

MPEG-7: un nouveau standard pour le multimédia



Alors qu'hier encore, la principale préoccupation était la compression des documents audiovisuels pour en permettre le stockage et l'échange, la question qui se pose aujourd'hui est la suivante : comment gérer toute cette information pour y accéder efficacement ?

Par **Valérie Gouaillier**
Agente de recherche
Vision et imagerie, R-D, CRIM

Télécharger des mégaoctets d'audio ou de vidéo de l'Internet, capter une foule de chaînes télé et radio numériques ou archiver des centaines de photos de famille font désormais partie de notre quotidien. À l'ère numérique, on assiste à une fulgurante prolifération de contenus multimédias, favorisée par la diminution des coûts d'archivage et l'amélioration de la vitesse de transmission dans les réseaux fixes et mobiles. Alors qu'hier encore, la principale préoccupation était la compression des documents audiovisuels pour en permettre le stockage et l'échange, la question qui se pose aujourd'hui est la suivante : comment gérer toute cette information pour y accéder efficacement ?

De plus, alors que le nombre de documents audiovisuels augmente, les nouvelles applications exigent un traitement et une interprétation de plus en plus poussés de l'information que ces contenus renferment. L'utilisateur de demain ne se contentera plus de rechercher par mots-clés dans une banque de documents. Il voudra, par exemple, retrouver une pièce dans ses archives musicales en fredonnant quelques notes du refrain. Il achètera un enregistreur vidéo numérique lui permettant de filtrer les reportages sportifs de son bulletin télévisé. Ou encore, il s'attendra à ce que son navigateur Internet bloque le téléchargement de tout document à contenu offensant pour ses enfants. Ces applications nécessitent que les systèmes analysent et « comprennent » le contenu du matériel audiovisuel pour y rechercher ou filtrer l'information pertinente.

Description du standard

C'est dans ce contexte que le Moving Picture Experts Group (MPEG) a développé MPEG-7, devenu standard ISO/IEC 15398 à l'automne 2001. Cette norme, formellement appelée *Multimedia Content Description Interface*, a été élaborée collectivement par des centres de recherche universitaires et des joueurs majeurs de l'industrie, dont IBM, Sony et Siemens. À la différence de ses frères MPEG-1, MPEG-2 et MPEG-4, qui standardisent l'encodage même des documents audiovisuels, MPEG-7 régit la description de ceux-ci et non pas l'encodage. La navigation, la recherche et le filtrage peuvent donc s'opérer sur ces descriptions plutôt que sur les fichiers sources, ce qui évite de devoir décompresser ces derniers et de les traiter à chaque requête.

Une description MPEG-7 est un fichier de métadonnées détaillant différents aspects du contenu et de la gestion d'un document audiovisuel. Par exemple, en ce qui concerne le contenu, on pourra retrouver, dans le fichier de description, des informations sur les caractéristiques de bas niveau (couleurs, textures, mouvement, contenu fréquentiel, mélodie, etc.) ou des concepts sémantiques de plus haut niveau (objets, événements, interactions entre les objets, lieux de l'action, etc.).

Les outils MPEG-7 permettent de décomposer un document audiovisuel en différents segments spatiaux, temporels ou spatio-temporels, offrant ainsi plusieurs niveaux de description. On pourra, par exemple, diviser

«L'utilisateur de demain ne se contentera plus de rechercher par mots-clés dans une banque de documents. Il voudra, par exemple, retrouver une pièce dans ses archives musicales en fredonnant quelques notes du refrain. Ou encore, il s'attendra à ce que son navigateur Internet bloque le téléchargement de tout document à contenu offensant pour ses enfants.»

un extrait vidéo en séquences et associer une description à chacune d'elles, ou encore décrire différentes parties d'une seule image. Les relations entre les divers segments peuvent également être exprimées.

L'information relative à la gestion des contenus multimédias est aussi assurée par un vaste ensemble d'outils. Outre les renseignements pertinents aux applications d'archivage, tels que les détails de création du document audiovisuel (titre, réalisateur, date de création, etc.), ses paramètres d'utilisation (heures de diffusion, prix, droits d'auteur, etc.) et son format d'encodage, MPEG-7 permet des descriptions plus spécialisées. Par exemple, un contenu peut être décrit en différents niveaux hiérarchiques ou séquentiels, afin de créer des résumés audiovisuels plus ou moins détaillés. Grâce à ce type de représentation, un usager pourrait demander à un système de ne lui présenter que les hauts faits d'un match sportif, comme les buts marqués. Il est aussi possible de décrire diverses variations d'un document audiovisuel ou encore d'élaborer des modèles statistiques de l'information. De surcroît, pour permettre d'adapter la diffusion aux habitudes de chaque usager, les préférences de celui-ci et l'historique d'utilisation du contenu peuvent être consignés dans la description.

Cependant, malgré sa grande richesse, MPEG-7 ne standardise que le format de description de toutes ces informations. Il ne spécifie pas comment obtenir ces caractéristiques automatiquement à partir des documents multimédias, ni comment utiliser les descriptions produites. Ceci laisse pleine latitude à l'industrie pour développer des applications novatrices dont l'interopérabilité est

```
<?xml version="1.0" encoding="UTF-8" ?>
<mpeg7 xmlns="urn:mpeg:mpeg7:schema:2001"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="urn:mpeg:mpeg7:schema:2001, MPEG7-2001.xsd">
  <Description@Mpeg7>
    <PrivateIdentifier>Exemple_de_description MPEG7/PrivateIdentifier </PrivateIdentifier>
  </Description@Mpeg7>
  <Description@Mpeg7 xsi:type="ContentEntityType">
    <MultimediaContent@Mpeg7 xsi:type="ImageType">
      <Image>
        <MediaLocator>
          <MediaURI>payes10.jpg@MediaURI</MediaURI>
        </MediaLocator>
        <TextAnnotation>
          <FreeTextAnnotation>Jardin de Clezard</FreeTextAnnotation>
        </TextAnnotation>
        <VisualDescriptor@Mpeg7 xsi:type="ColorLayoutType">
          <ColorQuant = "1">
            <Values>
              255 0 0 22 255 0 7 45
              255 255 255 184 255 254 108 27
              34 132 128 34 3 39 87 83
              18 24 98 99 93 80 50 38
            </Values>
          </VisualDescriptor>
        </Image>
      </MultimediaContent>
    </Description>
  </mpeg7>
```

Exemple de description MPEG-7 (format texte) d'une image.



assurée par la représentation standard des contenus qu'offre MPEG-7.

Structure du standard

Le principal objectif de MPEG-7 est de fournir un cadre normatif de description suffisamment générique et flexible pour s'appliquer à différents types de médias (images, vidéo, parole, musique, etc.) et convenir à un large éventail d'applications. C'est pourquoi ce nouveau standard, très modulaire, comprend une panoplie d'outils élémentaires qui, agencés, permettent de décrire les contenus à différents niveaux d'abstraction.

Les principaux outils sont les descripteurs qui constituent des représentations de caractéristiques, c'est-à-dire d'attributs distinctifs de l'information multimédia. Un descripteur définit la syntaxe et la sémantique de la représentation d'une caractéristique. Par exemple, la répartition des couleurs dans une image est un attribut pouvant être

mesuré et exprimé en valeurs numériques dans un format spécifié par un descripteur (*Color Layout*).

Afin d'élaborer des descriptions complexes, les descripteurs peuvent être combinés selon différentes structures appelées schèmes de description. Ceux-ci spécifient les relations entre les différentes composantes d'une description. Ainsi, un extrait musical pourra être décomposé en différents segments temporels dont chacun sera représenté par des descripteurs, tels que la tonalité (*Harmonic Instrument Timbre*), et des schèmes de description, tels que l'enveloppe mélodique (*Melody Contour Description Scheme*).

MPEG-7 comprend des descripteurs et des schèmes de description prédéfinis. Cependant, le langage de définition de description permet de créer de nouveaux schèmes de description et, si possible, de nouveaux descripteurs. Il s'agit du langage XML Schema, développé par le W3C,