Chapter 1

Introduction

1.1 Introduction

Research in Human-Computer Interface (HCI) deals with methods to design efficient and friendly ways of interacting between computers and humans [1] [2]. Research in this area is voluminous. A complete branch of computer science is devoted to this topic. Scientists and Researchers are trying to develop machines like human. Efforts are to develop autonomous machines which will be trained itself like human and take decision. HCI is the study how people interact with computers and so what extent computers are developed for successful interaction with human beings. As the human beings are biological systems (analog) and computers are digital systems, these two different types of technologies need to be bridged. There is a need to make advances in understanding biological signals to train computers for interfacing [3].

1.2 Human Computer Interfacing

1.2.1 Components of Human Computer Interfacing (HCI)

As mentioned above, HCI is a multidisciplinary field which is concerned with the application of Computer Science, Cognitive Psychology, Neuroscience Ergonomics (Human Factors), Engineering, Design, Anthropology, Sociology, Philosophy, and Linguistics Artificial Intelligence. This is illustrated figure 1.1. Out of these, we have only considered Speech (Linguistics), and Electroencephalographic Signal (Neuroscience and Cognitive Psychology) for our study. The background in HCI is concerned about numbers of ways user interact with the computers. The users interact with the computer using traditional methods of a key board and mouse as main input devices and monitor as the main output device. Due to one reason or another some users cannot be able to interact with machines using a mouse and keyboard device. The natural way to interact with computers is the way in which human beings interact with each other i.e. through biological signals. For this purpose, speech, brain and
others biological signals have been used for interfacing with computer. Appropriate technologies need to be developed [4].

![Multidisciplinary Research in Human Computer Interaction (HCI)](image_url)

**Figure 1.1: Multidisciplinary Research in Human Computer Interaction (HCI)**

### 1.2.2 Significance of HCI

**Easy Interfacing:** The people, basically unable to use mouse and keyboard (Physically Handicapped), can use the HCI technology to interact with computer. As HCI involves human ways of communication, it can have applications in medicines, education etc.

**National Security:** HCI–based command, control, communications and intelligence system are at the heart of our military infrastructure, as it covers security features very effectively. Interfaces between criminals and investigators have also been useful to find truth related to crime. Also in cockpits, on the bridge, and in the field to be effective systems must have high-quality human computer interfaces [5].

**Growth of HCI industries:** Powerful, interesting, and usable applications are the fuel for continuous growth of these industries. The resulting market supports commodity pricing for both hardware and software. Future growth cycles will in part be driven by current HCI research.

**Speech and Brain Signals:** The human speech is inherently a multi modal process that involves the analysis of the uttered acoustic signal and includes higher level
knowledge sources such as grammar semantics and pragmatics [6]. Brain signals contain information through multiple ways. This research intends to focus on the speech and brain signals.

1.3 Human – Computer Intelligent Interaction (HCII)

The intelligent human–computer interaction (HCI) technologies play important roles in the development of advanced communication. In contrast to the conventional mechanisms of passive manipulation, intelligent HCI integrates versatile tools such as perceptual recognition, machine learning, affective computing, and emotion recognition to enhance the ways humans interact with computers. Nonverbal information such as facial expression, posture, gesture, brain signals, and eye gaze is suitable for behavior interpretation. In recent years there has been a growing interest in improving all aspects of the interaction between humans and computers. It is argued that to truly achieve effective human computer intelligent interaction [7], there is a need for the computer to be able to interact naturally with the user, similar to the human-human interaction. Humans interact with each other mainly through speech, but also through body gestures, to emphasize certain part of the speech and display of emotions. Emotions are displayed by visual, vocal, and other physiological means [8].

1.4 Techniques in HCI

1.4.1 Non – Bio Signal Approach

Several devices have been developed beside the use of biomedical signals to implement a convenient solution of HCI for the disabled persons. For example, the device ‘head mouse’ is pointing device that transforms head movement into cursor movement on the screen. This uses infrared distance measurement to monitor the head motion. The mouse pointer movement on the screen is then proportional to the user’s head movement, which are used to trigger a switch through which the user can control various system functions. A specific problem with head mouse systems is the required motor skills. The approaches have potential disadvantages for some categories of
INTRODUCTION

users. A user with cerebral palsy may not have the fine motor abilities in the tongue to operate the Tonguepoint device. Similarly, a user with spinal vertebrae fusion may not be able to turn his or her head, so the Headmouse would be of little benefit [9]. In recent years, vision-based gesture recognition is achieved by using video cameras, image processing and visual tracking algorithms.

1.4.2 Bio Signal Based Approaches

1.4.2.1 Electrooculography (EOG) Signals

Some bio-signals have been developed for creation of a new communication interface between humans and computer. The position of the eye directly relates with the visual information of interest. It is possible to provide vary intuitive device by using the position of eyes. The position of eyes can be measured optically, mechanically, and electrically. The electrical method of measurement, the EOG, is least invasive method of determining the eye position [10]. Eye movement research is of great interest in the study of neuroscience and psychiatry, as well as ergonomics, advertising and design. Since eye movements can be controlled volitionally, to some degree, and tracked by modern technology with great speed and precision, they can be used as a powerful input device.

1.4.2.2 Electromyography (EMG) Signals

EMGs are considered to be the source of a new means of HCI, i.e. an alternative input mechanism. The electrical activity induced by the human’s arm muscles movements can be interpreted and transformed into computer’s control commands. Electromyography (EMG) measures electrical currents that are generated in muscles during its contraction and represent neuromuscular activities. EMG signals can be used for a variety of applications including clinical applications. Multi-channel EMG sensors have been used at the same time to detect relevant EMG patterns by a combined signal analysis [11]. The EMG signals have different signatures depending on age, muscle development, motor unit paths, skin-fat layer, and gesture style. The
external appearances of two peoples’ gestures might look identical, but the characteristic EMG signals are different.

1.4.2.3 Electroencephalographic (EEG) Signals

The BCI is an emergent multidisciplinary technology that allows a brain to control a computer directly, without relying on normal neuromuscular pathways [12]. The most important applications of the technology for the paralyzed people who are suffering from severe neuromuscular disorders, as BCI potentially provides them with communication, control, or rehabilitation tools to help compensate for or restore their lost abilities. Among various brain signal acquisition methods, the EEG is of particular interest to the BCI community [13] [14] [15] [16]. The EEG records the electrical brain signal from the scalp, where the signal originates from postsynaptic potentials, aggregates at the cortex, and transfers through the skull to the scalp [17]. EEG based device that requires extracting raw EEG data from the brain and converting it to device control commands through suitable signal processing techniques. The cerebral electrical activities of the brain are recorded via the EEG using electrodes that are attached to the surface of the skull. These signals measured by the electrodes are amplified, filtered and digitized for processing in a computer where feature extraction is performed, classification is done and a suitable control command is generated [18].

EEG based BCI technology as shown in figure 1.2, has seen much development in recent years. Specifically, EEG based BCI technologies that do not depend on peripheral nerves and muscles have received much attention as possible modes of communication for the disabled [19]. Various EEG phenomena, such as slow cortical potentials, P300 potentials, and mu and beta rhythm control can provide opportunities for severely disabled individuals to further interact with their environment. One of the popular phenomena utilized for BCI control is the modulation of mu (8-12 Hz) and beta (18-25 Hz) rhythms via motor imagery. Actual or imagined motor movements result in a de-synchronization (decrease in amplitude) of these rhythms over the sensorimotor cortex. Users are thus directly able to control a BCI by
modulating the magnitude of these rhythms by switching between motor imagery tasks [20].

The EEG bears merits as it is noninvasive, technically less demanding, and widely available at relatively low cost. On the other hand, it also brings great challenges to signal processing and pattern recognition, since it has relatively poor signal-to-noise ratio and limited topographical resolution and frequency range [21]. However non-invasive data acquisition makes automated feature extraction challenging. It is because the signals of interest are 'hidden' in a highly noisy environment. It was demonstrated that the spatial filtering operations improve the signal-to-noise ratio [22]. Unfortunately, the intensive training time (several months) involved for a user to gain a high degree of control (>80% accuracy) may be a deterrent for practical
applications of BCIs such as prosthetic control and daily computer use for disabled individuals [23].

1.4.3 Speech Based Approaches

Speech technology is the technology with a developing number of methods and tools for better implementation. Speech recognition has a number of practical implementations for both fun and serious works. Automatic speech recognition has an interesting and useful implementation in expert systems, a technology whereby computers can act as a substitute for a human expert. An intelligent computer that acts, responds or thinks like a human being can be equipped with an automatic speech recognition module that enables it to process spoken information. Medical diagnostic systems, for example, can diagnose a patient by asking him a set of questions, the patient responding with answers, and the system responds with what might be a possible disease.

1.4.3.1 Template Based Approach

Template based approaches matching [24], Unknown speech is compared against a set of pre-recorded words (templates) in order to find the best match. This has the advantage of using perfectly accurate word models. But it also has the disadvantage that pre-recorded templates are fixed, so variations in speech can only be modeled by using many templates per word, which eventually becomes impractical. Dynamic time warping is such a typical approach [25]. In this approach, the templates usually consists of representative sequences of features vectors for corresponding words. The basic idea here is to align the utterance to each of the template words and then select the word or word sequence that contains the best. For each utterance, the distance between the template and the observed feature vectors are computed using some distance measure and these local distances are accumulated along each possible alignment path. The lowest scoring path then identifies the optimal alignment for a word and the word template obtaining the lowest overall score depicts the recognized word or sequence of words.
1.4.3.2 Statistical Based Approach

In Statistical Based Approach variations in speech are modeled statistically, using automatic, statistical learning procedure, typically the Hidden Markov Models, or HMM. The approach represents the current state of the art. The main disadvantage of statistical models is that they must take priori modeling assumptions which are liable to be inaccurate, handicapping the system performance. In recent years, a new approach to the challenging problem of conversational speech recognition has emerged, holding a promise to overcome some fundamental limitations of the conventional Hidden Markov Model (HMM) approach [26,27]. This new approach is a radical departure from the current HMM-based statistical modeling approaches. Rather than using a large number of unstructured Gaussian mixture components to account for the tremendous variation in the observable acoustic data of highly co-articulated spontaneous speech, the new speech model that [28] have developed provides a rich structure for the partially observed (hidden) dynamics in the domain of vocal-tract resonances.

1.4.3.3 Dynamic Time Warping

Dynamic time warping (DTW) is being widely used in various pattern recognition and time series data mining application. It is an algorithm that calculates an optimal warping path between two time series and the distance between them. In a system of speech recognition containing words, the recognition requires the comparison between the entry signal of the word and the various words of the dictionary. The problem can be solved efficiently by a dynamic comparison algorithm whose goal is to put in optimal correspondence the temporal scales of the two words. An algorithm of this type is Dynamic Time Warping [29].
1.5 Objective of the Present Research

The main thrust of the present work was to design techniques to interface computer with speech and EEG signals, both biological signals. The speech signals were in Marathi which is a local regional language. The objective was to see feasibility of interfacing techniques which may be used to communicate with computer with spoken words or through EEG signals induced by a specific mental task.

1.6 Organization of the Thesis

The work in this thesis is described in six chapters. The present chapter is introductory. Chapter 2 gives advances in current research in speech and brain signal processing. The other five chapters are as follows: Chapter 3 devotes to design, development of various speech databases. The results related to feature extraction and recognition system are also described and discussed. The chapter 4 describes the experimental details regarding acquisitions of EEG signals under the different controlled conditions of mental tasks. The signals were analyzed and classified by statistical methods. The chapter five deals with experimental results related to song induced mental conditions. The results are also analyzed by different statistical methods. The chapter six has provided conclusion of the work along with future scope.
Reference:


