

CARBINI Sébastien
France Telecom R&D
2 av. P. Marzin - 22307 Lannion - France

Tel. : +33 (0)4 76 01 85 55
+33 (0)6 61 89 45 05
sebastien.carbini@rd.francetelecom.com

25 years – Single

SECOND YEAR PhD STUDENT IN MULTIMODAL INTERFACE

RESEARCH EXPERIENCES

- Since October 2003: PhD Multimodal speech-gesture interface, France Telecom R&D 22300 Lannion, France. Supervisors: J. E. Viallet and P.Y. Coulon.
- March - June 2003: Lip segmentation with active shape and appearance model Laboratoire d'Images et Signaux, Grenoble, France
- April - September 2002: Integration of real time lips analysis softwares (visual C++) Laboratoire d'Images et Signaux, Grenoble, France

PUBLICATIONS

- 1) S. Carbini, J.E. Viallet, O. Bernier, "*Reconnaissance de gestes de pointage dans le cadre d'interaction avec un grand ecran*", CORESA, Lille, France, 24-25 may 2004.
- 2) S. Carbini, J.E. Viallet, O. Bernier, "*Pointing Gesture Visual Recognition for Large Display*", Pointing'04 ICPR Workshop, Cambridge, United Kingdom, 22 August 2004.
- 3) S. Carbini, J.E. Viallet, O. Bernier, "*Pointing Gesture Visual Recognition by Body Feature Detection and Tracking*", ICCVG (International Conference on Computer Vision and Graphics 2004), Warszawa, Poland, 22-24 September 2004. **Best Paper Presentation Award.**
- 4) S. Carbini, J.E. Viallet, O. Bernier, "*Simultaneous Body Parts Statistical Tracking for Bi-Manual Interactions*", ORASIS, Fournol, France, 24-27 may 2005.
- 5) S. Carbini, J.E. Viallet, L. Delphin-Poulat, "*MOWGLI: Multimodal Oral With Gesture Large display Interface*", GW (Gesture Workshop), Berder Island, France, 18-20 may 2005.
- 6) S. Carbini, L. Delphin-Poulat, L. Perron, J.E. Viallet, "*From a Wizard of Oz Experiment to a Real Time Speech and Gesture Multimodal Interface*", (prospective submission) Special issue of Signal Processing (ELSEVIER) on Multimodal Human Computer Interfaces, January 2006.

EDUCATION

- 2002 – 2003: D.E.A. Signal, Image, Speech and Telecommunications in Grenoble, France (equivalent to M. S.)
- 1999 – 2002: Engineer degree in industrial computer science and instrumentation, speciality: image and signal processing, delivered by "Polytech Grenoble" in Saint. Martin d'Hères, France.

SKILLS

- Languages: French, English (TOEIC level 2+), Italian (Spoken, Read).
- Programming: C, C++, Java, Matlab, Linux and Windows environments.
- Areas of Interest: Multimodal Interfaces, Gesture Interfaces, Image processing, Computer science, Speech processing, Neural networks.
- Teaching experiences: C++ and Assembly language at the Computer Science Department, I.U.T. of Lannion, France, (2003-2005).
- Music: Jazz guitarist.

Preferred projects to be work on

1. (P7). A Multimodal (Gesture+Speech) Interface for 3D Model Search and Retrieval Integrated in a Virtual Assembly Application.
2. (P6). Speech Conductor.
3. (P3). Biologically-driven musical instrument.

Wireless connection laptop available

Projects Related Skills

Multimodality

Design and development of a Speech and Gesture Multimodal interface:

- Including multimodal feedback.
- No user specific calibration, learning or background.
- Interaction with large display: user can move freely without wearing any equipment.

Application context dependant fusion of speech and gesture.

Image processing

Body parts detection

Current work:

- Head detection by a Neural Network.
- Hands detection as skin-colour moving zone in front of the head.
- Arm detection as continuous depth zone in front of the head and including the hand.

Previous work:

- Lip segmentation with Active Contour Models and Active Appearance Models. Models parameters calculated by Principal Component Analysis (P.C.A.).

Body parts tracking

Current work:

- Statistical tracking: Head and hands are tracked simultaneously and are modelled by 3d Gaussian functions, exponential functions and colour histograms.
- Model parameters adaptation with Expectation Maximisation (E.M.) algorithm.

Previous work:

- Lucas-Kanade tracker: tracking of key points on the lip edge.

Bimanual continuous gesture recognition

Estimation of pointed location based on head and hands tracking.

Control of a third axis (zoom) or object selection with the second hand.

- Suited both for left handed and right handed person, without prior knowledge.
- Suited for interacting moving user in the field of the camera.

Speech recognition

Vocabulary and syntax definition. Push to talk controlled by gesture.

Selection among N best using recognition score, gesture and application context.

Music

Jazz guitarist (10 years practice).

Harmony theory, solfa, scales, arpeggio.