

Cascaded Subband Energy-Based Emotion Classification

Senaka Amarakeerthi¹, Chamin Morikawa², Ma Tin Lay New³,
Liyanage C. De Silva⁴ and Michael Cohen⁵

¹Spatial Media Group, University of Aizu, Aizu-Wakamatsu, Fukushima-Ken 965-8580, Japan, ²Interfaculty Initiative in Information Studies, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033; Japan, ³Institute for Infocomm Research, Fusionopolis Way, Connexis 138632; Singapore, ⁴Faculty of Science, University of Brunei Darussalam, Jalan Tungku Link BE 1410; Brunei, Darussalam

Abstract

Since the earliest studies of human behavior, emotions have attracted attention of researchers in many disciplines, including psychology, neuroscience, and lately computer science. Speech is considered a salient conveyor of emotional cues, and can be used as an important source for emotional studies. Speech is modulated for different emotions by varying frequency- and energy-related acoustic parameters such as pitch, energy, and formants. In this paper, we explore analyzing inter- and intra-subband energy variations to differentiate six emotions. The emotions considered are anger, disgust, fear, happiness, neutral, and sadness. In this research, Two-Layered Cascaded Subband Cepstral Coefficients (TLCS-CC) analysis was introduced to study energy variations within low and high arousal emotions as a novel approach for emotion classification. The new approach was compared with Mel Frequency Cepstral Coefficients (MFCC) and Log Frequency Power Coefficients (LFPC). Experiments were conducted on the Berlin Emotional Data Corpus (BEDC). With energy-related features, we could achieve average accuracy of 73.9% and 80.1% for speaker-independent and-dependent emotion classification respectively.

Keywords: Emotion Classification, Hidden Markov model, Sentiment Analysis