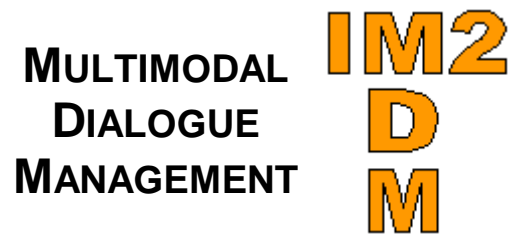




<http://www.im2.ch>



<http://www.issco.unige.ch/projects/im2/mdm/>

ISSCO/TIM/ETI, Université de Genève

Outils informatiques de manipulation de la vidéo et du son : une introduction

Maria Georgescu
Andrei Popescu-Belis
Marianne Starlander

Document interne IM2.MDM – Août 2002

Outils informatiques de manipulation de la vidéo et du son : une introduction

Plan du rapport

- Introduction : motivation
- Logiciels de traitement vidéo et multimédia
- Logiciels de traitement du son
- Logiciels de sauvegarde de flux multimédia
- Annexe : liste de pointeurs Internet vers les logiciels

1. Introduction : motivation

Le projet IM2 vise, dans sa partie applicative, à construire un « démonstrateur » qui puisse traiter et donner accès à des enregistrements de réunion. Le module IM2.MDM s'occupera plus particulièrement du traitement linguistique des transcriptions de parole, mais cela doit se faire en liaison avec les autres modalités enregistrées, notamment l'image et le son. Ainsi, pour les démonstrations locales qui sont en préparation dans IM2.MDM, il nous est apparu nécessaire de manipuler des fichiers contenant des enregistrements vidéo et/ou des fichiers contenant du son.

De nombreux formats et outils existent dans le monde du multimédia, et la vidéo numérique introduit chaque jour de nouvelles technologies. Aussi, il existe une certaine dépendance entre les plates-formes utilisées (ordinateur, système d'exploitation) et les outils disponibles. Le prix des outils est aussi un facteur important, sachant que notre utilisation serait en fin de compte peu intensive.

Pour l'application qui est décrite, les compétences et fonctionnalités suivantes ont été identifiées comme nécessaires (en gras les plus importantes) :

- **Visualisation des fichiers vidéo** aux formats courants (en particulier ASF/WMV)
- **Sauvegarde de flux vidéo/audio**, diffusés sur Internet, en un seul fichier (afin de constituer des ressources)
- Numérisation d'enregistrements sur bande (magnétoscope ou caméscope)
- **Segmentation de fichiers vidéo** et extraction d'images
- Annotation de séquences vidéo
- **Extraction de la bande son d'un fichier vidéo+son**
- Conversion entre différents formats de vidéo
- **Conversion entre les différents formats de son**
- **Lecture/manipulation des fichiers son**
- **Extraction de fragments d'un fichier son**
- **Transcription des paroles depuis un fichier son**
- **Mixage de plusieurs fichiers son**
- Compression d'un fichier son vers des formats de son plus compacts
- Diffusion du son sur Internet

Les programmes qui sont décrits dans ce qui suit ont été considérés intéressants du point de vue de ces fonctionnalités. Certains ont été testés, puis souvent utilisés pour la préparation des données. D'autres figurent ici pour mémoire.

Nota bene : de nombreux programmes cités ci-dessous contiennent des noms déposés par les compagnies qui les ont créés, mais nous ne l'indiquons pas pour chaque nom en particulier. Toute remarque est la bienvenue, par courrier électronique à Andrei.Popescu-Belis@issco.unige.ch.

2. Logiciels de traitement vidéo et multimédia

Adobe Premiere 6

Ce logiciel d'édition vidéo, disponible gratuitement dans une version démo, semble contenir deux outils intéressants :

- Mixer audio : cet outil ressemble à une table de mixage multi-canaux traditionnelle. Il fonctionne conjointement avec la fenêtre de contrôle qui permet de faire des ajustements tout en suivant la vidéo synchronisée sur l'écran.
- Enhanced Monitor and Timeline windows : permet une meilleure édition grâce à un contrôle à 3 points et à 4 points, et à des capacités d'édition et de division (« split-edit ») utilisant l'affichage source et programme.

Blaze Media Pro

Blaze Media Pro est un logiciel multimédia général supportant tous les formats média les plus répandus. Il permet aussi la conversion audio et vidéo, la construction et le décodage vidéo. Le logiciel permet la conversion de différents formats audio et vidéo (CD en .wav, CD à mp3, CD à wma, wav en mp3, mp3 en wav...). Cependant, l'extraction du son d'une vidéo ne peut se faire qu'à partir d'un fichier .avi¹.

Des outils d'édition vidéo, tels que le convertisseur vidéo, le constructeur vidéo et le décompilateur de vidéo, sont intégrés au logiciel.

Le logiciel permet aussi de segmenter la bande son, en utilisant l'Audio Editor, simplement en sélectionnant le fragment voulu, en le coupant et le collant dans un nouveau fichier. Cependant, cela est relativement difficile et long à utiliser, car il faut le réaliser manuellement: la sélection des segments semble particulièrement fastidieuse.

Convert-util

L'utilitaire convert.exe, disponible gratuitement sur Internet, permet de convertir des fichiers .wmv en .avi. Il s'agit d'un outil de commande en ligne.

DirectX de Microsoft

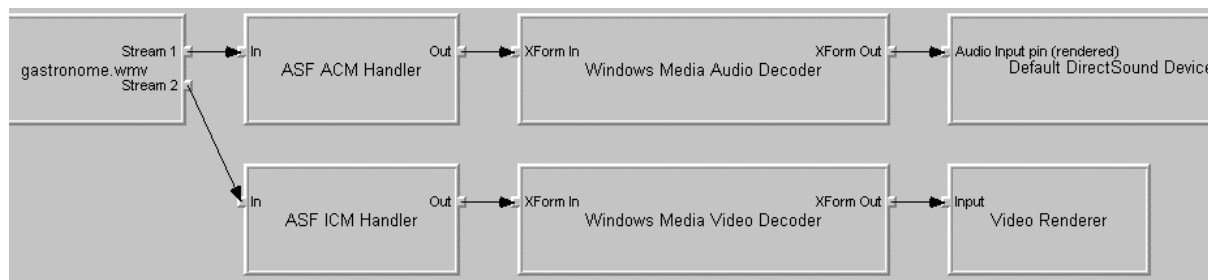
Microsoft DirectX est une interface de programmation de Microsoft regroupant des bibliothèques de fonctions destinés au développement d'applications multimédia sous Windows. Indépendant du matériel, DirectX fait le lien entre différents types de périphériques (carte graphique, carte son ou encore manette de jeu) et le système d'exploitation. Nous avons utilisé ainsi GraphEdit qui est écrit grâce à cette bibliothèque. Le logiciel Direct Show fait également partie de DirectX.

GraphEdit

Ce logiciel convivial, fondé sur Direct X, permet de réaliser de nombreuses opérations de conversion et de filtrage (dans la mesure des filtres et codecs disponibles sur l'ordinateur). Il peut être utilisé pour séparer le son de l'image dans un fichier vidéo+son. Le logiciel peut être téléchargé librement (cf. le site en Annexe).

Le logiciel est fondé sur la visualisation de filtres combinés dans des graphes, permettant de convertir un fichier d'entrée en un fichier de sortie. GraphEdit représente le graphe de la manière suivante:

¹ Dans notre cas, cela demande donc une étape intermédiaire de conversion de .wmv vers .avi.



Pour réaliser une conversion, il faut créer un nouveau graphe, puis (dans le menu 'File') choisir 'Render Media File', et choisir le fichier vidéo d'entrée sur lequel on veut travailler.

Par exemple, pour extraire le son d'un fichier vidéo+son de type WmV, il faut dans le graphe par défaut affiché ci-dessus enlever tout sauf le 'ASF ACM Handler' et le 'Windows Media Audio Decoder'. Il faut ensuite ajouter deux filtres en choisissant 'Insert Filters' dans le menu 'Graph', puis en sélectionnant sous le type 'Direct Show Filters' deux filtres 'WAV Dest' et 'File Writer'. Il faut enfin indiquer le nom du fichier sortie sans oublier de mettre l'extension .wav. Le graphe résultant doit avoir l'aspect suivant :



Les performances semblent acceptables, mais le processus de conversion reste assez long. L'extraction du son sur un fichier WMV d'environ 15 minutes (env. 12 MB, image de taille assez faible) prend environ 1/2 heure ; le fichier .wav résultant est assez volumineux (env. 43 MB). Les sites de téléchargement contiennent plus d'informations sur ce logiciel utile et gratuit.

iMovie

Puissant et convivial logiciel de traitement vidéo et multimédia, surtout destiné au Macintosh et au format multimédia QuickTime.

MovieXone

Logiciel permettant l'édition de films: extraction, ajout d'effets, de sous-titres, etc. Ce logiciel intègre un outil de reconnaissance automatique de scènes et de « batch capturing ». Toutefois, il ne semble pas très intéressant pour les tâches que nous voulons effectuer.

RealPlayer

Logiciel très populaire permettant d'écouter/voir des fichiers vidéo ou son au format propriétaire « real », mais surtout d'écouter/voir des flux transmis via Internet.

Windows Media Encoder

Ce logiciel permet de convertir des flux vidéo et audio, en direct, en format Windows Media (format Microsoft compressé qui permet l'envoi sous forme de flux des données audio et vidéo). On peut donc convertir une *entrée* audio ou vidéo depuis une caméra, en créant une *sortie* sur ordinateur, en fichier son ou vidéo aux formats suivants:

- Windows Media : .wma (windows media audio), .wmv (windows media video), .asf (advanced system format)
- Windows : .avi, .wav (waveform audio file)
- Moving Picture Expert Group: .mpg, .mp3

Le logiciel permet également de capturer des flux vidéo ou de faire des enregistrements de l'écran.

Xine

Plutôt qu'un éditeur, xine est surtout un outil libre pour Unix qui permet de visualiser un grand nombre de formats vidéo, mais aussi d'ajouter des sous-titres, extraire des images, etc. De nombreux formats sont acceptés, en particulier WMV, MPEG, etc.

3. Logiciels de traitement du son

Logiciels et commandes pour Unix / Solaris

Pour mémoire, nous rappelons quelques commandes utiles sous Unix / Solaris :

- *audioplay*, *audioconvert*, *audiorecord* (commandes en ligne, assez basiques) ;
- *audiocontrol* affiche une fenêtre de contrôle du volume, permettant aussi l'enregistrement ;
- *sdtaudio* est un outil avec une GUI très simple pour écouter des fichiers son (WAV ou AU) sans passer par la ligne de commande ;
- *sdjmplayer* est un autre outil simple qui utilise le 'Java Media Framework' (ses performances et son aspect son assez décevants) ;
- *xplay* est enfin encore un autre outil permettant d'écouter un fichier son donné en argument lors du lancement – quelques boutons de contrôle sont disponibles.

En conclusion, nous recommandons l'outil *sdtaudio*, qui paraît être le plus robuste et le plus léger.

Snack (2.2a2)

Il s'agit plutôt d'un ensemble d'outils, auxquels une interface graphique donne accès par un ensemble de boutons. Les outils permettent différentes opérations de traitement du signal audio, conversions, etc. On peut notamment découper assez facilement la bande son en segments grâce à l'outil XS et à sa fonction « Sound Cutter ».

Sox (12.17.3)

Commande en ligne pour Unix qui permet très facilement de convertir les formats des fichiers sons et d'appliquer divers effets sonores. Utile notamment pour extraire, dans la ligne de commande, un fragment d'un fichier son délimité par un intervalle de temps. Par exemple, on peut ainsi grâce à un script CGI qui appelle sox, diffuser le fragment de son correspondant à une phrase de la transcription.

Transcriber

Ce logiciel est un outil pour la transcription de la parole contenue dans des fichiers son. Il permet de faire un lien entre le son et les paroles transcrites en segmentant le signal en sections, avec indications des tours de paroles et des superpositions entre locuteurs (le cas échéant). Une version pour la transcription d'enregistrements multi-locuteurs et multi-canaux est utilisée à l'ICSI (Berkeley) mais ne semble pas disponible publiquement.

L'installation de Transcriber nécessite l'installation préalable de Snack (donc de WaveSurfer) et de tclLex et de Tcl/Tk. L'installation semble plus facile sur Unix que sur PC/Windows. A titre d'exemple, la transcription en trois canaux séparés d'un débat télévisé avec deux intervenants, d'une durée de 15 minutes, nécessite environ 8-10 heures (incluant la vérification).

WaveSurfer (1.4.2)

Cet outil permet d'écouter des fichiers son et de réaliser un certain nombre de transformations assez simples. Il est lui-même intégré à des outils plus puissants.

4. Logiciels de sauvegarde de flux multimédia

Ces logiciels sont utiles pour capturer un flux vidéo et/ou audio diffusé en temps réel via Internet. On peut ainsi stocker toute l'information reçue dans un seul fichier, la reconsulter hors-ligne, et la modifier selon ses souhaits. Cette capacité se heurte à l'idée de propriété intellectuelle (autant qu'un magnétoscope qui permet d'enregistrer des films diffusés à la télévision) et beaucoup de logiciels commerciaux ne l'autorisent pas. Il convient donc de limiter l'usage des documents ainsi capturés à des fins personnelles ou de recherche, et éviter tout usage commercial.

ASF Recorder

Ce programme très simple ('open-source' pour toute plate-forme) reçoit une adresse Internet qui diffuse un flux au format ASF ou WMV ou WMA, se connecte et stocke l'ensemble du flux dans un fichier. Une interface simple existe, mais on peut l'utiliser aussi en ligne de commande. Un message (dans 'About...') avertit l'utilisateur qu'il peut ne pas être légal de faire une copie du document, et suggère d'instaurer un mécanisme équitable de copyright numérique.

Streambox VCR

Streambox est un outil de capture plus puissant que ASFRecorder, mais fonctionnant selon le même principe. Le logiciel permet d'enregistrer en fait n'importe quelle partie d'un site Internet, de lancer des requêtes multiples pour optimiser le téléchargement, etc. Un logiciel additionnel permet de supprimer certaines protections des flux par leurs propriétaires.

5. Annexe

Tutoriels utiles sur Internet

Sur GraphEdit

<http://www.media-video.com/sections.php?op=listarticles&secid=12>

<http://digilander.iol.it/digivideoweb/downloadgb.htm>

Diffusion de flux

<http://archive.dstc.edu.au/RDU/staff/jane-hunter/video-streaming.html>

Sauvegarde de flux

<http://digilander.libero.it/digivideoweb/videocapturingwindows1gb.htm>

Outils Microsoft : Windows Media

<http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnwmt/html/introwmt7-2.asp>

Pointeurs vers les logiciels et outils mentionnés

Adobe Premiere 6

<http://www.adobe.com/products/premiere/main.html>

ASFRecorder

<http://asfrecorder.virtualave.net>

Convert-util

<http://www.digtv.ws/html/convert/convert-util.php>

DirectX

<http://www.microsoft.com/windows/directx/downloads/drx81.asp>

GraphEdit

<http://digilander.iol.it/digivideoweb/downloadgb.htm>

MovieXone

http://www.aist.com/international/register/reg_new/free_reg.asp
<http://download.com.com/3000-2194-8952847.html?legacy=cnet>

Snack

<http://www.speech.kth.se/snack>

Sox

<http://home.sprynet.com/~cbagwell/sox.html>

Streambox

<http://www.streambox.com>

Transcriber

<http://www.etca.fr/CTA/gip/Projets/Transcriber/>

WaveSurfer

<http://www.speech.kth.se/wavesurfer>

Windows Media Encoder

<http://www.microsoft.com/windows/windowsmedia/wm7/encoder.asp>

Xine

<http://xine.sourceforge.net>