

III.4.1 Vidéo algébrique

On observe actuellement un intérêt croissant pour l'utilisation de la vidéo et du son dans des applications informatiques. Mais, malgré l'apparition du support matériel et système qui permet de présenter la vidéo et le son sur une station de travail, peu d'applications bénéficient de ce nouveau moyen d'expression. En effet, ces données multimédia n'ont pas encore été intégrées comme un type de données nouveau qui peut être facilement manipulé par des applications. Ceci est d'autant plus important que la nature de données multimédia est très différente de celle des données traditionnelles comme le texte ou les données numériques. La différence réside essentiellement dans la dimension temporelle et spatiale de la vidéo. Le caractère temporel est souvent pris en compte de façon traditionnelle où la vidéo est vue comme un médium linéaire – une séquence d'images animées. Mais, la vidéo possède une structure logique dont les éléments comme la prise, la scène, la séquence sont hiérarchiquement liés. Les représentations existantes ne permettent pas de représenter cette structure de données. L'intégration de la vidéo est aussi difficile à cause de son volume important : la réutilisation des données et les moyens d'accès associatif sont essentielles pour supporter des applications voulant utiliser la vidéo. Le but de ce travail a donc été de proposer un nouveau type de données pour la vidéo qui permet :

- de modéliser la structure imbriquée de la vidéo,
- d'exprimer la composition temporelle des segments de vidéo,
- d'associer de l'information sur le contenu aux segments de vidéo,
- de faciliter la coexistence de vues multiples et des annotations des mêmes données,
- de supporter l'accès associatif selon le contenu, la structure et les relations temporelles,
- de définir les caractéristiques de présentations multi-fenêtres.

Pour illustrer les applications possibles de ce nouveau type de données, prenons l'exemple de un utilisateur qui voudrait composer une présentation sur les dernières nouvelles à propos des réformes économiques. D'abord, il fait une recherche dans une collection de segments vidéo provenant de différentes émissions télévisées. Ensuite, il veut examiner le contexte dans lequel les segments trouvés apparaissent : dans un journal télévisé ou dans le *Bebête show*. Enfin, l'utilisateur compose les segments choisis pour former une présentation qui peut être visualisée, stockée ou envoyée sur le réseau. Un autre exemple est celui de quelqu'un qui voudrait voir tous les matins un résumé d'événements importants du jour précédent. L'utilisateur spécifie ses préférences, par exemple en demandant que les informations sportives soient présentées d'abord dans une grande fenêtre en parallèle avec la carte actuelle de la météo et après, les informations politiques et économiques en parallèle avec les

cours de la bourse. Une fois ce squelette défini, l'ordinateur construit tous les jours une présentation à partir d'une collection de la vidéo enregistrée.

Pour pouvoir représenter du multimédia, on a besoin d'un modèle de données qui à la fois représente bien la structure, la composition et les relations temporelles entre des composants et qui a une puissance d'expression adéquate. Actuellement, les données multimédia sont souvent représentées et stockées dans des formats propriétaires ce qui ne facilite pas leur échanges et ne permet pas de référencer différents documents. Pour adresser ces problèmes, le standard *HyTime* (un sous ensemble de *SGML*) a été proposé comme un moyen de spécification du format, de la structure et de l'adressage des documents multimédia [Newcomb 91]. Il définit un méta-langage pour la définition et l'adressage des documents. Le contenu de différents média est laissé non spécifié et peut être encodé en utilisant n'importe quel format ; en particulier, l'encodage d'un document peut se faire dans un autre standard – MHEG [Price 93]. Il définit un modèle de composition orienté objet destiné à la distribution finale des documents avec des contraintes temporelles. D'autres formats existants incluent *OpenDoc* d'Apple, *ScriptX* de *Kaleida Labs* et *Open Media Framework* d'*Avid* [Lu 93].

L'aspect temporel dans la composition des documents multimédia commence à être pris en compte par des formateurs temporels qui permettent de spécifier et de valider des relations temporelles complexes entre des éléments multimédia. Les éditeurs de documents multimédia comme *CMIFed* [Rossum 93] ou comme le système décrit dans [Hamakawa 93] offrent à l'utilisateur une gamme d'opérations qui permettent de composer et de structurer des présentations multimédia. Quant à la vidéo, les outils existants comme *Adobe Premiere* [Adobe 91], *AVID VideoShop* [AVID 91] et *MacroMind Director* [MacroMind 90] sont basés sur le même paradigme que celui utilisé pour le montage des films: des objets représentant des scènes vidéo ou des séquences sonores sont arrangés sur un axe de temps et liés par des transitions. Ces outils ne permettent pas d'exprimer la structure complexe de données vidéo, leur relations temporelles ni d'associer des descriptions nécessaires au support d'accès associatif. Certains systèmes expérimentaux comme *Media Streams* [Davis 92] et *Electronic Scrapbook* [Bruckman 91] facilitent l'association des descriptions à la vidéo et permettent de l'annoter et d'y faire des recherches. Aussi, un système d'annotation de la vidéo basé sur le concept de la stratification a été développé [Aguierre Smith 91], [Aguierre Smith 92]. Gibbs *et al.* ont proposé une approche orientée objet à la conception et au développement de bases de données vidéo [Gibbs 93].

Le volume important de données, leur complexité et l'aspect temporel de la présentation multimédia font que des mécanismes de recherche, de navigation et de composition sont de première importance. La recherche des séquences d'intérêt sans outils adéquats peut durer trop longtemps parce qu'il est nécessaire de suivre linéairement toute une présentation. Avant de pouvoir accéder de manière associative

au multimédia, il faut d'abord pouvoir identifier son contenu, le représenter et l'indexer. Contrairement aux données textuelles ou aux données numériques, l'identification du contenu des données multimédia reste encore très difficile et la plupart des systèmes existant utilisent les descriptions textuelles [Little 93]. La détection et l'indexation des traits caractéristiques d'images commencent à apparaître dans des bases de données d'images. Par exemple, le système *QBIC* d'*IBM Almaden* permet de rechercher des images selon leur couleurs, textures ou contours [Nilblack 93]. La vidéo pose encore plus de problèmes parce qu'il faut identifier le contenu des séquences d'images. Très peu de travaux ont été consacrés à l'indexation de la structure et des relations temporelles dans des documents multimédia, et encore moins à l'identification du contenu au vol pour pouvoir les distribuer aux utilisateurs intéressés.

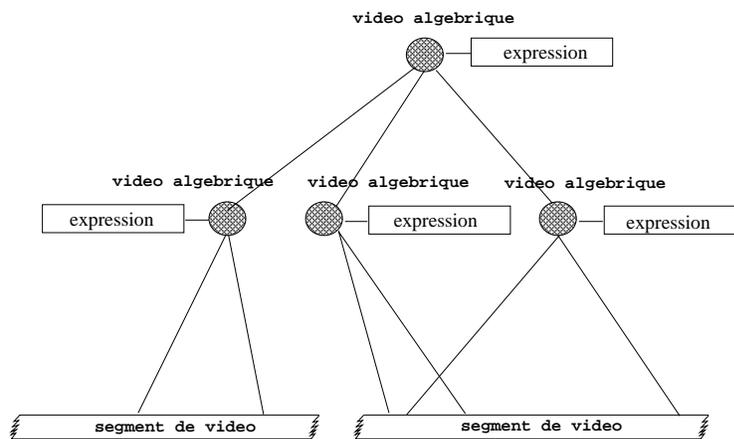


Fig. 3.8 : Vidéo algébrique

Nous avons conçu et développé un modèle appelé *vidéo algébrique* qui consiste en une composition hiérarchique d'*expressions vidéo* (cf. Fig. 3.8) [Weiss 94]. Ce nouveau type de données permet la structuration, la composition et la recherche des séquences vidéo. Il intègre de l'information sur le contenu, la structure, l'ordonnement temporel et la présentation sur l'écran avec des séquences vidéo. Une expression vidéo désigne une *présentation vidéo* qui est une combinaison spatiale, temporelle et concurrente des segments vidéo. Les expressions peuvent être composées à l'aide des opérateurs. Nous avons défini un *algèbre d'opérations* qui permettent de construire des expressions à partir de segments de vidéo, de les composer temporellement, de spécifier les caractéristiques d'une présentation et d'associer des descriptions du contenu.

Par exemple, nous définissons une opération d'union (désigné par \cup) qui permet de composer deux expressions vidéo séquentiellement l'une après l'autre de

telle sorte que leur parties communes ne sont pas répétées. La Fig. 3.9 présente comment l'expression vidéo ci-dessous est interprétée (l'opérateur **create** crée une expression vidéo à partir d'une séquence de vidéo désignée par son début et la fin ; l'opérateur **description** associe à une expression une description en forme d'attribut *nom = valeur*). Pour la définition des autres opérateurs on peut consulter [Weiss 94].

```

C1 = create CNN.HeadlineNews.rv 10 30
C2 = create Cnn.HeadlineNews.rv 20 40
C3 = create Cnn.HeadlineNews.rv 32 65

(description
 (C1  $\cup$  C2)  $\cup$  C3
 (title = CNN Headline News
  text = Smith proposes economic reform...))

```

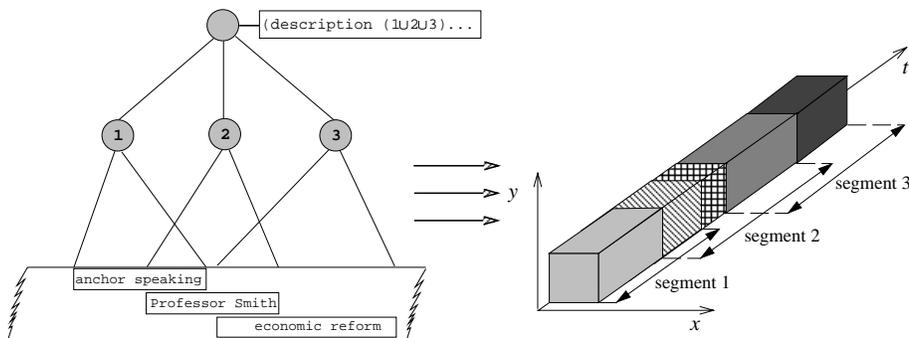


Fig. 3.9 : Représentation graphique de l'union.

Les *opérations d'interface* définissent les méthodes d'accès à une collection de la vidéo algébrique. On peut effectuer une *recherche* en spécifiant une requête, on peut *jouer* une présentation et on peut *naviguer* dans la collection en explorant la structure hiérarchique des expressions vidéo.

La vidéo algébrique a été prototypée en TCL [Ousterhout 91] en utilisant le support de la vidéo numérique du projet *ViewStation* [Tennenhouse 94], le support d'indexation du *Système de Fichiers Sémantiques* [Gifford 91b], le support d'accès à distance du *WWW* et l'interface utilisateur de *Mosaic*. En accédant à distance à travers le *WWW*, l'utilisateur peut composer des présentations multi-fenêtres en utilisant une base de séquences vidéo. Une présentation peut être stockée, visualisée et réutilisée dans une autre présentation. On peut rechercher des séquences en spécifiant une requête et naviguer dans la base en suivant des liens vers des segments référencés dans un autre segment. Ce prototype est accessible à partir de *Mosaic* en ouvrant le document [URLb].

III.5 Conclusion

L'analyse du phénomène des autoroutes de l'information nous porte à croire que la mise en place des réseaux rapides changera notre vision des systèmes répartis. Ces systèmes devront prendre en compte le facteur d'échelle et l'apparition des données multimédia. Un support de transmission à haut débit est une condition nécessaire au développement de futurs réseaux. Mais il est aussi important d'avancer sur le problème du développement de l'infrastructure de l'information qui comporte tous les services et outils nécessaires à la construction des applications nouvelles coopérantes et concourantes. À la suite de notre travail sur l'accès à l'information répartie et la vidéo algébrique, nous proposons dans le chapitre suivant une action qui va dans ce sens.

III.6 Bibliographie

- [Adobe 91] *Adobe Premiere User Guide*, Adobe Systems Incorporated, Mountain View, CA, 1991.
- [Aguierre Smith 91] Aguierre Smith, T.G. et N.C. Pincever, ‘‘Parsing Movies in Context’’, *Proc Summer 1991 Usenix Conference.*, pp. 157—168, Nashville, Tennessee, Jun 1991.
- [Aguierre Smith 92] T.G. Aguierre Smith et G. Davenport, ‘‘The Stratification System: A Design Environment for Random Access Video’’, *Proc. 3rd International Workshop on Network and Operating System Support for Digital Audio and Video*, La Jolla, CA, novembre 1992.
- [Alberti 92] F. Alberti et al., ‘‘The Internet Gopher Protocol: A Distributed Document Search and Retrieval Protocol’’, *University of Minnesota Microcomputer and Workstation Networks Center*, printemps 1992.
- [Andreessen 93] M. Andreessen, *NCSA Mosaic Technical Summary*, NCSA, mai 1993.
- [AVID 91] *AVID VideoShop*, AVID Corporation, Cambridge, MA, 1991.
- [Barbara 92] D. Barbara et C. Clifton, *Information Brokers: Sharing Knowledge in a Heterogeneous Distributed System*, (MITL-TR-31-92), Matsushita Information Technology Laboratory, Princeton, NJ, octobre 1992.
- [Berners-Lee 92] Berners-Lee et al., T., ‘‘World-Wide Web: The Information Universe’’, *Electronic Networking*, 2(1), pp. 52—58, 1992.
- [Berners-Lee 93] Berners-Lee, T., *Hypertext Transfer Protocol*, Internet Draft, novembre 1993.
- [Bruckman 91] Bruckman, A.S, *Electronic Scrapbook: Towards an Intelligent Home-Video Editing System*, Masters Thesis, Massachusetts Institute of Technology, septembre 1991.
- [Danzig 92] Danzig, P. B et al., ‘‘Distributed Indexing of Autonomous Internet Services’’, *Computing Systems*, 5(4), pp. 433—459, 1992.
- [Clark 92] Clark, D. et al., ‘‘The Aurora Gigabit Testbed’’, *Computer Networks and ISDN Systems*, pp. 599—621, février 1992.
- [Davis 92] M. Davis, ‘‘Media Streams: An Iconic Visual Language for Video Annotation’’, *Proc. IEEE Symposium on Visual Languages*, pp. 196—202, Bergen, Norway, 1993.
- [Dertouzos 91a] Dertouzos, M.L., ‘‘Building the Information Marketplace’’, *Technology Review*, janvier 1991.

- [Dertouzos 91b] Dertouzos, M.L., ‘‘Communications, Computers and Networks’’, *Scientific American*, septembre 1991.
- [Duda 94] Duda, A. et Sheldon, M.A., ‘‘Content Routing in Networks of WAIS Servers’’, *Proc. 14th IEEE Int. Conference on Distributed Computing Systems*, Poznan, Pologne, juin 1994.
- [Emtage 92] A. Emtage et P. Deutsch, ‘‘Archie – An Electronic Directory Service for the Internet’’, *USENIX Association Winter Conference Proceedings*, pp. 93—110, San Francisco, janvier 1992.
- [Etzioni 94] Etzioni, O. et Weld, D., ‘‘A Softbot–Based Interface to the Internet’’, *CACM*, 37(7), juillet 1994.
- [Gibbs 93] S. Gibbs, C. Breiteneder et D. Tschritzis, ‘‘Audio/Video Databases: An Object–Oriented Approach’’, *Proc. 9th IEEE Int. Data Engineering Conference*, pp. 381—390, 1993.
- [Gigabit 91] Special Report on Gigabit Networks, *IEEE Computer*, septembre 1991.
- [Gigabit 92a] Special Issue on Gigabit Networks, *IEEE Network*, mars 1992.
- [Gigabit 92b] Special Issue on Gigabit Networks, *IEEE Communications*, avril 1992.
- [Gravano 93] L. Gravano et al, *The Efficacy of GLOSS for the Text Database Discovery Problem*, (STAN–CS–TR–93–2), Stanford University, octobre 1993.
- [Hahn 94] H. Hahn et R. Stout, *The Internet Complete Reference*, Osborne McGraw–Hill, Berkeley, California, 1994.
- [Hamakawa 93] R. Hamakawa et J. Rekimoto, ‘‘Object Composition and Playback Models for Handling Multimedia Data’’, *Proc. First ACM International Conference on Multimedia.*, pp. 273—281, Anaheim, CA, Aug 1993.
- [Houh 94] Houh, H. et al., ‘‘Actives Pages: Intelligent Nodes on the World Wide Web’’, *Proc. First Int. WWW Conference*, pp. 265–270, Genève, mai 1994.
- [ARPA 94] *Intelligent Query Access to the National Information Infrastructure*, contrat ARPA, MIT, mai 1994.
- [I–95 93] *I–95: The Information Market*, Proposition de contrat, MIT, 1993.
- [Internet 94] Special Issue on Internet Technology, *Communications of the ACM*, août 1994.
- [Kahle 91] Kahle, B. et Medlar, A., *An Information System for Corporate Users: Wide Area Information Servers*, (TMC–199), Thinking Machines, Inc., avril 1991.

- [Little 93] Little, T.D.C et al., ‘‘A Digital On–Demand Video Service Supporting Content–Based Queries’’, *Proc. First ACM International Conference on Multimedia.*, pp. 427—436, Anaheim, California, Aug 1993.
- [Lu 93] Lu, C., ‘‘Publish It Electronically’’, *Byte*, pp. 94—110, septembre 1993.
- [MacroMind 90] *Director Version 2.0*, MacroMind, 1990.
- [Newcomb 91] S.R. Newcomb et al., ‘‘HyTime, Hypermedia/Time–based Document Structuring Language’’, *CACM*, 34(11), novembre 1991.
- [Neece 94] J. Neece, *Network Strategies at SUN*, séminaire IMAG, avril 1994.
- [Ordille 93] L. Ordille et B. Miller, ‘‘Distributed Active Catalogs and Meta–Data Caching in Descriptive Name Services’’, *Proc. 13th Int. Conference on Distributed Computing Systems*, pp. 120—129, 1993.
- [Ousterhout 91] Ousterhout, J.K., ‘‘An X11 Toolkit Based on the TCL Language’’, *USENIX Association 1991 Winter Conference Proceedings*, pp. 105—115, Dallas, TX, Jan 1991.
- [Niblack 93] W. Niblack et al., *The QBIC Project: Querying Images by Content Using Color, Texture and Shape*, (RJ 9203 (75655)), IBM Almaden Research Center, San Jose, février 1993.
- [NII] *The National Information Infrastructure: Agenda for Action*, <ftp://ftp.ntia.doc.gov/niiagenda.asc>.
- [Ousterhout 91] J.K. Ousterhout, ‘‘An X11 Toolkit Based on the Tcl Language’’, *Proc. USENIX Association 1991 Winter Conference*, pp. 105—115, janvier 1991.
- [Price 93] R. Price, ‘‘MHEG: An Introduction to the Future International Standard for Hypermedia Object Interchange’’, *Proc. First ACM Int. Conference on Multimedia*, pp. 121—128, Anaheim, août 1993.
- [Reinhardt 94] A. Reinhardt, ‘‘Building the Data Highway’’, *Byte*, pp. 46–74, mars 1994.
- [Rooholamini 94] R. Rooholamini et al., ‘‘Finding the Right ATM Switch for the Market’’, *IEEE Computer*, pp. 16–29, avril 1994.
- [Rossum 93] G. van Rossum et al., ‘‘CMIFed: A Presentation Environment for Portable Hypermedia Documents’’, *Proc. First ACM Int. Conference on Multimedia*, pp. 183—188, Anaheim, août 1993.
- [Sheldon 94] M.A. Sheldon, A. Duda, R. Weiss, J.W. O’Toole et D.K. Gifford, ‘‘A Content Routing System for Distributed Information Servers’’, *Proc. Fourth International Conference on Extending Database Technology*, Cambridge, England, mars 1994.

- [Simpson 89] P. Simpson et R. Alonso, *Querying a Network of Autonomous Databases*, (CS-TR-202-89), Princeton University, Princeton, NJ, janvier 1989.
- [Stein 91] R. M. Stein, ‘‘Browsing Through Terabytes: Wide-area Information Servers Open a New Frontier in Personal and Corporate Information Services’’, *Byte*, pp. 157—164, mai 1991.
- [Tennenhouse 94] D.K. Tennenhouse et al., ‘‘A Software-Oriented Approach to the Design of Media Processing Environments’’, *Proc. IEEE Int. Conference on Multimedia Computing and Systems*, pp. 435–444, Boston, MA, mai 1994.
- [URLa] URL, <http://www-psrg.lcs.mit.edu/content-router.html>.
- [URLb] URL, <http://www-psrg.lcs.mit.edu/Projects/AVS/intro.html>.
- [URLc] URL, <http://www.stir.ac.uk/jsbin/js>.
- [Verjus 94] J-P. Verjus, *Les autoroutes de l’information*, note IMAG, mai 1994.
- [Weiss 94] R. Weiss, A. Duda et D.K. Gifford, ‘‘Content-Based Access to Algebraic Video’’, *Proc. IEEE Int. Conference on Multimedia Computing and Systems*, pp. 140–151, Boston, MA, mai 1994.
- [Z39.50 88] *Information Retrieval Service Definition and Protocol Specification for Library Applications*, ANSI/NISO Standard Z39.50, janvier 1988.

Chapitre III

Autoroutes de l'information

III.1 Introduction	95
III.2 Autoroutes de l'information	95
III.3 Recherche de l'information sur réseaux	99
III.3.1 Routage par contenu	100
III.4 Multimédia réparti	108
III.4.1 Vidéo algébrique	109
III.5 Conclusion	113
III.6 Bibliographie	114