Center for Visual & Decision Informatics

A National Science Foundation IUCRC

Executive Summary

Title: Co-Botics – Intelligent Cooperating robots and humans – Phase II		**Project ID: 7a.028-TUT**
Today's Date: 14 Feb 2018	**Estimated Start Date:** 1 Aug 2018	**Type:** [ ] New [x] Continuing
Principal Investigator: Moncef Gabbouj	**University:** TUT	**Email:** moncef.gabbouj@tut.fi
Other Project Participants: Co-PIs Jenni Raitoharju and Alexandros Iosifidis, Researcher Fahad Sohrab		

Project Description: Computational models designed to automate and monitor interactions between humans and robots that provide humans and robots with an appropriate mental model of how the others (humans and robots) will react to various behaviors, data quality, instructions and environmental changes are of extreme importance for collaborative robotics, referred here as "co-botics". The planned research will focus on the application of advanced machine learning and pattern recognition methodologies for facilitating shared intelligent cooperation between robotic units and humans. Advanced multi-modal (or otherwise called multi-view) data analysis aiming at describing cues from the real world (including humans) from multiple information sources will be developed and applied to this end. Based on that technology, online visual information analysis will be combined with sensor data analysis for decision making that will be interpreted in the entire system as suggestion-based cooperation through shared intelligent interactions. The project will continue towards enhancing the performance of multi-modal visual/sensor data analysis methods for efficient robot-human interaction in efficient scheduling applications. Moreover, it will focus on creating data visualizations that combine information coming from various types of sources (visual, depth, audio) in order to provide insights on the way robots perceive their environment. We believe that such visualizations will allow us a better understand of how to enhance the overall operation and increase intelligence of robotic units in the targeted scenarios.

Experimental Plan: During the project period, the focus will work towards:

- Enhancing the multi-modal (visual, audio, depth, label) data analytics performance.
- Developing sensors' signal analysis and integration for visualization.
- Developing data visualizations for assisting decision making in shared intelligent human-machine systems.
- Efficient implementations using parallel computing.

Related Work: Making use of multiple data sources has been one of the primary efforts towards Human-Computer Interaction. Several approaches have been proposed, mainly by setting assumptions related to the underlying distributions of data coming from different modalities, such as Multi-view Discriminant Analysis and Multi-view Canonical Correlation Analysis. Recent advances in Deep Learning have led to extensions of such approaches to achieve a better performance in difficult problems requiring multi-modal data embedding. We recently devised a unified framework to exploit multiple types of input data and jointly optimize a combined representation in a latent space, which has been shown to enhance performance in image and text retrieval and cross-modal recognition. The advantage of our model is that it allows flexible class representation between different modalities, thus negating the need of strict assumptions related to the underlying data distributions from various data sources.

How this project is different: The novelty of this project (as a continuation of last year's project) lies in the proposal and implementation of an integrated system that takes into account all real-world information made available to the robotic systems through sensors. In this way, we aim at creating an augmented world representing the objects appearing in the direct environment of the robotic unit and the cooperating person(s), through multiple data types (e.g. depth, size, color, audio, etc.). This second year of the project will allow us to visualize this imaginary world in order to better understand the way each robotic unit perceives its environment and lead to better decision-making methodologies. We expect that such an approach can lead one step closer to real-life cooperation between robots and humans, allowing efficient and intelligent human-robot interactions.

Milestones:

3 months: Improvement of the existing multi-modal data analysis methods

6 months: Testing of state-of-art multi-modal data visualization models

9 months: Design, implementation and testing of new multi-modal data visualization methodologies

12 months: Prototype creation.

Deliverables:

1. New multi-modal data (videos/images, audio, depth) analysis methodologies.
2. Multi-modal data visualization models
3. Visualizations of combined data identities in latent space
4. Efficient implementation and integration to prototype.

Proposed Budget:

Researcher Salary	€34,400
Indirect salary costs	€18,232
Overhead	€50,527
Travel	€8,000
Material and other costs	€3,841
Total	**€115,000**

How this Project may be transformative? Cooperative robotics gained much attention during the last years due to the high overseen impact in commercial and everyday living scenarios. Multi-modal data analysis and fusion can be exploited by a large variety of products exploiting information collected by various types of sources, including but not limited to cooperative robotics targeted in this project. Such models can be exploited, e.g. for combining multi-spectral images patient monitoring with experts' descriptions and quantitative (e.g. ECG and EEG) measurements, or autonomous cars.

Potential Member Company Benefits: The output of the project would be very beneficial for companies that develop smart systems, such as autonomous machines, smart-building environments and multi-sensor based decision making. Furthermore, this approach would benefit companies processing data using various types of information, e.g. visual, spectral, attribute, in order to enhance data analytics performance.

Progress to Date: A rich set of algorithms were developed by our group in the past 5 years. So far, we have tested existing multi-modal data fusion (both shallow and deep) methodologies and developed novel optimization criteria for extending performance in classification and retrieval applications. Our work is currently focused on devising efficient implementations and new optimization criteria to further enhance efficiency.

Estimated Knowledge Transfer Date: 12 months