


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.

## Des images qui en disent long

 Vous souvenez-vous du film culte *Le fabuleux destin d'Amélie Poulain*? Et de la scène où Nino de Quincampoix suit les flèches bleues sur la butte de Montmartre pour récupérer son album photo? Scène mémorable, mais qui ne peut être appréciée pleinement par un non-voyant, à moins qu'on ne la lui décrive! Fort heureusement, de plus en plus de films offrent aujourd'hui un outil précieux pour les personnes aveugles : la vidéo description.

À l'instar du sous-titrage, qui fournit des informations visuelles pour les malentendants, la vidéo description donne des informations verbales sur l'action, les personnages ou les lieux clés d'un film, pendant les moments de silence. La description narrative facilite ainsi la visualisation mentale d'une scène. Toutefois, le travail de postproduction nécessaire est long et coûteux. « Pour chaque heure de film, la vidéo description nécessite plus de 20 heures de travail et coûte de 2 500 à 6 000 \$ US, explique **Langis Gagnon**, chercheur et responsable du thème « Interaction et extraction de contenu audiovisuel » au Centre de recherche informatique de Montréal (CRIM). Il s'agit d'un travail manuel : on visionne de nombreuses fois un film pour en détecter le contenu visuel, ce qui est fort ardu. »

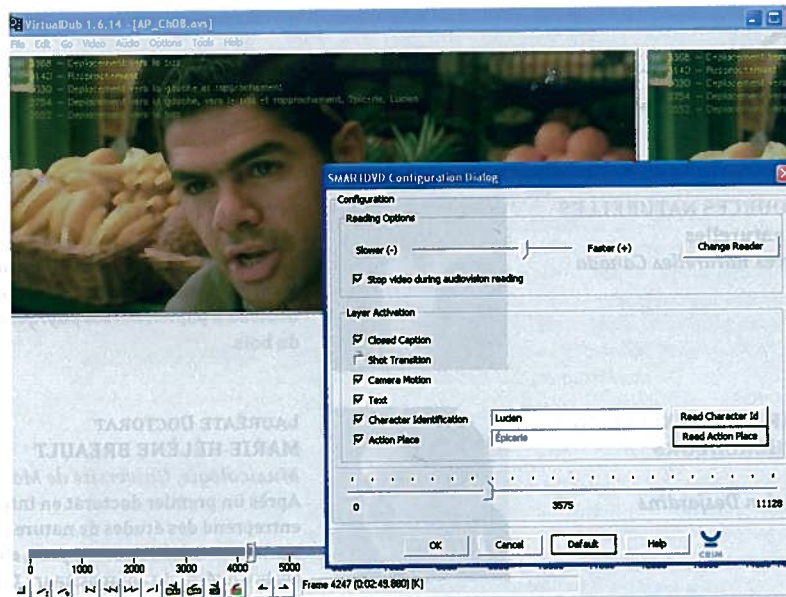
Pour faciliter le processus, l'équipe de Langis Gagnon tente de développer un logiciel qui assisterait les producteurs de vidéo description. Ce travail est fait en collaboration avec les professeurs Denis Laurendeau,

autres éléments visuels utiles à la compréhension d'un film. « Un visage, par exemple, se reconnaît par certaines caractéristiques faciales telles que sa couleur, sa texture et sa forme, explique M. Gagnon. Ces infor-

Les éléments relevés dépendent du film, des besoins des producteurs et des utilisateurs. C'est pourquoi l'équipe consacre une partie de son projet à consulter des producteurs de vidéo description pour trouver

les scénarios les plus utiles et à interviewer des personnes aveugles pour comprendre leur expérience d'écoute. Ces dernières n'expriment pas toutes les mêmes besoins. Certaines préfèrent avoir une description verbale détaillée alors que d'autres ne désirent obtenir des informations que sur les actions ou les personnages. En ce sens, les futurs DVD pourraient contenir plusieurs niveaux de vidéo descriptions.

Bien qu'il ne soit pas encore complètement au point, le prototype est déjà en mesure d'extraire de nombreux éléments visuels et de les décrire à l'aide d'une voix synthétisée. D'ici 2009, l'équipe vise la création d'un logiciel précommercial plus robuste et capable d'extraire plus de contenu visuel.



Le logiciel prototype permet l'extraction de certains contenus visuels et la génération de descriptions textuelles. Ces dernières peuvent être lues par une voix synthétique à l'aide du logiciel « frère », montré ici, ou simplement servir pour un résumé vidéo grossier lors de la production de la vidéo description.

du Laboratoire de vision et système numérique de l'Université Laval, et James Turner, de l'École de bibliothéconomie et des sciences de l'information de l'Université de Montréal. L'ordinateur serait en mesure de repérer et de répertorier automatiquement les personnages principaux, la couleur dominante d'une scène, la segmentation de scénarios et certains

mations sont « cachées » dans les pixels d'une image et représentent une véritable signature mathématique. Étant donné qu'on trouve plusieurs milliers de pixels par image, il s'agit de faire des liens entre toutes les informations et d'en extraire les plus pertinentes. »

SOURCE : CRIM