

MPEG-7

Interface de description de contenu multimédia, la norme MPEG-7 sera finalisée en 2001. Elle définira une syntaxe et un vocabulaire de description du contenu de documents audiovisuels, et constituant une indexation pour la recherche et la sélection de l'information associée.

Les années 1990 ont été marquées par une révolution des technologies de l'information. L'avènement du numérique et des normes associées, la baisse des prix des équipements de traitement et de stockage de l'information, la mise en place de nouveaux réseaux en sont quelques faits marquants. Dans ce nouveau contexte, les quantités d'informations multimédia disponibles sont en nombre sans cesse croissant, éparpillées ou transitant sans ordre dans des sites et sur des réseaux du monde entier. Simultanément, leur usage se démocratise, s'étendant à de nouvelles catégories socioprofessionnelles, ainsi qu'au grand public. Il devient donc crucial d'organiser l'accès à ces informations. Idéalement, il s'agirait d'avoir un moyen universel, efficace, s'adaptant à des applications comme à des utilisateurs ou à des contenus très divers, qui permette à chacun de trouver ou de sélectionner l'information souhaitée dans son propre environnement applicatif : c'est le but de MPEG-7.

Le groupe de normalisation ISO/MPEG (*Moving Pictures Experts Group*) a une expertise reconnue dans le développement de normes pour les nouvelles technologies de l'audiovisuel. MPEG-1, MPEG-2 et MPEG-4 permettent de représenter l'information audiovisuelle sous forme numérique respectivement pour le stockage, la télévision numérique, ou tous types d'applications multimédia [5]. Avec MPEG-7, ISO/MPEG normalise maintenant le niveau des métadonnées : c'est-à-dire une représentation de la représentation de l'information.

Le premier paragraphe donne une vue d'ensemble du contexte et des objectifs de MPEG-7. Les deux suivants décrivent un peu plus finement les caractéristiques de la future norme. Les éléments constituants de MPEG-7, leur organisation globale et la terminologie associée sont exposés dans le paragraphe 2, et les futures fonctionnalités de la norme ainsi que les enjeux techniques impliqués sont résumés dans le paragraphe 3.

MPEG-7 : le contexte, les objectifs

Objectifs généraux

Il existe à l'heure actuelle des méthodes de description de contenu textuel, à base d'indexation par mots-clés. Mais rien n'existe encore sur la base des caractéristiques propres aux informations audiovisuelles (rythmes d'une musique, couleurs d'une image, mouvements dans une vidéo,...). De telles informations ne peuvent être indexées que par mots-clés, ce qui doit être fait manuellement, et limite grandement les possibilités offertes par leur description. MPEG-7

se fixe donc comme objectif de normaliser un langage et un vocabulaire de description de contenu multimédia qui ne soit pas uniquement textuel, mais qui comprenne comme *mots* des éléments directement représentatifs du contenu, tels les éléments donnés plus haut (rythmes...) [1].

Cette description doit être flexible, afin de pouvoir s'adapter à de nouveaux types de contenus et à de nouvelles applications. Ainsi, elle sera indépendante du format de l'information (par exemple analogique, MPEG-2...) car représentative de son contenu sémantique. En outre, le langage de description pourra non seulement manier du vocabulaire standard, mais aussi définir lui-même de nouveaux mots de manière standard. Ceci permettra notamment d'indexer efficacement sans pour autant sortir de la norme un contenu aux caractéristiques particulières, ou à utiliser dans un domaine bien spécifique.

MPEG-7 ne normalisera ni les mécanismes d'extraction des descriptions, ni les moteurs de recherche ou les mécanismes de filtrage d'information qui géreront l'accès aux informations (figure 1).

En effet, la normalisation de la description suffit à assurer l'interopérabilité de différents outils de recherche avec différentes données. Les outils de recherche et d'extraction associés aux descriptions standardisées seront bien sûr développés par MPEG, mais ils resteront des annexes informatives de la norme. Ainsi la concurrence pourra s'exercer dans ces domaines, qui pourront en outre profiter des futurs progrès des techniques concernées.

Le standard sera finalisé en novembre 2001 (tableau 1).

Applications

Le spectre des applications existantes pouvant bénéficier de MPEG-7 est très large [2], et de nouvelles applications en ayant l'usage se créeront sûrement dans les prochaines années. De fait, MPEG-7 peut offrir une

Tableau 1 – Échéances de la normalisation MPEG-7

Octobre 1998	Appel à propositions
Mars 1999	Départ de la phase collaborative du développement de la norme
Novembre 2001	Norme internationale

valeur ajoutée à toute application pour laquelle il est important de caractériser des informations multimédia et d'y accéder. En conséquence, MPEG-7 se fixe comme objectif de ne pas être assujetti à une application particulière, mais de pouvoir décrire les données multimédia pour tous types d'environnements applicatifs.

Cette généréricité peut en particulier se décliner selon diverses caractéristiques d'applications, que nous passons en revue dans la suite :

— **utilisateurs** : du grand public (non-experts ou connaisseurs) aux professionnels ;

— **média** : multimédia en général, qu'il s'agisse d'images fixes, d'audio (musique, dialogues...), de vidéo, ou de contenu audiovisuel ;

— **mode d'accès** : du *push*, dans le cas où l'utilisateur reçoit des descriptions MPEG-7 à partir desquelles il filtre l'information, au *pull*, dans le cas où l'utilisateur va consulter des descriptions MPEG-7 pour rechercher une information ;

— **position dans la chaîne de traitement du contenu** : tout au long de la chaîne, du début lors de la création, l'indexation et l'annotation du contenu, jusqu'à la fin lors de sa recherche dans une base de données ou de sa diffusion ;

— **types de contenu** : amateur ou professionnel, tous types de contenu (journaux télévisés, séquences de vidéosurveillance, photos de vacances, sitcoms...).

Pour citer concrètement quelques applications, on peut envisager d'utiliser MPEG-7 pour le stockage et la recherche dans des bases de données vidéo, la sélection et le filtrage de programmes télévisés, l'accès sélectif à des vidéos dans le cadre de la vidéosurveillance, l'indexation d'images médicales, l'édition de contenu multimédia, etc.

Etat de l'art

Recherches et produits existants

Les États-Unis sont tout particulièrement à la pointe de la recherche en indexation des informations audiovisuelles et recherche par le contenu [4]. L'État y finance depuis plusieurs années un programme de recherche

universitaire très important dans ce domaine. Les universités de Columbia (système VideoQ), Berkeley (système Chabot) Carnegie Mellon (projet Informedia) et d'Urbana Champaign en Illinois (système MARS) sont, pour ne donner que quelques exemples, particulièrement actives.

De nombreuses sociétés ont également des activités de recherche dans ce domaine, et suivent activement le développement de MPEG-7 : Philips, Sony, Sharp, IBM, Canon, Kodak, ou encore NHK, en sont quelques exemples.

En ce qui concerne les produits, certaines compagnies commercialisent déjà des produits permettant l'indexation de contenu audiovisuel puis sa recherche par des caractéristiques non textuelles. La plupart de ces produits ont d'abord été conçus pour des images, puis sont passés aux séquences vidéo. Parmi les plus connus, nous pouvons citer le système QBIC (*Query By Image Content*) d'IBM, et VIRAGE.

Autres normes

Rappelons ici que MPEG-7 n'est nullement une amélioration ou un concurrent des normes MPEG précédentes. En effet, MPEG-1, MPEG-2 et MPEG-4 étaient toutes trois des normes de compression de l'information audiovisuelle sous forme numérique [5], et non de représentation de ses caractéristiques sémantiques.

D'autres normes auront en revanche des liens plus complexes avec MPEG-7. Nous l'avons mentionné, il existe déjà des normes relatives à l'indexation de documents textuels. MPEG-7 ne s'attachera donc pas à définir quoi que ce soit de nouveau dans ce domaine. Cependant, afin de pouvoir indexer efficacement de tels contenus s'il s'en présente (sous-titres, script ou tout type de texte), MPEG-7 choisira l'une de ces normes et spécifiera un mécanisme d'interfaçage de sa syntaxe avec celle de la norme sélectionnée.

Il existe également certaines normes qui décrivent des informations multimédia pour des applications spécifiques, le plus souvent par des mots-clefs. Un exemple en est la spécification DVB-SI, qui définit les descripteurs de programmes télévisés pour la diffusion. MPEG-7 se place comme on l'a vu dans un

cadre plus vaste. Pour une application donnée, il sera donc possible d'avoir le choix entre deux standards de représentation mais dont les fonctionnalités et les philosophies seront sûrement différentes.

Enfin, MPEG-7 offrira une bonne complémentarité avec la norme FIPA, qui définit pour des sociétés d'agents intelligents un langage de communication et des structures adaptées à certaines applications : MPEG-7 devrait offrir à ce langage une ontologie adaptée aux applications audiovisuelles.

Architecture et éléments d'un système MPEG-7

Terminologie MPEG-7 : éléments à considérer

Pour bien comprendre le fonctionnement d'un système MPEG-7 il faut considérer les éléments suivants, présentés en utilisant la terminologie définie et utilisée par MPEG :

— les **données** : elles sont l'ensemble des contenus indexés par MPEG-7. Comme nous l'avons vu, ces données seront de préférence audiovisuelles, et dans ce cas disponibles sous n'importe quel format (VHS, MPEG...), mais peuvent être d'une autre nature, textuelles par exemple ;

— la **caractéristique** : on appelle caractéristique toute caractéristique de l'information à indexer pouvant avoir un sens pour certains utilisateurs ou pour certaines applications. La couleur d'une image, l'auteur d'un film, le rythme d'une musique en sont quelques exemples ;

— le **descripteur** (D) : un descripteur est un modèle permettant d'associer une valeur (ou un ensemble de valeurs) à une ou plusieurs caractéristiques, pour une même donnée. Un exemple classique est l'histogramme, pour les couleurs ;

— le **schéma de description** (DS) : c'est un modèle de plus haut niveau : il définit une structure et une sémantique pour les descripteurs et les relations entre descripteurs. On peut l'assimiler à une sorte de formulaire, dont les champs seraient des descripteurs ;

— la **description** : elle est l'entité décrivant la donnée. Elle est formée d'un schéma de description et des instantiations des descripteurs correspondants ;

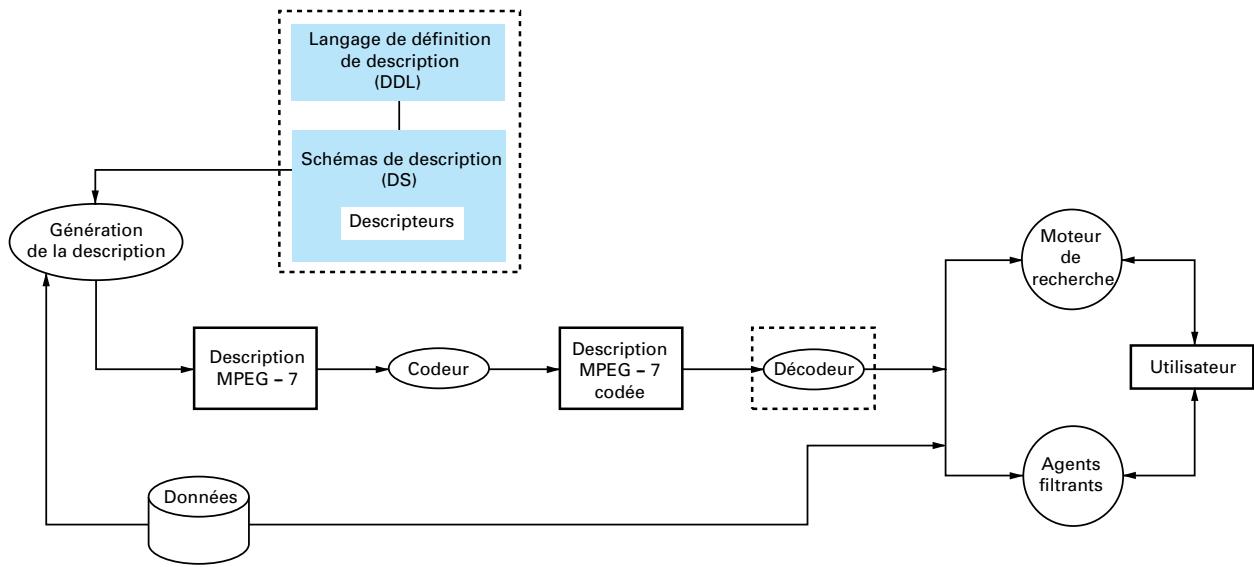


Figure 2 – Vue d'ensemble d'un possible système MPEG-7

- la **description codée** : elle est une représentation de la description optimisée pour le stockage et la transmission ;
- le **langage de définition de descriptions (DDL)** : le DDL est le langage utilisé pour spécifier descripteurs et schémas de description, normalisés ou non (cf. § *Objectifs généraux*).

également exprimées par le DDL, selon une syntaxe normalisée par MPEG-7.

Exemple d'architecture de systèmes MPEG-7

La figure 2 donne une vue d'ensemble, aussi générique et abstraite que possible, d'une éventuelle application MPEG-7. Ceci n'étant qu'un exemple, la présence de tous les éléments représentés ne sera pas nécessaire pour chaque application utilisant MPEG-7. Réciproquement, pour certaines applications, la présence d'autres éléments ou d'autres flux de données que ceux décrits ici pourrait être nécessaire.

Les pointillés entourent les éléments normatifs de ce système.

On peut notamment retrouver sur cette figure les deux types de mode d'accès, *push* et *pull*, décrits en début d'article. Il est en outre implicite que la description MPEG-7 peut être soit stockée, avec les données ou indépendamment, soit diffusée de façon synchrone avec les données qu'elle représente. Pour conclure, rappelons enfin que la description MPEG-7, exprimée à l'aide du DDL, pourra contenir des descripteurs ou des DS non normalisés. Dans ce cas, les méthodes nécessaires à l'utilisation de ces éléments (extraction, mesures de similarités...) seront

Fonctionnalités, enjeux techniques

Fonctionnalités

Les fonctionnalités de la future norme ont été définies d'après l'étude des applications qui pourraient bénéficier de MPEG-7 et de ce que MPEG-7 devrait pouvoir leur proposer pour leur offrir une valeur ajoutée significative. Elles sont nombreuses et très diverses. Le lecteur voulant les détailler peut se référer à [3]. Outre celles que nous avons déjà évoquées dans les précédents paragraphes, nous pouvons citer :

- la capacité à représenter les informations à différents niveaux de précision : disposer d'une description échelonnable suivant différents critères (niveaux de priorité parmi les caractéristiques présentes dans les schémas de description, niveaux de précision des sémantiques des descripteurs, niveaux de précision des valeurs d'instanciation des descripteurs, etc.) ;

- la capacité à offrir un support performant pour différents types de recherche : par similarités, par navigation, ou avec accès à des données complémentaires par exemple.

Enjeux techniques

De nombreux enjeux techniques seront à relever pour permettre la mise en œuvre d'applications utilisant efficacement MPEG-7. En analyse audiovisuelle, les principales études porteront sur la segmentation automatique des informations audiovisuelles en entités temporelles cohérentes à partir d'une donnée de départ quelconque, et surtout sur l'extraction dans ces entités de caractéristiques haut ou bas niveau, menant à des descripteurs. Du côté de l'accès à l'information, des études devront être menées sur les stratégies d'indexation et de recherche, notamment pour prendre en compte le volume des descriptions, qui, dès qu'il s'agit de vidéo par exemple, devient sans commune mesure avec celui associé aux traditionnelles données textuelles.

En outre, pour concevoir la norme elle-même, il faudra combiner les savoir-faire de deux communautés jusqu'alors relativement distinctes : celle de l'audiovisuel et celle des bases de données, afin de concevoir et de choisir le langage de définition des descriptions d'une part, et d'autre part les descripteurs et les schémas de description qui figureront dans la norme.

Sylvie JEANNIN

Ingénieur de recherche

Responsable du projet MPEG-7 au sein du groupe Image et Communication, Laboratoires d'Électronique Philips SAS.

DEMAIN DANS L'INDUSTRIE

Références bibliographiques

Documents [1] à [3] réactualisés en permanence sur
<http://drogo.cselt.stet.it/mpeg> :

- [1] MPEG-7 Context and Objectives.
- [2] MPEG-7 Applications.

[3] MPEG-7 Requirements.

[4] Special issue on Content-Based Image Retrieval Systems, IEEE Computer Magazine, vol. 28, n° 9, sept. 1995.

[5] FERT (E.), – MPEG-1, MPEG-2, MPEG-4. E 5 360 Traité Électronique. Techniques de l'Ingénieur (à paraître).

Normalisation

ISO/IEC JTC1/SC29/WG11 MPEG
<http://drogo.cselt.stet.it/mpeg>