

Human Computer Interaction: Smarter way of Communication

Sanya Harneja, Kunal Pahuja, Jatin Malhotra

Abstract— As the name suggests, Human Computer Interaction (HCI) basically refers to the interaction that take place between human beings and the computer system. It deals with the study, design, evaluation, implementation and use of interactive computer system by humans and its various phenomenon's'. It's not just limited to use of computers but is also concerned with the tasks performed jointly by humans and machines (computers), new interactive techniques that can be used to perform these tasks, structure of communication between human being and the machine and how capable are human beings to make use of these machines. This research paper gives an overview of all the methods, applications and various advances that has been made in this field.

Index Terms— Human Computer Interactions, programming interfaces, machines.

I. INTRODUCTION

Human Computer Interaction, sometimes also referred to as Computer Human Interaction (CHI), Man-machine interaction (MMI) or Human-machine interaction (HMI) is a multidisciplinary subject. It's an area of research and practice that became popular with the emergence of computer, more generally the machine itself. It also involves the use of various algorithms, tools that can be used to design different programming interfaces, process that are followed to implement these interfaces, presentation of information by computer system as requested by user and how well the user can control and monitor the computer. Various aspects such as design, science, engineering, human psychology etc are associated with it. Since HCI, studies both human and machine in conjunction, thus supporting knowledge is drawn from it on both machine and human side. If we talk of machine, different techniques in computer graphics, programming, operating system, and development environments are relevant. Whereas on human side, communication theory, graphics, industrial design disciplines, social sciences, linguistics, cognitive psychology, human performance, engineering and design methods are some of the relevant facts.

In past few decades, HCI has expanded rapidly and steadily because of increasing use of computers, thus attracting professionals from various domains and disciplines and incorporating diverse concepts and approaches. As we know that many of the sophisticated machines are useless if they are not used properly by men. If a human machine interface is

poorly designed it can lead to various unexpected problem. Thus in order to ensure proper functionality and usability of computer between the users, HCI was designed. Functionality of a system is defined by the set of services and actions that it provides to its users and this functionality only becomes visible when it is efficiently utilized by users. On the other hand, usability of a computer system or machine with certain sets of functionality is the range or degree up to which a system can be used efficiently used in order to accomplish certain goals for certain users. Thus if there is proper balance between the functionality and usability of a system then only the actual effectiveness of a system can be achieved. [1]

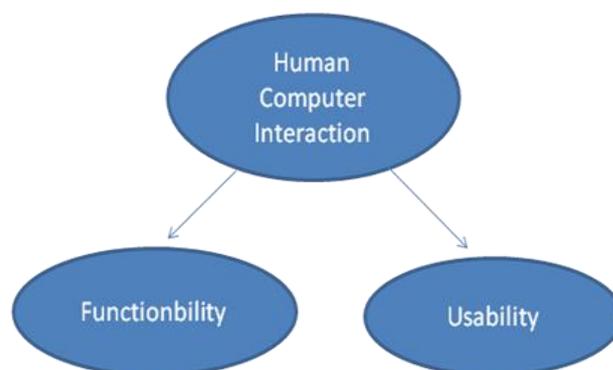


Fig 1: Main components of HCI Design

II. OVERVIEW OF HCI

In recent years, several advances have been made in the field of HCI. Out of these, some are fictional and are on the verge of development, while others are real. In the first part of this section, an overview of human computer interaction along with human characteristics has been explained. The second part of the section describes the various existing technologies and the direction to which the HCI research is heading.

A. Nature of human computer interaction

Human computer interaction is a discipline that is mainly concerned with type of interaction between human beings and computer and how effective computer systems are developed based on these interactions for the users. HCI basically act as mode of communication between user and the machine. It is kind of agent paradigm, tool paradigm, and work centered point of view. Its main objective is productivity and user empowerment. Science, engineering, journal, literature, design etc are various aspects that are associated with it [2].

Manuscript received April 18, 2014.

Sanya Harneja, Information Technology, Dronacharya College of Engineering, Gurgaon, Hr, India-123506

Kunal Pahuja, Electronics & Comm.Engineering, Dronacharya College of Engineering, Gurgaon, Hr, India-123506

Jatin Malhotra, , Electronics & Comm.Engineering, Dronacharya College of Engineering, Gurgaon, Hr, India-123506

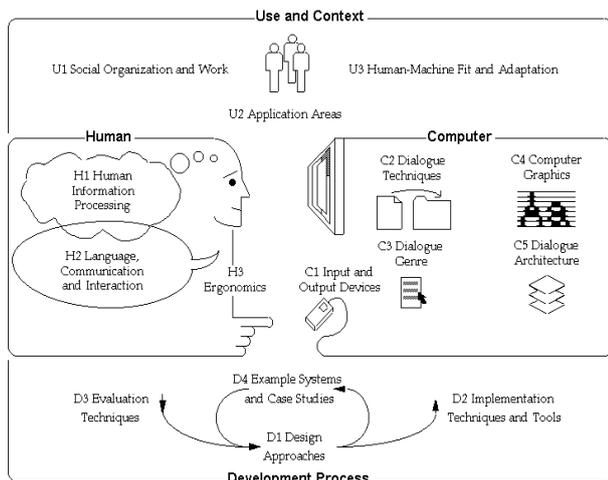


Fig 2: Human Computer Interaction [2]

B. Human Characteristics

The various characteristics of human as a processor of information are as follows:

- Actions of human beings are structured as models of cognitive architecture: connectionist models, symbol-system models, engineering models.
- It also consist of various phenomena and theories of perception, phenomena and theories of memory
- Various theories of motor skills
- Also the phenomena and theories related to problem solving
- Users' conceptual models
- As theories of attention and vigilance
- Phenomena and theories of motivation
- Phenomena and theories of learning and skill acquisition
- It describes various human diversities, including disabled populations.

C. Existing HCI Technology

Whatever HCI design we create, it must fulfill all the aspects of human behavior and should be a useful one. As compared to simplicity of interaction, the degree of complexity of human computer interaction is sometimes invisible. Some of the existing HCI interfaces are as follows:

- Command Line Interface (CLI)
- Menu Driven Interface
- Graphical User Interface(GUI)
- Natural Language Interface

All these existing interfaces have different degree of complexity based on their functionality as well as usability. Therefore, before designing a HCI, the degree of activity that involves both user and machine should be thoroughly studied [3]. The user activity is basically divided into three different levels: physical, cognitive and affective. The three most important human senses namely: vision, audio and touch are used to categorize the existing physical technologies for HCI.

Almost all the input devices rely on vision and are most commonly used. It's because of them we are able to communicate with the system. These devices are either switch based or pointing devices. They make use of switches and buttons like keyboard. Examples of pointing devices are mouse, joystick, trackballs, graphic tablets etc. Auditory devices are the more advance devices that need some kind of speech recognition technique. The main aim of these devices is to facilitate the much needed interaction and is therefore difficult to build. Beep, alarms, turn by navigation are some of the examples of these devices. Haptic devices are the most difficult and costly devices to build. Haptic devices generate sensation to skins and muscles through touch rigidity, weight. These kinds of devices are becoming popular for virtual reality or disability assistance applications [3].

D. Advances in the field of HCI

The field of human computer interaction has advanced at a fast pace and has become a well known area of interest. Since it deals with the design, evaluation, adoption, and use of information technology (IT), more and more people are indulging into this field. Intelligent and adaptive interfaces along with ubiquitous computing are the most recent advances in the field of HCI.

1) *Intelligent and Adaptive HCI*: Talking of today's scenario, majority of users in the world is making use of different devices in order to accomplish their task. However, the devices used by most of the people are either plain command/action setups, lacking much of the sophisticated and intelligent interfaces. Although technology is advancing rapidly at a very fast pace, there is a need for development of effective, efficient and natural interfaces that can provide support access to various kinds of information, applications and people. There is a need for design of some intelligent and adaptive interfaces that can provide a number of additional benefits to users. To accomplish this goal, interfaces are getting more natural to use every day. Traditional interfaces such as typewriters, keyboards, mouse etc have now been replaced by touch screen tablets, smart phones, PC's which more learnable, usable and transparent [4].

Intelligent HCI are human computer interfaces, whose basic aim is to improve the effectiveness, efficiency and naturalness of human machine interaction and make use of intelligence to design the interface. They incorporate at least some kind of intelligence for perceiving and making response back to user. For example, graphics, speech enabled interfaces that make use of natural language, gestures etc. On the other hand, Adaptive HCI design unlike Intelligent HCI, make use of interface in order to continuously interact with the user. A website that makes use of attractive GUI for selling its products can be good example of adaptive HCI. Such kind of adaption deals with both cognitive and affective levels of security. PDA or tablet PC is another example that makes use of both intelligent and adaptive interfaces.

2) *Ubiquitous Computing and Ambient Intelligence*: Ubiquitous computing, which is also known as ubicomp is an advance computing technique in which computing can be done everywhere and anywhere. Ubiquitous computing, in contrast to desktop computing can be done using any device, in any location and in any format. Sometimes it is also known

as pervasive computing or ambient intelligence. The human computer interaction exist in different forms including laptops, computers etc. The underlying technology that support ubiquitous computing internet, microprocessors, sensors, new i/o devices, operating systems and many more. The idea of ubiquitous computing was first introduced by Mark Weiser in 1998, during his tenure as chief technologist at Computer Science Lab in Xerox PARC.

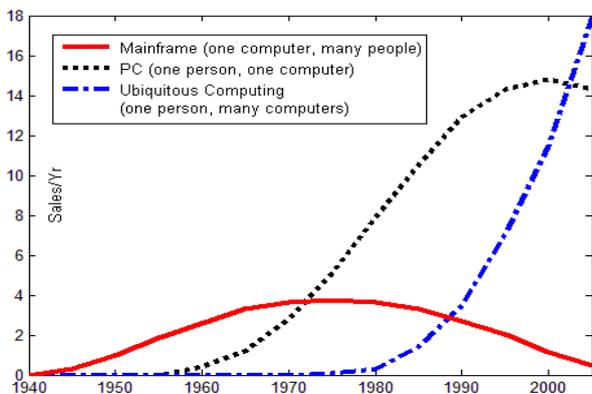


Fig 3: Major Trends in Computing [13]

Ubiquitous computing is sometimes also described as pervasive computing or ambient intelligence. Ambient Intelligence is a new paradigm of information technology that is sensitive and responsive to the presence of people and is also adaptive to their needs, habits, gestures, and various emotions shown by them. Ubiquitous computing has been regarded as third wave of computing that allows one person to make use of more than one computer. They key technology includes miniature hardware, seamless communication and dynamic device network respectively [5].

III. HCI ARCHITECTURE

Configuration in HCI plays an important role, where the number and diversity of output and input generally defines the interface. Architecture of human computer interface shows the basic working of the system on the basis of inputs and outputs obtained and the way in which they interact with the system. The most important parts of the architecture are Unimodal HCI System and Multimodal HCI System. With the help of various configurations and designs the interface is hence explained [6].

A. Unimodal HCI System

As the number and diversities of the inputs and outputs which are mainly dependent on interface, basically are communication channels that enable the users to interact or provide real time processing with the help of the interface. A modality is defined as the condition in which each of the different independent single channels is present. Thus a system consisting of only one modal is defined as unimodal system. Unimodal system is divided into three categories –

- Visual based
- Audio based
- Sensor based

1) *Visual based HCI*: It is the most widespread area in HCI field. Taking in mind the extent of application and variety of open problems as well as approaches, researchers tried to tackle different aspect of human responses which can be recognized as visual signals, for which some of the main research areas in the section are as following-

- Facial expression analysis
- Body movement tracking(large scale)
- Gesture recognition
- Gaze detection(Eyes movement tracking)

2) *Audio based HCI*: This area deals with the interaction between a computer and a human which is in audio form. It deals with the information required by different audio signals. It is termed as a unique provider of information as the information gathered by audio signals can be more truthful, as the nature of audio signal may not be as variable as video signal. Research area in this section can be divided into following parts:

- Speech recognition
- Speaker recognition
- Musical interaction
- Auditory emotion analysis
- Human made noise/sin detection

3) *Sensor based HCI*: This area is a combination of variety of areas with a wide area of applications. At least one physical sensor is used between user and provider which provide the interaction. These are some of the sensors which can be primitive or very sophisticated:

- Mouse and key
- Pen based interaction
- Joy sticks
- Haptic sensors
- Pressures sensors
- Motion tracking centers and Digitizers
- Taste/ smell sensors

B. Multimodal HCI System

Multimodal is defined as the combination of different modalities. In MMHCI systems, modalities focus on the way that the system responses on the input, i.e. communication channel. Via two or more nodes of input that goes beyond the traditional keyboard and mouse, the multimodal acts as facilitator of human –computer interaction. Input modes, their types and the manner in which they work vary from one multimodal system to another. It incorporate different combination of speech, gesture, gaze facial expressions and other unconventional modes of input, in which gesture and gaze are the most commonly supported combination of input methods [3].

Correlatively intractable signal modalities should be present in an ideal HCI system. The fusion of different modalities (open problems and practical boundaries) leads to different limitations. In most of the existing multimodal, the modalities are still treated separately and at the end, the results of different modalities are combined together. Important aspect of multimodality is the collaboration of different modalities to assist the functioning of recognition. For example, lip movement tracking can help speech recognition methods and speech recognition can assist command acquisition.

IV. APPLICATIONS

With the development of new multi-sensory user interfaces such as speech, sound, haptics etc and metaphors like gestures, avatar in augmented or virtual reality world, shared cognitive spaces, the field of HCI has undergone a tremendous change. Increasing use of technology has now made this interaction even simpler. Nowadays large interactive displays, smart devices and embedded systems have become more and more pervasive. Earlier the interaction was just limited to traditional keyboard and mouse. But nowadays the multimodal interfaces have been developed that are offering a huge amount of advantages over these traditional interfaces. A multimodal interface provides the facility for human computer interaction, by making use of two or more modes input. However the exact number of input modes, their types and their working may vary widely from one multimodal system to another, depending upon its design and implementation. These interfaces incorporate different combinations of speech, gaze, gesture and facial expressions and other non-conventional modes of input. Another striking feature of multimodal systems is that they can accommodate different people and different circumstances much easily [1]. One of the most classic examples of multimodal system is the “Put That There” demonstration system. Basically this system allows the user to move an object to a new location on map of screen by just saying “put that there” while it points to the object itself and then points to the desired location [1]. Some other examples of applications of multimodal systems are listed below:

- Intelligent Games
- Smart Video Conferencing
- E-Commerce Intelligent Homes/Offices
- Helping People with Disabilities
- Driver Monitoring

In the following sections, some of important applications of multimodal systems have been presented with greater details.

A. Gesture Recognition

Gesture Recognition basically involves interfacing with the computer system, using different gestures of human body such as fingers, hands, arms, and head in three dimensions through use of camera or via device with embedded sensors. In gesture recognition technology, a camera reads various movement of human body and in turn transfers that data to computer that uses the gestures as input to various control devices or applications. At present the primary application of gestural interface is in gaming and home entertainment market. It is also being used to help the physically impaired people to interact with computers such as interpreting the sign language. Data visualization and analytics and interaction with Large Group Displays (LGD's) are some of its other applications

B. Multimodal Systems for Disabled people

One of the best applications of multimodal system is to help and assist physically disabled people to interact with machines, especially the computers. Normally such people require a complete different kind of interface to interact than ordinary people. The disabled people can interact with machines either by using their voice or head movements. The two main modalities used are speech and head movement. Both these modalities work continuously. The position of the head indicates the coordinates of cursor in current time moment on the screen. On the other hand, speech provides the needed information that is required to perform the action with an object selected by the cursor. As far as synchronization is concerned, it is performed by calculating the position of cursor at the beginning of speech detection [3].

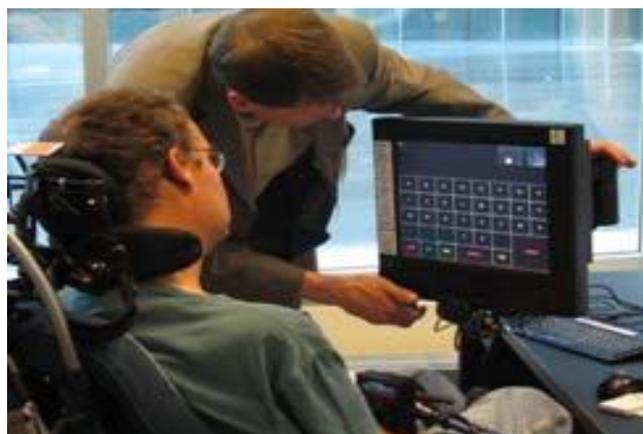


Fig 4: Gaze detection pointing for disabled people [1]

C. Speech Recognition and Translation

Speech recognition is the translation of spoken words into machine readable inputs such as text. Sometimes it is also known as automatic speech recognition or computer speech recognition. Voice recognition on the other hand, is a system that is trained for the particular user. It is a technology that converts the spoken words, phrases, sounds produced by human beings into electrical signals and then these signals are converted into meaningful patterns. It simply recognizes the speech of user unique vocal sound. Speech recognition is used in car systems, healthcare, military, especially in high performance fighter aircrafts, helicopters, training air traffic controllers, telephony and various other domains. Further applications include aerospace, automatic translation, video games, robotics etc. Translation involves communication of the meaning of source language text by means of a translator. We translate the text into format that is easily understood by the user. Whatever we give as input to the system is converted into a format that can be easily understood by it and vice versa [5].

D. Multitouch

Touch sensing is common for single point of contact. Multi touch on the other hand, enables the user to interact with the system by using more than one finger at a time. It is a touch

screen interaction technique in which we can simultaneously touch multiple points and can even control the movement of objects in user interface or application. The surface is able to recognize the presence of more than one point of contact made by user. Multi-touch user interface has become an important feature of smart phones, tablets, laptops, pad's and many more electronic devices where multi-touch gestures are used to interact with the devices. For example, we can zoom in or zoom out a picture or web page with the help of our thumb and index finger [7].

E. Emotion Recognition Multimodal Systems

In case of emotion recognition multi-modal systems, people are able to perceive one's emotional state based on their observations about one's face, body, voice etc. out of all these modalities face modality produces the best prediction [8].

Since we are moving towards a world in which computers are more and more ubiquitous, it is essential that machines perceive and interpret all the clues that are provided to them by user both implicitly and explicitly. A human computer interaction cannot be based solely on commands explicitly delivered by user. Computers will have to find out an alternate way to detect various behavior signals on the basis of which they can infer one's emotional state. Various researches in multimodal system have been conducted so far in order to infer one's emotional state. On the other hand, a bimodal system that is based on complete fusion of facial recognition and acoustic information provided an accurate classifier of 89.1 percent in terms of recognizing different kinds of emotions such as 'sadness, anger, happiness, and neutral state'. It's also been seen that emotion recognition system that was based on acoustic information only gave an overall performance of 70.9 percent, as compared to facial recognition system, that gave overall performance of around 85 percent.

F. Map-Based Multimodal Application

For expressing different messages, different modalities such as speech, gestures, head movements etc. can be used. Map based multimodal systems greatly improves the user experience as they support multi modes of input. One of the oldest and widely known map based application that make use speech and pen gesture is Quickset [9]. It's a military- training application that allows the user to express command by using either one of the two modalities or both simultaneously. For example, users can draw a predefined symbol for platoons with a help of a pen at a given location on the map, thus creating a new platoon at that location. Alternatively, user can also make use of speech for specifying their intent for creating a new platoon and can also make use of their vocals to specify the coordinates where the platoon has to be placed. A recently developed map-based application is Real Hunter. Real Hunter is a real estate interface that allows the user to select an object or region with touch input while making queries using speech [10]. Similar to Quickset, MATCH-Kiosk is another type of map based application. It's an interactive city guide or precisely a tour guide that have great potential to provide benefit to multimodal interface.

V. CONCLUSIONS

Thus from the above study it is clear that human computer interaction has become an integral part of system design. Since the rise of this field in 1980's, a number of diverse methodologies and ample amount of techniques have evolved which has made this interaction even more simpler. New trends in ubiquitous communications and computing will help people to interact with the technology that surrounds them in an intuitive and less restrictive ways. Ambient Intelligence on the other hand is also trying to embed new technology into the environment, thus making it more natural and invisible at the same time. Also the dramatic changeover from traditional keyboards and mouse interface to touch screen devices like smart phones, tablets, PC's etc has given a new face to this to this interaction. In this paper we also came across the various applications of multimodal systems that provided the user with multi modes of interface. This research paper gave an overview of all the existing HCI technologies and the various advances that have been made in this field so far.

REFERENCES

- [1] Fakhreddine Karray, Milad Alemzadeh, Jamil Abou Saleh and Mo Nours Arab, "Human-Computer Interaction: Overview on State of the Art" VOL. 1, NO. 1, (MARCH 2008)
- [2] Nature of Human Computer Interaction, Retrieved from <http://old.sigchi.org/cdg/cdg2.html>
- [3] Kuntal Saroha1, Sheela Sharma, Gurpreet Bhatia, "Human Computer Interaction: An intellectual approach" Vol. 11, Issue 02, (Aug 2011).
- [4] M.T. Maybury and W. Wahlster, Readings in *Intelligent User Interfaces*, Morgan Kaufmann Press, San Francisco (1998).
- [5] José Bravo, Xavier Aleman, Teresa Riesgo, "Computing and Ambient Intelligence: New Challenges for Computing" Volume 12, Issue 3
- [6] *HCI Architecture*, Retrieved from <http://www.slideshare.net/gasantosh031/1204-hci-new>
- [7] Denis Gouin, Valérie Lavigne "Trends in Human-Computer Interaction to Support Future Intelligence Analysis Capabilities" Paper 130, Track 5, (June 21-23, 2010)
- [8] M. Pantic, A. Pentland, A. Nijholt and T. Huang, "Human computing and machine understanding of human behavior: a survey" Proceedings of the 8th International Conference on Multimodal Interfaces, Banff, Alberta, Canada, pp 239-248 (2006).
- [9] M. Johnston, P.R. Cohen, D. McGee, S.L. Oviatt, J.A. Pittman and I. Smith, "Unification-based multimodal integration", Proceedings of the Eighth Conference on European Chapter of the Association for Computational Linguistics, pp 281-288 (1997).
- [10] J. Y. Chai, P. Hong and M. X. Zhou, "A probabilistic approach to reference resolution in multimodal user interfaces", Proceedings of the 9th International Conference on Intelligent User Interfaces, Funchal, Madeira Portugal, pp 70-77 (2004).
- [11] *HUMAN COMPUTER INTERACTION* (Internet Sources: Wikipedia; Network world; Webopedia).
- [12] http://www.interaction-design.org/encyclopedia/human_computer_interaction_hci.html
- [13] G.Riva, F.Vatalaro, F.Davide and M.Alaniz, *Ambient Intelligence: The Evolution of Technology, Communication and Cognition towards the Future of HCI*, IOS Press, Fairfax (2005).