ERMIS / IST-2000-29319

humaine
emotion-research.net

Concertation with related EU project "ERMIS"

HUMAINE Workshop
"From Signals to Signs of Emotion and Vice Versa",
organized by WPe 4 of the HUMAINE Network of Excellence ,
Santorini, Greece; September 19th, 2004

Martin Nelke; MIT GmbH Aachen, Germany

ERMIS / IST-2000-29319

humaine
emotion-research.net

Outline

- Introduction and Overview
- Consortium
- Project Results
- Running Research Issues

ERMIS / IST-2000-29319

Introduction to ERMIS

humaine emotion-research.net

- "Emotionally Rich Man-machine Intelligent System"
- Keywords:
  - System,
  - Man-machine Interface
  - Intelligent System
  - Emotionally rich
- HUMAINE: development of systems that can register, model and/or influence human emotional and emotion-related states and processes - 'emotion-oriented systems'.

Sept. 19th 2004 3

ERMIS / IST-2000-29319

Scope

humaine emotion-research.net

- The development of a prototype system
  - for human computer interaction than can
  - interpret its users' attitude or emotional state, e.g., activation/interest, boredom, and anger,
  - in terms of their
    - speech and
    - facial

  gestures and expressions

Sept. 19th 2004 4

ERMIS / IST-2000-29319

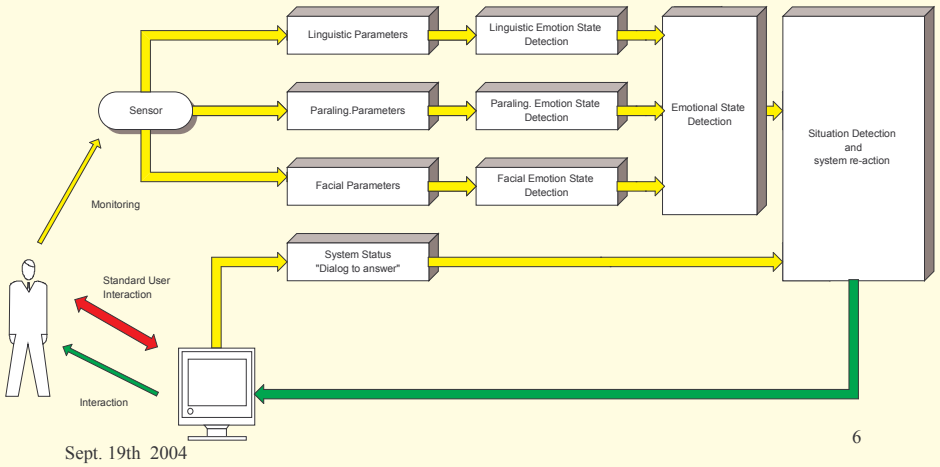
Tasks

- speech recognition
- facial / gestures recognition
- feature extraction and generation
- linguistic and paralinguistic speech analysis and, facial expression analysis,
- interpretation of the user's emotional state using hybrid and neurofuzzy techniques
- Related system reaction

Sept. 19th 2004 5

ERMIS / IST-2000-29319

System Structure

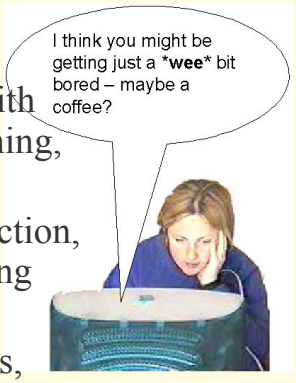


Sept. 19th 2004 6

ERMIS / IST-2000-29319

humaine emotion-research.net

Applications



- multi-modal interaction of PC users with their multimedia PCs (interactive learning, consultant services, entertainment)
- e-commerce type of multimedia interaction, e.g., with a call center or booth (booking holiday trips)
- dialog services like directory inquieries, intelligent personal assistants to locate and filter communication, managing time and communication

Sept. 19th 2004 7

ERMIS / IST-2000-29319

humaine emotion-research.net

Sensitive Artificial Listener

- which chats to a user on his/her own PC; key challenge: to handle spontaneous dialogue
- ELIZA is an early AI program that 'chats' with the user via text I/O. It has no real understanding, only tricks (stock responses, rephrasing the user's last comment); but people enjoy the quirky interaction.
- SAL will be ELIZA-like, but with input coming from speech; and able to use signs of emotion in the user's voice and facial expressions like key words/phrases.

Sept. 19th 2004 8

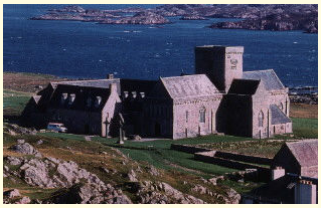
ERMIS / IST-2000-29319

Finding a holiday you feel good about

humaine emotion-research.net

System: How about scuba diving in the groovy Bermudas?

User: Oh, gruesome. Try something completely different.



System: Well, how about a retreat on a remote historic island?

User: Ah, now you're talking.

Sept. 19th 2004 9

ERMIS / IST-2000-29319

Consortium

humaine emotion-research.net

ALTEC S.A.	Greece
ILSP	Greece
IVML-ICCS-NTUA	Greece
EYETRONICS	Belgium
KUL	Belgium
QUB	UK
KCL	UK
MIT	Germany
FRANCE TELECOM	France
BRITISH TELECOM	UK
NOKIA	Finland

Sept. 19th 2004 10

ERMIS / IST-2000-29319

Consortium

humaine emotion-research.net

- The multidisciplinary consortium includes
  - big industrial IT and telecommunication partners,
  - well known research centres and universities in
    - the speech analysis,
    - image analysis,
    - psychological and computational emotion analysis fields and in intelligent multimedia systems.
- The Consortium expertise in the above fields will be used
  - for achieving its technological goals and developments,
  - for performing reliable market and exploitation analysis.

Sept. 19th 2004 11

ERMIS / IST-2000-29319

Synergy impacts by the project consortium

humaine emotion-research.net

- Cooperation between visual and speech recognition and analyis experts, data mining experts and business end users is essential for a multi-purpose approach
- Transferring the user requirements to the available methodologies or the ones to be developed
- Bringing together distributed know how about EIMS data (for different business cases), data mining algorithms, data modelling and exploitation for a product development
- European links between university and business

Sept. 19th 2004 12

ERMIS / IST-2000-29319

humaine
emotion-research.net

Visual Signal Analysis

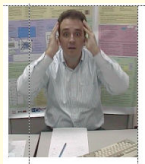


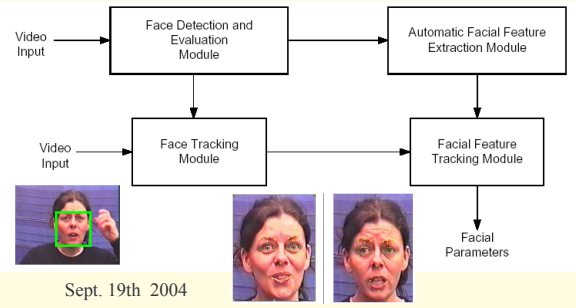
Figure 7: The original frame from the input sequence

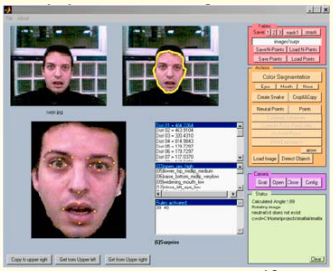
Figure 8: Detected primary facial features

Figure 9: The apex of the expression

Figure 10: Detected facial features

Figure 11: The neutral expression



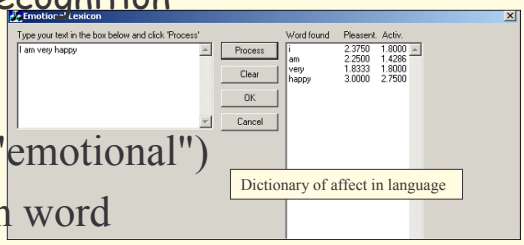


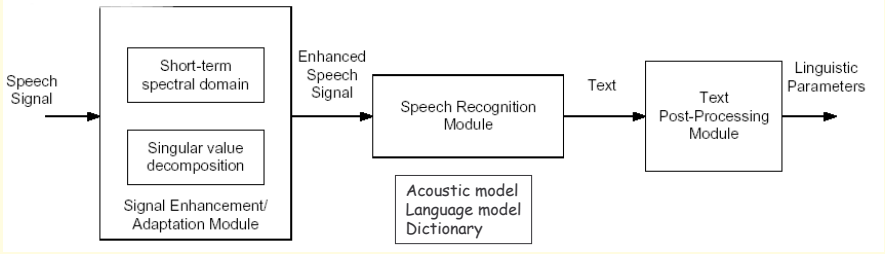
ERMIS / IST-2000-29319

humaine
emotion-research.net

Linguistic Speech Recognition

- transcription of a spoken utterance ("emotional")
- probability for each word



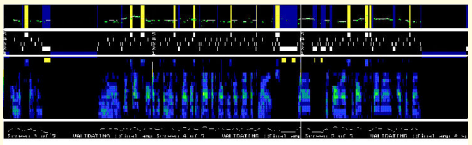


ERMIS / IST-2000-29319

Paralinguistic Speech Recognition

humaine emotion-research.net

- extract significant variations in the way words are produced
- Description of speech at multiple levels that are potentially relevant to emotion - intensity & spectrum; edits, pauses, frication; raw pitch estimates & a smooth fitted curve; rises & falls in intensity & pitch (=features)
- examining emotional dynamics of speech episodes over time



15

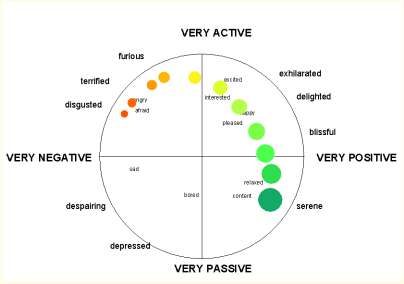
Sept. 19th 2004

ERMIS / IST-2000-29319

Emotional Representation

humaine emotion-research.net

- dimensional representation, based on activation and evaluation, as well as emotionality (i.e., the distance from the origin).



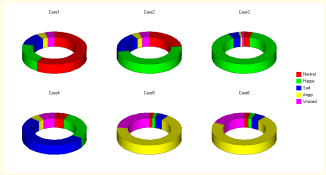
16

Sept. 19th 2004

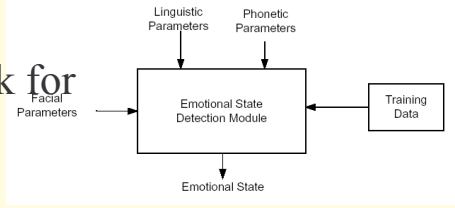
ERMIS / IST-2000-29319

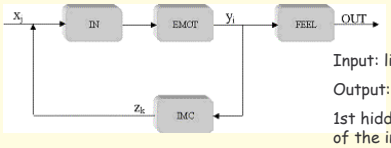
Emotional Recognition

humaine emotion-research.net



- Fuzzy Clustering of features compared with emotional patterns
- Artificial Neural Network for Attention to Emotion





Input: linguistic, paralinguistic and facial features

Output: label of the emotional state.

1st hidden layer: representing the emotional content of the input

2nd hiden layer: attention control inverse model, which learns to attend to the best input features of an input

Sept. 19th 2004 17

ERMIS / IST-2000-29319

Running research issues

humaine emotion-research.net

- Signal Enhancement in order to address the problem of robust Automatic Speech Recognition
- Parameterised approach to facial expression synthesis that is compatible with the MPEG-4 standard using Facial Animation Parameters (FAP's) using as rule-based technique for analysing both archetypal and intermediate expressions
- Emotional Recognition Architecture based on Human Brain Structure

Sept. 19th 2004 18