SAgE Doctoral Training Awards

[Multimodal Techniques for Monitoring Emotional State in Elderly People and Their Carers]

Theme: [Ageing]

[School of Electrical and Electronic Engineering (SEEE), Newcastle University, UK & Singapore]

In partnership with [Nanyang Technological University (NTU), Singapore & Ageing and Vitality, FMS]

| Supervisory Team | Mohsen Naqvi, Jonathon Chambers, Satnam Dlay, Lok Woo SEEE & Singap. [http://www.ncl.ac.uk/eee/staff/profile/{mohsen.naqvi}]
| Lynne Corner, Ageing and Vitality, FMS [http://www.ncl.ac.uk/medicalsciences/research/groups/profile/lynne.corner]
| Zhiping Lin, NTU [http://research.ntu.edu.sg/expertise/academicprofile/Pages/StaffProfile.aspx?ST_EMAILID=EZPLIN]
| The Lead Supervisor is Early Career or newly hired Staff ☒

Key Words | Healthcare, Multimodal Processing, Machine Learning, Assisted Living, Ageing

Overview

In the UK there are 1.4M over 85 and by 2035 this will increase to 3.6M. In another statement by the Telegraph, it is stated the UK population is ageing rapidly and it is estimated that 10 million alive today will live to be more than 100 years old [http://www.telegraph.co.uk/health/8230749/10m-alive-in-Britain-today-will-live-to-be-more-than-100-years-old.html].

China’s over 60 population will double by 2040 and will then be larger than the entire population of Western Europe.

The above mentioned positive change in the living age will also bring challenges to the existing healthcare systems, particularly, the already overstretched NHS UK. The Engineering and Physical Sciences Research Council (EPSRC) UK already highlighted these challenges which also directly maps onto the new Healthcare Technologies Theme strategy under the umbrella of “Transforming Community Health and Care”. In EU based research directives under the framework of HORIZON 2020, the societal challenges have been given significant importance.

Newcastle University aspires to be a leading Civic University i.e. helping societies and has already taken steps to contribute world leading research to the global challenges and this project will be under one of the three over-arching research themes on Ageing, Social Renewal and Sustainability. This specific work will be performed in the Faculty of Science, Agriculture and Engineering (SAgE) and the research area is: Ageing and Digital Technologies.

The main focus of this project is to monitor emotional state of elderly people and their carers in a smart home environment, specifically, designed for assisted living based healthcare system. Multimodal (audio-visual) processing based algorithmic solutions will be contributed in this project. In signal and information processing, the main technical challenges to achieve the desired targets are:

1. Measure and quantify human behaviour
2. Process overtly expressed and directly observable signals
3. Deal with varying degrees of existing state-of-the-art technology
4. Acquire multimodal data for stress and cognitive behaviour in real time
5. Effective computing for meaningful analysis and timely decision making for actions

Methodology

The postgraduate researcher (PGR) will initially perform a thorough literature review and explore the available multimodal datasets e.g. IEMOCAP, VAM to select sequences which contain audio and video recordings for emotion recognition for human behaviour analysis. More challenging datasets will be
The uncertainty in emotional state itself is manifested through multimodality such as: vocal, language and visual behavioural cues. It is highlighted that in this project overtly expressed and directly observed signals captured through audio-visual sensors (non-intrusive sensors i.e. microphones and video cameras) will be used. Moreover, the multimodality can provide additional strength with covertly observed signals such as heart rate, brain activity, electrophysiological response, which will be considered in parallel work. The audio processing will begin with background interference and noise cancellation. The separated speech signals will be processed for keyword spotting and prosody modelling in which an emotional state of a speaker will be expressed and characterized by variations of supra-segmental language features, such as pitch, syllable duration, and voice quality. Initially, exiting state-of-the-art techniques will be applied on the real datasets and variations will be analysed before new contributions.

In video processing, facial features such as eyebrows, lip movements, and different face regions will be mainly focused. Since a human is the most complex machine in this world, therefore in this project, for monitoring of emotional states such as: anger, sadness, neutral, happiness, etc., the face and head regions of an elderly or carer will be only considered in the video processing part. Finally, both audio-visual information will be fused for emotional state classification for decision making.