Emotive Dialog Between a Human and a Robot

A New Probabilistic Methodology
“People perceive and treat robots not just as machines, but also as their companions or artificial partners.”

Person–Robot Interactions From the Robopsychologists’ Point of View

Dr. Elena V. Libin - 2004
“Emotion affects many different aspects of cognitive functions including memory and intelligence”
Summary of this presentation

- Introduction
- Visual analysis
- Auditory analysis
- Visual synthesis
- Auditory synthesis
- Emulating Personality (Emotional Vector)
Introduction: Contributions

- Visual and Auditory Emotion Analysis and Synthesis
- Emulation of personality in the robot.
- Assessments for human engagement during a conversation.
- A robotic prototype platform
Introduction: Motivation / Problem Context

- As PCs were introduced into ordinary people life in the 70’s and 80’s; In near future, personal robots will become part of our life, helping in daily activities and interacting with us.

- Contactless interfaces shall be used to reduce the estrangement between humans and robots.
Introduction: Motivation / Applicability

- The importance of emotive characteristic become more special for robots supporting everyday life of persons. (e.g.: elderly or children).

- **Applicability 1:** With no one around, a person is injured and helpless. The robot can detect an abnormal emotional situation and trigger an alert to someone to help this person.
Introduction: Motivation / Applicability

- **Applicability 2**: An elderly that lives in an elderly house, has his/her daily morning exercises at the agenda. The robot can guide exercises.

- Sad person --- happy robot --- motivation.
- Happy person --- neutral robot --- stabilization.

- How the robot shall select the emotion to synthesize?
- It is context dependant, that’s why we have created the Emotional Vector, to make the robot configurable to different scenarios.
Introduction: Devices

Devices for the robotic solution:

- Cameras
- Microphone
- Robotic Base
- Robotic expressive face
  (screen and/or physical head)
Introduction: Proposed Solution

Strategy summarized schema (How):
Introduction: Proposed Solution

Scope of possible expressions (FE, VE): Angry, Sad, Neutral, Fear, Happy
Introduction: Proposed Solution

Strategy summarized schema (How):
Introduction: Proposed Solution

Scope of possible expressions (RES): Angry, Sad, Neutral, Fear, Happy
Introduction: Proposed Solution

Strategy summarized schema (How):
Introduction: Proposed Solution

Scope of possible personalities (SBP):
Neurotic, Agreeable, Conscientious, Extroverted, Humorous
Introduction

Visual emotion analysis
  • Auditory emotion analysis

Visual emotion synthesis
  • Auditory emotion synthesis

Emulating Personality (Emotional Vector)
Visual Emotion Analysis:
Sensory Processing / Feature Extraction

- **Facial Action Coding System (FACS) defines a total of 52 Action Units.**

- “**Action Units (AUs) are small distortions over the face that characterizes an facial expression**”

- **8 are related to the head pose, 44 are related to facial expressions.**

- **From the 44 we selected 11 that are recognizable by computer vision.**

Visual Emotion Analysis:
Sensory Processing / Feature Extraction: Action Units

AU1  AU4  AU6  AU7

AU12  AU15  AU17

AU20  AU23  AU24  AU25
Visual Emotion Analysis:
Sensory Processing / Feature Extraction: Action Units

Since some Action Units are mutually exclusive, we grouped the AUs into 7 variables.

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Stands for</th>
<th>Variable Scope</th>
</tr>
</thead>
<tbody>
<tr>
<td>EB</td>
<td>EyeBrows</td>
<td>AU4, AU1, none</td>
</tr>
<tr>
<td>Ch</td>
<td>Cheeks</td>
<td>AU6, none</td>
</tr>
<tr>
<td>LE</td>
<td>Lower Eyelids</td>
<td>AU7, none</td>
</tr>
<tr>
<td>LC</td>
<td>Lip Corners</td>
<td>AU12,AU15,none</td>
</tr>
<tr>
<td>CB</td>
<td>Chin Boss</td>
<td>AU17,none</td>
</tr>
<tr>
<td>MF</td>
<td>Mouth Form</td>
<td>AU20,AU23,none</td>
</tr>
<tr>
<td>MA</td>
<td>Mouth Aperture</td>
<td>AU24,AU25,none</td>
</tr>
</tbody>
</table>
Visual Emotion Analysis:
Sensory Processing / Feature Extraction

Pseudo-face

AU1 or 4

AU12 or 15

AU7

AU6

AU17

AU20 or 23

AU24 or 25

Original Input image

Zoom and crop nose

Upper Face ROI

Static ROI

Lower Face ROI

Dynamic Thresholding

Region Eroding

Region Segmentation

Mirror the best side
Visual Emotion Analysis:
Selecting the proper Modelling Approach

Unable to deal with uncertainty

Deterministic
Ex.: Linear control

Generative Stochastic
Ex.: HMM, GMM, DBN

Discriminative Stochastic
Ex.: NBN, BN, SVM, NN, A*

Descriptive Stochastic
Ex.: PCA

Dynamic Bayesian Networks

Does not allow prediction
Focus on data’s intrinsic structure

Fail on novel inputs and complex multimodal data

Learning Action Primitives
Dana Kulic, Danica Kragic, Volker Kruger
Visual Emotion Analysis:
Bayesian Classification - DBN

EB {AU1, AU4, none}
Ch {AU6, none}
LE {AU7, none}
LC {AU12, AU15, none}
CB {AU17, none}
MF {AU23, AU20}
MA{AU24, AU25}
Visual Emotion Analysis:
Bayesian Classification - Inference

The prior $P(FE)$ is initially an uniform distribution

Marginalization Rule:

$$P(EB, Ch, LE, LC, CB, MF, MA) = \sum_{FE} P(EB|FE) \cdot P(Ch|FE) \cdot P(LE|FE) \cdot P(LC|FE) \cdot P(CB|FE) \cdot P(MF|FE) \cdot P(MA|FE) \cdot P(FE)$$
Visual Emotion Analysis:
Bayesian Classification - Results
Visual Emotion Analysis: Demo - Video

Web Link
Visual Emotion Analysis: Experiments

Hit-Rate of our Facial Expression (FE) classifier:

570 tests --- Hit Rate = 89% of correct classifications
Visual Emotion Analysis:
Comparison with other works

Our hit rate is 89%

Is it enough to compare just the hit rate?

Automatic features extraction or manual?

Different databases

Different processing time

Useful or not for HRI

All these issues on comparison were addressed by Pantic in [Towards an Affect-Sensitive Multimodal Human-Computer Interaction --- 2003]

Sixteen (16) standard questions were proposed as assessments for this type of system
Visual Emotion Analysis:
Comparison with other works

1. Is the input image provided automatically?
2. Is the presence of the face assumed?
3. Is the performance independent of variability in subject’s sex, physiognomy, age, and ethnicity?
4. Can variations in lighting be handled?
5. Can rigid head movements be handled?
6. Can distractions like glasses and facial hair be handled?
7. Is the face detected automatically?
8. Are the facial features extracted automatically?
9. Can inaccurate input data be handled?
10. Is the data uncertainty propagated throughout the facial information analysis process?
11. Is the facial expression interpreted automatically?
12. How many interpretations categories (labels) have been defined?
13. Are the interpretation labels user profiled?
14. Can multiple interpretation labels be scored at the same time?
15. Are the interpretation labels quantified?
16. Is the input processed in fast or real time?

This questionaty assessments were defined by Pantic’2003
## Visual Emotion Analysis:
### Comparison with other works

<table>
<thead>
<tr>
<th>Reference</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>Test results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohen’ 02 [2]</td>
<td>●</td>
<td>●</td>
<td>U</td>
<td>U</td>
<td>X</td>
<td>X</td>
<td>●</td>
<td>●</td>
<td>X</td>
<td></td>
<td></td>
<td>7</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>12600 frames, 5 subjects, Correct:65%</td>
</tr>
<tr>
<td>Wuhan’ 04 [41]</td>
<td>X</td>
<td>●</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>●</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>7</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>213 frames, 10 woman, Correct 77%</td>
</tr>
<tr>
<td>Nicolaou’10 [26]</td>
<td>X</td>
<td>●</td>
<td>U</td>
<td>X</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td>6</td>
<td>X</td>
<td>●</td>
<td>●</td>
<td></td>
<td>Correct 91.76%</td>
</tr>
<tr>
<td>Pantic’09 [28]</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>X</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td>6</td>
<td>X</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gayatri’08 [27]</td>
<td>X</td>
<td>●</td>
<td>●</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>7</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>Hit rate = 77%</td>
</tr>
<tr>
<td><strong>Our Approach</strong></td>
<td>●</td>
<td>X</td>
<td>X</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>X</td>
<td>●</td>
<td></td>
<td></td>
<td>5</td>
<td>X</td>
<td>●</td>
<td>●</td>
<td></td>
<td>Hit rate = 94%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hit rate = 89%</td>
<td></td>
</tr>
</tbody>
</table>

Our input is processed in “real time” and has a higher hit-hate than the other “real time” method.

All other approaches assume the presence of a face, our do not. If there is no face, our system automatically waits until a face is present.

Gayatri’08 has a higher hit rate, but features were extracted manually, their system is not real time, and the face is not detected automatically.
Introduction

- Visual emotion analysis
- Auditory emotion analysis

- Visual emotion synthesis
- Auditory emotion synthesis

- Emulating Personality (Emotional Vector)
The person can speak the same phrase with different intonations, according to each emotion in the voice.

We focused on detecting three (3) parameters from the audio.

- Pitch (SPit)
- Sentence Duration (SDur)
- Volume Level or Energy (SEne)

We used the tool praat to detect these three variables in the sound.

*Praat is a reliable tool, developed by: Paul Boersma and David Weenink
Phonetic Sciences, University of Amsterdam
http://www.fon.hum.uva.nl/praat/
Auditory Emotion Analysis:
Sensory Processing / Feature Extraction

Description of each variable.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SDur</strong></td>
<td>Stands for Sentence Duration: Since we know the sampling frequency ((sfreq)) of the acquired sound, we also know the beginning and the end of each sentence, and consequently the number of samples ((nsam)); and then it is simple to determine the duration in seconds by (SDur = nsam/sfreq).</td>
</tr>
<tr>
<td><strong>SPit</strong></td>
<td>Stands for Sentence Pitch: Pitch represents the perceived fundamental frequency of a sound. The pitch extraction was done by autocorrelation method.</td>
</tr>
<tr>
<td><strong>SEne</strong></td>
<td>Stands for Sentence Energy: This variable is actually the energy or intensity of the signal, which for a theoretically continuous-time signal (x(t)) is given by (SEne = \int</td>
</tr>
</tbody>
</table>
Auditory Emotion Analysis:  
Feature Extraction Tests in Different Environments

Three (3) different environmental conditions were established for tests.

<table>
<thead>
<tr>
<th>Environment</th>
<th>Descr</th>
<th>Background Noise</th>
<th>Voice/Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>abs(SPit)</td>
<td>σSPit</td>
<td>SPit</td>
<td>SPit</td>
</tr>
<tr>
<td>GE</td>
<td>10.7</td>
<td>124.5</td>
<td>133</td>
</tr>
<tr>
<td>ME</td>
<td>15.4</td>
<td>126.2</td>
<td>133</td>
</tr>
<tr>
<td>BE</td>
<td>103.2</td>
<td>181.7</td>
<td>139</td>
</tr>
<tr>
<td>SDur</td>
<td>σSDur</td>
<td>SDur</td>
<td>SDur</td>
</tr>
<tr>
<td>GE</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>ME</td>
<td>0.3</td>
<td>3.10</td>
<td>3</td>
</tr>
<tr>
<td>BE</td>
<td>1.77</td>
<td>4.36</td>
<td>3</td>
</tr>
<tr>
<td>SEnε</td>
<td>σSEnε</td>
<td>SEnε</td>
<td>SEnε</td>
</tr>
<tr>
<td>GE</td>
<td>105.1</td>
<td>1613.9</td>
<td>1590</td>
</tr>
<tr>
<td>ME</td>
<td>153.1</td>
<td>1623</td>
<td>1590</td>
</tr>
<tr>
<td>BE</td>
<td>235.4</td>
<td>1777.5</td>
<td>1707</td>
</tr>
</tbody>
</table>

Standard deviation of Pitch, Sdur and Sene are acceptable for:  
Good Environment (GE) and Medium Environment (ME). For Bad Environment (BE) the system did not work properly.
Auditory Emotion Analysis:
Bayesian Classification - DBN

PT \{low, normal, high\}
SD \{short, normal, long\}
VL \{low, normal, high\}
Auditory Emotion Analysis:
Bayesian Classification - Inference

The prior $P(VE)$ is initially an uniform distribution

Marginalization Rule:

$$P(PT, SD, VL) = \sum_{VE} P(PT|VE) \times P(SD|VE) \times P(VL|VE) \times P(VE)$$
Auditory Emotion Analysis: Bayesian Classification - Results
Auditory Emotion Analysis:
Experiments - Graphical Interface

[Image of graphical user interface for Auditory Emotion Analysis]

- **Human:**
  - Ang: 20%
  - Fea: 20%
  - Sad: 20%
  - Hap: 20%
  - Neu: 20%

- **Human Emotional State:**
  - Start qjackctl

- **Robot Emotional State:**
  - Start Virtual World
  - Use Virtual World
Auditory Emotion Analysis: Experiments

Hit-Rate of Our Vocal Expression Classifier (VE)
Four batteries of tests were performed with the same phrase and randomly changing the emotions on voice.

Hit Rate = 81% of correct classifications
Auditory Emotion Analysis:
Comparison with other works

Is enough to compare only the hit rate?

The issues on comparison about audio emotional detection were also addressed by Pantic in [Towards an Affect-Sensitive Multimodal Human-Computer Interaction --- 2003]

Another set of sixteen (16) standard questions was proposed, by Pantic’03, as standard assessments for this type of system.
Auditory Emotion Analysis:
Comparison with other works

1. Can non professionally spoken input samples be handled?
2. Is the performance independent of variability in subject’s sex, physiognomy, age, and ethnicity?
3. Are the auditory features extracted automatically?
4. Are the pitch-related variables utilized?
5. Is the vocal energy (intensity) utilized?
6. Is the speech rate utilized?
7. Are pitch contours utilized?
8. Are phonetic features utilized?
9. Are some other auditory features utilized?
10. Can inaccurate input data be handled?
11. Is the extracted vocal expression information interpreted automatically?
12. How many interpretation categories (labels) have been defined?
13. Are the interpretation labels scored in a context-sensitive manner (application, user, task-profiled manner)?
14. Can multiple interpretation labels be scored at the same time?
15. Are the interpretation labels quantified?
16. Is the input processed in fast or real time

This questionnaire assessments were defined by Pantic’2003
Auditory Emotion Analysis: Comparison with other works

<table>
<thead>
<tr>
<th>Reference</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>Test results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimitrius' 06 [38]</td>
<td>✗</td>
<td>U</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>U</td>
<td>×</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>Hit rate = 76%</td>
</tr>
<tr>
<td>Nicolau' 10 [26]</td>
<td>✗</td>
<td>T</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>●</td>
<td>×</td>
<td>●</td>
<td>3</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>U</td>
<td>Hit rate = 76%</td>
</tr>
<tr>
<td>Our Approach</td>
<td>✗</td>
<td>T</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>●</td>
<td>×</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>5</td>
<td>Hit rate = 81%</td>
</tr>
</tbody>
</table>

Our input is processed in “real time” and has a higher hit-hate than the others.

Our scope of analyzed emotions is higher than the others.
Introduction

Visual emotion analysis
Auditory emotion analysis

Visual emotion synthesis
Auditory emotion synthesis

Emulating Personality (Emotional Vector)
**Visual Emotion Synthesis**

**Facial expression synthesis:** An inverted Bayesian Network was designed to re-use the information from the learning phase for classification. By re-using the likelihood table in the inverted BN it is possible to come from Expression to Features.

From the likelihood histogram, the correct AU’s to be synthesized for each expression are obtained.

Here an example of which AU’s are needed to synthesize the *Angry* facial expression.
Visual Emotion Synthesis

An 3D avatar was created, inside a virtual world. One avatar mimics the expressions of the person, the other represents the robot and it is also capable to synthesize expressions.
Visual Emotion Synthesis:
Video of Virtual World avatar synthesizing facial expressions.
Visual Emotion Synthesis

Translucent Mask:

- Mini-Projector
- Mirror
- Translucent mask
- Pan-Tilt Unit
- Hidden-Camera

Rigid-body
Visual Emotion Synthesis

Lips Synchronization:

Beyond the five emotions that robot can represent, we add capability of lips synchronization to endow a better human-machine interaction.

These 9 visemes were morphed according to each of our 5 emotions.
Introduction

Visual emotion analysis
Auditory emotion analysis

Visual emotion synthesis
Auditory emotion synthesis

Emulating Personality (Emotional Vector)
Audio Emotion Synthesis

Different emotions along the same sentence can be produced with different wave features.

Sad – speech is slow and paused

Anger – speech is fast and loud

Happy – speech is fast and the Pitch is higher
Audio Emotion Synthesis:
Video of Lips Synchronization

Lips Synchronization test with random phrases
Introduction

Visual emotion analysis
Auditory emotion analysis

Visual emotion synthesis
Auditory emotion synthesis

Emulating Personality (Emotional Vector)
Emulating Personality:

1. Robot’s answer carries different emotions according to the selected personality.

2. Adjusting the personality of the robot to the profile of the user will reduce the estrangement between them, consequently improving the interaction.

3. The scope of possible personalities is: Neurotic, Extraverted, Conscientious, Agreeable, Humorous
Emulating Personality:

To test our personality emulation system, a story board of a short dialog was used:

<table>
<thead>
<tr>
<th>Phrase Number</th>
<th>Who says</th>
<th>What is said</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Robot</td>
<td>Exercise, do it <strong>20</strong> times.</td>
</tr>
<tr>
<td>2</td>
<td>Human</td>
<td>I finished</td>
</tr>
<tr>
<td>3</td>
<td>Robot</td>
<td>How many times did you do?</td>
</tr>
<tr>
<td>4</td>
<td>Human</td>
<td>I did <strong>10</strong> times.</td>
</tr>
</tbody>
</table>

Pause and wait for confirmation that the exercises ended.

From this point in time, robot answers can be different, according to the selected SBP as shown in table below. If the number is correct, robot trigger phrase 7.

<table>
<thead>
<tr>
<th>Social Behaviour Profile (SBP)</th>
<th>Robot Possible Phrases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuroticism</td>
<td>No, I told you to do more, you should do it as I say!</td>
</tr>
<tr>
<td>Extraversion</td>
<td>Common, you are good, you are strong. Just try again and I am sure you will accomplish the task.</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>Please try again.</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>Ok, no problem, do as you want.</td>
</tr>
<tr>
<td>Humorous</td>
<td>Ha ha ha. Really? My grandma can do better than this!</td>
</tr>
</tbody>
</table>

This conversation will be in loop changing the *Exercise* and the *amount* according to the list of exercises given for this context.
Emulating Personality:
Dialogue with the robot set to Conscientious personality
Emulating Personality:
Same dialogue with robot set to Neurotic personality
Emulating Personality:

Our proposed assessments aims to collect 3 variables during the dialog:

- **TBP** – Time Between Phrases
- **AH** – Amount of Happiness
- **TT** – Total Time

This variables represents the engagement of the user during the interaction.

Graphical User Interface for experiments
Emulating Personality:
Results

Time Between Phrases

Total Time (TT)

Amount of Happiness (AH)
Impact

Social Impact:
- Ordinary people are likely to use a social robot with emotions in daily life. The interaction among the persons and the robot will change our society in near future. Emotive robots are more easily accepted in human society.
Recent Publications

Publications

2012 - 2011 - 2010 - 2009 - 2008

2012

Journal
http://dx.doi.org/10.1007/s11367-011-0134-7

IJSR

2011

Conference

RA-IASTED

Conference

DOCEIS

2010

Conference

RA-IASTED

Conference
J. Prado, J. Lobo, J. Dias - SoPHIE: Social Robotic Platform for Human Interactive Experimentation, - 4th International Conference on Cognitive Systems, COGSYS 2010, ETH Zurich, Switzerland

COGSYS

2009

Conference
http://dx.doi.org/10.1109/IROS.2009.5354564

IROS

Conference
http://dx.doi.org/10.1109/ROMAN.2009.5320393

RO-MAN

Conference

ICAR

2008

Conference
J. Prado, L. Santos, J. Dias - Heron based Dynamic Background Segmentation applied to an Interactive Mobile Robot - Proceedings of the ICAR 2008 - 13th International Conference on Advanced Robotics - Munich, Germany, June 22-26, 2008

ICAR

Conference
R. Menendez, J. Dias, J. Prado, H. Alakbarpour, S. Andino, “Multidase brain computer interface based on visual attention”, European Symposium on Artificial Neural Networks Advances in Computational Intelligence and Learning, Bruges, Belgium, 2009

ESANN

2007

Conference
J. Prado, J. Dias - “Visuovestibular-Based Gaze Control Experimental Case” - RECPAD 14a Conferência Portuguesa de Reconhecimento de Padrões, Coimbra, 2008

RECPAD

Conference
http://dx.doi.org/10.1109/IROS.2008.4651258

IROS

Past Work: Human-Robot-Interaction was extensively studied

Ondas cerebrais dirigem robô a 1590 quilômetros

Em Genebra, o investigador concentra-se na luz do lado direito do olho do computador e, na Universidade de Coimbra, um pequeno robô com rodas viaja à direita, e viaja à esquerda, se o olhar se fixa na luz da esquerda.

A experiência foi feita por dois cubanos, do Hospital Universitário de Genebra, e por um português e um iraniano, da Faculdade de Ciências e Tecnologia da Universidade de Coimbra (FCTUC). Separados por mais de 1590 quilômetros, os investigadores - Sara González e Rolando Grava, físicos ligados à Medicina, e José Prado e Hadi Aliakbarpour, engenheiros - testavam a possibilidade de ondas cerebrais controlarem, a grande distância, um dispositivo electromédico.

Um dos objetivos da investigação é facilitar a vida de deficientes motores, com cegueira, e com tetraplegicos, em tarefas quotidianas como abrir uma porta ou atender um telefone. Segundo o supervisor do Laboratório do Instituto de Sistemas e Robótica da FCTUC, Jorge Dias, é um 'muito real' a hipótese do mercado olerese, que por quatro anos, soluções tecnológicas desenvolvidas por aquele projeto.

A novidade da parceria entre Coimbra e a cidade suíça de Genebra é a transmissão da mensagem à distância - a FCTUC bem outro projeto em que se pretende que o 'deficiente' dirija uma cadeira de rodas através de ondas cerebrais -, com recurso à tecnologia de baixo custo.

Ontem, em Genebra, cada cubano experimentou um capacete, ligado ao cérebro por elétricos e conectado, também, a um computador. Num escritó, observaram uma luz de cada lado, mais as imagens transmitidas em direito por uma câmera. Instalada na robô da FCTUC, a câmera filmava a percurso, assinalado por lixa adesiva, a seguir. Se a lixa indicava uma curva à direita, o cubano deveria concentrar o olhar na luz da direita - e vice-versa. O processamento dos sinais do cérebro, para a esquerda ou para a direita, era feito na Suíça e remetido por Internet para Coimbra.

NELSON MORAIS

publicado o 2009-04-22 às 00:30


GRUPO CONTROLE/INVESTIG.

Copyright © Todos os direitos reservados
Discussion and Future Work

It was proposed, implemented and tested:

1. **PhD work**
   1. Human-Robot-Interaction was extensively studied
   2. Auditory Emotional Perception (Detectors and Classifiers)
   3. Visual Emotional Perception (Detectors and Classifiers)
   4. A Bayesian Fusion
   5. The response is synthesized with expressions and lips synchronization
   6. Toolboxes for benchmark were developed

2. **For the Social Robot Project**
   1. Integrate the emotive modules over a social robot
   2. Navigation ideas – parking place
   3. Sensor ideas
     1. Blood pressure by pharmacy sensors
     2. Body Temperature from thermo-cameras
     3. Movements by Kinect and cameras
   4. Task ideas:
     1. Morning facial exercises
     2. Daily movement measuring and report (to be developed – with Kinect?)
     3. Daily advices about medicines and feeding (advices, location of medicine and possibly taking the medicine to the person [shell device])
Discussion and Future Work

Collaborative project is funded by the European Commission within the Seventh Framework Programme FP7, by People Programme, Industry-Academia Partnerships and Pathways (IAPP), under grant agreement 285870.
Questions?

José Prado – jaugusto@isr.uc.pt
  jorge@isr.uc.pt

Jorge Dias – jorge.dias@kustar.ac.ae

Mobile Robotics Laboratory

Instituto de Sistemas e Robótica

Departamento de Engenharia Electrotécnica e de Computadores
Universidade de Coimbra
Pinhal de Marrocos - Polo II
3030-290 COIMBRA - Portugal

Thank you!