MUSCLE

Application for a Showcasing Project

Title of the proposed showcase project:

ACADI: Automatic Character (in Audiovisual Document) Indexing

Objective of the proposed showcase project (about 10 lines):

We propose a system which permits to describe and structure audiovisual documents without training, nor corpus knowledge, and to visualize with an interface the principal interventions. It posts the most significant person list of the processed documents (news, TV games, variety programs, film, etc.). A person will be considered as significant if she/he speaks or appears on the screen during a minimum time lapse. We will have to specify this threshold in order to keep only the principal and significant characters. This list is then presented with representative labels of the character (face or/and sound extract for example). Thanks to this person list, it is possible to listen and/or to view all interventions of each character by clicking on the representation of the selected one. The system will be based on the INRIA/Texmex's face detection tool and the IRIT's speaker and costume segmentation tools. The interface allows to visualize (and/or to listen) the only segments where the character of interest appears, without *a priori* knowledge.

Possible industrial application: measure of the audience of each character on a TV program.

Showcase project Leader (name, e-mail and institute):

Julien Pinquier, pinquier@irit.fr - UPS-IRIT

Names of additional institutes involved (not more than 3):

Patrick Gros, Patrick.Gros@irisa.fr - INRIA-Texmex

Description of work within the showcasing project (about 20 lines):

The work to be achieved can be split into three work packages.

WP1: The tool development. This part of the work will lead to format the three existing tools in an exchange format. This WP will aim at the:

- study and choice of the exchange format,
- adaptation of the face detection tool by INRIA/Texmex,
- adaptation of segmentation tools (speakers and costumes) by IRIT,
- choice of an experiment corpus to validate this package.

WP2: The fusion/integration system. This part will be carried out mainly by IRIT. This work is the core of this project: we will have to combine tools results to recover the most important errors of each one. It includes:

- independent tool evaluation (starting point, reference),
- fusion of the descriptors,
- integration of the tools in the system,
- evaluation of the system (with result improvements compared to reference results).

WP3: The Visualization module will be developed by IRIT. We will produce an interface which presents each character by a representative label (face and/or sound extract of the person). This WP consists of:

- choice of different corpora for the demo,
- development of a visualization interface,
- statistics of speaking times and appearance times,
- tests and validation.

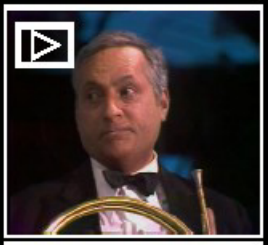
Showcasing results form (Demo / Video / other):

The result is an audiovisual demo which permits to identify the principal characters of any mpeg movie file (TV program, film, etc.). Thus, we listen and/or view the only parts of the content which concern our preferred character. We will also have the statistics over the speaking times and appearance times of each character.

Use case:

- from an audiovisual document, the system identifies the principal speakers and their audio and video occurrences (indexing phase),
- a list of the principal character is presented to the user who could give their names,
- the system indexes the contents according to these names,
- it becomes possible to find each intervention of a given person, without manual labelling, nor *a priori* information (only the names provided by the user at the end of the indexing).

Interface:

Audiovisual document	Variety program: “Le grand Echiquier”		
Principal Characters			
Names	Character1 Edit	Dutronc Edit	Character3 Edit
Speaking times	18 min	25 min	3 min
Appearance times	32 min	23 min	5 min

Character name	Character1
Entire part	
Principal interventions