Introduction to EEG- and Speech-Based Emotion Recognition

nka A. Abhang, Bharti W. Gawali, and Suresh C. Mehrotra

INTRODUCTION TO EEG- AND SPEECH-BASED EMOTION RECOGNITION

PRIYANKA A. ABHANG
BHARTI W. GAWALI
SURESH C. MEHROTRA
Department of Computer Science and Information Technology,
Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Academic Press is an imprint of Elsevier

Academic Press is an imprint of Elsevier
125 London Wall, London EC2Y 5AS, UK
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, USA
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

ISBN: 978-0-12-804490-2

For information on all Academic Press publications
visit our website at https://www.elsevier.com/

Publisher: Mara Conner
Acquisition Editor: Natalie Farra
Editorial Project Manager: Kristi Anderson
Production Project Manager: Edward Taylor
Designer: Mark Rogers

Typeset by TNQ Books and Journals
www.tnq.co.in

Preface

The human brain is the most complex organ. It controls intelligence, senses, body movements, and behavior. The brain is an association of many parts which work in coordination with each other but are functionally unique and special. It is the place where emotions are generated and controlled. Emotions play an important role, as they unite us irrespective of our country, caste or creed. Humans use different ways — speech, face, gesture, and text — to understand and recognize emotions.

This book provides a far-reaching examination of research with different modalities of emotion recognition. The discussion of key topics starts with the physiology of the brain, followed by brain rhythms, speech parameters, acquisition and analysis tools, features of the modalities, and a case study. It concludes with applications for brain—computer interfacing.

A special feature of this book lies in its inclusion of two very important modalities of emotion: electroencephalography (EEG) signals/images and speech signals, used together. It also presents features to be considered when researchers experiment with recognition with different modalities.

It is hoped that this introduction to EEG- and speech-based emotion recognition will be of great interest to young researchers and postgraduate students. It provides basic knowledge regarding the domain, and focuses on various computational techniques to be implemented. It proposes many research problems to young researchers to prepare them for their own research. Emotion recognition has been proven to be a significant research area which will be important in designing and developing human—computer interface systems. The material provided in this book is referenced for further reading, and is presented in eight chapters.

Chapter 1, "Introduction to Emotion, Electroencephalography, and Speech Processing" outlines the fundamental aspects of the book's theme. It introduces the basic aspects of the physiology of the brain, EEG, the human auditory system, and speech emotion recognition.

Chapter 2, "Technological Basics of EEG Recording and Operation of Apparatus" presents technological basics of EEG and speech. It gives details of EEG and speech acquisition tools, and describes the Acquire and Analysis software available with the equipment.

Chapter 3, "Technical Aspects of Brain Rhythms and Speech Parameters" introduces brain frequencies along with the speech features considered during analysis. It also explains the preprocessing, feature extraction, and classification techniques which can be implemented in emotion recognition research.

The data acquired through EEG and computerized speech laboratories is in the frequency and time domains. Chapter 4, "Time and Frequency Analysis" illuminates mathematical techniques related to time and frequency analysis transformation. It also provides examples related to the transformation described in the chapter.

Chapter 5, "Emotion Recognition" sheds light on modalities researched in emotion recognition systems. Modalities like face, gesture, speech, text, and brain imaging are individually explained, along with their features.

To recognize emotion properly, much information is needed, along with the tone of speech, how the face looks and how gestures are used add to the recognition of emotion. This information raises the concept of multimodality. Chapter 6, "Multimodal Emotion Recognition" explains the concept of multimodal emotion recognition, and different models and theories of emotion. It also reviews the earlier efforts in multimodal emotion recognition systems, and provides information about online databases available for multimodal emotions. The chapter concludes by discussing challenges for these systems.

Chapter 7, "Proposed EEG/Speech-Based Emotion Recognition System: A Case Study" outlines a case study of an EEG/speech-based emotion recognition experiment with volunteers from the Department of Computer Science and IT, Dr. Babasaheb Ambedkar Marathwada University. The experimental analysis for happy and sad emotional states is discussed. EEG brain images and speech signals are explored in the chapter. Features considered for the analysis and its numeric data are also described. The correlation between the modalities is justified with linear discriminate analysis classification.

The ultimate goal of emotion recognition is to utilize it for the human (brain) and computer interfacing. Chapter 8, "Brain–Computer Interface Systems and Their Applications" discusses the types and applications of brain–computer interfacing, and also highlights the challenges.

We hope that this book will not only be a useful reference source of information for the scientific community, but will also be of help to the general community interested in the subject.

Priyanka A. Abhang
Bharti W. Gawali
Suresh C. Mehrotra

Acknowledgments

We would like to thank all the people who were directly and indirectly involved in supporting us in the accomplishment of this book, and to express our sincere gratitude to all of them.

Firstly, we would like to acknowledge the System Communication and Machine Learning Research Laboratory for all their fundamental support.

A colossal thanks to our youthful and energetic research colleagues Ganesh Manza, Ganesh Janvale, Rakesh Deore, Santosh Gaikwad, Shashibala Rao, Kavita Waghmare, Sangramsing Kayte, Reena Chaudhary, and Arshiya Khan, who have always been full of life and willing to help in the preparation of technical and nontechnical preliminary materials related to the book.

We would like to thank all the authors of the references cited in the book.

We would like to express our thanks to the management of Elsevier, our publishers, for their guidance in various administrative matters related to the book.

Lastly, we would like to thank all our family members for their support throughout the writing of the book.

Priyanka A. Abhang
Bharti W. Gawali
Suresh C. Mehrotra

Contents

7. Proposed EEG/Speech-Based Emotion Recognition System: A Case Study

8. Brain—Computer Interface Systems and Their Applications

CHAPTER

1

Introduction to Emotion, Electroencephalography, and Speech Processing

OUTLINE

1.1 INTRODUCTION

In a very real sense, we have two minds, one that thinks and one that feels. ***Daniel Goleman, Emotional Intelligence.***

Emotions are intrinsically connected to the way that people interact with each other. Emotion constitutes a major influence for determining human behaviors. A human being can read the emotional state of another human, and behave in the best way to improve their

Introduction to EEG- and Speech-Based Emotion Recognition
http://dx.doi.org/10.1016/B978-0-12-804490-2.00001-4

communication at that moment. This is because emotions can be recognized through words, voice intonation, facial expression, and body language.

Emotions have been studied in scientific disciplines such as physiology, psychology, speech science, neuroscience, psychiatry, communication, and so on. As a result, distinctive perspectives on the concept of emotions have emerged, appropriate to the complexity and variety of emotions. It is important to see these different perspectives as complementary to each other.

In humans, emotions fundamentally involve physiological arousal, expressive behaviors, and conscious experience. Thus in that respect there are three views of emotions.

- Psychological (what one is thinking).
- Physiological (what one's body is doing).
- Expressive (how one reacts) in nature.

It is thought that emotions are predictable, and are settled in the different regions in the brain depending on what emotion is invoked. An emotional reaction can be separated into three major categories: behavioral, automatic, and hormonal.[1]

In psychology, expression of emotion is viewed as a reaction to stimuli that involves characteristic physiological changes. According to psychology, an emotion is seen as a disturbance in the homeostatic baseline. Based on these changes, the properties of emotions can be represented as a three-dimensional construct. Essential dimensions of emotional states are measured by the features of activation.

- Arousal: measured as an intensity.
- Affect: valence of pressure, measured as positive or negative feeling after emotion perception.
- Power (control): measured as dominance or submissiveness in emotion expression.

Thus the psychology of emotions can be viewed as a complex experience of consciousness (psychology), bodily sensation (physiology), and behavior (active speech). Orientational dimensions of emotional states are captured by the features of activation, affection, and power.

The emotions generally represent a synthesis of subjective experience, expressive behavior, and neurochemical activity. In the physiology of emotion production mechanisms, it has been found that the nervous system is stimulated by the expression of high-arousal emotions such as anger, happiness, and fear. This phenomenon causes an increased heart rate, higher blood pressure, changes in respiratory pattern, greater subglottal air pressure in the lungs, and dryness in the mouth.[2]

Emotions are generally expressed in positive and negative ways. The positive emotions such as happiness, excitement, joy, etc. are pleasant and are seen as constructive in an individual, whereas negative emotions such as sadness, anger, fear, etc. are considered unpleasant and may be considered to be destructive for an individual.

According to Robert Plutchik,[3] any emotion is based upon one of six primary emotions. Happiness, sadness, anger, disgust, fear, and surprise are considered as the main or basic emotions by most researchers and are known as archetypal emotions.[4]

1. **Happiness** is the emotion that expresses various degrees of positive feelings, ranging from satisfaction to extreme joy.

2. **Sadness** is the emotion that expresses a state of loss or difficulty. Sadness causes individuals to be slow at processing information.
3. **Anger** is the emotion that expresses dislike or opposition toward a person or thing that is causing aversion. Anger is sometimes displayed through sudden and overt aggressive acts.
4. **Disgust** is the emotion that expresses a reaction to things that are considered dirty, revolting, contagious, contaminated, or inedible. Disgust is associated with a distinct facial expression and a drop in heart rate.
5. **Fear** is the emotional reaction to an actual and specific source of danger. Fear is often confused with anxiety, which is an emotion that is often exaggerated and experienced even when the source of danger is not present or tangible.
6. **Surprise** is the emotion that arises when an individual comes across an unanticipated situation. A surprise emotion can be a positive, neutral, or negative experience. A human being can understand the emotional state of another human being and behave in the best manner to improve the communication in a certain situation. This is because emotions can be recognized through various modalities such as words, voice intonation, facial expression, body language and by brain signals.[5]

1.2 BRAIN PHYSIOLOGY

The brain is the central controlling organ of the human being. Various scientific studies have proved that some regions of the brain are involved in thinking of emotions, responding to extreme emotional stimuli, and viewing emotional situations. Nearly all vital activities necessary for survival, as well as all emotions, originate inside the brain. The brain also receives and interprets a multitude of signals sent to it by other parts of the body and the environment.

1.2.1 Major Brain Areas

The brain is composed of a number of different regions, each with specialized functions. Fig. 1.1 shows the major parts of the brain.

The central core of the brain is made up of the brain stem and midbrain. The cerebral cortex is a covering layer for this central core. The central nucleus is moderately elementary and older, and its activity is mainly unconscious. In contrast, the cerebral cortex is extremely developed and capable of deliberation and functions, while the older parts of the brain remain relatively stable.[6]

The Brain Stem

The stem is in between the spinal cord and the rest of the brain. It is made up of the medulla that controls breathing, heart rate, and digestion, and the cerebellum which coordinates sensory input with muscle movement. The functions of the brain stem govern respiration, blood pressure, and some reflexes. The brain stem is further distributed into several distinct sections: the midbrain, pons, and medulla oblongata.

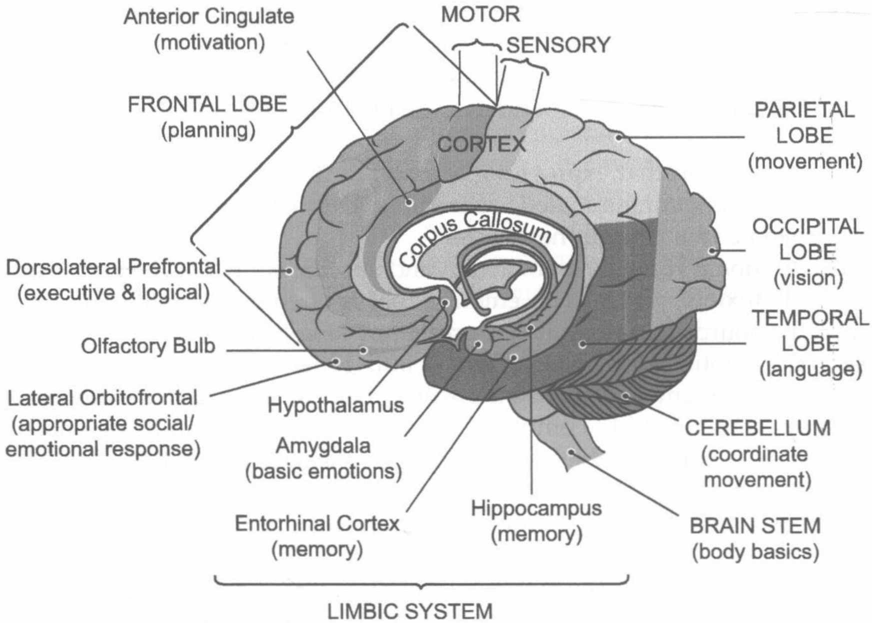

FIGURE 1.1 Major parts of the brain.

The Midbrain

The midbrain consists of features that are intimately connected to human emotion and the shaping of long-term memory via neural connections to the lobes of the cortex. The structures contained also link the lower brain stem to the thalamus to relay information from the senses to the brain, and back out to muscles and the limbic system.

The Limbic System

The limbic system holds the hypothalamus, amygdala, and hippocampus.

- The hypothalamus is responsible for driving actions.
- The amygdala is connected to aggressive behavior.
- The hippocampus plays a crucial role in processing various types of information to form long-term memories.

One key feature of the midbrain and limbic system is the reticular activating system (RAS). It is this area that keeps us awake and aware of the world. The RAS acts as a master switch that alerts the brain to incoming data—and to the urgency of the message.

The Cerebral Cortex

This is the outermost layer of the brain and is the starting point of thinking and voluntary monuments. In humans, the cerebral cortex has evolved into two symmetrical cerebral hemispheres.

The Basal Ganglia

The basal ganglia is a cluster of structures in the center of the brain that coordinates messages between multiple brain regions.

The Cerebellum

The cerebellum is below and behind the cerebrum, at the base and back of the brain, and is tied to the brain stem. It controls motor function and the body's ability to interpret information sent to the brain by the eyes, ears, and other sensory organs. It is responsible for coordination and balance. It is separated into right and left hemispheres; the functions of these are shown in Fig. 1.2.[7]

The brain is divided down the middle into two symmetrical and equal parts, considered as right and left hemispheres. Although equal in size, these two sides are not the same and do not carry out the same functions. Both hemispheres are connected by the corpus callosum and serve the body in different ways.

1. Functions of the left brain.
   - a. The left side of the brain is responsible for controlling the right side of the body.
   - b. It also performs logical tasks such as those found in science and mathematics.
   - c. The left hemisphere is dominant in language.
   - d. The left hemisphere is important for preprocessing social emotions.
2. Functions of the right brain.
   - a. The right hemisphere is responsible for controlling the left side of the body.
   - b. It is responsible for creative awareness.
   - c. The right hemisphere is dominant in emotional expression.
   - d. It is also dominant in the perception of facial expression, body posture, and prosody.

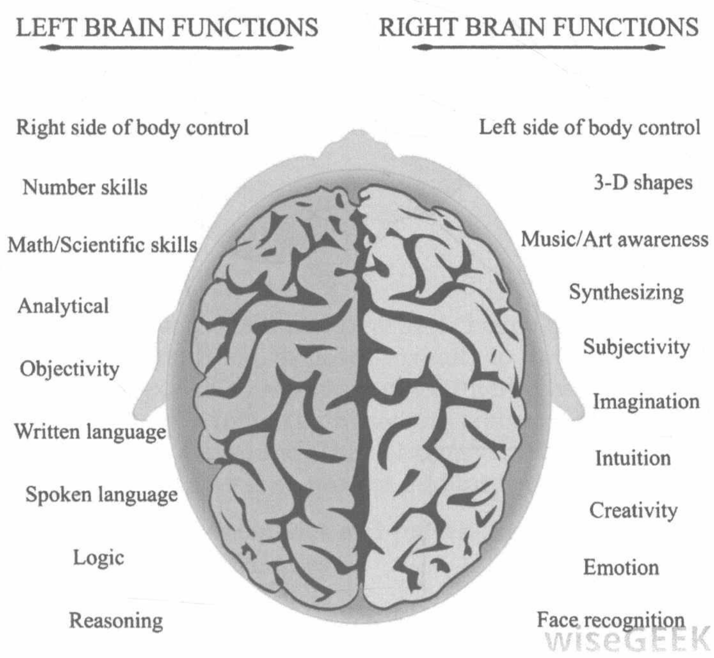

FIGURE 1.2 Functions of left and right hemispheres of the brain.

The Cerebrum

The cerebrum is the largest part of the brain. It is responsible for memory, speech, the senses, and emotional response. It is divided into four sections called lobes: the frontal, temporal, parietal, and occipital. Each handles a specific segment of the cerebrum's jobs.

The diencephalon is inside the cerebrum above the brain stem. Its tasks include sensory function, food intake control, and the body's sleep cycle. As with the other parts of the brain, it is divided into sections. These include the thalamus, hypothalamus, and epitheliums.

The brain is protected from damage by several layers of defenses. Outermost are the bones of the skull. Beneath the skull are the meninges, a series of sturdy membranes that surround the brain and spinal cord. Inside the meninges, the brain is cushioned by fluid.[8]

1.3 LOBES OF THE BRAIN AND THEIR FUNCTIONS

The brain is the most complex organ in the human body. It comprises the frontal, occipital, temporal, and parietal lobes, as shown in Fig. 1.3.

The four lobes have different locations and functions that support the responses and actions of the human body.

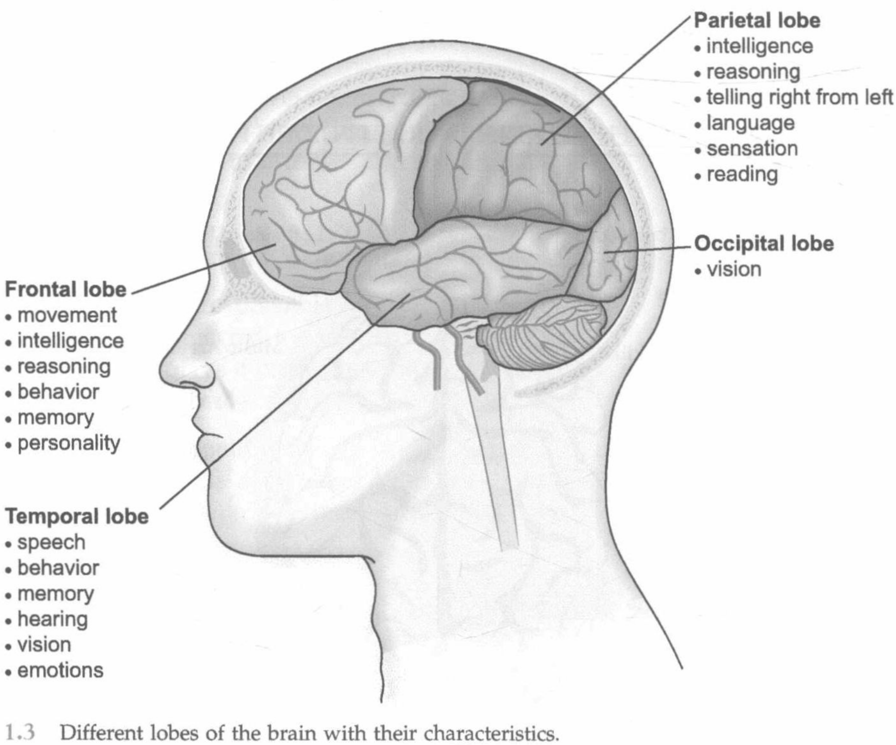

FIGURE 1.3 Different lobes of the brain with their characteristics.

1.3.1 The Frontal Lobe

The *frontal lobe* is the emotional control center of the brain, responsible for forming our personality and influencing out decisions. It is located at the front of the central sulcus, where it receives information signals from other lobes of the brain. The frontal lobe is responsible for problem solving, judgment, and motor function; it controls thinking, planning, organizing, short-term memory, and movement. Most of its functions center on regulating social behavior. Some of the important functions of the frontal lobe include

- cognition, problem solving, and reasoning
- motor skill development
- parts of speech
- impulse control
- spontaneity
- regulating emotions
- regulating sexual urges
- planning.

1.3.2 The Parietal Lobe

The *parietal lobe* processes sensory information for cognitive purposes and helps coordinate spatial relations. It resides in the middle section of the brain behind the central sulcus, above the occipital lobe. The parietal lobe is responsible to manage sensation, handwriting, and body position. It interprets sensory information, such as temperature and touch, and is responsible for processing sensory information from various parts of the body. Some of the functions of the parietal lobe include

- sensing pain, pressure, and touch
- regulating and processing the body's five senses
- movement and visual orientation
- speech
- visual perception and recognition
- cognition and information processing.

1.3.3 The Temporal Lobe

The *temporal lobe* is located at the bottom of the brain below the lateral fissure; there is one temporal lobe on each side of the brain, in close proximity to the ears. This lobe is the location of the primary auditory cortex, which is important for interpreting the sounds and language we hear. The temporal lobes are involved with memory and hearing, and process information from our senses of smell, taste, and sound. They also play a role in memory storage. The primary function of the temporal lobes is to process auditory sounds. Other functions include

- help in formation of long-term memories and processing new information
- formation of visual and verbal memories
- interpretation of smells and sounds.

1.3.4 The Occipital Lobe

The *occipital lobe* is located in the back portion of the brain behind the parietal and temporal lobes, and is primarily responsible for processing visual information. The occipital lobe contains the brain's visual processing system: it processes images from our eyes and links that information with images stored in memory. The occipital lobe, the smallest of the four lobes, is located near the posterior region of the cerebral cortex, near the back of the skull. It is the primary visual processing center of the brain; other functions include[9,10]

- visual-spatial processing
- movement and color recognition.

1.4 ELECTROENCEPHALOGRAPHY

Electroencephalography (EEG) is a non-invasive brain-imaging method that records the brain's electrical activity at the surface of the scalp. EEG was first used in 1929 by Hans Berger, who recorded brain activity beneath the closed skull and reported changes during different states. In 1957 Gray Walter was the first person to record the brain with electrodes, and showed that brain rhythms changed according to different mental tasks.

EEG equipment usually comes in the form of a cap or headset that has several electrodes or sensors which are designed to fix to the surface of the head.[11] An EEG device records electrical signals from the brain, specifically postsynaptic potentials of neurons, through electrodes attached either to the subject's scalp, the subdural (ie, beneath the dura matter—the outermost, toughest, and most fibrous of three membranes covering and protecting the brain and spinal cord), or even the cortex itself (these latter two cases are relatively rare). EEGs are based upon the theory of volume conduction of ionic current through nonempty extracellular space. The recording is obtained by placing electrodes on the scalp, usually after some abrasion and with a conductive gel to create a better contact. The measured EEG activity is the sum of all the synchronous activity of all the neurons in the area below the electrode that have the same approximate vertical orientation to the scalp. Fig. 1.4A and B shows the activity in the brain after electrodes are placed on the scalp using the illustrated EEG cap.

An EEG is used to diagnose certain brain disorders. The measurements given by an EEG are used to provide information about disorders such as:

- seizure disorders, including epilepsy
- head injury
- encephalitis, or inflammation of the brain
- brain tumor
- encephalopathy, or brain dysfunction resulting from various causes
- memory problems
- stroke
- sleep disorders.[12,13]

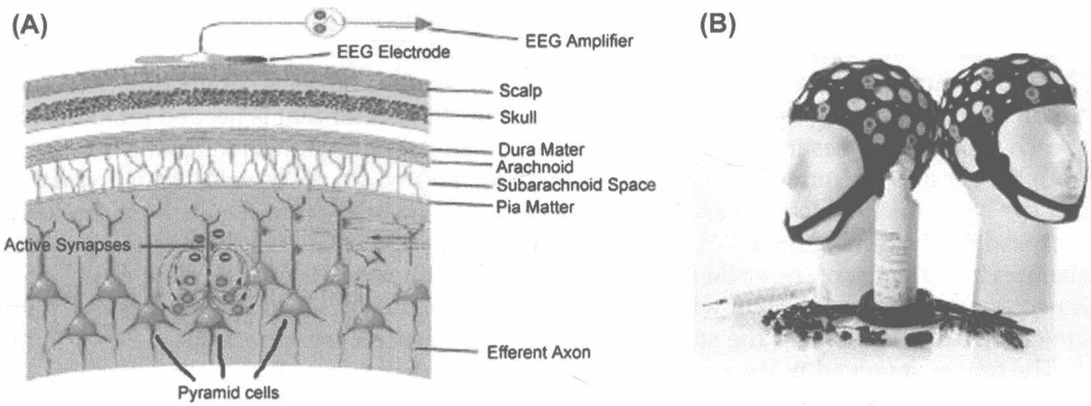

FIGURE 1.4 (A) Activity in brain after placing the electrodes; (B) EEG cap.

1.5 HUMAN AUDITORY SYSTEM

To plan a speech-based interface system, it is significant to comprehend the working of the human auditory system. At the linguistic level of communication, an idea is first formed in the mind of the speaker, and then produced in the form of speech. The idea is transformed into words, phrases, and sentences according to the grammatical rules of the language. At the physiological point of communication, the brain creates electric signals that travel along the motor nerves. These electric signals activate muscles in the vocal tract and vocal cords. This vocal tract and vocal cord movement results in pressure changes within the vocal tract and in particular at the lips, initiating sound waves that propagate into space. Finally, at the linguistic level of the listener, the brain performs speech recognition and understanding.[14–16] Besides interactive command and control signals, it is efficient in expressing emotions. This effectiveness can be examined to build robust emotion recognition systems.

1.5.1 Speech Production Mechanism

Human speech is brought out by complex interactions between the diaphragm, lungs, pharynx, mouth, and nasal cavity. The procedures which control speech production are phonation, resonation, and articulation. Phonation is the process of converting air pressure into sound via the vocal folds, or vocal cords. Resonation is the process by which certain frequencies are emphasized by resonances in the vocal tract. Articulation is the procedure of altering the vocal tract resonances to produce distinguishable sounds.

Air enters the lungs via the normal breathing mechanism. As air is released from the lungs through the trachea, the tensed vocal cords within the larynx are made to oscillate by the air stream. The air-flow is chopped into quasiperiodic pulses, which are modulated in frequency by passing through the pharynx, mouth cavity, and nasal cavity. Depending on the status of various articulators, different sounds are made. The lungs and associated muscles act as the