010. His area of interests is Artificial Intelligence, Android development, Advance Java and C++.

Ms. Ekta Palsingh has completed her M.Sc. in Computer Application & Information Technology from Hemchandracharya North Gujarat University, Patan with First-Distinction in 2011. Her area of interest includes Artificial Intelligence, System, Analysis and Design, JAVA and Software Engineering.