Chapter 92
Emotional State Recognition Using Facial Expression, Voice, and Physiological Signal

Tahirou Djara
Université d'Abomey-Calavi, Benin

Abdoul Matine Ousmane
Université d'Abomey-Calavi, Benin

Antoine Vianou
Université d'Abomey-Calavi, Benin

ABSTRACT

Emotion recognition is an important aspect of affective computing, one of whose aims is the study and development of behavioral and emotional interaction between human and machine. In this context, another important point concerns acquisition devices and signal processing tools which lead to an estimation of the emotional state of the user. This article presents a survey about concepts around emotion, multimodality in recognition, physiological activities and emotional induction, methods and tools for acquisition and signal processing with a focus on processing algorithm and their degree of reliability.

INTRODUCTION

For twenty years, the computer modeling of emotion is a theme increasingly recognized, particularly in the field of human-machine interaction (Picard, 1997). The term "emotion" is relatively difficult to define from a scientific point of view. Indeed, the phenomenon of emotion is based at the same time on physical, physiological, mental and behavioral considerations. Thus, many areas such as affective computing and image processing are interested in human emotional dimensions. For ten years, the emotional component has been taken into account and developed significantly in the fields of robotics, human-machine interaction, and more particularly in the context of interaction with animated conversational agent (ACA).

DOI: 10.4018/978-1-7998-1754-3.ch092

The growing maturity of the field of emotion recognition is creating new needs in terms of engineering. After a replication phase, during which numerous works have been proposed with recognition systems (Jaimes & Sebe, 2007), we are gradually entering an empiricism phase (Clay, Couture, & Nigay, 2009), where models for the design are developed (Jaimes et al., 2007). Most designed systems allow passive recognition of emotions. To define emotion, we base ourselves on Scherer's theory (Scherer, 2000). An emotion is characterized by a highly synchronized expression: the whole body (face, limbs, physiological reactions) reacts in unison and the human emotional expression is clearly multimodal. Indeed, a large number of studies have been carried out in order to define the relationship between emotion and physiological signal. These have allowed to highlight a significant correlation between this type of signal and certain emotional states.

This article is a state of the art about emotions and emotion recognition systems. We discuss some concepts about emotion, its representation and characteristics, and the multimodal approach in recognition systems. We present an analysis of physiological activity and emotional activation and review methods and tools for acquisition and processing of physiological signals, in particular classification methods. Finally, we focus on evaluation criteria and induction techniques (images, videos).

EMOTION RECOGNITION SYSTEM ARCHITECTURE

The analysis of existing emotion recognition systems reveals a decomposition into three levels, each fulfilling a specific function: Capture, analysis and interpretation levels. Figure 1 shows the emotion recognition system architecture.

At the capture level, the information is captured from the real world and in particular from the user through devices (camera, microphone, etc.). This information is then analyzed in the analysis level, where emotionally relevant characteristics are extracted from the captured data. Finally, the extracted characteristics are interpreted to obtain an emotion. This division into three level - capture, analysis and interpretation - is classic in emotions recognition and form a functional motif on which we rely to develop a model.

This architecture model offers five component types (Figure 2). Each component subscribes and issue one or more data stream. The capture unit has the role of interfacing with a physical device for capturing data. The feature extractor analyzes input data in order to extract one or more emotionally relevant characteristics. An interpreter receives the values of several characteristics. Its role is to interpret emotion. This interpretation is subject to the emotion model considered (discrete model, continuous, or componential) as well as the computer algorithm used (e.g., neural network, hidden Markov model, etc.).

The model also has two types of components unrelated to an "emotion recognition" logic. The role of the adapter is to modify a data flow. It can be a simple modification of format as a heavy processing

Figure 1. Emotion recognition system architecture