



Presentation of the funded projects in 2010 for the
« Contents and Interactions » program

Multimedia evaluation challenge: Person identification in
multimedia content (REPERE)

ACRONYM and project title	Page
PERCOL – Person recognition in audiovisual content	2
QCOMPERE – Quaero Consortium for Multimodal Person Recognition	4
SODA – perSon recOgnition in Debate and broAdcast news	6

REPERE – CONTINT Program

2010 edition

Project title	PERCOL – Person recognition in audiovisual content
Abstract	<p>The goal of the PERCOL proposal is to participate to the REPERE evaluation program by building a consortium including partners from the industry (Orange Labs) and academic labs (LIA, LIF, LIFL). The proposal is structured in 3 mains sections:</p> <ol style="list-style-type: none">1) Using existing technology to cover the different modalities and tasks included in the REPERE evaluation program This section is based on the use of modules already existing within the partners of the PERCOL project:<ul style="list-style-type: none">- Speaker segmentation and recognition- Automatic Speech Recognition- Optical Character Recognition of text within video- Named Entity recognition- Visage recognition- Person recognition2) Developing research, within the REPERE framework, allowing combining different scientific fields (audio, text and video) and moving forward the state of the art in each of them for performing the main task of REPERE. The two main scientific directions that are going to be addressed are:<ul style="list-style-type: none">- Multimodal fusion algorithms for automatic person recognition in video broadcast.- Improving information extraction from speech and images thanks to a combine decoding using both modalities to reduce decoding ambiguities.3) Introducing ergonomic analysis in the definition of the task. <p>Thanks to tests with Orange Labs video retrieval platform users, we will work on the metrics that are going to be defined for the REPERE evaluation program.</p>
Partners	<p>Laboratoire d'Informatique Fondamentale de Marseille (LIF) Université d'Avignon et des Pays de Vaucluse (UAPV) Laboratoire d'Informatique Fondamentale de Lille (LIFL) France Télécom</p>

Coordinator	Frédéric Béchet - LIF Frederic.Bechet@lif.univ-mrs.fr
ANR funding	397 800 €
Starting date and duration	November 2010 - 36 months
Cluster label	Images & Réseaux SCS (Solutions Communicantes Sécurisées) Industries du Commerce

REPERE – CONTINT Program

2010 edition

Project title	QCOMPERE – Quaero Consortium for Multimodal Person Recognition
Abstract	<p>The QCOMPERE project aims at building a consortium of QUAERO partners for participating to the REPERE challenge. The QUAERO program, started in May 2008, is composed of French and German research organizations and the goal of the QUAERO CTC project is to improve the state-of-the-art in automatic multimedia document structuring and indexing. Several of its tasks are directly related to the REPERE challenge call, and each of the partners has demonstrated state of the art technologies in their respective areas, making the participation to the REPERE challenge a natural extension to the work already planned for QUAERO. The project will be organized in 6 task, with four tasks addressing a single modal questions: who is seen, who is speaking, whose name is written on screen, whose name is pronounced; a fifth task will address the multimodal fusion and the sixth task will provide technology coordination and annotated data resources for the project. The primary goal of the project is to participate to the REPERE challenge, thus the main deliverables will be submissions to the first and second official REPERE evaluations for each modalities. Other outcomes of the project will be scientific collaboration and publications. The collaboration with the German partner, not funded by the project, will be mainly covered through missions.</p>
Partners	<p>CNRS – Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI) Centre de Recherche INRIA Grenoble Rhône-Alpes Laboratoire d'Informatique de Grenoble (LIG) YACAST Vocapia Research Groupe de Recherche en informatique, image, automatique et instrumentation de Caen (GREYC)</p>
Coordinator	<p>Claude Barras - LIMSI CNRS barras@limsi.fr</p>
ANR funding	400 000 €

Starting date and duration	November 2010 - 36 months
-----------------------------------	---------------------------

Cluster label	-
----------------------	---

REPERE – CONTINT Program

2010 edition

Project title	SODA – perSon recOgnition in Debate and broAdcast news
Abstract	<p>The call for projects REPERE is an evaluation campaign of people recognition technologies for audio-visual French TV shows. The competitors will have to propose systems based on various information sources present in the shows to determine, which will be used to determine who appears in the images, who speaks, which people names are pronounced or appear on the screen, and to which people those names correspond. Addressing all these points will require a mix of various skills: in speaker and face recognition (detection, segmentation, regrouping) to determine people biometric characteristics (voice, face); in speech recognition; in character recognition; and in natural language processing, in order to extract people names and to correctly associate them with people.</p> <p>There are two laboratories involved in this proposal: the Laboratoire d'Informatique de l'Université de Maine (LIUM), which will be the coordinator, and the Swiss research institute IDIAP. The competences of the partners make it possible to cover all the topics of the challenge. LIUM has been developing since 2004 a powerful speaker diarization system (2nd at the ESTER 2005 evaluation campaign, 1st in ESTER 2 campaign in 2008). LIUM has also been working since 2006 on speaker identification using speech transcription to extract the names from the recordings. On the other side, IDIAP has developed since many years competences in automatic processing of audio and video data. Within this project, IDIAP will focus mainly on people detection and recognition, as well as on character recognition (OCR). IDIAP regularly took part in NIST evaluation campaigns. It organized and participated in one of the tasks of the CLEAR (Classification of Events, Activities, and Relationships) evaluation in 2006 and 2007, and is currently in charge of a face and speaker recognition evaluation campaign in the international conference ICPR 2010.</p> <p>The project will benefit from the results previously obtained by the partners, but it will require integration efforts in order to build a complete system to follow the main speakers throughout a show and to name them. Research</p>

will focus on the combination of information coming from the various sources (acoustic signal, images, words and text). In addition, an original aspect of this project will be the recognition of people roles (presenter, journalist or regulator, guest), of their relations and their interactions (for example, who talks to whom), as well as the exploitation of these data in order to improve the performance of speaker/face diarization. For example, the interactions will help linking the speaker to the people on screen; while the roles will be able to facilitate speaker identification by allowing to target regions which contain more useful information.

Partners

Laboratoire d'Informatique de l'Université du Maine (LIUM)
Idiap Research Institute

Coordinator

Sylvain Meignier - LIUM
sylvain.meignier@lium.univ-lemans.fr

ANR funding

432 380 €

**Starting date
and duration**

November 2010 - 36 mois

Cluster label

Images & Réseaux