

Body Gesture Analysis and Emotion Recognition for Healthcare Application

Amal El Fallah Seghrouchni, Fatima zahra Salmam

Key-words: emotion recognition, machine learning, facial expression, body gesture, speech recognition.

1 Research Motivation

Human emotional states are conveyed by a set of nonverbal and verbal communication. The nonverbal communication includes facial expression, body movement, posture, and gesture. The verbal communication includes speech, tone of voice, and speaking style. The combination of these different modalities yields to interpret human behavior, in which non-verbal communication plays a vital role [1].

Automated body gesture analysis and emotion recognition are useful in various application areas, such as health-care, online chat, and computer-mediated communication. For example, online chat systems focus on detecting people's moods, i.e. whether they are happy or not, while health applications focus on identifying potential signs of developmental disorder such as Autism spectrum disorder (ASD) or Attention Deficit Hyperactivity Disorder (ADHD), or mental disorders such as depression or panic attacks[2]. For example, body gesture analysis and emotion recognition can be used to propose an automatic system to help doctors early screen ASD.

The project covers different disciplines such as computer vision, voice recognition more particularly automatic tone recognition, and machine learning.

2 Problem Statement

Facial expression and voice have been widely studied in the field of Human-computer recognition. However, the development of real-life applications in natural environments is still far from ideal. Body gesture learning is still an unexplored and unsolved area compared to facial expression or audio learning tasks; many issues about emotional body gesture remain unknown or still under discussion for psychologists[1]. To better understand human behavior, additional studies and investigations in body gesture recognition must be performed in combination with automatic facial expression and tone recognition. Automatic tone recognition is also an important task; it represents a classification process that determines the tone identities of individual syllables based on speech signal. Tone is described by the pitch contour of a syllable that can be measured acoustically in terms of frequency (F0). The major challenge comes from the great variability of F0 in natural speech, which is caused by multifarious linguistic and non-linguistic factors [3]. In fact, the emotional feature extraction represents the main issue in the automatic tone recognition

system. Several studies [4], [5] have proposed many features that contain emotion information, such as energy, pitch, formant frequency, Linear Prediction Cepstrum Coefficients (LPCC), Mel-frequency cepstrum coefficients (MFCC), and modulation spectral features (MSFs) [6]. Hence, most researchers prefer to use several combined feature types containing more emotional information [7] that may give rise to high dimension and redundancy of speech features.

3 Research scope

The objective of this thesis is to analyze verbal and non-verbal communication to detect the developmental disorder. The goal is to develop and propose a set of combined features extracted from facial expression, speech signal, and body gesture, and show how the combined features contribute to detecting developmental disorder.

4 Admission Criteria

The PhD position is proposed by the International Center of Artificial Intelligence of Morocco, of the Mohammed VI Polytechnic University.

Applicants with excellent cursus must be holders of a Master's, an engineering or an equivalent recognized degree in Computer Science. In addition, they should have skills in Programming (Python and C++) and good communication skills in English. Particular attention will be given to the suitability of this research project with the applicant's background.

References

- [1] Caifeng Shan, Shaogang Gong, and Peter W McOwan. "Beyond Facial Expressions: Learning Human Emotion from Body Gestures." In: *BMVC*. Citeseer. 2007, pp. 1–10.
- [2] Zhengyuan Yang et al. "Pose-based Body Language Recognition for Emotion and Psychiatric Symptom Interpretation". In: *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE. 2021, pp. 294–301.
- [3] Tan Lee, Yao Qian, and CH Lee. "Tone modeling for speech recognition". In: *Advances in Chinese spoken language processing (2007)*, pp. 179–200.
- [4] Meshach A Martin et al. "AUTOMATIC SPEECH EMOTION RECOGNITION USING MACHINE LEARNING." In: *International Journal of Advanced Research in Computer Science* 12 (2021).
- [5] Siqing Wu, Tiago H Falk, and Wai-Yip Chan. "Automatic speech emotion recognition using modulation spectral features". In: *Speech communication* 53.5 (2011), pp. 768–785.
- [6] Leila Kerkeni et al. "Automatic speech emotion recognition using machine learning". In: *Social media and machine learning*. IntechOpen, 2019.
- [7] Siqing Wu. *Recognition of human emotion in speech using modulation spectral features and support vector machines*. Queen's University, 2009.