Multimodal Focus Attention Detection in an Augmented Driver Simulator

eNTERFACE
The Similar NoE Summer Workshop on Multimodal Interfaces
July 18, August 12 2005
Faculté Polytechnique de Mons, Belgium

Coordinators :

Prof. B. Macq, TELE Lab, UCL Louvain La Neuve
Profs. Laurent Bonnaud & Alice Caplier, INPG_LIS, Grenoble

Abtract:

This project proposes to use the analysis of facial expression as well as tracking of the user's state (i.e. using physiological signals and eye tracking) in an augmented driver simulator able to appropriately react to critical situations (such as hypovigilence and stress conditions)

1. Project objective

The main goal of this project is to use multimodal signals processing analysis to provided an augmented user's interface for driving. The term augmented here can be understood as such kind of attentive interface supporting the user interaction. So far at most basic level, the system should contain at least four components:

1. sensors for determining user state
2. an inference engine or feature extractor to evaluate incoming sensor information
3. an adaptive user interface based on the results from step 2
4. and an underlying computational architecture to integrate these components.

In reality a fully functioning system would have many more components, but these are the most critical for inclusion as an augmented cognition system and will be covered during this project implementation.

2. Background information

a brief review of the related literature, so as to let potential participants prepare themselves for the workshop

3. Detailed technical description

Technical description

The project is divided in 2 parts:

The first part concerns data acquisition, analysis, fusion and interpretation

- 4 types of data are considered with associated captors:
  - video sequences focusing on the face of the driver,
  - physiological signals: heart rate (HR), EEG measures, galvanic skin response (GSR), blood pressure, blood volume pulse (BVP)...
  - speech if available (we think of an opportunistic system that would use speech produced for other interaction purposes)
  - ambiant temperature (changes in the ambient temperature has great influence to put the user into a hypovigillence state).
- A priori description of hypovigilence state for each type of data. The aim of this part is to define what kind of features need to be extracted from each kind of signal and how they should be interpreted to perform hypovigilence detection. For example on the video signal, fast eye blinking, thick head, yawning are typical of hypovigilence. The EEG signals can be used as additional data to detect eye blink. Other physiological signals can be explored to identify stress. For instance under driver stress the GSR, HR increase, whereas BVP decreases.
- Targeted signal analysis in order to extract to relevant feature. In the previous example, eye localisation, lip motion and head motion analysis are needed.

- Multimodal data fusion for hypovigilence and/or stress state detection.

The second part concerns the development of the adaptive augmented user interface based on the previous analysis. In order to build an augmented reality driving simulator we have two solutions:

- Solution1 - Use a programmable game PC version. The user interactions in the game are captured in a video sequence. In this case user interactions are performed using keyboard and mouse or a joystick. The augmented view can be done by overlapping to the video some relevant textual information. Audio feedback (speech) could be also provided. The extra information provided to the user will be based on the data analysis. A head mounted display can not be used to visualize the scene because it prevents us to have video information about face and especially about eyes.
- Solution 2 – Driving simulator with a vehicle cabin and programmable software.

Resources needed: facility, equipment, software, staff etc.

- Equipment :
  - One fast computer per participant with IEEE 1394 interface and 3D video card with OpenGL Linux driver
  - The considered captors are low-cost consumer devices : webcam for video acquisition (INPG-LIS will also provide IEEE1394 digital cameras), microphone for speech acquisition, heartbeat captor (as used by sportsmen/sportswomen). Ucl will provide 19 sensors channels for capture EEG signals with the respective AD (Analog-Digital) box.
  - Video tripod
  - Simulator : force feedback wheel and video-projector for panoramic vision ?
- Facility : large room with network, projection screen and controlled light conditions
- Software : Linux system (preferably Debian), or at least OS with C and C++ compilers
- Staff : system administrator with knowledge about device interfacing

Project management

- July 18 to July 29 : INPG-LIS (L.Bonnaud, A.Caplier)
- August 1st to August 12 : UCL (B.Macq)

4.Work plan and implementation schedule

a tentative timetable detailing the work to be done before and during the workshop

- April 16 : just after the team selection, start of a data acquisition campaign, in order to have data ready for the workshop start.
- July 18 : kick-off day with all participants.

- Parallel work of all sub-teams (video analysis team, physiological signals analysis teams, simulation team, ...)
- Coordination points every few days.
- Adaptation of software tools to local conditions and application
- Software integration
- HCI Evaluation Framework - Tests
- Software cleanup and packaging for later reuse
- Redaction of report and preparation of presentation
- Final presentation

5. Benefits of the research

expected outcomes of the project NB : we insist on the fact the all the software components used for this project, and all the software built during the project, should be free for use, and available as such to all participants (including after the workshop).

6. Profile of team

a. Leaders

Laurent Bonnaud was born in 1970. He graduated from the École Centrale de Paris (ECP) in 1993. He obtained his PhD from IRISA and the Université de Rennes-1 in 1998. Since 1999 he is teaching at the Université Pierre-Mendès-France (UPMF) in Grenoble and is a permanent researcher at the Laboratoire des Images et des Signaux (LIS) in Grenoble.His research interests include segmentation and tracking, human motion and gestures analysis and interpretation.

Alice Caplier was born in 1968. She graduated from the École Nationale Supérieure des Ingénieurs Électriciens de Grenoble (ENSIEG) of the Institut National Polytechnique de Grenoble (INPG), France, in 1991. She obtained her Master's degree in Signal, Image, Speech Processing and Telecommunications from the INPG in 1992 and her PhD from the INPG in 1995. Since 1997, she is teaching at the École Nationale Supérieure d'Électronique et de Radioélectricité de Grenoble (ENSERG) of the INPG and is a permanent researcher at the Laboratoire des Images et des Signaux (LIS) in Grenoble. Her interest is on human motion analysis and interpretation. More precisely, she is working on the recognition of facial gestures (facial expressions and head motion) and the recognition of human postures.

Benoit Macq

b. Staff proposed by the leader (with brief Cvs)

Alexandre Benoit phD student at LIS_INPG lab

c. Other researchers needed (describing the required expertise for each)

Experts in physiological signals acquisition and analysis
Experts in speech recognition and analysis
Experts in face tracking and analysis
Experts in data fusion

7. References