

Le CRIM, inauguré en 1985, a été créé pour renforcer les liens entre les universités et les entreprises, petites, moyennes ou grandes, du secteur des technologies de l'information.

À la recherche des sons et des images

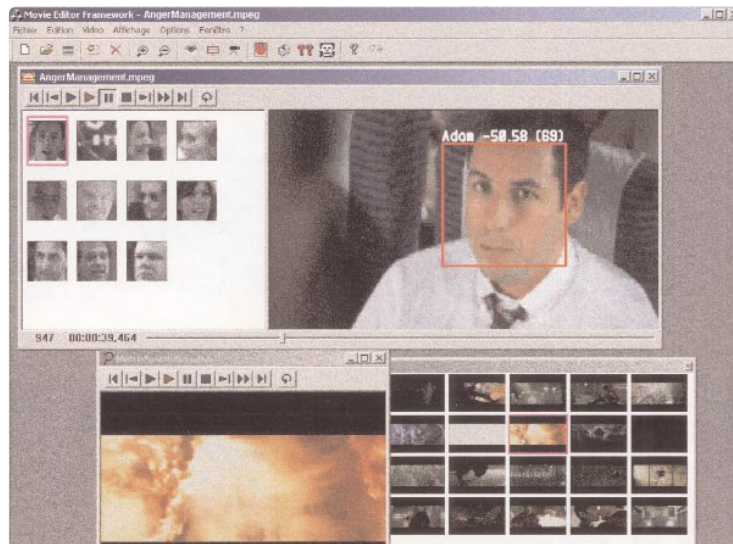


« Pouvez-vous me donner tous les extraits audiovisuels où René Lévesque parle de souveraineté? » Si vous posez aujourd'hui cette question au préposé d'une banque de documentaires, il vous répondra probablement : « Voici la liste des films où figure M. Lévesque. » Et vous en serez quitte pour des heures et des heures de visionnement!

Heureusement, dans quelques années, la réponse sera probablement différente. En effet, il vous suffira d'introduire un mot-clé dans un logiciel et hop!, vous aurez la liste des titres des vidéos pertinentes et l'on vous fournira aussi des repères permettant de visionner seulement les extraits voulus.

La différence entre ces deux réponses réside dans la façon d'indexer les documents. Dans le premier cas, c'est la méthode « ancienne » : on note toutes les métadonnées telles que le titre de la vidéo, le réalisateur, un résumé du contenu, etc. Dans le deuxième, il s'agit de la méthode « futuriste », celle à laquelle travaillent les chercheurs du CRIM, dans le cadre du projet MADIS soutenu financièrement par le réseau canadien CANARIE.

Les spécialistes du CRIM développent actuellement des



Capture d'écran d'un logiciel de traitement vidéo en cours de développement par l'équipe Vision et Imagerie. Il s'agit d'un des outils utilisés dans le cadre du projet MADIS pour l'encodage MPEG-7 du contenu audio visuel.

SOURCE : CRIM

technologies d'encodage du contenu audiovisuel directement tiré de bandes d'images et de sons. Cet encodage, qui se doit d'être entièrement automatique car on traite de grandes quantités d'informations, touche tout élément de la bande vidéo qu'on voudrait repérer plus tard.

« Notre premier défi est de trouver les algorithmes qui permettent de transformer le contenu audio et visuel en informations sémantiques fiables », explique Langis Gagnon, chef de l'équipe Vision et Imagerie au CRIM. Par exemple, dans le cas ci-dessus, il faut associer le nom de René Lévesque à une diversité de visages montrant aussi bien le jeune journaliste

que le politicien fatigué. Autre exemple : il faut que les mots-clés « scène d'hiver » renvoient à un paysage enneigé, et non à un désert africain ou à une image surexposée. Chaque type d'information fait l'objet du développement d'outils d'encodage particuliers : bruits, paroles, formes, visages, etc.

Le second défi des chercheurs consiste à intégrer ensuite tous ces outils dans un système cohérent et fonctionnel. « Tout ce qu'un humain fait spontanément, dit M. Gagnon, nous devons trouver la façon de l'enseigner... à un ordinateur! »

Le résultat de ces encodages? Une banque de données et,

dans le cas du projet MADIS, des documents conçus selon la nouvelle norme internationale d'indexation MPEG-7. Une fois encodée, toute l'information audiovisuelle d'intérêt se retrouve dans un fichier XML, qu'on consulte avec un engin de recherche spécialisé.

En décembre 2003, l'équipe du CRIM devrait avoir en main l'ébauche d'un tel système expérimental, développé à partir d'une banque de films numérisés fournie par l'Office national du film (ONF). La partie « requête » du système sera accessible sous forme d'un démo interactif sur le site Cinéroute de l'ONF [www.cineroute.ca].

« Cette étape vise à établir la faisabilité d'une indexation de l'image et du son pour une application concrète, précise Langis Gagnon. Si tout va bien, nous travaillerons ensuite à mettre au point un prototype commercialisable. »

Pour de telles technologies, le marché est d'importance. Les maisons de production, comme l'ONF, s'y intéressent de près car elles cherchent toujours à améliorer la disponibilité de leur matériel. Les banques de photos, les centres d'archives et toute institution possédant des bandes vidéo en grande quantité pourraient également bénéficier de cette découverte.