Chapitre 8

Systèmes interactifs pour la construction et l'utilisation de la vidéo hyperliée

Avoir une bonne idée ne suffit pas !!!
Le maître du jeu, au final, c'est
toujours le client.

Après la présentation des méthodes de fabrication de la structure hiérarchique de la vidéo de bas niveau (i.e. plans, images-clés) et de haut niveau (i.e. groupes d'objets, scènes, ...), ce chapitre a pour objectif de résumer les niveaux d'interactions entre *Homme* (utilisateur), *Machine* et *Vidéo* que nous avons défini et de les intégrer dans nos systèmes interactifs de construction et d'utilisation de la vidéo hyperliée, suite au travail initial conduit par Pascal Bertolino.

8.1 Interactions dans un système de vidéothèque hyperliée

Par analogie avec le standard *MPEG-4* [112] on distingue deux types d'utilisateurs de la vidéo hyperliée : *auteur* et *utilisateur final*. La figure 8.1 illustre les interactions entre Utilisateur/Vidéo, Utilisateur/Machine et Machine/Vidéo dans un système de vidéothèque hyperliée. D'abord, l'utilisateur auteur spécifie un film vidéo à traiter, il fixe les paramètres des méthodes d'indexations automatiques et/ou semi-automatiques à appliquer sur ce film. Ensuite, la machine exécute ces modules et produit une "vidéo hyperliée" qui sera utilisée manuellement par l'utilisateur. A un deuxième niveau d'interaction l'utilisateur final joue la vidéo hyperliée.

La liste suivante résume les tâches de l'auteur de la vidéo hyperliée :

- *Paramétrage* des méthodes d'indexation et de structuration de la vidéo présentées auparavant. A titre d'exemple, lors de la fabrication semi-automatique du niveau "groupe d'objets" l'auteur désigne les "objets suivis modèles" (voir figure 4.1), ensuite tout autre objet de la séquence vidéo traitée sera classé automatiquement dans l'un d'eux.

- *Édition* des résultats obtenus par les méthodes d'indexation automatique comme nous l'avons précisée dans les chapitres précédents. A titre d'exemple, il corrige le mauvais classement d'un objet suivi par une opération de genre *Drag* & *Drop* par la souris.

- *Description* textuelle des entités de la vidéo structurée. L'utilisateur auteur intervient pour décrire la sémantique ou l'histoire d'une scène ou d'une classe d'objets du film vidéo.

- *Pointage* d'un objet du film vers un lien extérieur comme l'adresse d'une page web ou l'exécutable d'un programme spécifique (clip de musique, modèle 3D de l'objet d'intérêt, ...).

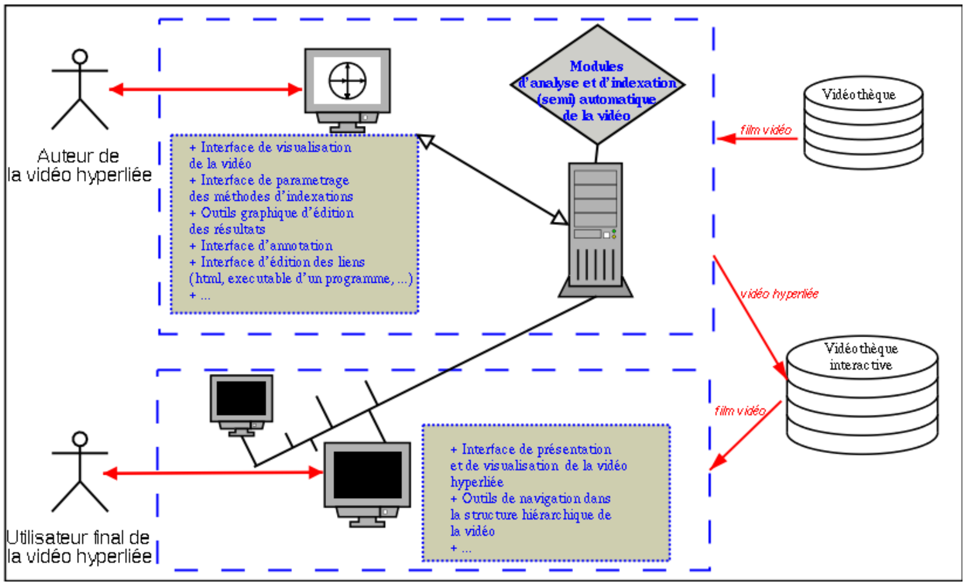

FIG. 8.1: *Nivaux d'interactions entre Utilisateur/Machine/Vidéo dans un système de vidéothèque hyperliée.*

Afin de faciliter les tâches ci-dessus des interfaces d'acquisition des entrées, de visualisation et de présentation des résultats d'indexation ainsi que des outils graphiques simples pour éditer manuellement ces résultats sont indispensables. Aussi, la vidéo hyperliée ainsi construite est présentée à l'utilisateur final via des interfaces graphiques simples et intuitives. C'est le niveau d'interaction Utilisateur/Machine. Les interfaces sont illustrées dans les pages suivantes.

8.2 Résumé de "Interactive Tools for Constructing and Browsing Structures for Movie Films" – ACM'2000

Selon le schéma de la vidéothèque hyperliée que nous venons de présenter, deux interfaces graphiques nommées `VideoPrep` et `VideoClic` ont été développées. La première est conçue pour préparer une vidéo interactive tandis que la seconde est destinée aux utilisateurs finaux pour consulter et naviguer dans la vidéo préparée.

Une première version de ces interfaces a été achevée vers la fin du contrat de collaboration entre Alcatel CRC et l'INRIA (équipes VISTA et MOVI) en septembre 1998, par l'ingénieur expert P. Bertolino. Bertolino a été recruté dans le projet MOVI sur ce contrat [8]. Ma participation à cette version a porté principalement sur l'intégration des modules d'indexation par la couleurs, la conception et le développement de leurs interfaces graphiques. L'indexation des objets est effectuée sur la base des images-médianes représentant les plans comme une solution simple des problèmes abordés dans cette thèse (voir chapitre 3).

Dans le cadre du notre contrat [1] nous avons repris le développement de `VideoPrep` et `VideoClic` en apportant des améliorations pertinentes sur leurs fonctionnements et leurs structures d'indexation, et en intégrant les approches de structuration discutées dans les chapitres précédents de ce mémoire [35]. Une démonstration de la deuxième version de ces systèmes a été présentée à ACM Multimédia [56]. Les détails sont décrits dans les pages suivantes.

Le défaut technique de `VideoPrep` est principalement son traitement des séquences vidéo décompressées en plusieurs formats d'images individuelles (PPM, PGM, GIF) (voir section 2.5). Ceci conduit à l'explosion de l'espace mémoire réquis même pour une séquence vidéo de quelques minutes. Il faudrait repenser à toute l'architecture dans un travail ultérieur; l'utilisation d'un espace disque cache permet vraisemblablement de résoudre ce problème par décompression en temps réel de la bande vidéo MPEG. En ce qui concerne `VideoClic` le téléchargement de toutes les structures de la vidéo hyperliée (plans, classes d'objets, scènes). Pour une large séquence vidéo ceci conduit également à une occupation mémoire trop large. Là encore, un astucieux chargement à la demande permettrait de résoudre ce problème; il faudra juste veiller à ce que ce processus ne ralentisse pas la navigation dans la vidéo.

Malgré leurs défauts techniques nos systèmes sont complets du point de vue recouvrement de la totalité de la structure hiérarchique de la vidéo, automatisation du processus de structuration, souplesse de l'interaction entre l'utilisateur, la vidéo et la machine lors de la création et la présentation de la vidéo hyperliée. Celles-ci les rendent comparables à plusieurs produits concurrents dans le marché comme Mvshots et MoVideo [2] (voir page 13).

1. Contrat sur trois ans entre Alcatel CRC et MOVI démarré le premier Janvier 1998, numéro 198098G
2. http://www.artsvideo.com