
Des hypervidéos pour créer et échanger des analyses de documents audiovisuels

Olivier Aubert — Yannick Prié

LIRIS - FRE 2672 CNRS
Université Claude Bernard Lyon 1
Bâtiment Nautibus, 43, Boulevard du 11 Novembre 1918
F-69622 Villeurbanne cedex
prenom.nom@liris.univ-lyon1.fr

RÉSUMÉ. Les possibilités offertes par l'hypertexte ne sont pas encore très exploitées dès lors que l'on travaille avec des documents audiovisuels. Le besoin de métadonnées y est plus immédiatement perceptible, la temporalité des documents pose des problèmes nouveaux en termes de manipulation et d'interaction, et la question des droits étant loin d'être réglée. Afin de pouvoir analyser les expérimentations novatrices actuelles dans ce domaine et éclairer la construction de nouveaux systèmes, nous présentons la notion d'hypervidéo comme cadre général mettant en évidence différentes dimensions des hypermédias à base de documents audiovisuels. Le projet open-source Advène nous permet de mener des expérimentations de mise en application des hypervidéos et fournit un support à l'innovation par le bas propice à l'émergence de nouveaux modes d'interaction audiovisuelle.

ABSTRACT. The possibilities offered by hypertext are not very well utilised for audiovisual documents. Associated metadata is even more essential, the temporality of the documents introduces new issues in interaction and processing, and rights managements issue are more stringent. In order to analyse existing innovative uses of audiovisual documents and to provide guidelines for the building of new systems, we present the notion of hypervideo as a general framework identifying various dimensions of audiovisual-based hypermedia documents. The open-source Advène project provides a testbed to experiment with hypervideos, as well as a support for user-lead innovation, which is a potential source of new audiovisual interaction modalities.

MOTS-CLÉS : Audiovisual, Hypervidéo, Annotation, Temporalités, Échange

KEYWORDS: Audiovisual, Hypervideo, Annotation, Temporalities, Sharing

1. Introduction

Les changements induits par les nouvelles technologies ont produit des changements dans la forme et la représentation des documents et par là même dans leur usage et leur perception. L'hypertextualité [KOL 97], apportant une rupture de la linéarité des documents ainsi qu'une expression plus immédiatement perceptible de l'interconnexion de différents documents, transforme particulièrement les rapports de l'auteur et du lecteur à une œuvre. De plus, les possibilités de transformation des documents électroniques induisent une multiplicité de présentations possibles, dont la prise en compte peut elle-aussi influencer la conception des documents. Cette évolution est potentiellement applicable à l'audiovisuel, nonobstant deux enjeux à surmonter : 1/ un aspect simplement technologique, où les outils permettant cette rupture de linéarité doivent être disponibles (ainsi que les ressources pour visualiser les documents AV, mais c'est le cas maintenant) ; 2/ un aspect plus fondamental, prenant en compte le caractère temporel des documents pour proposer de nouvelles modalités d'interaction. Une des grandes sources d'innovation en matière d'exploitation novatrice des possibilités fournies par l'hypertexte vient de son appropriation par un large public, grâce à la large diffusion assurée par le web. Le web offre un simple support d'expression, une plate-forme, qui rend possible la création de projets plus évolués.

Tirant leçon de cette expérience, nous allons voir dans quelle mesure et par quels moyens une telle innovation est également possible pour les documents audiovisuels. Comment peut-on favoriser l'innovation par le biais des utilisateurs, et quelles sont les limites des outils actuels ? Pour répondre à cela, nous allons nous attacher à étudier quelques exemples existants d'utilisations innovantes du médium audiovisuel. Nous présenterons ensuite le modèle des hypervidéos, qui nous permettra d'analyser ces différents exemples pour en envisager des extensions.

2. Quelques expériences innovantes existantes

On peut étudier certains systèmes existants suivant la finesse du grain avec lequel ils manipulent les documents audiovisuels. Certains projets considèrent les documents audiovisuels dans leur intégralité, et visent simplement à les articuler entre eux, ou à proposer des informations complémentaires globales.

Les sites d'archives vidéo tels que Internet Movie Archive¹, la base InforMedia [HAU 02] ou le récemment introduit Google Video proposent ainsi de stocker des documents audiovisuels, accompagnés éventuellement de métainformations concernant le document dans son intégralité (auteur, date, etc). Internet Movie Archive a de plus proposé un concours de montage sur le sujet de la seconde guerre mondiale, utilisant les documents disponibles. On voit apparaître ici un nouvel usage rendu possible par la disponibilité des documents, tant en termes techniques (réseau) que légaux. Ces sites institutionnels trouvent leur contrepoint dans les *video blogs* ou

1. <http://www.archive.org/details/movies>

vlogs [HOE 04], qui reprennent le principe des carnets électroniques (*blogs*), en utilisant la vidéo comme support d'expression principal. Bien souvent, un commentaire textuel est associé à chaque document audiovisuel, ainsi qu'un espace de commentaires, que l'on peut assimiler à un système d'annotations.

D'autres projets s'attachent toujours aux document AV dans leur intégralité, mais en modifiant la substance par l'ajout ou le remplacement d'une composante (audio, texte, image, etc). La linéarité temporelle du document audiovisuel est cependant conservée.

Ainsi, la possibilité de sous-titrage optionnel et multiple (choix de la langue) offert par le support DVD transforme le document initial en offrant le choix de sous-titrer ou non un film, ainsi que le choix de la langue. De nombreux lecteurs vidéo logiciels permettent par ailleurs, lors de la visualisation d'une vidéo numérique, d'incruster des sous-titres tirés d'un fichier distinct, spécifiés de manière plus ou moins précise (mise en forme, etc). Cette forme de sous-titrage, à la limite de la légalité, apporte plus de souplesse à l'utilisateur, au détriment du contrôle effectué par l'auteur ou l'éditeur sur l'utilisation d'une œuvre. Il est également possible de transformer une autre composante du document audiovisuel, la bande son, pour la remplacer ou l'enrichir avec une information externe. Cette transformation pose également le statut du document obtenu, notamment vis-à-vis du droit d'auteur : quid par exemple de l'intégrité artistique de l'œuvre ? Une fois encore, le support DVD offre cette possibilité, et permet au spectateur de choisir la langue de doublage. Cette latitude offerte à l'utilisateur reste cependant très contrainte par les choix effectués par l'éditeur du DVD et négociés avec les ayants-droits. La diffusion d'un commentaire sur un film par le biais d'une bande-son alternative calée sur le film lui-même n'est donc pas aisée (problèmes de calage, etc), mais est tout de même pratiquée avec les moyens du bord².

Il est en outre possible d'apporter une rupture de la temporalité du document AV, en introduisant des liens hypermédias soit à l'intérieur du document lui-même, soit d'une partie du document vers d'autres documents. Des problèmes se posent alors en termes de représentation de ces liens, de gestion du temps de réaction de l'utilisateur, etc.

Les possibilités de navigation offertes par les DVD sont une illustration courante de ce principe, que ce soit le chapitrage présentant un accès direct à différentes séquences, ou le remontage dynamique présentant différentes versions d'un même film (comme dans le DVD du *Rocky Horror Picture Show* par exemple). Ces fonctionnalités sont d'ailleurs reprises dans les nouveaux formats audiovisuels numériques tels que DivX 6. Cependant, l'utilisateur reste encore une fois contraint dans ses choix par les volontés des éditeurs. L'utilisation des documents audiovisuels numériques

2. <http://www.renegadecommentaries.co.uk/>, <http://www.hurricaneandy.com/comment/>

permet de nouvelles formes de manipulation : les extensions interactives de Quick-Time [App04], ainsi que le projet Annodex [PFE 03] permettent d'insérer des liens de navigations dans une vidéo. Le projet Hyperfilm [HYP 03] offre des possibilités similaires, avec des modalités différentes d'affichage des liens (internes ou externes) : ces derniers sont présentés dans une fenêtre différente de la vidéo, afin de pouvoir éventuellement rester plus longtemps affichés. Les outils d'annotation de documents audiovisuels utilisés dans des domaines spécialisés tels que Anvil [KIP 01] ou ELAN fournissent également des possibilités de navigation à l'intérieur de documents, par le biais des annotations. Les recherches sur les représentations alternatives de documents audiovisuels (*surrogates* [DIN 99]) telles que la présentation de *story-board* sous la forme d'images ou une visualisation en accéléré [WIL 03, GRA 03] offrent par ailleurs un accès fin au contenu des documents.

Enfin, les techniques d'animation apportent une vision différente des médiums audiovisuels. Leur codage (programmation) lui-même ne traduit pas forcément leur temporalité, et ils peuvent par essence s'accommoder de ruptures de temporalité. Le format le plus connu est Macromedia Flash, qui apporte de nombreuses possibilités en termes d'animation et d'interactivité mais donne plus de pouvoirs au créateur qu'au lecteur. Le document, ou du moins son rendu, peut être perçu comme un document audiovisuel très avancé en termes d'interactivité. Cependant, cette latitude offerte aux concepteurs limite fortement de l'autre côté les possibilités de manipulation (d'enrichissement/personnalisation) par l'utilisateur.

Ces différents exemples, proposant une intégration plus ou moins fine de données dans des documents audiovisuels, présentent tous certains attributs communs. Tout d'abord, la présence, explicite ou implicite, de méta-données plus ou moins structurées. Ensuite, le besoin d'outils spécialisés pour la visualisation des nouveaux documents audiovisuels, afin d'offrir des fonctionnalités de rendu ou de navigation spécifiques. Enfin, une certaine forme de latitude offerte à l'utilisateur, souvent limitée toutefois. Pour analyser ces différents exemples, nous allons introduire le modèle des hypervidéos et voir dans quelle mesure il permet de tirer des leçons et d'envisager des évolutions des systèmes existants suivant quelques dimensions d'analyse.

3. Le modèle des hypervidéos

Présentons rapidement le modèle des hypervidéos, décrit plus précisément dans [AUB 04] : soit un document audiovisuel D_{AV} , auquel on adjoint une structure d'annotation S_A . Cette structure de données contient des annotations liées spatio-temporellement au document audiovisuel, des annotations non liées, ainsi qu'une structure de mise en relation des annotations. On appelle *Document Audiovisuel Annoté*, noté DAA , l'ensemble formé par le document audiovisuel et la structure d'annotation.

On appelle alors *vue d'un DAA* toute manière de visualiser le DAA. La spécification des vues peut s'effectuer de diverses manières : par le biais de modèles, par le

biais de feuilles de style, voire codées de manière programmatique dans l'outil de visualisation. Les différentes vues peuvent être discriminées suivant deux critères : d'une part, la répartition des sources d'information, et d'autre part, l'accès à la temporalité du flux audiovisuel. Lors de la visualisation, la proportion d'informations provenant du Document Audiovisuel et provenant de la structure d'annotation peut être variable. La simple visualisation du document AV est donc une vue, tout comme la génération d'un document à partir uniquement de données de la structure d'annotation. Cependant, la plupart des vues utiliseront un mélange des deux sources, comme par exemple l'affichage de sous-titres (annotations) lors de la visualisation du document AV. De plus, la temporalité du document AV peut ou non être accessible à partir de la vue : une simple table des matières, générée à partir de la structure d'annotations, sans lien vers le document audiovisuel ne possède plus de lien avec la temporalité du document AV.

En utilisant les deux critères évoqués ci-dessus, nous définissons une *hypervidéo* comme toute vue d'un *DAA* qui d'une part utilise à la fois la structure d'annotation S_A et le document audiovisuel D_{AV} , et d'autre part donne accès au flux temporel issu du DAA.

Voyons pour illustrer cette définition comment ce modèle s'applique à travers l'analyse de deux exemples cités précédemment : les DVD et le format Annodex. Dans une utilisation classique, un DVD est articulé autour d'un document AV (le film), qu'il va enrichir de diverses informations. Dans le cadre des hypervidéos, la structure d'annotation contiendra de multiples informations de types divers. On trouvera tout d'abord les méta-données de chapitrage, liées temporellement aux séquences du découpage. Les sous-titres sont constitués d'annotations textuelles (conceptuellement parlant, car dans la réalité ils sont stockés sous forme d'images) liés temporellement au document AV. Enfin, les bandes son alternatives sont également des méta-données, de type sonore, qu'on peut choisir de lier en intégralité à l'ensemble du film. Au niveau de la visualisation, l'ensemble des spécifications de la manière de visualiser ces informations se trouve codée dans les lecteurs de DVD.

Annodex[PFE 03] est un standard visant à étendre les fonctionnalités de liens hypertextes, de recherche et de composition de documents actuellement proposée par le web aux documents temporalisés. Pour cela, le langage CMMML (*Continuous Media Markup Language*) permet de coder des méta-données associées temporellement au document audiovisuel. Ces méta-données sont ensuite intégrées pour produire un DAA contenant à la fois le document AV et les méta-données. On voit ici apparaître un point important qui n'est pas directement traduit par le modèle, mais plutôt induit : l'importance du mode de stockage de chaque élément de l'hypervidéo (document AV, structure d'annotation, spécification des vues). Dans le cas d'Annodex, on utilise deux supports bien distincts que l'on rassemble en un seul, à des fins de diffusion (*streaming*).

Il est possible de généraliser la notion d'hypervidéo à un ensemble de documents audiovisuels, auquel on adjoint une structure d'annotation dont les données sont liées à des fragments spatio-temporels des différents documents AV. Une telle structure

permet alors d'établir des correspondances à la fois à l'intérieur et entre des documents AV. Le modèle des *video blogs* en offre une illustration, où chaque document AV est accompagné de méta-données (date, auteur, commentaire, mots-clés, etc.) liées à l'ensemble du document. La navigation dans un blog s'effectue alors par le biais de cette structure d'annotation qui articule les différents éléments AV [HOE 04].

4. Étude du modèle des hypervidéos

Le modèle des hypervidéos proposé ci-dessus est assez générique pour permettre une analyse de nombreux exemples existants, et offre également des pistes de travail intéressantes. En précisant le caractère distinct des documents AV, des méta-données et des modes de visualisation, il permet d'envisager des variantes dans l'utilisation de l'un ou l'autre de ces éléments. D'autre part, en les considérant comme un ensemble articulé et cohérent, il clarifie les relations entre ses différents éléments. Nous allons donc voir les apports de ce modèle suivant deux axes : d'une part, l'hypertextualité et son rapport à la temporalité ; d'autre part, le mode de stockage des éléments et son influence sur les possibilités d'échange et de collaboration.

4.1. *Hypertextualité et temporalité*

La structure d'annotation permet d'envisager la mise en place de liens hypermédiés. Une première distinction apparaît déjà entre liens profonds (vers l'intérieur d'un document) et liens non profonds (vers l'intégralité d'un document). La plupart des pratiques actuelles d'annotation de documents proposent des liens non profonds, apportant une information globale sur le document. Ainsi, Internet Movie Archive permet d'associer des notices et commentaires aux vidéos qu'il propose. Les liens profonds sont plus généralement proposés dans le cadre d'une utilisation spécialisée : les sous-titres ajoutés à des documents audiovisuels pointent vers l'intérieur du document, et pourraient être utilisés pour naviguer dans le document. Les liens profonds sont également une manière d'explicitier ou de relier les diverses temporalités présentes dans un document audiovisuel (temps du flux, temps du document, temps sociaux) [AUB 04]. L'introduction de liens profonds permet donc d'envisager une rupture de la temporalité des documents AV. De la même manière que l'hypertexte a gommé les frontières classiques du texte [KOL 97], les hypervidéos brisent la linéarité temporelle des documents, et par là la perception de leur durée. On se retrouve donc avec des objets sans limite temporelle. Cependant, la seule exploitation généralement envisagée de ces liens est l'affichage sur l'écran des informations associées. Annodex propose toutefois une forme de généralisation, à travers un format d'encapsulation de données audiovisuelles et de méta-données permettant la navigation, la recherche et la visualisation d'informations.

4.2. Stockage, réutilisation et échange

Comme on peut le retrouver dans le modèle de contenus de Whitehead [WHI 02], le mode de stockage/codage des trois éléments du modèle des hypervidéos (document AV, structure d'annotation et spécification des vues), résultant d'un choix technique, a également un grand impact sur les utilisations possibles des hypervidéos. Voyons comment la séparation des modes de transfert/stockage des éléments permet de considérer différents points de vue et différentes visualisations des mêmes données et enfin facilite la réutilisation et l'enrichissement de travaux.

La dissociation entre document et méta-données, par rapport à une approche d'intégration, affranchit de tout point de vue prédominant sur le document. Les méta-données permettant l'interprétation et l'exploitation du document, elles peuvent être multiples suivant l'utilisation envisagée. Il est donc primordial de fournir un support d'expression de cette multiplicité. Cette dernière est bien présente dans le cadre des projets traitant uniquement des documents intégraux comme les archives ou les *video blogs*, chacun pouvant ajouter une information liée au document AV pour peu qu'il dispose d'un identificateur (URI ou autre). Dans le cadre des systèmes à liens profonds, la spécialisation des systèmes actuels ne permet pas encore d'envisager les mêmes variations : les sous-titres et les informations de chapitrage d'un DVD sont intégrés dans le flux de données, et difficilement manipulables ; le format Annodex vise à encapsuler les méta-données et le document AV en une seule entité, les liant ainsi fortement. Le format de méta-données MPEG7 [SAN 02] fournit bien un stockage indépendant des méta-données, permettant d'envisager de multiples dimensions d'annotation, mais les outils permettant de l'exploiter [MU 03, PFE 00, SCH 03, TRA 03] ne sont pas encore largement diffusés.

De plus, au delà de la multiplicité des méta-données associées à un même document, il convient de considérer la variété des exploitations possibles de la même méta-donnée. Les utilisations courantes des méta-données sont généralement très spécialisées, comme l'affichage des sous-titres, et codées en dur dans l'application de visualisation. Le format SMIL [W3C 01] n'aborde que la problématique de la visualisation et de la synchronisation, mais peut servir de base pour spécifier une visualisation particulière. Ainsi, l'outil de création audiovisuel Mdefi [TRA 03] combine MPEG7 pour le stockage des méta-données avec SMIL pour leur représentation. Par ailleurs, si on peut bien isoler des visualisations classiques des méta-données, elles ne font pas généralement pas l'objet d'une spécification réutilisable. Concevoir par exemple un DVD comme une instance d'hypervidéo permet d'envisager, à partir du document AV et de la structure d'annotation (sous-titres liés), d'autres manières de visualiser l'information : génération d'une transcription, recherche textuelle permettant de localiser des séquences audiovisuelles, etc. La spécification de la visualisation peut être intégrée à l'outil de visualisation (lecteur DVD et sous-titres, etc), ou spécifiable (feuilles de style), mais cette dernière forme ne dispose pas encore de ses moyens d'expression.

Enfin, la séparation des éléments du modèle leur permet d'évoluer séparément, et d'être échangés indépendamment les uns des autres. Dans le cadre de la création et de l'échange de données, séparer la structure d'annotation (source de l'hypertextualité) du document audiovisuel lui-même, et offrir au lecteur l'opportunité de manipuler les méta-données, apporte un support à un travail collaboratif, ou au moins à une réappropriation des données pour les visualiser différemment ou les transformer.

Cet échange suppose deux pré-requis : d'une part, que les modes de stockage du document AV et des méta-données soient différents, de manière à pouvoir transmettre l'un indépendamment de l'autre et de ne pas voir appliquées aux méta-données les mêmes restrictions de diffusion qu'au document AV. D'autre part, que le contenu des méta-données soit au moins accessible (lisible, dans un format ouvert), voire modifiable, afin de fournir au lecteur l'opportunité de le faire évoluer. Ces nouvelles possibilités d'interaction offertes au lecteur modifient son statut (de lecteur à lecteur-auteur) et celui du document. On retrouve alors les questionnements sur le statut du document numérique qu'aborde R.T. Pédaque [PéD 04]. On constate toutefois dans la quasi-totalité des cas, et plus fortement encore dans le domaine de l'audiovisuel, une non-ouverture des données ou méta-données qui rend difficile leur réexploitation. Les sous-titres de DVD sont stockés dans un format rendant impossible leur indexation en tant que texte, et donc des réutilisations potentielles. Les formats récents tels que Flash, qui apportent plus de pouvoir à l'auteur d'un document, le font au détriment des possibilités accordées au lecteur : il devient impossible de spécifier un facteur de zoom, d'effectuer une recherche dans un texte, etc. La réexploitation des méta-données n'est donc pas désirée, ni même envisagée.

5. Enjeux soulevés

Le modèle des hypervidéos offre donc un cadre d'analyse d'exemples existants, afin d'en tirer une expérience et d'en envisager une évolution. Il permet également, en différenciant les différents éléments, de mettre en valeur des enjeux dans le cadre du partage et de l'évolution des informations.

Dans le domaine audiovisuel, les méta-données sont cruciales, et reconnues comme telles. Cependant, elles sont peu diffusées, et considérées comme une entité intermédiaire dans la production des nouveaux objets audiovisuels, le lecteur/spectateur étant lui-même considéré dans un rôle passif de consommation. Les considérer comme des entités de premier ordre, devant être accessibles par l'utilisateur, permet d'en envisager une réappropriation et une transformation par les utilisateurs. L'importance des questions relatives au droit d'auteur s'en trouve accrue. Autant il peut être considéré comme acceptable et naturel de partager des notes (sous forme d'annotations) sur un document AV, autant l'utilisation de ces annotations pour effectuer des transformations automatiques ou semi-automatiques (possibilités de navigation) sont actuellement plus sujettes à caution. Un simple chapitrage de film sur support DVD par exemple suppose l'établissement d'un contrat d'exploitation.

Comme le souligne von Hippel [HIP 05], dans le monde immatériel des documents numériques, les utilisateurs peuvent prendre un rôle actif et devenir source d'innovations. Cependant, cette évolution va à l'encontre des habitudes et du droit actuel (droit d'auteur notamment). Jusqu'où veut-on aller dans l'accroissement des possibilités des utilisateurs ? Peut-on fixer des limites à la manipulation de contenu sans limiter les possibilités d'expression ? Les projets tels que Creative Commons et wikimedia offrent de nouveaux terrains d'expérimentation, où on trouve une volonté de partage et d'innovation, accompagnée d'une réflexion sur les enjeux de l'appropriation des nouveaux moyens de communication, et pourraient bien être le terreau des innovations futures.

6. Advene

Le projet Advene [AUB 05] vise à développer une plate-forme *open-source*³ de création d'hypervidéos, permettant de 1/ annoter des documents audiovisuels ; 2/ offrir des visualisations enrichies des documents annotés ; 3/ échanger les annotations et leurs modes de visualisation indépendamment du document AV original. L'objectif du projet est de faciliter l'émergence de nouveaux usages des documents audiovisuels, en permettant aux utilisateurs d'expérimenter rapidement de nouvelles idées de visualisation basées sur des méta-données pré-existantes ou créées spécialement dans ce but. Le modèle des hypervidéos est donc explicitement utilisé, afin de profiter de ses potentialités en termes de visualisation et d'échange des données.

Le principe, présenté dans la figure 1, est donc d'intégrer les méta-données et la spécification de leurs visualisations dans un document appelé *recueil*. Un recueil peut ensuite être partagé de diverses manières (envoyé par courriel, mis à disposition sur un site web, etc.) et réutilisé par d'autres personnes. Cette réutilisation peut consister simplement en l'accès aux informations contenues par le biais des visualisations stockées dans le recueil. Mais il est également possible d'apporter des modifications ou des ajouts aux méta-données ou aux spécifications des visualisations. De plus, il est possible de référencer explicitement des éléments d'un autre recueil, afin de pouvoir y accéder sans en faire une copie. Ce mécanisme appelé *import* permet d'envisager la mise en place de dépôts (*repositories*) de recueils permettant de centraliser des schémas ou visualisations communes, à la manière des dépôts de DTD. Dans l'état actuel, les problématiques de gestion de version ne sont pas abordées par le prototype.

En résumé, les possibilités d'échange offertes par Advene reposent essentiellement sur 1/ l'association des métadonnées structurées et des visualisations associées, avec toujours la possibilité d'enrichir et de modifier les unes ou les autres ; 2/ la réutilisation d'éléments (schémas d'annotation, spécifications des visualisations, métadonnées) par simple copie ou import (référence).

Le prototype permet actuellement de spécifier deux types de vues : d'une part, des vues statiques correspondant à des documents XHTML générés suivant des modèles à

3. Le prototype est téléchargeable sur <http://liris.cnrs.fr/advene/>

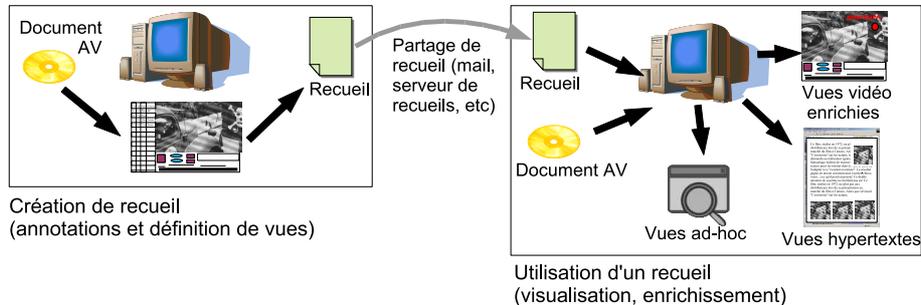


Figure 1. Le principe de fonctionnement d'Advene

partir des annotations ; d'autre part, des vues dynamiques, enrichissant la visualisation du document audiovisuel par le biais des annotations. Ces vues enrichies sont spécifiées par le biais de règles similaires à celles que l'on peut trouver dans les logiciels de courrier électronique, afin d'être au maximum accessibles aux utilisateurs. Elles permettent de modifier le cours de la lecture du document, pour par exemple effectuer un résumé dynamique ne visualisant que certains fragments, ou d'introduire une interactivité plus grande par le biais d'outils de navigation supplémentaires utilisant les annotations (fenêtres *popup* d'information et de navigation, historiques de navigation, etc.).

Par le biais de ce prototype, nous disposons à la fois d'un outil pour expérimenter de nouvelles modalités d'interaction avec les documents audiovisuels, et d'une manière de valider et enrichir nos réflexions sur le modèle des hypervidéos. Ce travail en cours se nourrit également de collaborations avec des utilisateurs potentiels d'hypervidéos, comme par exemple des chercheurs en sciences sociales étudiant les interactions humaines et utilisant des corpus vidéos pour leurs analyses. En proposant un système ouvert, capable d'exploiter de manière variée un large corpus (films sur support DVD notamment) et de faciliter les échanges d'informations structurées, nous cherchons à faciliter le développement par les utilisateurs [HIP 05] de nouvelles formes d'interaction.

7. Conclusion

L'immatérialité des documents numériques offre de nombreuses nouvelles potentialités d'interaction, allant notamment vers plus d'interactivité et de créativité de la part des lecteurs. Si le web repose largement aujourd'hui sur des supports textuels, l'extension de ses mécanismes à d'autres types de documents est déjà bien amorcé. Le développement des technologies numériques offre aux utilisateurs de plus en plus de possibilités, qui étaient jusqu'alors réservées à des professionnels (possibilités de manipulation, de stockage, de transfert, de diffusion, etc.). Cet élargissement du nombre de *créateurs* potentiels de contenu entraîne un fort potentiel d'innovation [HIP 05],

mais également des problèmes nouveaux en termes de droit (d'auteur, de copie, etc). Il est important d'accompagner ce mouvement en fournissant des outils permettant une telle innovation par le bas, tout en réfléchissant à l'impact de ces nouveaux outils.

Nous avons dans ce cadre réfléchi sur les documents audiovisuels et la manière dont on peut envisager leur évolution, à la lumière d'exemples existants d'utilisations nouvelles rendues possibles par la numérisation. Pour analyser ces exemples, nous avons proposé le modèle des hypervidéos, qui distingue clairement trois éléments : le document audiovisuel, la structure d'annotation qui l'accompagne, et les manières de visualiser les deux éléments précédents.

L'utilisation de ce modèle nous a permis de dégager quelques lignes directrices pour l'évolution des exemples sus-cités. Nous avons souligné l'importance de l'indépendance des méta-données par rapport aux documents audiovisuels auxquels elles se rapportent, ainsi que la nécessité de fournir des moyens de spécifier des visualisations alternatives. Ces pistes ne sont cependant viables que si l'on revoit également notre conception de document et d'auteur.

8. Bibliographie

- [App04] Apple, « Quicktime HREF Tracks », 2004, http://www.apple.com/quicktime/tools_tips/tutorials/hreftracks.htm.
- [AUB 04] AUBERT O., PRIÉ Y., « Documents audiovisuels instrumentés - Temporalités et détemporalisations dans les hypervidéos », *Document Numérique : numéro spécial Temps et document numérique*, , 2004, p. 143-168.
- [AUB 05] AUBERT O., PRIÉ Y., « Advene : active reading through hypervideo », *ACM Hypertext'05*, Sep 2005.
- [DIN 99] DING W., MARCHIONINI G., SOERGEL D., « Multimodal surrogates for video browsing », *Proceedings of the fourth ACM conference on Digital libraries*, Berkeley, California, United States, 1999, p. 85-93.
- [GRA 03] GRAHAM J., EROL B., HULL J., LEE D.-S., « The Video Paper Multimedia Playback System », *Proceedings of the eleventh ACM international conference on Multimedia*, Nov 2003, p. 94-95.
- [HAU 02] HAUPTMANN A., BARON R., LIN W., CHEN M., DERTHICK M., JIN M. C. R., YAN R., « Video Classification and Retrieval with the Informedia Digital Video Library System », *Proceedings of the Tenth Text REtrieval Conference*, Gaithersburg, MD, 2002.
- [HIP 05] VON HIPPEL E., *Democratizing innovation*, MIT Press, Apr 2005, ISBN 0-262-00274-4.
- [HOE 04] HOEM J., « Videoblogs as Collective Documentary », Jul 2004.
- [HYP 03] HYPERFILM, « Video Hyper-Media Production », Oct 2003.
- [KIP 01] KIPP M., « ANVIL - A Generic Annotation Tool for Multimodal Dialogue », *Proceedings of Eurospeech 2001*, Aalborg, Sep 2001, p. 1367-1370.
- [KOL 97] KOLB D., « Scholarly hypertext : self-represented complexity », *Proceedings of the eighth ACM conference on Hypertext*, Southampton, United Kingdom, 1997, p. 29-37.

- [MU 03] MU X., MARCHIONINI G., « Enriched Video Semantic Metadata : Authorization, Integration, and Presentation », *Proceedings of the Annual Meeting of the American Society for Information Science and Technology*, 2003, p. 316-322.
- [PFE 00] PFEIFFER S., SRINIVASAN U., « TV Anytime as an application scenario for MPEG-7 », *Workshop on Standards, Interoperability and Practice, ACM Multimedia 2000*, Los Angeles, Oct 2000.
- [PFE 03] PFEIFFER S., PARKER C., SCHREMMER C., « Annodex : a simple architecture to enable hyperlinking, search and retrieval of time-continuous data on the Web », *5th ACM SIGMM International workshop on Multimedia information retrieval*, 2003, p. 87-93.
- [PÉD 04] PÉDAUQUE R. T., « Le texte en jeu, Permanence et transformations du document », STIC-SHS-CNRS - Working paper - http://archivesic.ccsd.cnrs.fr/sic_00000511.html, Apr 2004.
- [SAN 02] SANCHEZ J. M. M., KOENEN R., PEREIRA F., « MPEG-7 : The Generic Multimedia Content Description Standard, Part 1. », *IEEE Multimedia Journal*, vol. 9, n° 2, 2002, p. 78-87.
- [SCH 03] SCHROETER R., HUNTER J., KOSOVIC D., « Vannotea - A Collaborative Video Indexing, Annotation and Discussion System For Broadband Networks », *Workshop on "Knowledge Markup and Semantic Annotation"*, 2003.
- [TRA 03] TRAN-THUONG T., ROISIN C., « Multimedia Modeling Using MPEG-7 for Authoring Multimedia Integration », *Proceedings of the 5th ACM SIGMM international workshop on Multimedia information retrieval*, 2003.
- [W3C 01] W3C, « Synchronized Multimedia Integration Language (SMIL 2.0) », W3C, 2001, <http://www.w3.org/TR/smil20/>.
- [WHI 02] WHITEHEAD J., « Uniform comparison of data models using containment modeling », *Thirteenth ACM conference on Hypertext and hypermedia*, College Park, Maryland, USA, june 2002, p. 182-191.
- [WIL 03] WILDEMUTH C. M., MARCHIONINI G., YANG M., GEISLER G., WILKENS T., HUGHES A., GRUSS R., « How fast is too fast ? : evaluating fast forward surrogates for digital video », *Proceedings of the third ACM/IEEE-CS joint conference on Digital libraries*, Houston, Texas, 2003, p. 221-230.