Wednesday 20 December 2017
10h30
UT3 Paul Sabatier, IRIT, Auditorium J. Herbrand

François-Xavier DECROIX
Team SAMOVA - IRIT

Online learning of audiovisual signatures for people recognition and tracking within a network of ambient sensors

Jury:
Frédéric Lerasle – Professeur, LAAS-CNRS, directeur de thèse
Julien Pinquier – Maître de Conférences, IRIT - Université Paul Sabatier, co-encadrant de thèse
Patrick Lambert – Professeur, LISTIC - Polytech'Savoie, rapporteur
Michel Vacher – Ingénieur de Recherche Hors Classe CNRS, Laboratoire d'Informatique de Grenoble (LIG) - UMR 5217, rapporteur
Serge Miguet – Professeur, LIRIS - Université Lumière Lyon 2, examinateur
Isabelle Ferrané – Maître de Conférences, IRIT - Université Paul Sabatier, examinateur

Abstract: In a Smart City context, scaled at a university campus, reducing the ecological footprint and improving user comfort are two major issues in this transition. The intelligence we want to bring to the campus of the future requires to provide to its buildings a perception of the users and a re-identification process, computed during their movements within the building. Indeed, optimizing the energy resources needs a characterization of their actimetry so that the building can automatically adapt its actuators to it. Human activity being open to multiple levels of interpretation, our work is focused on using audio and visual data to ensure a "real time" tracking of their movements.

Characterizing users activities, in terms of behavior, uses data extracted from cameras and microphones distributed in a room, forming a sparse network of heterogeneous sensors. From these data, we then seek to extract audiovisual signatures of each individual, constituting the basic element of re-identification, and rough localizations of the people transiting through the network. While protecting person privacy, signatures must be discriminative, to distinguish a person from another one, and compact, to optimize computational costs and enables the building to adapt to itself. The objective of this multimodality is the complementarity of the audio and visual components in this ambient intelligence context. Faced with the challenge of combining the sound and visual components of the signature, new source location indices as well as an audiovisual adaptation of a Bayesian multi-target tracking method have been introduced, representing the main contributions of this work.