

## **Indexation de séquences vidéo en temps-réel et transmission d'extraits pertinents relatifs à une problématique donnée.**

Sébastien Lefèvre

Directeurs de thèse : Nicole Vincent, Christian Proust

Ce travail se déroule dans le cadre d'une bourse CIFRE avec la société AtosOrigin

Laboratoire d'Informatique, Université de Tours / E3i

64, Avenue Jean Portalis - 37200 Tours

e-mail : lefevre@univ-tours.fr

La société de l'information se caractérise par de nouveaux usages et de nouveaux supports d'information. En particulier, la vidéo est de plus en plus utilisée. Les destinataires se trouvent donc souvent confrontés au tri de cette information. Il apparaît alors comme nécessaire de fournir des outils d'indexation permettant de ne transmettre que les informations pertinentes. Dans le cadre de la vidéo et pour une problématique donnée, des extraits de séquences vidéo doivent être sélectionnés. La diffusion de ces extraits peut alors s'effectuer via de nouveaux moyens de communication tels qu'Internet, les téléphones mobiles, les assistants personnels, ou encore la télévision interactive. Nous présentons ici l'architecture d'une solution d'indexation de séquences vidéo appliquée à la diffusion des buts inscrits lors de matchs de football. Cette architecture peut être adaptée à d'autres problématiques, telles que la télésurveillance de sécurité ou médicale.

Une fois l'acquisition de la séquence vidéo effectuée, les différents média (son et image) sont analysés séparément. Dans le domaine de l'analyse du son, une séparation de sources peut être réalisée si la bande sonore est diffusée sur plusieurs voies. Il est ensuite possible d'analyser les différentes sources sonores pour détecter certains événements. L'analyse d'image ou de séquence d'images comprend tout d'abord une segmentation temporelle de la vidéo obtenue par détection des changements de plans. Chaque plan ainsi obtenu peut être analysé en utilisant des méthodes de suivi d'objet en disposant ou non de connaissance a priori sur l'objet. Cette détection des objets en avant-plan doit s'accompagner de l'analyse et de la modélisation de la scène en arrière plan en utilisant des approches contour (via par exemple la détection de lignes) ou région (via par exemple une segmentation multi-résolution). Il est de plus intéressant d'effectuer une détection puis une analyse des zones de texte présentes dans les images ou encore de détecter les séquences vidéo représentant un ralenti. Tous les outils décrits précédemment permettent d'obtenir un ensemble de décisions indépendantes quant à la présence d'un extrait pertinent (modification du score) ou non. Une étape de fusion est finalement nécessaire pour obtenir la décision finale. Si la détection de l'événement recherché est obtenue, la dernière étape consiste en la diffusion de l'extrait pertinent de la séquence vidéo. Les destinataires peuvent être soit sédentaires (micro-ordinateur connecté à Internet, télévision interactive), soit nomades (téléphones mobiles, assistants personnels).