

Les normes MPEG-7 et MPEG-21 pour la description des contenus multimédias : indexation et réutilisation en postproduction cinématographique

Titus Zaharia, Alain Vaucelle, Thomas Laquet et Françoise Prêteux

Institut TELECOM ; TELECOM SudParis ; Département ARTEMIS
9, rue Charles Fourier, 91011 Evry, France

{titus.zaharia, alain.vaucelle, thomas.laquet, francoise.preteux}@it-sudparis.eu

Résumé Dans cet article, nous proposons une nouvelle plateforme d'indexation de contenus multimédias, appelée INVENIO (*INdexing Visual ENvironment for multimedia Items and Objects*). Fondée entièrement sur la norme ISO/MPEG-7, la plateforme INVENIO offre, dans un système intégré, moteurs d'extraction de métadonnées MPEG-7, outils d'annotation, moteur de requête, outils de gestion de bases de données multimédias, ainsi que des interfaces utilisateurs appropriées et ergonomiques. Pour valider la plateforme INVENIO, nous avons considéré une application industrielle d'indexation liée à une problématique de réutilisation de contenus numériques au sein d'une chaîne de production audiovisuelle. Les expérimentations réalisées concernent différentes productions audiovisuelles, incluant des contenus aussi bien naturels que de synthèse (*i.e.* dessins animés). Les solutions proposées démontrent notamment que l'exploitation des technologies MPEG-7 à travers la plateforme d'indexation INVENIO permettent de réaliser une économie significative de temps de travail, ainsi qu'une réutilisation optimale des contenus numériques en cours de production. Enfin, nous montrons également comment l'intégration de la norme MPEG-21 peut être utilisée afin d'étendre les possibilités d'échange de contenu de la plateforme INVENIO dans un environnement de production en ligne et partagé.

Mots clés indexation d'images, MPEG-7, MPEG-21, indexation par le contenu, moteur de recherche, interfaces, ergonomie.

1 Introduction

Aujourd'hui les méthodologies d'indexation par le contenu [1] visent à proposer des solutions alternatives pour l'indexation des documents multimédias. Le principe consiste à associer aux contenus des métadonnées non plus textuelles, mais liées intrinsèquement au contenu audiovisuel lui-même. Cela est possible puisque les représentations mathématiques permettent la description de manière automatique (ou semi-automatique) et discriminante du flux audiovisuel. Ce sont des caractéristiques perceptuelles associées classiquement aux attributs visuels comme la forme, la couleur, la texture ou encore le mouvement, qui rendent cette indexation possible. Le principe de la recherche de contenus dans une base de données change alors de façon fondamentale. L'utilisateur présente à l'entrée de son moteur de recherche non plus des mots-clés mais un exemple (*e.g.* une image/vidéo ou une partie d'une image/vidéo), ou une ébauche (*e.g.* un dessin réalisé à la main). Les mesures de similarités associées aux descripteurs images permettent alors de réaliser des requêtes automatiques et de retrouver des images similaires.

Le domaine de l'indexation d'images par le contenu n'a cessé de connaître une effervescence spectaculaire depuis le milieu des années 1990, comme en témoignent l'impressionnant volume de méthodes et techniques proposées dans la vaste littérature scientifique [1, 2] (pour des synthèses critiques) consacrée à ce sujet. Dans ce cadre, une étape marquante a été la sortie officielle au début des années 2000 de la norme ISO MPEG-7 [3]. Officiellement appelée *Multimedia Content Description Interface*, la norme MPEG-7 propose un large éventail de technologies de description de documents multimédias, intégrant des approches aussi bien textuelles que par le contenu.

En particulier, la spécification ISO MPEG-7 propose aujourd'hui un riche ensemble de descripteurs visuels [4, 5] et de schémas de description, exprimés dans un langage de description de données fondé sur une approche XML Schema.

Si MPEG-7 traite exclusivement des aspects de description de contenus multimédias, la norme MPEG-21 [6] élargit ce périmètre en proposant un environnement fédérateur de description, applicable à la fois aux contenus, aux utilisateurs, aux droits d'usage, aux réseaux et aux terminaux. Cette tentative de gestion globale de la chaîne de consommation numérique est sans commune mesure avec les normes précédentes. MPEG-21 devient ainsi la norme universelle d'intégration du multimédia tout en restant prioritairement conçue et organisée pour l'échange de l'audiovisuel.

Toutefois, la pertinence des normes MPEG-7/21 pour des applications industrielles grandeur nature reste à démontrer. Pour cela, il est indispensable d'élaborer et de mettre en œuvre des outils efficaces et ergonomiques permettant de faciliter les processus d'annotation, de consultation, de navigation dans des bases de données et de requête de l'information.

La plateforme d'indexation INVENIO (*INDEXING VISUAL ENVIRONMENT FOR MULTIMEDIA ITEMS AND OBJECTS*) proposée dans cet article s'attaque notamment à cette problématique. Fondée entièrement sur la norme MPEG-7, la plateforme INVENIO offre, dans un système intégré, moteurs d'extraction de métadonnées MPEG-7, outils d'annotations, moteur de requête, outils de gestion de bases de données, ainsi que des interfaces utilisateurs ergonomiques.

Pour valider la plateforme, nous avons développé une application industrielle d'indexation d'images couleur/texture pour la production de films d'animation. L'objectif est de favoriser la réutilisation de contenus numériques au sein d'une chaîne de production audiovisuelle. Ce développement est lié au projet structurant HD3D-IIO du pôle de compétitivité CapDigital (www.capdigital.com).

La suite de cet article est structurée comme suit. Tout d'abord nous énonçons les objectifs généraux du projet HD3D-IIO en mettant notamment l'accent sur la problématique de réutilisation de contenus au sein d'une chaîne de production audiovisuelle (paragraphe 2). Le cadre normatif MPEG est ensuite exposé (paragraphe 3), en précisant notamment les éléments de la norme MPEG-7 retenus pour nos développements, ainsi que la pertinence des technologies MPEG-21. La quatrième partie décrit les aspects d'interface utilisateur et d'ergonomie ainsi que les fonctionnalités supportées. Le paragraphe 5 analyse comment MPEG-21 pourrait offrir une base normative pour tous les contenus multimédias. Enfin, la partie 6 conclut l'article et ouvre quelques perspectives de travail futur.

2 Contexte HD3D-IIO : réutilisation de contenus

Le projet HD3D-IIO concerne la création et la fabrication des contenus numériques pour les industries techniques de l'audiovisuel et du cinéma. Son ambition est de doter le secteur de moyens technologiques nouveaux, dans un environnement de travail en mutation permanente et en confrontation avec les exigences de la compétitivité à l'échelle mondiale. Ces moyens doivent être conçus collectivement, dans la perspective d'une industrie ouverte du point de vue des échanges numériques entre diverses entreprises du secteur. Le consortium HD3D-IIO regroupe des acteurs

professionnels majeurs de la production audiovisuelle (publicité, dessin animés 2D/3D, films de cinéma, effets spéciaux 2D/3D...) en Ile de France, comme les sociétés, Duran Duboi, Eclair, LTC, Mac Guff, Mikros Images, Teamto, ou 2 Minutes.

À long terme, l'ambition de HD3D-IIO est d'augmenter la compétitivité de l'industrie francilienne du film numérique face à des concurrents anglais ou américains. Pour cela, de nombreux verrous technologiques nécessitent d'être levés. Cela concerne la spécification de formats d'échanges de contenus 2D et 3D, l'élaboration de plateformes de production collaboratives avec une maîtrise des développements réalisés sur des sites distants, la spécification de méthodes de réutilisation de contenus, ou encore la prise en compte des aspects de protection et de traçabilité des contenus nécessaires pour leur transmission sécurisée.

Dans ce cadre, un des objectifs majeurs de HD3D-IIO concerne la réutilisation des contenus images en environnement de production. Ce besoin fait écho à la constatation suivante : chaque production (*e.g.* dessin animé, clip vidéo ou film) conduit à des dizaines de milliers d'images. Or, de nombreux éléments (*e.g.* parties de décor d'un film, personnages/accessoires d'un dessin animé 2D ou 3D...), peuvent être ré-exploités lors d'une nouvelle création, à condition toutefois de pouvoir les retrouver facilement dans les collections d'images précédemment produites.

Aujourd'hui, aucun outil efficace n'est disponible pour accéder à de telles bases d'images pour faciliter/accélérer les modes de production multimédias en capitalisant sur des créations antérieures. Le seul recours possible est d'exploiter le savoir-faire et l'expérience des professionnels et artistes impliqués dans la chaîne de production. Cela est consommateur en ressources humaines et à un impact sur le temps et donc sur le coût de production.

L'indexation automatique de texture et donc des contenus des bases de données d'images manipulées lors d'une production peut apporter un gain de temps significatif lors de la production d'un dessin animé, en offrant des mécanismes de recherche d'images afin d'optimiser la réutilisation de contenus numériques. Aujourd'hui le partage et la réutilisation des contenus dans le cadre d'une production est difficile, les solutions existantes présentant de nombreuses limitations :

- l'indexation manuelle est fastidieuse et peu utilisée à cause des contraintes de temps pour l'indexer,

- la construction et le classement des éléments nécessaires à la production de dessins animés (décors, textures, objets, personnages,...) dans des fichiers à l'intérieur d'une arborescence au sein d'un disque dur partagé sur l'intranet est utile mais peu commode puisque les utilisateurs doivent maîtriser la logique de sa construction (classement, arborescence).

- la base de données avec une recherche par mots clés ou par nom ou par extension de fichiers est fastidieuse. Souvent, les mots clés sont absents. De plus, même lorsqu'ils existent, ils sont trop limitatifs. Par exemple, les noms de fichiers suivent une logique définie en début de production et la nomenclature utilisée varie fortement d'une production à une autre.

L'indexation automatique par le contenu d'images pourrait apporter des éléments de réponse à cette problématique. Au minimum trois contraintes doivent être néanmoins respectées pour garantir une solution viable. La première concerne l'efficacité des requêtes (*i.e.* capacité de retrouver l'information pertinente dans des grandes bases d'images). La seconde est liée au temps de calcul associé qui doit être compatible avec des réponses quasi instantanées à des requêtes interactives. Enfin, la troisième concerne l'élaboration de moteurs de recherche adaptés et aisés d'utilisation, supportant un large éventail de formats d'images et intégrables dans les chaînes de production existantes.

La plateforme d'indexation INVENIO proposée dans cet article et décrite dans la section suivante, se propose notamment de répondre à ces différentes contraintes pour assurer une réutilisation optimale des contenus numériques en cours de production.

3 Les normes MPEG-7 et MPEG-21

Au niveau de l'ISO (*International Standardization Organization*), la normalisation de l'image et du son est réalisée au sein du sous-comité technique 29 (ISO/IEC JTC 1/ SC29 - Codage du son, de l'image, de l'information multimédia et hypermédia) dont MPEG (*Moving Picture Experts Group*) est un groupe de travail (*Working Group 11*). Après avoir lancé, avec le large succès que l'on sait, les normes bien connues de codage vidéo MPEG-1/2/4, le groupe MPEG s'est attaché à l'indexation documentaire des contenus, en proposant la norme MPEG-7 [3]. Enfin, le dernier produit de la famille est la norme MPEG-21 [7], qui normalise à la fois la description des contenus, mais aussi tous les éléments intervenant dans la chaîne de production et d'échange d'objets numériques complexes. La panoplie des normes MPEG avec les fonctionnalités correspondantes est illustrée dans la figure 1.

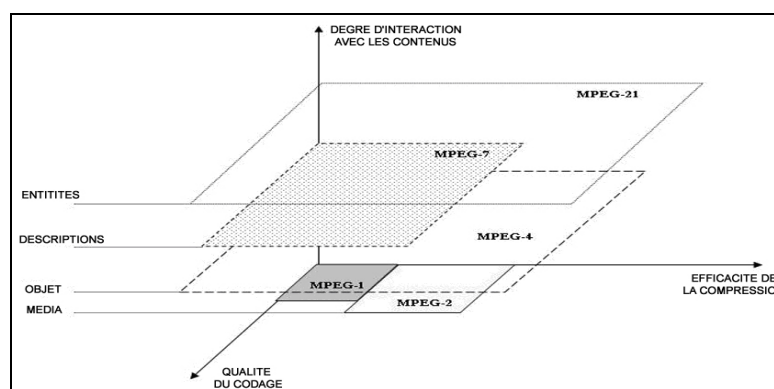


Figure 1 : Les différentes normes MPEG et leurs caractéristiques en termes de qualité de codage, d'efficacité de compression et d'interactivité.

Dans la suite, présentons synthétiquement les éléments des normes MPEG-7 et MPEG-21 retenus pour nos développements.

3.1 Cadre normatif MPEG-7 : descripteurs retenus

La norme MPEG-7 regroupe un riche ensemble de descripteurs et schémas de description aussi bien par le contenu (audio et visuel) que sémantiques/textuels. Elle propose également une architecture qui combine descripteurs et schémas de description pour décrire des concepts liés à la structure de l'image et de plus haut niveau sémantique. Le langage de description est fondé sur XML Schema [8].

Parmi les différentes parties de la norme, la partie 3, dite MPEG-7 Visual [9], propose un éventail de descripteurs adaptés et optimisés pour des applications de recherche par le contenu selon différentes caractéristiques visuelles comme la forme, le mouvement, la texture ou la couleur.

Pour l'application HD3D-IIO, seuls les attributs de couleur et de texture ont été jugés pertinents dans la phase actuelle du projet. C'est notamment cette partie qui a été retenue pour intégration dans la

plateforme INVENIO. Le standard MPEG-7 propose actuellement quatre descripteurs de couleur et trois descripteurs de texture, présentés succinctement dans le tableau 1.

Pour une description plus approfondie des descripteurs, avec définition mathématique, mesures de similarité associées et propriétés, le lecteur est invité à se rapporter à [5, 10].

Tableau 1 : Descripteurs MPEG-7 retenus pour intégration dans INVENIO.

Descripteurs de couleur	
<i>Nom</i>	<i>Principe</i>
Espace de couleur	Spécification de l'espace de couleur de représentation des couleurs (e.g. RGB, HSV, Luv, HMMD...)
Quantification de couleur	Spécification d'une quantification uniforme de l'espace de couleur (composante par composante)
Histogramme de couleur scalable	Histogramme de couleur indexé et représenté de manière multi-résolution par transformée de Haar
Descripteur par couleurs dominantes	Représentation d'une image par un nombre relativement réduit (maximum 8) de couleurs dites dominantes, obtenues par algorithmes de <i>clustering</i>
Descripteur couleur-structure	Histogramme de couleur enrichi d'information spatiale à l'aide d'un élément structurant
Distribution spatiale de couleur	Description de la distribution globale des couleurs dans l'image par transformée en Cosinus discrète
Descripteurs de texture	
<i>Nom</i>	<i>Principe</i>
Histogramme des orientations des contours	Classification grossière des orientations des contours en 5 catégories et construction d'un histogramme par rapport à ces classes
Texture homogène	Représentation multi-résolution par transformée de Gabor avec 6 orientations et 5 échelles

La norme MPEG-7 offre un cadre générique de description des contenus multimédias. Toutefois, son périmètre d'action concerne uniquement les contenus. Or, aujourd'hui, dans le contexte actuel de convergence technologique et avec l'explosion des réseaux et terminaux fixe ou mobile de tout type ainsi qu'avec les nouvelles fonctionnalités mises en avant par les technologies du Web 2.0, la consommation multimédia est en train de changer de manière spectaculaire. Cela implique la nécessité de disposer non plus uniquement de descriptions de contenus, mais d'une façon plus générale de tous les éléments susceptibles d'intervenir dans la chaîne de consommation, depuis la création, en passant par la diffusion et en allant jusqu'à l'utilisateur final. La norme MPEG-21 [6], dont le nom officiel est *Multimedia Framework*, se propose notamment de lever ce verrou technologique.

3.2 Cadre normatif MPEG-21 : la norme multimédia du XXI siècle

Le concept central mis en avant par la norme MPEG-21 est celui de DI - *Digital Item*, défini de façon générique et abstraite comme un produit numérique simple ou composite. Un exemple type est celui d'une page web, contenant différentes ressources multimédias comme du texte, des images, des vidéos, des éléments de mise en page (e.g. feuilles de style), des hyperliens, mais aussi des scripts de

programmation qui conduisent à une apparence dynamique, en fonction de l'interaction de l'utilisateur. MPEG-21 fournit les mécanismes de description de tels produits numériques complexes. En particulier, les parties 2 – *Digital Item Declaration* et 3 – *Digital Item Identification* permettent respectivement la spécification complète et structurée des DI et leur identification/localisation.

Soulignons également la partie 7 du standard, dite *Digital Item Adaptation*, qui standardise des descripteurs et des schémas de description permettant l'adaptation des contenus vis-à-vis des utilisateurs, des réseaux, des terminaux ou encore de l'environnement d'utilisation.

Enfin, un important travail a été consacré aux aspects de propriété intellectuelle et de droits d'usage (partie 4 – *Intellectual Property Management and Protection Components*). Pour cela, MPEG-21 standardise un langage d'expression de droits (partie 5 – *Rights Expression Language* - REL) et un dictionnaire terminologique correspondant (partie 6 – *Rights Data Dictionary* - RDD).

Le langage REL permet de décrire les droits et les permissions associés à un contenu, qui protègent et garantissent les conditions d'usage. Ces éléments fournissent les descripteurs d'accès et d'obtention des droits. C'est notamment à travers l'expression de l'autorisation d'accès que l'interopérabilité est assurée, en définissant les protocoles et en autorisant les droits d'accès aux contenus.

La partie 4 – IPMP propose un système de gestion de droits associés aux objets multimédias. Cela permet donc de pouvoir accéder et interagir avec les outils IPMP, d'échanger des données entre les outils (décryptage...), et d'authentifier les outils. Cette couche est nécessaire pour l'utilisation effective des technologies REL et RDD.

Le dictionnaire RDD est défini comme « un ensemble de mots clairs, cohérents, structurés, intégrés, et identifiés de manière unique pour permettre la génération d'expressions de droits » [11]. Il permet ainsi de décrire d'un point de vue sémantique les mots décrivant les droits, en facilitant les passerelles d'une terminologie à une autre dans le domaine des droits.

L'ensemble des travaux MPEG-21 s'inscrit en parfaite continuité avec ceux réalisés précédemment dans MPEG-7. Les descripteurs et schémas de description correspondants sont développés sous la responsabilité du même groupe MDS (*Multimedia Description Schemes*) et à l'aide du même langage de description fondé sur XLM Schema.

Le tableau 2 présente l'ensemble de parties de la norme MPEG-21.

Tableau 2 : Les 18 parties de la norme MPEG-21.

1. Vision, technologie, stratégie (définition du DI)	10. Digital Item Processing
2. Digital Item Declaration	11. Evaluation methods for persistent association Technologies
3. Digital Item Identification	12. Test bed for MPEG-21 Resource delivery
4. IPMP	13. Cette partie a été transférée dans MPEG-4
5. Rights Expression Language	14. Conformance testing
6. Rights Data Dictionary	15. Event Reporting
7. Digital Item Adaptation	16. Binary format
8. Reference software	17. Fragment Identification of MPEG Resources
9. MPEG-21 file format	18. Digital Item Streaming

Pour une description plus approfondie de la norme et de ses diverses parties, le lecteur est invité à se reporter à [11].

4 La plateforme INVENIO

L'approche d'indexation d'images par le contenu proposé par la plateforme INVENIO s'appuie entièrement sur les technologies de description de contenus multimédias proposées par la norme MPEG-7. Dans sa conception actuelle, INVENIO est conçu comme une application *Standalone*, à être utilisée en interne au sein des sociétés de production audiovisuelle. À terme, nous souhaitons intégrer dans INVENIO les éléments MPEG-21 mentionnés au paragraphe 3.2. Cela permettra notamment de convertir INVENIO en une plateforme d'indexation en ligne et partagée. Un tel outil serait particulièrement utile pour les co-productions impliquant plusieurs sociétés ou bien pour des objectifs de création collaborative à distance entre sites délocalisés.

4.1 Fonctionnalités supportées par la plateforme INVENIO

Etant essentiellement destiné à des graphistes, l'intégration dans un environnement graphique et ergonomique est déterminante pour son adoption par les utilisateurs. Afin de répondre à ces objectifs, INVENIO propose les fonctionnalités suivantes :

- la requête est réalisée sous forme de recherche par l'exemple. Une image entière ou une région d'une image est tout d'abord spécifiée. La requête en fonction des critères de recherche de couleur/texte spécifiés par l'utilisateur est ensuite effectuée.

- un deuxième mode de requête, cette fois par mots clés sélectionnés à partir d'une liste représentative des concepts de texture. Une mini-ontologie métier de type « production » a été constituée. Elle inclut des différents éléments relatifs aux matières (*e.g.* métal, bois, pierre, verre, ciel, nuage, liquide, végétale), motifs (*e.g.* damier, brique, hexagonal, radial, cubique, cellulaire, graduée, léopard, oignon, matelassé, aplat...), formes (*e.g.* ronde, ovale, rectangle, carré...), couleurs, fréquences, régularité de répétition et type d'objets.

- un système de description personnelle par « tags » que l'on peut ajouter permet d'annoter l'image de façon personnelle. Lors d'une production, les tags en « langage utilisateur » (*i.e.* annotation en texte libre) permettent d'indexer une image en vue de créer un « dictionnaire terminologique » lié à un projet.

Outre ces éléments de requêtes et d'annotation par descripteurs individuels, le système permet également d'hybrider de multiples critères de recherche et de description. La combinaison des descripteurs rend notamment possible la prise en compte de la polysémie de l'image en croisant un mode recherche lié à la structure de l'image (descripteur de couleurs) et un autre lié à un niveau de représentation sémantique (mots-clés et tags). Cette approche permet donc d'établir des relations entre les différentes composantes d'une description de l'image. L'interface déployée offre à l'utilisateur une relation efficace entre une indexation en langage naturel, et un langage de requête structurel afin d'organiser une représentation des modes d'indexation à partir de critères multiples.

4.2 Interface et ergonomie

La figure 2 illustre la présentation graphique adoptée par INVENIO. Les différents éléments constitutifs de l'interface, énumérés de 1 à 8 dans la Figure 2 sont explicités ci-dessous :

1 : La représentation graphique de la base de données est symbolisée par une spirale 3D déroulante d'images qui donne à l'utilisateur l'impression de profondeur et qui permet la navigation dans la base. À chaque requête, les images de la base sont triées et présentées par ordre décroissant de similarité avec l'image requête.

2 : Cadre de sélection de l'image sur laquelle on effectue la requête.

3 : Cadre d'affichage/mémorisation de l'image sélectionnée.

4 : Données relatives aux formats du fichier image utilisé et métadonnées textuelles (mots clés : matières, motifs, régularité, formes, fréquences, couleurs, types et/ou des tags).

5 : Menu d'aide de l'interface

6 : Liste des critères de recherche structurels : descripteurs de couleur et de texture.

7 : Bouton de lancement de la requête.

8 : Menu de sélection des serveurs de calcul. Menu de chargement d'une image externe à la base de données pour le lancement d'une requête.

Notons enfin que, pour augmenter davantage la sensation d'immersion 3D dans une base d'images, et une interaction plus directe, l'utilisation d'un écran tactile est envisagée.

Les résultats expérimentaux, obtenus sur un corpus de quelques milliers d'images à la fois naturelles et synthétiques fournies par les créateurs du projet HD3D-IIO (correspondant aux productions Yakari, Atomic Betty, Cajou, Step by Step, Faubourg 36), démontrent la pertinence de l'approche MPEG-7 proposée, avec des scores de requêtes de plus de 60-70% et des temps de réponse aux requêtes inférieurs à 0.5s.

Les Figures 2 et 3 présentent quelques exemples de recherche par similarité avec la plateforme INVENIO. Dans le premier cas, une image de la base a été donnée comme exemple. Dans le deuxième, l'utilisateur a présenté une ébauche dessinée à main levée. On observe qu'un personnage similaire est retrouvé.



Figure 2 : Interface et requête par similarité avec la plateforme INVENIO.

Validée par les artistes et les experts graphiques dans le cadre du projet HD3D-IIO, INVENIO offre ainsi un outil performant pour retrouver et réutiliser des images au sein de la chaîne de production.



Figure 3 : Requête par l'ébauche similarité avec la plateforme INVENIO.

5 Vers une convergence interdisciplinaire

Un aspect particulièrement intéressant dans l'essor du multimédia est la façon dont les contenus se développent d'un point de vue technologique. Ce déploiement nécessite une instrumentation de tous les composants médias (image, son, texte, hyperlien...), et doit s'inscrire dans un cadre normatif afin de garantir interopérabilité et réutilisation par le plus grand nombre. Toutefois, il convient d'inclure dans les objets multimédias, l'ensemble des ressources disponibles qui peuvent aider à la construction de la connaissance, notamment l'indexation et les conditions d'utilisation...

Or, cette logique référentielle du corpus de documents numériques est l'un des postulats de base pour les concepteurs des systèmes d'information pour donner naissance à une normalisation de l'ingénierie linguistique (XML, MPEG-7, MPEG-21). MPEG-21 est devenu le cadre modulaire de développement et donc un bon candidat à l'intégration globale de tous les documents multimédias. Cette affirmation peut paraître audacieuse, mais se justifie sous la forme d'un syllogisme assez simple. Le multimédia n'est pas né de rien. C'est une conséquence directe de la normalisation des pratiques numériques : les concepteurs de téléphonie, d'audiovisuel, d'informatique textuelle, se sont mis d'accord dans des instances de normalisation pour être interopérables et compatibles. Cela a conduit au multimédia numérique mondialement multilingue que nous connaissons. Le multimédia et la mondialisation numérique en réseaux se sont ainsi généralisés comme un effet direct de l'effort normatif. Dès lors, il est primordial de poursuivre cet effort même si cette normalisation s'attaque de plus en plus sur des domaines de moins en moins triviaux et matériels, comme ceux d'applications sociales, de traitement du savoir, de sémantique, de disparité culturelle, linguistique ou disciplinaire.

Faute d'un cadre global d'intégration du « commerce entre les hommes » que nous pouvons synthétiser sous le terme général d'*e-procurement*, nous risquons une babélisation numérique. Cette globalisation doit intégrer tous les aspects de diversité de l'information (mode de médiation, typologie d'usage, typologie d'accès, d'acheminement et d'échange, gestion référentielle, structurelle et sémantique ...).

6 Conclusion

La plateforme INVENIO intègre, dans une conception graphique innovante, outils d'annotation et de recherche d'images à partir de critères de description d'ordre textuelle, sémantique et structurelle. L'ensemble des outils de description structurelle est fondé sur la norme ISO MPEG-7, dont INVENIO propose une implantation efficace.

Pour valider la plateforme, nous avons développé une application de réutilisation des contenus au sein d'une chaîne de production audiovisuelle, liée au projet CapDigital HD3D-IIO. Les expérimentations réalisées concernent différentes étapes de production, incluant des contenus aussi bien naturels que de synthèse (*i.e.* dessins animés). Les solutions proposées démontrent notamment que l'utilisation de la norme MPEG-7 à travers la plateforme INVENIO est tout à fait pertinente pour optimiser la réutilisation des contenus au sein d'une chaîne de production audiovisuelle.

Les perspectives de travail futur concernent tout d'abord la fusion des critères de recherche textuels et structurels de description. Cela permettrait de combler le décalage sémantique caractéristique des approches d'indexation par descripteurs visuels, en exploitant une connaissance métier. À plus long terme, il serait intéressant d'étendre la plateforme à d'autres types de médias, comme les vidéos et les données graphiques 3D. Enfin, il serait utile de convertir la plateforme en un environnement de travail collaboratif à distance. Dans ce contexte, l'intégration des technologies MPEG-21 constitue une piste privilégiée de développement futur.

Références

- [1] Gupta A., Jain R., Santini S., Smeulders A. W. M., Worring M., *Content-Based Image Retrieval at the End of the Early Years*, IEEE Trans. on PAMI, Volume 22, Issue 12, Décembre 2000.
- [2] Datta R., Li J., Wang J., Wang, J.Z., *Content-based image retrieval: approaches and trends of the new age Proc, 7th ACM SIGMM international workshop on Multimedia information retrieval*, p. 253 – 262, 2005.
- [3] ISO/ IEC 15938-5, Information technology - MultimediaContent Description. Interface - Part 5: Multimedia Description Schemes. 2003
- [4] Manjunath B. S., Salembier P., Sikora T., *Introduction to MPEG-7 : Multimedia Content Description Interface*, John Wiley & Sons, 2002.
- [5] Zaharia, T., Prêteux F., Descripteurs visuels dans le standard MPEG-7, chapitre dans Gestion des données multimédias (Chapitre 5), *Traité IC2 - Série Informatique et Systèmes d'Information*, Mostefaoui, A., Prêteux, F., Lecuire, V., Moureaux, J.-M. (Eds.), Editions Hermès-Lavoisier, Paris, 2004, p. 85-139.
- [6] ISO/IEC 21000. (2003). Information technology - Multimedia framework (MPEG-21).
- [7] Bormans, J., Hill, K. (2002). MPEG-21 Overview v.5. ISO/IEC JTC1/SC29/WG11/N5231, Security. ARES 2007, Vienna, Austria, 1235 – 1239.
- [8] XML Schema, Spécification disponible sur <http://www.w3.org/XML/Schema>.
- [9] ISO/IEC 15938-3: 2002, MPEG-7-Visual, Information Technology – Multimedia content description interface – Part 3: Visual, 2002.
- [10] Zaharia, T., Prêteux F., Normes de description des contenus multimédias. L'indexation multimédia - description et recherche automatique, *Traité IC2 - Série Traitement du Signal et de l'Image*, P. Gros, Editions Hermès-Lavoisier, Paris, 2007, p. 163-185.
- [11] Panorama de MPEG-21, disponible sur <http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm>.