
Support de transmodage de contenus multimédia dans MPEG-21

Etude et validation d'un outil de description

Mariam Kimiaei Asadi et Jean-Claude Dufourd

*Ecole Nationale Supérieure des Télécommunications
46 rue Barrault
75634 Paris, France
Département ComElec
mariam.kimiaei@enst.fr
jean-claude.dufourd@enst.fr*

RÉSUMÉ. Dans le processus d'adaptation de contenu d'un système de type Universal Multimedia Access, la possibilité de changer la modalité d'une ressource est un avantage important. Dans cet article nous présentons un outil de description de « transmodage » à la volée, dans le cadre de la norme MPEG-21. Nous montrons ensuite l'efficacité et la nécessité de cet outil de description par quelques exemples de cas d'utilisation et les applications correspondantes. Notre implémentation de ces applications est basée sur notre version du logiciel de référence de MPEG-21 à laquelle l'outil de description de transmodage a été ajouté. Nous décrivons également l'intégration de cette implémentation dans un projet européen IST. Enfin nous concluons en donnant une perspective de la future intégration des outils de description de ce type dans MPEG-21.

ABSTRACT. Transmoding is an important part in the content adaptation process of a Universal Multimedia Access system. In this paper we introduce a description tool for online transmoding in MPEG-21. We then give some use case scenarios for adaptation by transmoding. We also demonstrate the efficiency and the necessity of the transmoding description tool through some use cases and the corresponding applications. Our implementation of these applications is based on our approach of the MPEG-21 reference software to which the transmoding description tool has been added. Also we describe the integration of this implementation into an IST European project. Finally we end up by giving a perspective for future integration of transmoding-type description tools in MPEG-21.

MOTS-CLÉS : ressource multimédia, modalité, transmodage, conversion de modalité, Universal Multimedia Access.

Keywords: multimedia resource, modality, transmoding, transcoding, modality conversion, Universal Multimedia Access.

1. Introduction

Aujourd'hui, afin de pouvoir fournir et consommer des contenus multimédia de manière transparente, il est important de disposer d'une infrastructure d'adaptation. Sans une telle infrastructure, les créateurs de contenus et les fournisseurs de services font face à plusieurs problèmes pour fournir le contenu multimédia à leurs consommateurs. Ces problèmes sont la multiplication, souvent non réaliste, du nombre des versions de contenus à produire, à stocker et à servir, ainsi que l'impossibilité de servir à chaque client la version optimale pour sa configuration.

Une infrastructure complète d'adaptation aurait besoin des descriptions complètes de contenus, de contextes et des informations guidant et aidant les processus d'adaptation. Plusieurs cadres de standardisation comme W3C (World Wide Web Consortium) et MPEG (Moving Picture Experts Group) proposent des solutions pour aider et guider des infrastructures d'adaptation de contenus multimédia. Le cadre de travail de CC/PP (Composite Capabilities/ Preferences Profile) (Klyne G., *et al.*, 2004), défini par le W3C, représente un modèle pour la description des caractéristiques des terminaux et des préférences des utilisateurs pour guider les processus d'adaptation de contenu. MPEG-21 (Burnett I. *et al.*, 2003) est une norme de la famille MPEG dont un des buts est de normaliser un cadre pour l'adaptation de contenus multimédia. Aujourd'hui MPEG-21 et CC/PP restent tous les deux des solutions incomplètes. CC/PP ne peut pas répondre à tous les besoins d'une structure d'adaptation de contenus multimédia car il se limite uniquement à décrire un nombre restreint des caractéristiques de contexte telles que les caractéristiques des terminaux et des préférences des utilisateurs. MPEG-21 a également des restrictions qui concernent notamment la description des recommandations et suggestions de l'auteur de contenus sur l'adaptation d'une ressource ou d'un ensemble des ressources multimédia. Néanmoins MPEG-21 donne une solution plus complète qui permet de décrire au mieux les ressources, les contextes et les informations nécessaires à l'adaptation. C'est pour cette raison que nous avons choisi d'effectuer ce travail dans le cadre de MPEG-21. Dans les sections suivantes nous donnons une vision plus claire de CC/PP et MPEG-21.

L'infrastructure d'adaptation de MPEG-21 est basée sur plusieurs éléments, dont les contraintes et les caractéristiques de l'environnement dans lequel le contenu multimédia concerné est censé être utilisé. Les caractéristiques de l'environnement d'utilisation (en anglais UED: Usage Environment Description) contiennent les différentes contraintes telles que les capacités du terminal utilisateur, les caractéristiques du réseau, les préférences de l'utilisateur et de l'auteur ainsi que les contraintes imposées par n'importe quelle autre entité intermédiaire existant dans une chaîne de distribution de contenus multimédia. La décision optimale d'adaptation doit être prise en respectant toutes ces contraintes. L'adaptation peut consister en un transcodage, un transmodage, tout autre changement de paramètres du contenu original, ou une combinaison quelconque de ces types d'adaptation. Dans MPEG-21, il existe aussi des outils de description tels que AQoS (Adaptation

Quality of Service) et BSD (Bit Stream Description) qui aident à faciliter ou guider la tâche de l'adaptation d'une ressource par transcodage. Ces outils de description fournissent les paramètres et méta-données nécessaires au processus de transcodage d'une ressource multimédia. Actuellement dans MPEG-21 il n'existe pas d'outils de description pour guider les processus d'adaptation par transcodage.

Bien que l'adaptation par transcodage puisse être la solution à plusieurs problèmes fréquents d'adaptation de contenus multimédia, la plupart des recherches effectuées sur l'adaptation de contenu ont été limitées à l'adaptation par transcodage. En conséquence, le support de l'adaptation par transcodage n'a pas été suffisamment étudié. Récemment un travail, auquel nous avons contribué, a été effectué sur la normalisation des descripteurs donnant des informations sur l'adaptation par changement de la modalité. Ces descripteurs (en anglais : Modality Conversion Preferences) aident à prendre une décision optimale pour un changement de modalité d'une ressource, mais restent limités aux changements de modalité statique, c'est-à-dire à choisir une alternative existante d'une ressource dans une autre modalité et non pas le transcodage à la volée.

Le travail présenté dans cet article ne cherche pas à proposer ou optimiser des convertisseurs de modalité, mais a pour objectif de contribuer à l'adaptation par transcodage à la volée, en utilisant et exploitant les recommandations et informations fournies par l'auteur de contenu. Ces informations sont associées aux contenus concernés et sont nécessaires à l'adaptation par transcodage à la volée. Le concept de ce travail est comme suit : l'auteur fournit une ou des ressources pour un média donné, et pour étendre les possibilités d'adaptation, donne des informations supplémentaires rendant possible et guidant les transcodages, transmodages ou changements de paramètre. Ces informations supplémentaires sont présentées sous la forme de descripteurs auxiliaires des ressources. Dans cet article nous définissons un outil de description de transcodage à la volée dans MPEG-21 et nous montrons son efficacité et sa pertinence par plusieurs cas d'utilisation. Nous donnons tout d'abord une introduction courte sur CC/PP et MPEG-21 dans la section 2. Les sections 3 et 4 couvrent une étude sur des travaux existants dans ce domaine et des comparaisons avec ce travail. Nous définissons ensuite dans les sections 5 et 6 les notions de « modalité » et « transcodage » et le tableau des transmodages que nous avons considérés. Dans la section 7, en apportant plusieurs scénarios d'utilisation, nous discutons la nécessité des descripteurs de transcodage dans MPEG-21. Ensuite dans la section 8, nous présentons l'outil de description de transcodage et dans la section 9, nous présentons notre implémentation des descripteurs de transcodage, qui est basée sur notre version du logiciel de référence MPEG-21. Nous décrivons aussi l'ensemble des transmodages que nous avons étudiés et nous présentons l'architecture d'une application de transcodage dans le projet européen ISIS (Kamyab K. *et al.*, 2004) qui constitue une partie de notre recherche sur MPEG-21. Dans la section 10, nous donnons notre conclusion sur ce travail et finalement dans la section 11, nous traitons des perspectives de ce travail dans MPEG-21.

2. Les standards

Une chaîne de distribution de contenus multimédia consiste en la création, la production, la livraison et la consommation de contenus multimédia. Pour atteindre ces objectifs, le contenu doit être identifié, décrit, contrôlé et protégé. Le transport et la livraison du contenu s'établissent sur un ensemble hétérogène de terminaux et de réseaux dans lesquels des événements différents se produiront et peuvent donner lieu à l'envoi d'une trace (*event reporting*). Un tel mécanisme de distribution inclura une livraison fiable, une gestion des caractéristiques et des préférences des utilisateurs (tout en respectant la sécurité de ces données) et une gestion des transactions (financières).

Une infrastructure particulière est requise pour ce nouveau type d'utilisation de contenus multimédia. Une telle infrastructure permet d'assurer l'interopérabilité de systèmes fournissant des contenus multimédia, ainsi que la simplification (et si possible l'automatisation) des transactions. Ceci exige une vision partagée par tous les architectes de l'infrastructure. Cette infrastructure doit intégrer les technologies permettant la sécurité de la livraison de contenus, la sécurité de paiement ainsi que la gestion de droits.

Des cadres de standardisation du W3C (World Wide Web Consortium) et de MPEG (Moving Picture Experts Group) définissent des solutions pour le support d'une infrastructure d'adaptation de contenus multimédia. Dans les paragraphes suivants nous donnons des courtes introductions sur des solutions proposées par MPEG et W3C.

2.1. CC/PP et RDF

CC/PP est basé sur XML (Yergeau F., *et al.*, 2004) et RDF (Resource Description Framework) (Lassila O. *et al.*, 1999). CC/PP se limite à la description des contextes et permet uniquement de décrire des caractéristiques logicielles et matérielles d'un terminal et les préférences de l'utilisateur. RDF est également une norme du W3C, qui permet l'expression et l'utilisation plus intelligente des méta-données. Elle associe aux ressources un ensemble de méta-données les décrivant.

CC/PP et RDF restent des solutions incomplètes car elles se limitent à la description de très peu d'informations nécessaires à l'adaptation. Par exemples elles ne mettent pas à disposition des moyens pour la description des droits d'utilisation ou des paramètres nécessaires à l'adaptation. Des travaux basés sur CC/PP et RDF ont été effectués qui proposent des extensions à CC/PP. Dans la section 3, nous en donnons plus de détails.

2.2. MPEG-21

Le groupe MPEG a défini des standards qui donnent des moyens et fournissent des supports pour l'adaptation de contenus multimédia (Magalhães J., *et al.*, 2004).

MPEG-21 est une norme ISO de la famille MPEG, qui identifie et définit les éléments principaux d'une chaîne de livraison de contenus multimédia, les rapports entre ces éléments et les opérations effectuées. MPEG-21 définit également l'ensemble des interfaces.

La notion de base de MPEG-21 est le « Digital Item » (DI). Un Digital Item est une représentation numérique d'« un travail » et en tant que tel, c'est la chose sur laquelle on agit (contrôle, décrit, échange, adapte, etc.). Un DI contient à la fois des ressources multimédia et les descriptions de multiples facettes de ces ressources.



Figure 1. La chaîne de livraison de contenus multimédia de MPEG-21

La figure 1 présente l'infrastructure de MPEG-21 pour une chaîne de livraison de contenu multimédia. Le contenu original est fourni sous forme d'un DI. Des différentes entités MPEG-21 manipulent ce DI. Le DI adapté est finalement livré à l'utilisateur ayant un terminal conforme MPEG-21.

Quelques éléments principaux de MPEG-21 sont :

- DID (Digital Item Declaration) : un langage XML de description de DI (MPEG Group, 2003).
- DIA (Digital Item Adaptation) : un langage XML de description de contraintes et des adaptations possibles (Vetro, A., 2004). Dans DIA les descripteurs MPEG-7 sont utilisés à plusieurs reprises. MPEG-7 est une norme définie par MPEG pour la description des contextes et des contenus (Martínez José M. *et al.*, 2001).
- REL (Right Expression Language) : un langage XML d'expression des droits associés à un contenu (MPEG Group, 2004).
- RDD (Right Data Dictionary) : un langage XML pour la définition des droits, permettant en particulier de créer de nouveaux droits et ainsi d'étendre le domaine géré par REL (MPEG Group, 2004).
- IPMP (Intellectual Property Management and Protection) : des moyens fiables pour permettre au contenu d'être constamment contrôlé et protégé à travers une grande variété de réseaux et de dispositifs.

Un Digital Item est donc représenté par un document XML nommé DID, dans lequel il y a :

- des descripteurs de ressources, comme une vidéo ou un flux sonore
- des descripteurs DIA rendant possible des adaptations de ressources
- des descripteurs REL de droits d'accès pour le contenu complet et/ou pour chaque ressource
- des descripteurs RDD pour le cas où un droit non standard est exprimé.

Les caractéristiques du terminal sur lequel ce Digital Item doit être présenté, du réseau utilisé et de l'utilisateur avec ses droits et préférences, peuvent être représentés dans un ou plusieurs fragments de DID.

L'ensemble du DI et des contraintes est géré par un environnement qui doit obéir aux règles de l'IPMP.

3. Différentes approches et stratégies

Plusieurs travaux de recherche ont été effectués dans différents domaines de l'adaptation de contenus multimédia (Lei Z. *et al.*, 2001). Pour décrire les contenus multimédia et les contextes, la plupart de ces travaux sont basés sur MPEG-21, MPEG-7 ou CC/PP et RDF. Ces travaux ont pour but d'apporter et de proposer des supports pour les adaptations de type transcodage, adaptations de documents multimédia structurés ou adaptations statiques, et n'adressent pas la problématique du support de l'adaptation par transmodage à la volée. Dans cette section nous donnons une courte couverture de l'état de l'art de ces travaux, classifiée selon les types d'adaptation. Nous expliquerons finalement notre approche dans le cadre de travail du projet ISIS.

3.1. Adaptation statique par changement de modalité

Des travaux de recherche contribuant à MPEG-21 ont été effectués dans le but de trouver des supports nécessaires pour l'adaptation par changement de modalité (en anglais 'Modality Conversion') (Cong Thang T *et al.*, 2004). Les solutions proposées se limitent aux supports de la prise de décision et de l'adaptation statique et ne permettent pas des prises en charges de méta-données fournies par l'auteur et nécessaires au transmodage à la volée. Le projet européen AMEBICA (Viano G. *et al.*, 2000), quant à lui propose des architectures d'adaptation multimédia qui sont basées sur les systèmes d'*agents multiples* (multiple-agent) et utilisent des formats non-xml. Il implémente uniquement des outils d'adaptation multimédia statiques (*off-line*) ; ce qui signifie des outils qui sélectionnent la meilleure présentation d'un document multimédia selon les changements de statut de l'opérateur, de l'utilisateur et du système. Les points faibles d'AMEBICA

résident dans la difficulté de faire un modèle utilisateur efficace et exploitable et le manque d'adaptation à la volée.

Le modèle *InfoPyramid* (Mohan R. *et al.*, 1999) définit un système de transmission et d'adaptation de contenu multimédia composé d'objets vidéo, image, texte et audio. *InfoPyramid* manipule des variantes différentes des ressources média en différentes modalités et avec des niveaux de fidélité différents. Selon un nombre limité de contraintes, il sélectionne des alternatives optimaux. La ressource optimale dans *InfoPyramid* est la ressource qui satisfait les contraintes du contexte et qui a la *valeur de contenu* la plus élevée. Le support de contrainte du modèle est très faible et contient uniquement trois contraintes: la taille des ressources envoyées, le temps de chargement, et la taille de l'écran du dispositif.

3.2. Adaptation par transcodage à la volée

InfoPyramid définit aussi un système d'adaptation de contenu multimédia basé sur le transcodage (et aussi transmodage) à la volée. Le système essaie de définir un processus dynamique qui assure une sélection automatique des méthodes de transcodage (et de transmodage) pour satisfaire les contraintes du client.

La norme MPEG-21 met à disposition des moyens complets pour supporter des adaptations des contenus *scalables*. Les outils de description et d'adaptation AQoS et BSD aident et guident les processus de la prise de décision et d'adaptation. Concolato propose l'utilisation des moyens de MPEG-21 pour l'adaptation par transcodage de contenus graphiques BIFS (Concolato *et al.*, 2004). Le projet européen ISIS (Kamyab K. *et al.*, 2004), donne aussi des solutions de transcodage, basées sur MPEG-21, pour l'adaptation des contenus *scalables*. D'autres travaux sur l'adaptation par transcodage utilisent la base de CC/PP et RDF. NAC (Lemlouma, 2004) propose des solutions d'adaptation par transcodage à la volée. NAC utilise UPS (Lemlouma *et al.*, 2002) qui définit des extensions de type CC/PP et RDF afin de couvrir les points faibles de CC/PP.

3.3. Adaptation d'un ensemble structuré de ressources par XSLT

Des travaux basés sur XSLT sont effectués pour l'adaptation des documents multimédia structurés, c'est-à-dire des scènes ou des compositions multimédia qui contiennent un ensemble de ressources avec, entre elles, des relations spatiales et temporelles. NAC utilise des feuilles XSLT qui se mettent à jour selon des variations du contexte pour l'adaptation des documents multimédia structurés.

Un autre cadre de travail sur l'adaptation sémantique des documents structurés a été développé (Euzenat J. *et al.*, 2003). Ce travail se charge principalement de l'adaptation de documents multimédia en prenant en compte les relations temporelles entre les media.

Une autre approche développée pour l'adaptation sémantique des scènes (compositions) multimédia est d'introduire des informations sémantiques sur la structure du document lors de la création de la présentation (Rousseau F. *et al.*, 2001).

3.4. Notre approche dans le cadre du projet ISIS

Le travail de cet article a été effectué dans le cadre du projet européen ISIS (Intelligent Scalability for Interoperable Services) (Kamyab K. *et al.*, 2004).

Le but du projet ISIS est de définir, implémenter et valider un cadre de travail qui permet la création et ensuite l'adaptation des contenus audio-visuels à une large variété de scénarios de service ou de contextes. L'adaptation de contenu dans ISIS se fait selon différentes caractéristiques de transport, de capacités du dispositif et de préférences de l'utilisateur et du créateur du contenu.

En comparant ISIS avec les autres travaux cités dans la section précédente, ISIS apporte une vérification et une description plus complète d'un ensemble plus large de contraintes, une adaptation à la volée par transmodage ainsi que transcodage, et une prise en compte des recommandations et des suggestions du créateur de contenu lors de l'adaptation. Plusieurs types d'adaptation sont implémentés dans ISIS, parmi lesquels : transcodage de contenu vidéo, audio, graphiques 2D et 3D et transmodage. Les points faibles d'ISIS sont la non prise en compte des changements dynamiques de valeurs des contraintes et l'absence d'une adaptation efficace des documents multimédia structurés (ce que l'on appelle les scènes dans MPEG-4).

Dans ISIS le changement dynamique (au cours de l'adaptation) du contexte (l'ensemble des contraintes) n'est pas considéré. Autrement dit, au début de la procédure de l'adaptation, les contraintes sont prises en compte une fois par le moteur de la prise de décision. Dans les cas où un changement de statut des contraintes survient, les nouvelles valeurs des contraintes ne sont prises en compte que pour la prochaine transmission. Ceci est le point faible d'ISIS. Le projet européen DANAE (<http://danae.rd.francetelecom.com>) qui est la continuation d'ISIS et auquel nous participons est basé sur l'adaptation dynamique.

4. Modalités

Pour définir le transmodage, nous devons dans un premier temps adopter une définition exacte de «la modalité» d'une ressource multimédia ainsi qu'une liste de modalités.

La modalité a au moins deux significations. Au niveau perceptuel, des modalités sont attachées aux cinq sens humains, ainsi il y a seulement «une» modalité visuelle. Au niveau structurel, il peut y avoir plusieurs modalités dans une même

modalité perceptuelle. Par exemple la modalité perceptuelle « visuelle » peut inclure vidéo « bitmap » (c'est-à-dire décrite point par point), image « bitmap » et deux sortes d'images vectorielles (graphiques 2D et graphiques 3D).

Bernsen *et al.* ont défini un arbre de taxonomie pour les modalités représentables qui contient quatre niveaux (Bernsen N. O. *et al.*, 1994). Les modalités définies dans cette taxonomie sont extrêmement détaillées au niveau de la sémantique et sont donc des modalités sémantiques et non pas des modalités physiques (perceptuelles) ou structurelles.

Dans MPEG-21, le niveau de modalité significatif du point de vue de l'adaptation est le niveau structurel, puisque la structure impose les types d'algorithmes qui peuvent être appliqués aux ressources.

Nous proposons une approche hiérarchique pour la classification de modalité pour le transmodage. Dans cette classification, nous avons cinq modalités principales et quelques modalités secondaires pour certaines d'entre elles :

- Vidéo
- Audio (Audio2D, Audio3D, Parole)
- Image
- Graphiques (Graphiques2D, Graphiques3D)
- Texte

Depuis récemment, MPEG-21 utilise une version étendue du *ContentCS* de MPEG-7 pour la classification des modalités. Cette classification est basée sur notre approche mais contient aussi des modalités multiples comme la modalité « audiovisuelle ». Notre classification des modalités est un sous-ensemble de cette classification qui ne contient que les mono-modalités.

NOTE — Dans cet article, la modalité *graphique* se rapporte à des graphiques vectoriels et la modalité *image* se rapporte à des images bitmap. Nous distinguons les modalités graphiques 2D et 3D car les modèles utilisés dans ces deux types de descriptions et les algorithmes de rendu sont fondamentalement différents.

5. Transmodage

Le concept de *transmodage* se rapporte à la conversion d'une ressource multimédia en changeant sa modalité initiale en une autre modalité. Par conséquent, nous définissons le transmodage comme un type d'adaptation numérique de ressource qui change la modalité de la ressource originale.

Dans cet article, le transmodage fait référence à une adaptation à la volée ou à la demande, et non pas à la substitution d'une ressource multimédia par des versions alternatives existantes en d'autres modalités. Dans MPEG-21, cette substitution est souvent appelée conversion de modalité (*modality conversion*).

Dans le tableau de la figure 2, tous les transmodages possibles entre les modalités définies précédemment sont illustrés. Les cases blanches présentent les transmodages qui sont pertinents et ont des paramètres généraux qui ne sont pas dépendants d'algorithmes particuliers. Les cases grises présentent les autres transmodages. Les transmodages qui n'ont pas été étudiés dans ce travail ont besoin d'être étudiés par les experts de la modalité concernée : les cases en question contiennent un point d'interrogation.

Entrée \ Sortie		Audio			Image	Vidéo	Graphiques		Texte
		Parole	Audio 2D	Audio 3D			Graphiques2D	Graphiques3D	
Audio	Parole	■	?	?	■	■	■	■	?
	Audio 2D	?	■	?	■	■	■	■	■
	Audio 3D	?	?	■	■	■	■	■	■
Image		*	■	■	■	√	√	√	√
Vidéo		*	■	■	■	■	√	√	√
Graphiques	Graphiques 2D	*	■	■	■	√	■	√	√
	Graphiques 3D	*	■	■	■	■	■	■	?
Texte		?	■	■	■	■	■	■	■

* Ces transmodages peuvent être faits de différentes manières : une façon serait la conversion de la parole au texte puis du texte à l'image/vidéo/graphiques ; l'autre manière serait la conversion de parole à une représentation 3D de langage des signes.

Figure 2. Le tableau de transmodages à la volée

6. Quelques scénarios d'utilisation

Le transmodage est nécessaire dans divers scénarios d'adaptation de contenu multimédia. Voici un exemple de transmodage vidéo-vers-image et vidéo-vers-graphiques2D : l'auteur d'une ressource originale vidéo n'a pas fourni les versions

image ou séquence d'images, mais a donné ses intentions au sujet du transmodage vers image ou séquence d'images (c'est-à-dire la modalité graphiques2D) par le moyen du descripteur de transmodage. Pour la conversion de vidéo-vers-graphiques2D, l'auteur a déterminé les trames vidéo qui sont à employer en tant qu'images, les transitions entre ces images et les durées de présentation pour chaque image.

Un exemple concret de ce scénario est le suivant : la ressource vidéo originale est une bande annonce d'un film qui est conçu pour la transmission sur ADSL mais n'est pas appropriée pour la diffusion sur réseaux cellulaires. Dans l'expression correspondante de REL, le transmodage vidéo-vers-graphiques2D (séquence d'images) est autorisé. La décision optimale d'adaptation est la conversion de cette vidéo en une séquence d'images. Une fois cette décision prise, afin d'effectuer l'action de transmodage, les données nécessaires sont extraites du descripteur DIA. Par exemple, la résolution d'affichage et le format de l'image sont récupérés du descripteur *TerminalCapabilities* (ISO/IEC 21000-7:2004, 2004). Les contraintes sur la modalité et les priorités des modalités sont récupérées dans le descripteur *ModalityConversionPreference* ou *PresentationPriorityPreferences* (ISO/IEC 21000-7:2004, 2004). La bande passante disponible qui est nécessaire pour déterminer le niveau de compression de la séquence d'image est récupérée dans le descripteur *NetworkCharacteristics* (ISO/IEC 21000-7:2004, 2004). Et finalement la liste des trames à employer, leurs importances (à prendre en compte dans les cas où il y a des limitations de bande passante) et les durées relatives sont récupérées dans le descripteur *transmoding*.

Voici également un scénario de cas d'utilisation pour le transmodage texte-vers-image : la ressource à adapter est une ressource texte, par exemple, un message en langue Persane (ou avec une police spéciale). Le terminal est très simple et ne dispose pas de la police persane, mais peut présenter des images. Dans ce cas, si autorisé dans l'expression de REL, une solution est de transmoder le texte en une image. Le processus de transmodage exige la connaissance des caractéristiques de la police (afin d'afficher le texte dans une image), l'information sur les couleurs, etc. Le même cas d'utilisation peut être considéré pour un transmodage de type texte-vers- vidéo, quand le texte ne rentre pas dans l'écran du dispositif (dont la taille est extraite à partir du descripteur *TerminalCapabilities*), et a ainsi besoin de défiler sur l'écran. Ce processus de transmodage exige en plus la connaissance des préférences d'auteur sur le type de défilement du texte : défilement horizontal sur une ligne, vertical sur plusieurs lignes, présentation par paragraphes successifs...

7. Outil de description de transmodage

Nous présentons ici un outil de description pour transmodage. Nous nous intéressons plutôt au point de vue du fournisseur, mais n'importe quelle entité dans la chaîne d'adaptation peut exprimer ses préférences sur les paramètres d'une

conversion spécifique par le biais de ce descripteur. Dans ce cas, afin de gérer les valeurs des paramètres, un niveau de priorité peut être attribué à chaque entité. Le fournisseur, à partir de sa connaissance des ressources, peut fournir quelques conseils ou recommandations de transmodage pour guider l'adaptation; ceci sera fait par les descripteurs proposés. Les données de transmodage incluent les descriptions des paramètres de transformation les plus généraux, c'est-à-dire que les descripteurs ne sont liés à aucun algorithme particulier. Un descripteur de transmodage n'est pas non plus censé contenir d'expression des droits ou des permissions liés au transmodage. Les expressions correspondantes de droit et d'autorisation seront données par MPEG-21 REL. La figure 3 présente un exemple concret du descripteur *transmoding* pour la conversion texte-vers-image. Dans le descripteur *transmoding* de la conversion *text visualization*, sont exprimées, la modalité finale, les caractéristiques (recommandées) de la police du texte et la taille de l'image finale. Si, après vérification des contraintes, un transmodage texte-vers-image doit être appliqué, le convertisseur texte-vers-image utilisera ces informations afin de transformer le texte en une image.

```

<Transmoding>
  <TransmodingMethodParameters
    xsi:type="TextVisualizationParametersType" >
    <To href="urn:mpeg:mpeg21:cs:ContentCS:2003:2">
      <mpeg7:Name>Image</mpeg7:Name>
    </To>
    <FontParameters fontSize="20">
      <FontStyle href="urn:mpeg:mpeg21:cs:FontStyleCS:2003:2">
        <mpeg7:Name>BOLD</mpeg7:Name>
      </FontStyle>
      <FontFamily href="urn:mpeg:mpeg21:cs:FontFamilyCS:2003:2">
        <mpeg7:Name>ARIAL</mpeg7:Name>
      </FontFamily>
    </FontParameters>
    <TextColorInformation foregroundColor="0.2 0.5 0.6"
      backgroundColor="0 0 0"/>
    <Resolution horizontal="300" vertical="400"/>
  </TransmodingMethodParameters>
</Transmoding>

```

Figure 3. Exemple du descripteur de transmodage pour la conversion texte-image

La figure 4 présente l'utilisation du descripteur *transmoding* pour un exemple de conversion vidéo-vers-graphiques2D. Le descripteur *transmoding* de la conversion *video summarization*, contient la modalité finale, les trames (à extraire) et les durées recommandées, ainsi que l'importance relative de chaque trame. L'attribut *importance* peut guider le moteur d'adaptation à choisir les trames les plus importantes, au cas où la bande passante serait limitée. Après avoir vérifié les contraintes, si la décision optimale d'adaptation est un transmodage vidéo-vers-*slideshow*, le moteur d'adaptation utilisera ces informations pour convertir la vidéo en une séquence d'images.

Le rôle du descripteur de transmodage dans DIA est identique à celui des descripteurs de caractéristiques de réseau ou de terminal qui, selon les contraintes données, contiennent les données nécessaires pour le processus d'adaptation.

```

<Transmoding>
  <TransmodingMethodParameters
    xsiType="SummarizationParametersType">
      <To href="urn:mpeg:mpeg21:cs:ModalityCS:2003:5">
        <mpeg7:Name>Graphics</mpeg7:Name>
      </To>
      <Slide importance="high">
        <mpeg7:MediaTimePoint>T01:14:30:12F24</mpeg7:MediaTimePoint>
        <mpeg7:MediaDuration>PT01M</mpeg7:MediaDuration>
      </Slide>
      <Slide importance="low">
        <mpeg7:MediaTimePoint>T01:15:50:10F24</mpeg7:MediaTimePoint>
        <mpeg7:MediaDuration>PT01M</mpeg7:MediaDuration>
      </Slide>
      .
      .
      .
      <Slide importance="high">
        <mpeg7:MediaTimePoint>T01:11:34:19F24</mpeg7:MediaTimePoint>
        <mpeg7:MediaDuration>PT01M</mpeg7:MediaDuration>
      </Slide>
    </TransmodingMethodParameters>
  </Transmoding>

```

Figure 4. Exemple du descripteur de transmodage pour la conversion de vidéo vers graphiques2D

La figure 5 montre comment le descripteur *transmoding* de DIA peut être employé dans une instance de DID. Le descripteur *transmoding* est associé à une ressource.

L'outil de description de transmodage permet également d'attribuer les paramètres de transmodage à un ensemble de ressources. Pour faire cela, DIDL dispose des moyens d'avoir plusieurs *Components* dans un seul *Item* au quel on peut associer un seul *Descriptor* contenant les paramètres de transmodage qui s'appliquent à toutes les ressources mentionnées dans cet *Item*.

9. Implémentation

Afin de mettre en application l'outil de description de transmodage, dans une première phase nous avons développé un module de gestion des descripteurs de transmodage dans le logiciel de référence de MPEG-21. Dans une deuxième phase, un ensemble de convertisseurs de modalité a été intégré dans ce logiciel. L'ensemble du logiciel a été proposé à MPEG-21 comme résultat d'un « Core Experiment » sur le transmodage. Cette implémentation est intégrée dans le projet ISIS.

```

<DIDL>
  <Item>
    <Component>
      <Descriptor>
        <Statement mimeType="text/xml">
          <dia:ConversionInformation>
            <dia:Conversion>
              <dia:Transmoding>

<!-- ici il y aura les descripteurs transmodage exprimant les
recommandations et les conseils du fournisseur concernant la
conversion de cette vidéo en d'autres modalités !-->

              </dia:Transmoding>
            </dia:Conversion>
          </dia:ConversionInformation>
        </Statement>
      </Descriptor>
      <!-- ici il y aura la référence à la ressource initiale -->
      <Resource mimeType="video/video"
        ref="myInitialVideo.mp4" />
    </Component>
  </Item>
</DIDL>

```

Figure 5. Une instance de DID avec le descripteur transmoding

Il est évident qu'afin d'améliorer les performances du système, les *proxys* et *caches* peuvent être introduits dans l'architecture. Néanmoins le travail présenté dans cet article n'a pas pour but d'optimiser les performances du système au niveau d'implémentation, mais de proposer des supports descriptifs pour les adaptations de type de transmodage à la volée. Aussi, ce travail a été effectué dans le cadre du projet européen ISIS dont l'architecture logicielle est basée sur un système client-serveur, et dans lequel les requêtes des clients sont traitées par le serveur. En tant que suite du projet ISIS, le projet européen DANAE auquel nous participons, donne des solutions pour les systèmes d'adaptation distribués.

Le module Transmodage d'ISIS est présenté dans la figure 6 et se compose de deux blocs principaux : le Décideur et le Transmodeur de ressources.

Les entrées du module Transmodage sont :

- Le choix de contenu de l'utilisateur transmis par le Streaming Server, qui reçoit la requête initiale de l'utilisateur (étape 1).
- La description du contenu venant d'une base de données de descriptions de contenus (étape 2).
- La description des contraintes utilisateur et des caractéristiques de l'environnement d'utilisation (réseau, terminal, etc.) venant d'une base de données provisoire (étape 3).
- Le contenu à adapter, venant d'une base de données de ressources (étape 4). La ressource adaptée est finalement renvoyée au Streaming Server (étape 5).

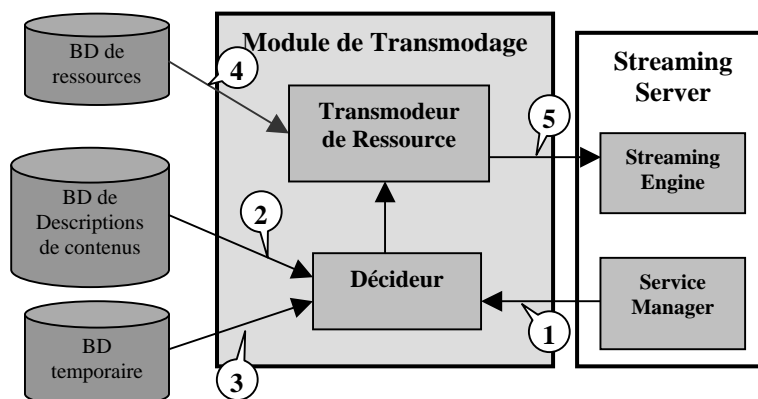


Figure 6. L'architecture du module de transcodage dans ISIS

Dans les paragraphes suivants nous décrivons le fonctionnement ainsi que l'architecture des modules décideur et transmodeur de ressources.

9.1. Décideur

Pour un média donné, le DID peut contenir plusieurs possibilités, à savoir des ressources réelles, comme une vidéo ou une image, et des ressources virtuelles, comme la définition d'une séquence d'images à construire à partir d'une vidéo. Le décideur implémente l'algorithme de prise de décision entre les diverses formes possibles d'un média selon les contraintes applicables. Il a été entièrement développé en langage Java, et comme présente la figure 7, comporte trois modules principaux :

- L'analyseur DID (*DID Parser*)
- L'analyseur DIA (*DIA Parser*)
- Le Moteur de Prise de Décision de Transcodage

Nous avons utilisé un *DID Parser* développé par Ghent University (<http://www.titr.uow.edu.au/cgi-bin/mpeg-ref-sw.pl>). Il prend comme entrée le DID décrivant le contenu et l'analyse afin de récupérer l'information concernant la(les) ressource(s) incluse(s) dans le DI. Cette information inclut le nombre de ressources, les références à la localisation des ressources et les descripteurs de transcodage associés aux ressources.

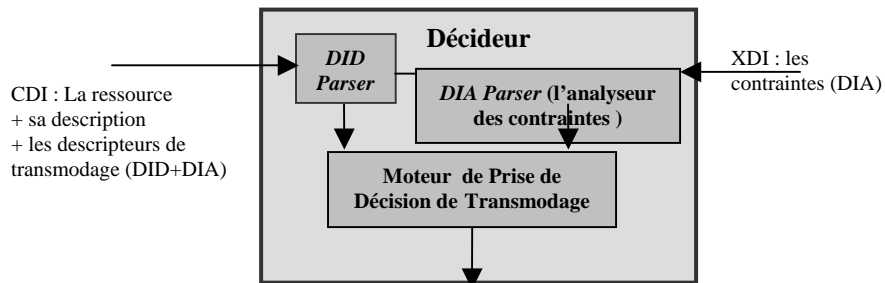


Figure 7. *L'architecture du Décideur*

Nous avons ensuite développé un *DIA Parser*, qui analyse les descripteurs DIA, auquel nous avons ajouté la gestion du descripteur *transmoding*. Ce logiciel a été développé en Java et avec une approche DOM (Document Object Model).

Le moteur de prise de décision de transmodage prend comme entrée les contraintes stockées dans une structure interne et les références aux ressources. Des priorités statiques sont attribuées à toutes les contraintes selon leurs valeurs sémantiques. Par exemple les capacités du terminal sont plus prioritaires que les préférences de l'utilisateur. Le moteur de prise de décision vérifie d'abord les contraintes de priorité les plus élevées.

Le processus de prise de décision comporte deux étapes principales. Dans la première étape sont prises en compte toutes les contraintes du contexte (appelé XDI : *contexT Digital Item*) selon leurs priorités, par exemple la non prise en compte d'une modalité dans le descripteur *TerminalCapabilities* ou l'incompatibilité d'une modalité avec la bande passante décrite dans le descripteur *NetworkCharacteristics*. Dans la deuxième étape, parmi les options (adaptations) techniquement faisables, la décision prise est optimisée en utilisant la description du contenu original (appelé CDI : *ContenT Digital Item*) qui contient aussi les descripteurs *transmoding*. Enfin, si la décision optimale est de faire un transmodage, le module décideur renvoie cette décision (le type de transmodage) ainsi que les paramètres nécessaires au transmodage (extraits du descripteur *transmoding*) au module transmodeur de ressources.

9.2. *Transmodeur de ressources*

Le transmodeur de ressources d'ISIS est une implémentation des outils de transmodage suivants :

- Convertisseur de texte vers image
- Convertisseur de graphiques2D vers vidéo

- Convertisseur de graphiques2D vers image
- Convertisseur de vidéo vers graphiques2D (slide-show)

Tous ces outils de transcodage ont été développés en Java ou C++. L'entrée du transcodeur de ressources est la décision de transcodage, la référence de la ressource à transcoder et les paramètres exigés pour le transcodage relatif.

Convertisseur de texte vers image : Ce module est un programme simple en Java qui fait le rendu d'un texte dans une image en utilisant des paramètres du transcodage du texte vers l'image, tels que la taille d'affichage, les paramètres de police, les paramètres de couleur. Un encodeur PNG a aussi été intégré dans ce module.

Convertisseur de graphiques2D vers vidéo : Ce module est un outil qui convertit une animation graphique MPEG-4 2D (sans audio ni image ni vidéo) vers une vidéo MPEG-4 Simple Profile avec une fréquence de trame donnée comme paramètre de transcodage. Cet outil a été développé en langage C++.

Convertisseur de graphiques2D vers image : ce module est un outil qui convertit une animation graphique MPEG-4 2D (sans audio ni image ni vidéo) en une image bitmap en utilisant le paramètre de numéro de trame (time point), donné comme paramètre de transcodage. Cet outil a été développé en langage C++.

Convertisseur de vidéo vers graphiques2D : Ce module convertit une vidéo MPEG-4 Simple Profile vers un contenu de la modalité graphiques2D contenant une séquence d'image. C'est très utile dans les cas où l'utilisateur n'est pas autorisé ou n'est pas en mesure de recevoir les contenus vidéo sur son terminal ; par exemple son lecteur MPEG-4 est seulement capable de jouer des contenus de type MPEG-4 2D contenant des médias de la modalité image, audio et graphiques2D.

L'architecture du module transcodeur est présentée dans la figure 8.

9.3. Quelques mesures

Nous avons mesuré les temps d'exécution de trois outils de transcodage. Au niveau du serveur, le système d'exploitation est Windows 2000 et le processeur est un AMD Athlon™ XP 2000. Au niveau du client, la dimension de l'écran du dispositif est 176x144 pixels. Les temps d'exécution moyens sont :

- 2800 ms pour la conversion d'un contenu graphique (15s, 12fps) en vidéo (MPEG-4)
- 420 ms pour la conversion d'un contenu texte en image (avec l'encodage en PNG)
- et 15 ms pour la conversion d'un contenu graphique en image.

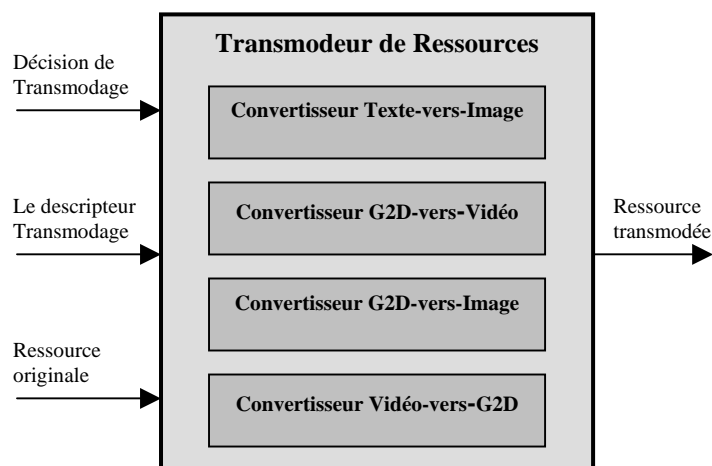


Figure 8. *Module de Transmodeur de Ressources*

NOTE. — Le but du travail présenté dans cet article est de proposer des solutions conformes à MPEG-21 pour la description des informations nécessaires aux transmodages à la volée et non pas de développer ni d’optimiser des convertisseurs de modalité ou le coût de transmodage. Nos outils de transmodage ne sont donc pas optimisés et ces chiffres sont uniquement fournis à titre indicatif. De même, les mesures plus étoffées sur le coût de transmodage ne peuvent pas être données.

10. Conclusion

Dans ce travail, l’outil de description de transmodage s’est avéré nécessaire pour réaliser l’adaptation de contenus au sein d’une chaîne de distribution MPEG-21. Cet outil de description, qui est un complément à la description des ressources faite par l’auteur, est conçu pour que l’auteur exprime ses préférences et conseils sur différents paramètres de transmodage.

Grâce à notre logiciel d’application, nous avons montré que la présence de l’outil de description de transmodage dans MPEG-21 DIA est cruciale pour permettre et guider la conversion des modalités de ressources dans le processus d’adaptation.

11. Perspectives

Dans MPEG-21 DIA jusqu'à présent, seules sont décrites les contraintes et les informations nécessaires à l'adaptation, par exemple le descripteur *AdaptationQoS* (ISO/IEC 21000-7:2004, 2004). La détermination des paramètres optimaux d'adaptation y est un processus complexe. Par exemple, il est possible de spécifier une table des paramètres d'adaptation à utiliser pour obtenir une liste de débits pour une vidéo. Mais il n'est pas possible de décrire un transcodage simple comme de GIF à JPEG.

Les descripteurs concrets que nous proposons mènent au contraire à un processus de décision simple et le type et les paramètres d'adaptation découlent simplement de nos descripteurs.

Suite à notre travail sur l'outil de description de transcodage, le besoin d'un descripteur concret pour la description des différents types de conversions a été récemment identifié dans MPEG-21. L'outil de description de transcodage a été considéré comme un exemple d'une telle description concrète et constitue la base d'une extension de DIA à n'importe quels autres types d'adaptation de ressource (transcodage, transcodage ou tout autre changement de paramètres).

Nous participons actuellement à un amendement de DIA, qui constitue la partie 7 de MPEG-21, sur les descripteurs de conversion.

Notre futur travail concernera donc ce nouvel aspect et inclura la recherche d'un outil de description plus complet que le descripteur transcodage ainsi que son implémentation.

Les auteurs tiennent à remercier les membres de la plate-forme MER de l'ENST pour leur soutien et plus particulièrement Cyril Concolato et Jean Le Feuvre pour leurs contributions à ce travail, ainsi que tous les partenaires du projet européen ISIS pour leur coopération.

12. Bibliographie

- Bernsen, N. O., « Foundations of multimodal representations. A taxonomy of representational modalities », *Interacting with Computers*, vol. 6 n° 4, 1994, p. 347-371.
- Burnett I., Van de Walle R., Hill K., Periera F. « Mpeg-21 goals and achievements », *IEEE MultiMedia*, vol. 10,n° 4, octobre-décembre 2003, p. 60-70.
- Concolato C., Dufourd J-C., « Adaptation de contenu MPEG-4 BIFS suivant la norme MPEG-21 », *Mcube Multimedia et Mobilité*, Montbéliard, 30-31 mars 2004.
- Cong Thang T., Ju Jung Y., Wook Lee J., Man Ro Y., « Modality conversion for universal multimedia services », *International Workshop on Image Analysis for Multimedia Interactive Services*, Lisbonne, 21-23 avril 2004.

- Euzenat J., Layaïda N., Dias V., « A semantic framework for multimedia document adaptation », *Proceeding of 18th International Joint Conference on Artificial Intelligence (IJCAI), Acapulco (MX), San-Mateo (CA US), 2003*, p. 31-36.
- MPEG Group, ISO/IEC TR 21000-1:2001 Information technology - Multimedia framework (MPEG-21) - Part 1: Vision, Technologies and Strategy, 2001.
- MPEG Group, ISO/IEC 21000-2:2003 Information technology - Multimedia framework (MPEG-21) - Part 2: Digital Item Declaration, 2003.
- MPEG Group, ISO/IEC 21000-3:2003 Information technology - Multimedia framework (MPEG-21) - Part 3: Digital Item Identification, 2003.
- MPEG Group, ISO/IEC 21000-5:2004 Information technology - Multimedia framework (MPEG-21) - Part 5: Rights Expression Language, 2004.
- MPEG Group, ISO/IEC 21000-6:2004 Information technology - Multimedia framework (MPEG-21) - Part 6: Rights Data Dictionary, 2004.
- MPEG Group, ISO/IEC 21000-7:2004 Information technology - Multimedia framework (MPEG-21) - Part 7: Digital Item Adaptation, 2004.
- Kamyab K., et al. « ISIS: Intelligent Scalability for Interoperable Services », *Conference on visual media production CVMP 2004*, Londres, Angleterre, 15-16 mars 2004.
- Klyne G., et al. (eds), *Composite Capability/Prefrence Profiles (CC/PP) : Structures and Vocabularies 1.0*, W3C Recommendation, 2004.
- Lassila O., Swick R.R. (eds), *Resource Description Framework (RDF) Model and Syntax Specification*, W3C Recommendation, 1999.
- Lei Z., Georganas N. D., « Context-based Media Adaptation in Pervasive Computing », *Proceeding of Canadian Conference on Electroic and Computer Engineering (CCECE'2001)*, Toronto, mai 2001.
- Lemlouma T., *Architecture de négociation et d'adaptation de services multimédia dans des environnements hétérogènes*, Thèse de doctorat, Institut Nationale Polytechnique de Grenoble, 2004.
- Lemlouma T., Layaïda N., « Universal Profiling for Content Negotiation and Adaptation in Heterogeneous Environments », *W3C Workshop on Delivery Context*, W3C/INRIA Sophia-Antipolis, France, 4-5 mars 2002.
- Magalhães J., Pereira F., « Using MPEG standards for multimedia customization », *Signal Processing: Image Communication*, vol. 19, n° 5, mai 2004, p. 437-456.
- Martínez José M., *Overview of the MPEG-7 Standard (ISO/IEC JTC1/SC29/WG11N5525)*, Singapore, mars 2001, <http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>.
- Mohan R., Smith J.R. et Li C.S., « Adapting Multimedia Internet Content for Universal Access », *IEEE Transactions Multimedia*, vol. 1, n° 1, mars 1999, p. 104-114.
- MPEG MDS Group, *Multimedia framework (MPEG-21) – Part7: Digital Item Adaptation, (Final Committee Draft)*, ISO/IEC JTC 1/SC 29/WG 11/N5845, juillet 2003, http://www.chiariglione.org/mpeg/working_documents.htm

Support de transmodage dans MPEG-21

- Rousseau F., García-Macías J.A., Valdeni de Lima J., Duda A., « User Adaptable Multimedia Presentations for the WWW », *Computer Networks*. vol. 31, n° 11-16, 1999, p. 1273-1290.
- Vetro, A., « MPEG-21 Digital Item Adaptation: Enabling Universal Multimedia Access », *IEEE Multimedia*, n° 11, n° 1, janvier - mars 2004, p. 84-87.
- Viano G., Parodi A. et al., « Adaptive User Interface for Process Control based on Multi Agent Approach », *Proceedings of Advanced Visual Interfaces 2000*, Italie, mai 2000.
- Yergeau F., Cowan J., Bray T., Paoli J., Sperberg-McQueen C. M., Maler E., *Extensible Markup Language (XML) 1.1*, W3C Recommendation, février 2004.



Mariam Kimiaei Asadi est ingénieur en télécommunications. Elle est actuellement doctorante à l'Ecole Nationale Supérieure de Télécommunications (ENST) de Paris où elle prépare sa thèse de doctorat au sein du groupe multimédia, sur l'adaptation de contenus multimédia. Elle a participé au projet européen IST ISIS en 2003, ainsi qu'à plusieurs contributions au groupe MPEG sur la norme MPEG-21 de 2002 à 2005. Elle contribue actuellement au projet européen IST DANAE.



Jean-Claude Dufourd est Directeur Scientifique de Streamezzo dont il est co-fondateur. Il était auparavant Directeur d'Etudes à l'Ecole Nationale Supérieure des Télécom. depuis 2002. Enseignant-chercheur à l'ENST depuis 1990, il a participé à de nombreux projets européens et industriels liés aux standards MPEG-4 et MPEG-21. Il a créé et animé une équipe "Outils et Services Multimédia", spécialisée dans les outils auteur et a fédéré les activités multimédia des différents départements de l'Ecole au sein de la plate-forme "Multimédia en Réseau". Impliqué dans la normalisation MPEG depuis 1995, il est actuellement Chairman du groupe MPEG Integration, en charge de la coordination des aspects conformité et logiciels de référence. Jean-Claude Dufourd est ancien élève de l'Ecole Normale Supérieure de la rue d'Ulm, diplômé de l'ENST et docteur en Informatique.