PhD Proposal 2017

School: CentraleSupélec	
Laboratory: Campus de Rennes	**Web site:** http://www.rennes.supelec.fr/ren/rd/fast/fast_accueil.php
Team: FAST	**Head of the team:** Renaud Séguier
Supervisor: Renaud Séguier	**Email:** renaud.seguier@centralesupelec.fr
Collaboration with other partner during this PhD: **In France:**	**In China:**

Title: Voice and Video Stress Analysis
Scientific field: Signal and Image processing, Man Machine Interaction
Key words: Facial, Signal, Image, Emotion detection, Lie detector

Details for the subject:

Context

FAST team of CentraleSupélec is based in Rennes and belong to the Image department of the CNRS research unit IETR (Institut d'Electronique et de Télécommunications de Rennes, UMR6164). We work since more than 15 years on face analysis and synthesis (lip-reading, face cloning, gaze detection) for Human Computer Interaction throughout an animated 3D character.
The past five years we dedicate a strong energy in multimodale (audio-video) emotion analysis. We were the leader of the Immemo project (French Research ANR found) and as such we won the following international challenges: first place in micro-expression detection (FERA 2011), first and second place in emotion detection (AVEC 2012). Furthermore our research work has lead to the creation of two startups (Dynamixyz and 3D Sound Labs) which stay privileged partners.

Research subject:

The objective of this thesis is to identify stress variations. This work aims two application contexts: lies detection and stress management in public talk. A simple webcam record can give extensive information on the emotional state of a person via facial expressions produced, orientation and movement of the head, gaze direction, the voice or the heartbeat directly detected on the color information of the face over time.

These multimodal information is essential to detect whether someone is lying or not. It is also useful in speech context. On the one hand, stress can decrease the performance of a person. But also stress, like all emotions, is contagious, that is to say, it is transmitted to people with whom you are in contact. Thus, during a speech, it is important to "hide" his visible stress in order not to contaminate the audience.

The team's work in recent years in analysis and synthesis of facial expressions resulting in an invariant representation of emotional facial expressions, independent of morphology of people. This representation is based on the use of deformable statistical models to automatic extraction of facial information. A Delaunay tesselation is used to extract a mathematical manifold in order to align and compare several analysis coming from different persons.

Work plan:

Through this thesis we wish extend our facial representation to other modalities from audio-video analysis such as face gesture, gaze direction or heart rate, and the voice. Then we want to study the stress-related multimodal path in this space.

References:

Jérôme Manceau, Catherine Soladie, Renaud Séguier, *Patches Detection and Fusion for 3D Face Cloning*, Advances in Image and Video Processing, Society for Science and Education, 2015

Catherine Soladie, Hanan Salam, Nicolas Stoiber, Renaud Séguier, *Continuous Facial Expression Representation for Multimodal Emotion Detection*, International Journal of Advanced Computer Science, 2013

Catherine Soladie, Nicolas Stoiber, Renaud Séguier, *Invariant Representation of Facial Expressions for Blended Expression Recognition on Unknown Subjects*, Computer Vision and Image Understanding, Elsevier, 2013

Thibaud Sénéchal, Vincent Rapp, Hanan Salam, Renaud Seguier, Kevin Bailly, Lionel Prevost, *Facial Action Recognition Combining Heterogeneous Features via Multi-Kernel Learning*, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Institute of Electrical and Electronics Engineers (IEEE), 2012