Detection and Identification of Rare Audiovisual Cues

Insperata accident magis saepe quam quae speres.
(Things you do not expect happen more often than
things you do expect) Plautus (ca 200(B.C.)

Project no: 027787

DIRAC

Detection and Identification of Rare Audio-visual Cues

Integrated Project
IST - Priority 2

DELIVERABLE NO: D6.7
Database of low-level incongruencies

Date of deliverable: 30.6.2009
Actual submission date: 11.8.2009

Start date of project: 01.01.2006 Duration: 60 months

Organization name of lead contractor for this deliverable: **FRA**

Revision: [1]

Project co-funded by the European Commission within the Sixth Framework Program (2002-2006)		
Dissemination Level		
PU	Public	
PP	Restricted to other program participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	X
CO	Confidential, only for members of the consortium (including the Commission Services)	

Detection and Identification of Rare Audiovisual Cues

Insperata accident magis saepe quam quae speres.
(Things you do not expect happen more often than things you do expect) Plautus (ca 200(B.C.)

D6.7 - DATABASE OF LOW-LEVEL INCONGRUENCIES

FRAUNHOFER INSTITUT DIGITALE MEDIENTECHNOLOGIE,
PROJECT GROUP HEARING, SPEECH AND AUDIO TECHNOLOGY (FRA)

Abstract:
One of the objectives of WP6 is "Recording audio, visual, and audio-visual databases that would support DIRAC research thrusts".

During the last review meeting it was concluded that for WP1 it was necessary to have recordings of so-called 'low-level incongruencies' to see whether the DIRAC framework could be used to detect e.g. camera misalignment.

For this purpose several recordings were made using the AWEAR II platform and supplied to the partners in WP1.

Detection and Identification of Rare Audiovisual Cues

Insperata accident magis saepe quam quae speres.
(Things you do not expect happen more often than things you do expect) Plautus (ca 200(B.C.)

Table of Contents

1. Introduction

During preparatory meetings for the review meeting held in Oldenburg at 6 and 7 April 2009, it was concluded that WP1 was in need of input data on so-called low-level incongruencies. One of the questions in WP1 was whether the DIRAC framework could be used to determine incongruencies down to the 'low level' of e.g. camera misalignment.
Together with the partners involved in WP1 several low-level incongruencies were defined. Most of these would be best applied to a moving camera platform. Also, in the case of a moving platform low level incongruencies would be more likely to occur. Therefore it was decided to use the AWEAR II platform for these recordings.

2. AWEAR II hardware and software

The AWEAR II hardware and software were described in the report on deliverable D6.6 and will not be repeated here.

Figure 2: The AWEAR II hardware

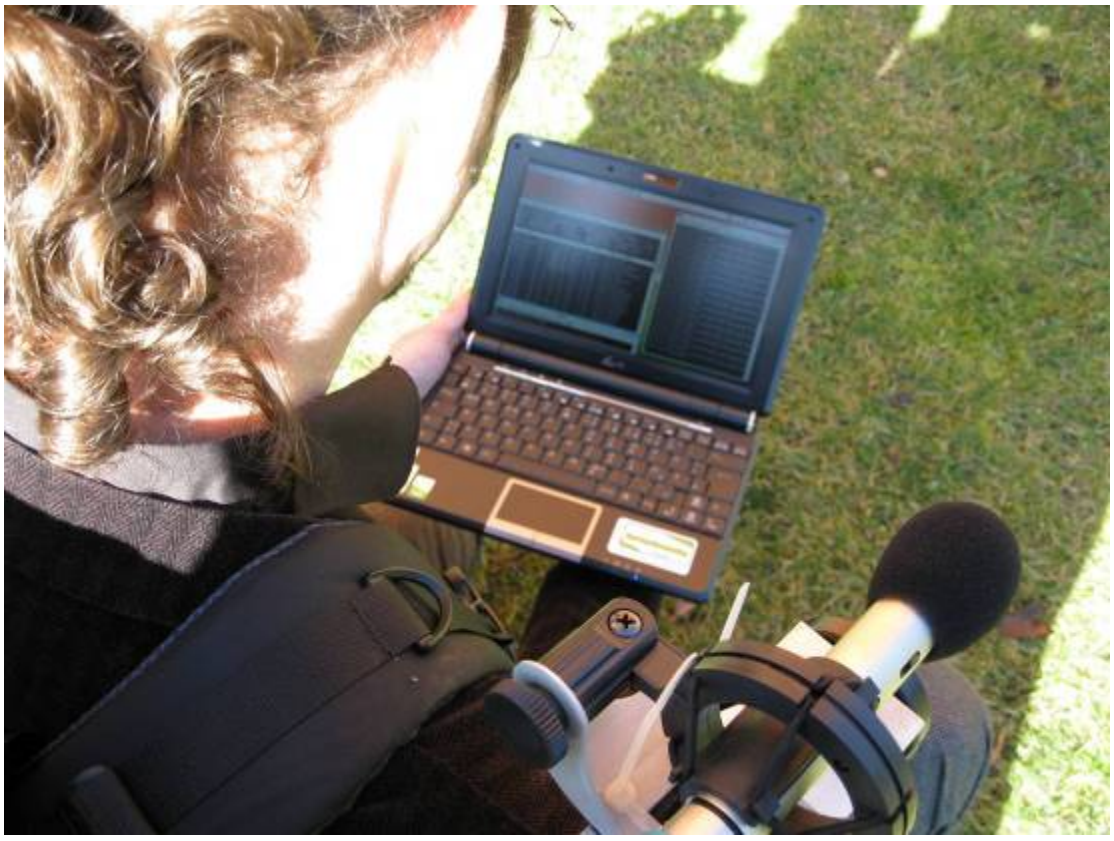

Figure 3: AWEAR user with EE PC for control of recording

3. AWEAR II Recording session I

Oldenburg, 7-8 April 2009

Disturbances in the camera system of the AWEAR II system were created by:

- disconnecting one of the cameras
- blurring the signal from one of the cameras by covering the lens with a piece of transparent foil
- changing the camera geometry by applying (gentle) force to the camera rig

Figure 4: car park of the House of Hearing

The recordings were made on the car park of the House of Hearing with the AWEAR II system worn by a person and feature the natural surroundings, several pedestrians and some vehicles.

4. AWEAR II Recording session II

Oldenburg, 23 July 2009.

Disturbances in the audio recording system were created by:

- damping the signal of one of the microphones
- changing the frequency response of one of the microphones by holding a flower pot in front of the membrane
- turning off one of the microphone signals
- disturbing a microphone signal by creating a noise (hissing) in front of one of the microphones
- placing one of the microphones in another room and speaking into it

The recordings were made in the offices of FRA in the House of Hearing with the AWEAR II system worn by a person and feature the TIMIT database played via a loudspeaker placed at 2 meters distance central in front of the system.

5. Conclusion

For both the video and the audio part of the AWEAR II platform recordings were made that can be used to study the possibility to use the DIRAC theoretical framework to detect low level incongruencies. These recordings were made available to the partners in WP1.

6. Reference

D6.6 First audio-visual data collection with AWEAR and OHSU system
Fraunhofer Institut Digitale Medientechnologie, Project Group Hearing, Speech and Audio Technology (FRA), Oregon Health & Science University (OHSU)