

LA RECHERCHE D'INFORMATION DANS LES ARCHIVES DU TÉLÉJOURNAL À LA TSR

LE PROJET COALA

NASTARAN.FATEMI@epfl.ch, EPFL/I&C/LITH,
OMAR.ABOUKHALED@eif.ch, EIA-FR/TIC/MISL
& GIOVANNI.CORAY@epfl.ch, EPFL/I&C/LITH



LE PROJET COALA

Le projet COALA (Content Oriented Audiovisual Library Access) concerne la recherche d'informations audiovisuelles. Il s'est déroulé en étroite collaboration entre le laboratoire d'informatique théorique (LITH) de l'EPFL, le laboratoire des systèmes d'informations mobiles (MISL) de l'EIA - FR et la Télévision Suisse Romande (TSR) et a débouché sur la réalisation d'une *plate-forme prototype* d'indexation et d'annotations du téléjournal de la TSR.

Le projet vise à concevoir et à mettre en œuvre un système de librairie audiovisuelle pour différents utilisateurs, tels que journalistes, documentalistes, producteurs, voire des chercheurs.

Les résultats de ce projet couronnent un travail de thèse récemment achevé au sein du LITH intitulé: **A Semantic Views Model for Audiovisual Indexing and Retrieval**. Cette thèse propose un modèle innovateur de recherche d'information audiovisuelle basé sur des vues sémantiques (Semantic Views Model) permettant de répondre au mieux aux besoins des utilisateurs d'un environnement de production de programmes télévisés (tels que journaliste, éditeur, archiviste, etc.). Ce modèle sémantique utilise un format de description interopérable (TVN Schema) défini dans le but d'intégrer les différentes ressources de description de programmes télévisés pour former un corpus enrichi. Il tient compte des besoins des utilisateurs et de leurs points de vue, d'où le nom de **Semantic Views Model**.

LA RECHERCHE D'INFORMATION AUDIOVISUELLE

La recherche d'information audiovisuelle devient aujourd'hui une nécessité absolue pour un nombre croissant d'utilisateurs qui gèrent des informations audio/vidéo diverses dans leurs activités. Beaucoup de travaux de recherche, émanant d'institutions académiques et industrielles, ont focalisé leurs efforts plutôt sur les aspects technologiques comme l'analyse et l'extraction automatique de l'information audiovisuelle et il existe très peu de travaux qui tiennent compte des vrais besoins des utilisateurs.

L'ÉTUDE DE LA TSR

Une étude récente que nous avons réalisée au sein de l'environnement de production de journaux télévisés de la TSR (Télévision Suisse Romande) a révélé que les systèmes actuels de recherche d'information audiovisuelle ne tiennent pas compte des besoins variés des utilisateurs dans leurs activités professionnelles et qu'ils ne sont pas adaptés pour

s'intégrer facilement avec les autres processus et les outils de travail de production télévisée.

L'utilisation principale du système d'archivage du téléjournal est le processus de la recherche dans les téléjournaux pour des futures productions (reportages, magazines, etc.). La production, l'archivage et la recherche dans les journaux télévisés constituent trois processus fortement interconnectés et complémentaires comme le montre la figure 1.

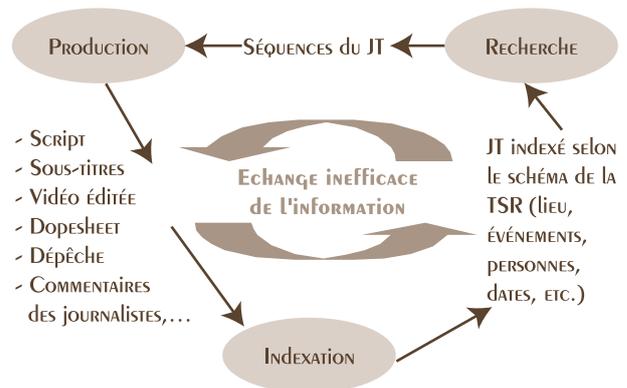


fig. 1 – LES TROIS PROCESSUS DE L'ENVIRONNEMENT DE PRODUCTION DES JOURNAUX TÉLÉVISÉS

Le problème majeur qui a été constaté est le manque d'intégration de ses trois processus et l'absence d'un format d'échange entre les différentes sources de l'information comme les scripts, les dépêches, les commentaires de journaliste et les *dopesheet*¹, etc (fig. 2).

Les autres problèmes sont liés à l'inadéquation des outils de recherche en vigueur, la complexité de la nature des requêtes utilisées, la richesse du contenu recherché (contenu audio, visuel, thématique, etc.), la diversité des utilisateurs avec leurs besoins et leurs différents points de vue. De plus, le langage de requête actuel est rudimentaire et ne permet qu'une expressivité limitée. Finalement, la vérification des résultats reste un travail très laborieux, à cause du manque d'outil de visualisation et de navigation.

La somme de tous ces problèmes fait que seuls les archivistes sont capables d'utiliser le système d'une manière optimale, les autres utilisateurs étant souvent incapables de l'exploiter à leur satisfaction (fig. 3).

¹ *Dopesheet*: les descriptions thématique et visuelles qui commentent des images reçues des agences de presse.



fig. 2.a – RÉCEPTION DES NOUVELLES DES DIFFÉRENTES AGENCES DE PRESSE INTERNATIONALE (SÉQUENCES VIDÉOS NON ÉDITÉES, DOPESHEET, DÉPÊCHE, ...)



fig. 2.b – PRÉPARATION DES REPORTAGES PAR DES JOURNALISTES (SÉQUENCES VIDÉOS ÉDITÉES, COMMENTAIRES DES JOURNALISTES, ETC.)



fig. 2.c – LECTURE DES SCRIPTES PENDANT LA DIFFUSION



fig. 2.d – L'INTÉGRATION DE TOUS LES COMPOSANTS DU TÉLÉJOURNAL PENDANT LA DIFFUSION

fig. 2 – INTERVENTION DE PLUSIEURS SOURCES D'INFORMATION DANS LA PRODUCTION D'UN TÉLÉJOURNAL

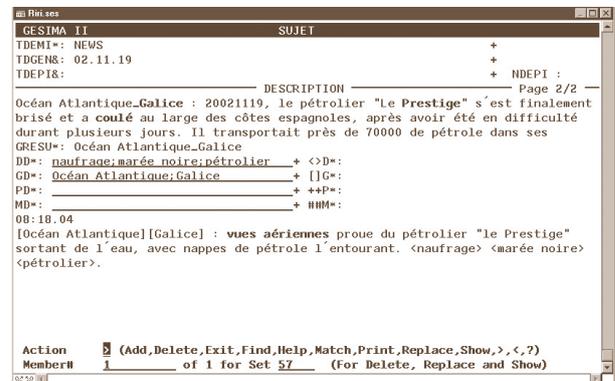
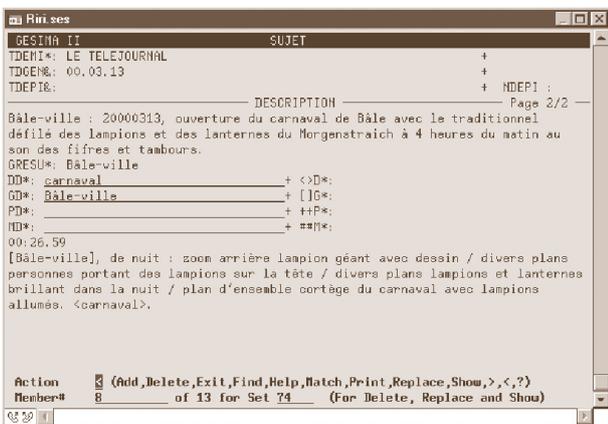


fig. 3 – LES INTERFACES D'INDEXATION ET DE RECHERCHE

LE MODÈLE SÉMANTIQUE DE RECHERCHE

Afin de pallier aux problèmes cités plus haut, il est nécessaire de proposer un modèle innovateur de recherche d'information audiovisuelle qui se caractérise par une attention particulière aux besoins de l'utilisateur pendant l'accomplissement de son métier (archiviste, journaliste, producteur, etc.). Nous allons expliquer le modèle sémantique proposé par l'exemple suivant, où un journaliste recherche **un sujet dans le contexte des jeux de football EURO 2000 contenant un plan qui dure au moins 5 secondes et qui montre un supporteur français qui dit que le meilleur gagne** (fig. 4).

Tous les types de vues possèdent une structure simple et commune qui se caractérise par les points suivants:

1. des unités fondamentales de descriptions (plan, reportage, etc.)
2. un ensemble de *descriptions* permettant de décrire des caractéristiques d'un segment audiovisuel selon le point de vue.
3. un ensemble de *relations intra vue* exprimant la structure interne d'un segment audiovisuel selon chaque point de vue.
4. un ensemble de *relations inter vue* permettant d'exprimer des requêtes portant sur plusieurs points de vues en même temps.

UN SUJET DANS LE CONTEXTE DES JEUX FOOTBALL EURO 2000 CONTENANT UN PLAN
 QUI DURE AU MOINS 5 SECONDES QUI MONTRE UN SUPPORTEUR FRANÇAIS QUI DIT
 "QUE LE MEILLEUR GAGNE".



fig. 4 – LES VUES SÉMANTIQUES

Pour interroger les données audiovisuelles, nous avons proposé un langage de requêtes **Semantic Views Query Language (SVQL)** adapté à la structure proposée plus haut (Semantic Views Model). Ce langage permet une expression à la fois riche et précise des requêtes complexes. Pour faciliter son utilisation, nous avons implémenté une interface de requêtes qui tient compte de ses spécificités (fig. 5).

De plus, pour accéder directement au contenu des segments audiovisuels, nous avons défini un modèle de navigation hypermédia basé également sur le modèle sémantique. Cette approche est très utile: elle permet une vérification aisée du contenu des résultats des requêtes et elle offre une représentation hiérarchique fidèle à la structure d'un journal télévisé (fig. 6).

Enfin, l'utilisateur a la possibilité de visualiser une séquence vidéo correspondante à chaque partie de la structure hiérarchique dans un style **browse and play**.



fig. 6 – LES INTERFACES DE NAVIGATION HIÉRARCHIQUE ET DE VISUALISATION

INDEXATION DES JOURNAUX TÉLÉVISÉS

Pour pouvoir exprimer des requêtes complexes et variées il faut disposer d'une description riche du contenu audiovisuel qui couvre la globalité des informations fournies par le processus de production et de l'archivage du journal télévisé. De plus, pour pouvoir garantir une interoperabilité au niveau d'échange entre les différentes sources d'informations, et au niveau du stockage des descriptions audiovisuelles dans un format neutre et extensible, nous avons proposé un schéma de description basé sur un standard. Dans cette optique nous avons défini **TVN Schema**, un schéma de description des journaux télévisés conforme au standard MPEG-7.

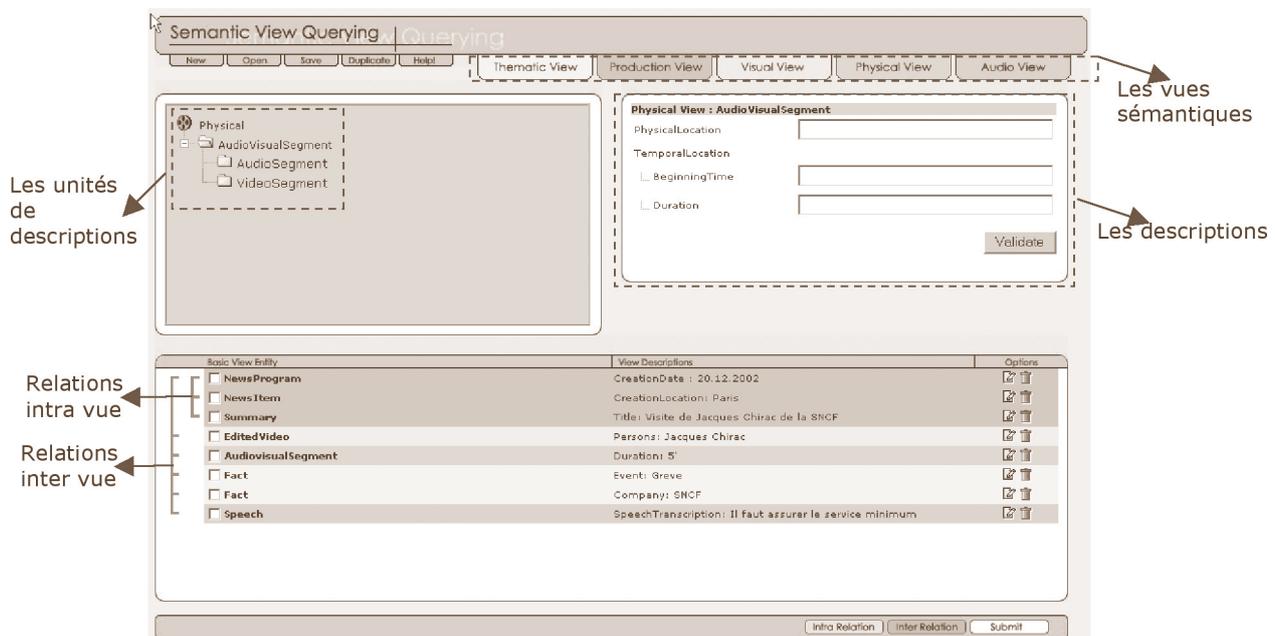


fig. 5 – L'INTERFACE DE RECHERCHE SÉMANTIQUE

Dans le cadre du projet COALA ce schéma est utilisé dans un système d'indexation de journaux télévisés (LogCreator) composé de deux interfaces: une interface de segmentation (fig. 7) et une interface d'annotation (fig. 8). L'interface de segmentation permet le découpage *a priori* des plans et une présentation de leurs images-clés représentant toute la séquence. Ensuite, l'utilisateur aura la possibilité de raffiner le découpage selon la structure hiérarchique de production (reportage, interview, etc.). L'interface d'annotation permet l'ajout des descriptions appropriées à chaque type de segment (ex. le nom du journaliste pour un reportage, etc.). Les résultats de l'indexation sont stockés en format MPEG-7.

MPEG-7 est le standards de description de contenu de documents audiovisuels.

Caractéristiques principales:

- standard de méta-données;
- basé sur XML;
- offre des représentations de contenu sous forme textuelle et binaire;
- librairie très riche de description des caractéristiques audiovisuelles



fig. 7 – LOGCREATOR – INTERFACE DE SEGMENTATION

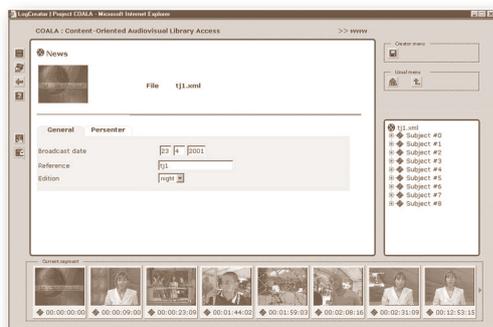


fig. 8 – LOGCREATOR – INTERFACE D'ANNOTATION

CONCLUSION ET TRAVAUX FUTURS

Le projet COALA a abouti à la mise en place d'une **plate-forme prototype** d'indexation et de recherche sémantique des journaux télévisés. L'approche tient compte du comportement des utilisateurs dans un environnement réel de production de journaux télévisés. Deux axes essentiels régissent cette plate-forme. Premièrement, la *réutilisation* grâce à un schéma de description standard conforme à la norme MPEG-7 (TVN Schema), et offrant un format d'échange neutre entre les différentes sources d'information. Deuxièmement, *l'expressivité* qui s'appuie sur un modèle formel (Semantic Views Model), un langage de requête (SVQL) et un modèle de navigation hiérarchique.

VERS UN PROTOTYPE/PRODUIT

Les résultats obtenus par le projet COALA ont suscité un sérieux intérêt auprès des équipes de production et d'archivage de la TSR. Dans cette optique, un projet expérimental d'intégration de COALA au sein de leur environnement est en cours de développement (collaboration EPFL/EIA-FR/TSR). De plus, une collaboration scientifique sur les problématiques traitées est établie avec l'institut National Audiovisuel (INA) de France.

APPEL AUX COLLABORATIONS FUTURES

La plate-forme COALA représente actuellement un environnement intéressant de test et de validation d'expériences d'indexation et de recherche des données audiovisuelles. Toutefois, il est intéressant d'enrichir cette plate-forme par des nouvelles expériences scientifiques basées sur des approches d'indexation automatiques utilisant des techniques de traitement de signal, de reconnaissance de la parole et de l'analyse de langage naturel.

Dans cette optique nous sommes prêts à partager notre expérience en faveur d'une collaboration avec des laboratoires et des institutions intéressés.

REMERCIEMENTS

Nous tenons à exprimer notre gratitude aux équipes de la production et de l'archivage du journal télévisé de la TSR. Egalement nous remercions Ronald Schroeter et tous les étudiants de l'EPFL et de l'EIA-FR qui ont participé très activement à la réalisation de la plate-forme COALA. Enfin, nous remercions nos collègues du **Center for Digital Video Processing, Dublin City University** pour leurs conseils et leur précieuse collaboration scientifique.

RÉFÉRENCES

- Le site du projet: <http://coala.epfl.ch>
- Chapitre du livre: N.Day, N. Fatemi, O. Abou Khaled. *Search and Browsing. Introduction to MPEG-7: Multimedia Content Description Interface*. Wiley. England: B.S. Manjunath | Philippe Salembier | Thomas Sikora, 2002. ISBN 0-471-48678-7, p. 335-352.

Articles scientifiques et rapport de thèse:

- N. Fatemi and O. Abou Khaled, *COALA: Content Oriented Audiovisual Library Access*, Sent for publication, 8th International Conference on Multimedia Modeling (MMM'2001), Amsterdam, Netherlands, November, 2001.
- N. Fatemi and O. Abou Khaled, *Indexing and Retrieval of TV News Programs Based on MPEG-7*, in Proceedings of the IEEE International Conference on Consumer Electronics (ICCE'2001), Los Angeles, CA, June 2001.
- N. Fatemi, J. A. DeBlasio, and G. Amaudruz, *Adding closed-caption and subtitle to VideoText DS*, MPEG-7 contribution: m7092, Singapore, Mars 2001.
- N. Fatemi, J. A. DeBlasio, and G. Amaudruz, *Some Remarks and Propositions to the MDS CD Resulting from the Study of Archive Applications*, MPEG-7 contribution: m6696, Pisa, Italy, January 2001.
- N. Fatemi, *A Semantic Views Model for Audiovisual Indexing and Retrieval*, these N° 2738, EPFL mars 2003. ■