

TD séance n° 13

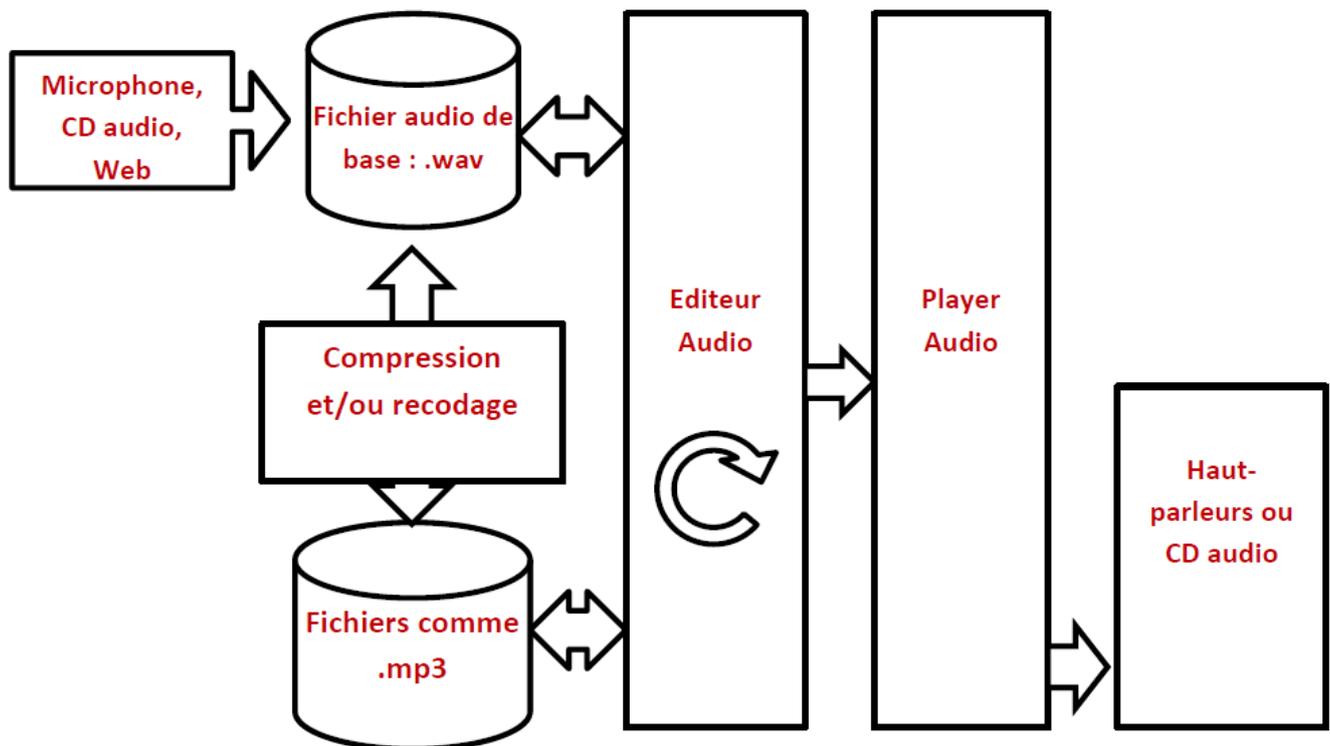
Multimédia Son

Ce TD est dédié aux chaînes de traitements audio, un des objets multimédia de base. Il se déroulera sous Windows, bien que les logiciels utilisés ou équivalents soient aussi disponibles sous Unix.

La quasi majorité des logiciels que vous utilisez sur votre PC, sur votre box chez vous, sur mobile, sur le Web ..., pour écouter de la musique, travaillent avec les notions que nous allons illustrer ici.

1 Chaîne de traitements pour le son

Une chaîne de traitement pour le son se décline de la même manière que la chaîne de traitement du texte que nous avons étudié la semaine dernière.



1.1 Mais c'est quoi un son ?¹

Le son est une vibration de l'air, c'est-à-dire une suite de surpressions et de dépressions de l'air par rapport à une moyenne, qui est la pression atmosphérique. D'ailleurs pour s'en convaincre, il suffit de placer un objet bruyant (un réveil par exemple) dans une cloche à vide pour s'apercevoir que l'objet initialement bruyant n'émet plus un seul son dès qu'il n'est plus entouré d'air ! Comme quoi les bruits de combats spatiaux n'est qu'une pure invention d'Hollywood !

Commençons par regarder les extrémités de la chaîne de traitement à savoir l'acquisition et la restitution d'un son.

1.2 Acquisition et restitution d'un son

1.2.1 Acquisition

Il y a différentes façon de récupérer un objet multimédia audio. La première, la plus simple, est d'utiliser une source audio déjà sous forme numérique (un fichier audio comme celui stocké sur un CD audio en utilisant un

¹ Source de la partie cours : « Comment ça marche.net »

TD séance n° 13

Multimédia Son

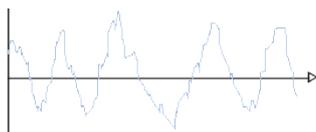
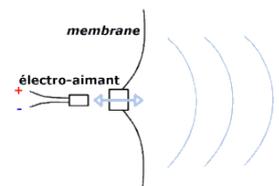
logiciel approprié, un fichier que l'on récupère sur Internet, en ayant vérifié qu'ils sont libres de droits). Ceci est très pratique pour ajouter des bandes sonores à vos objets multimédia composites. Nombreux sont les sites qui vous le permettent.

La deuxième façon consiste à capturer un son du monde réel, grâce à un microphone branché sur la carte audio de votre machine, et de la stocker dans l'ordinateur. Mais comme se déroule le passage du son physique au son numérique ?

1.2.2 Restitution

La façon la plus simple de reproduire un son actuellement est de faire vibrer un objet. De cette façon un violon émet un son lorsque l'archet fait vibrer ses cordes, un piano émet une note lorsque l'on frappe une touche, car un marteau vient frapper une corde et la fait vibrer.

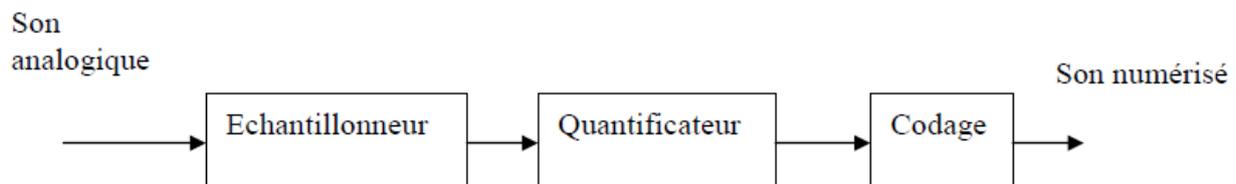
Pour reproduire des sons, on utilise généralement des haut-parleurs. Il s'agit en fait d'une membrane reliée à un électroaimant, qui, suivant les sollicitations d'un courant électrique va aller en avant et en arrière très rapidement, ce qui provoque une vibration de l'air situé devant lui, c'est-à-dire du son !



De cette façon on produit des ondes sonores qui peuvent être représentées sur un graphique comme les variations de la pression de l'air (ou bien de l'électricité dans l'électroaimant) en fonction du temps. On obtient alors une représentation comme montré sur le graphique ci-contre.

1.3 Représentation du son sous forme numérique

Le passage d'un son analogique (du monde physique) à un son numérisé se déroule suivant 3 étapes que nous allons détailler pour mieux les comprendre.

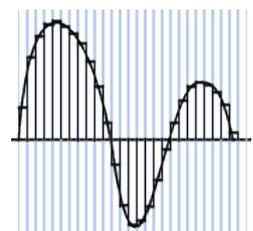


1.3.1 Echantillonnage

Pour pouvoir représenter un son sur un ordinateur, il faut arriver à le convertir en valeurs numériques, car celui-ci ne sait travailler que sur ce type de valeurs. Il s'agit donc de relever des petits échantillons de son (ce qui revient à relever des différences de pression) à des intervalles de temps précis. On appelle cette action l'échantillonnage ou la numérisation du son.

A chaque échantillon (correspondant à un intervalle de temps) est associée une valeur qui détermine la valeur de la pression de l'air à ce moment, le son n'est donc plus représenté comme une courbe continue présentant des variations mais comme une suite de valeurs pour chaque intervalle de temps.

L'intervalle de temps entre deux échantillons est appelé taux d'échantillonnage. Etant donné que pour arriver à restituer un son qui semble continu à l'oreille il faut des échantillons tous les quelques 100 000^{èmes} de seconde, il est plus pratique de raisonner sur le nombre d'échantillons par seconde, exprimés en Hertz (Hz).



TD séance n° 13

Multimédia Son

Voici quelques exemples de taux d'échantillonnage et de qualités de son associées :

Taux d'échantillonnage	Qualité du son
44 100 Hz	Qualité CD
22 000 Hz	Qualité radio
8 000 Hz	Qualité téléphone

La valeur du taux d'échantillonnage, pour un CD audio par exemple, n'est pas arbitraire, elle découle en réalité de la théorie de Shannon dont nous avons parlé lors du précédent cours. En effet, la fréquence d'échantillonnage doit être suffisamment grande, afin de préserver la forme du signal. Notre oreille perçoit les sons environ jusqu'à 20 000 Hz, il faut donc une fréquence d'échantillonnage au moins de l'ordre de 40 000 Hz pour obtenir une qualité satisfaisante.

1.3.2 Quantificateur et codage de l'information

Nous avons vu la semaine dernière que l'ordinateur travaille avec des bits pour stocker toute information. Il faut donc déterminer le nombre de valeurs que l'échantillon peut prendre. Cela revient à fixer le nombre de bits sur lequel on code les valeurs des échantillons.

- Avec un codage sur 8 bits, on a 2^8 possibilités de valeurs, c'est-à-dire 256 valeurs possibles
- Avec un codage sur 16 bits, on a 2^{16} possibilités de valeurs, c'est-à-dire 65536 valeurs possibles

Avec la seconde représentation, on aura bien évidemment une qualité de son bien meilleure, mais aussi un besoin en mémoire beaucoup plus important. L'inconvénient de ce format (qui est celui utilisé en mémoire) est qu'il n'est pas compact et produit des fichiers très gros.

1.3.3 Plusieurs pistes sons

Enfin, la stéréophonie nécessite deux canaux sur lesquels on enregistre individuellement un son qui sera fourni au haut-parleur de gauche, ainsi qu'un son qui sera diffusé sur celui de droite.

Et bien entendu dans le cas d'un son 5.1, nous aurons 5 pistes pour envoyer le son sur le haut-parleur avant gauche, avant droit, arrière gauche, arrière droit, celui du centre (ou la voie), ce qui fait 5 canaux plus une dernière piste pour le caisson de basse, que qui nous donne 5 + 1, noté 5.1.

1.3.4 Conclusion

Un son est donc représenté (informatiquement) par plusieurs paramètres :

- la fréquence d'échantillonnage
- le nombre de bits d'un échantillon
- le nombre de voies (une seule correspond à du mono, deux à de la stéréo, etc.)

Il est simple de calculer la taille d'une séquence sonore non compressée. En effet, en connaissant le nombre de bits sur lequel est codé un échantillon, on connaît la taille de celui-ci (la taille d'un échantillon est le nombre de bits...).

Pour connaître la taille d'une voie, il suffit de connaître le taux d'échantillonnage, qui va nous permettre de savoir le nombre d'échantillons par seconde, donc la taille qu'occupe une seconde de musique. Celle-ci vaut :

$$\text{Taille Voie} = \text{Taux d'échantillonnage} \times \text{Nombre de bits}$$

Ainsi, pour savoir l'espace mémoire que consomme un extrait sonore de plusieurs secondes, il suffit de multiplier la valeur précédente par le nombre de secondes :

$$\text{Taille Extrait sonore} = \text{Taux d'échantillonnage} \times \text{Nombre de bits} \times \text{Nombre de secondes}$$

Enfin, la taille finale de l'extrait est à multiplier par le nombre de voies (elle sera alors deux fois plus importante en stéréo qu'en mono...).

TD séance n° 13

Multimédia Son

La taille en bits d'un extrait sonore est ainsi égale à :

```
Taille fichier = Taux d'échantillonnage x Nombre de bits x Nombre de secondes x  
Nombre de voies
```

1.3.5 Recodage et compression

La manière d'encoder les informations de sons conduit à générer une très grande quantité de données. Mais il est possible de modifier l'encodage de l'information pour obtenir une information plus compacte sans perte (format FLAC) ou avec une perte plus ou moins grande perte d'information (MP3, AC3, WMA, OGG vorbis). Nous étudierons ces différents formats à l'aide des exercices et nous comparerons.

Mais sachez d'ores et déjà que le fait de compresser les sons dans le cas de la compression avec perte va changer la dynamique du son. La compression consiste à remonter tous les sons faibles, compris entre 5 et 30 décibels (dB), pour en faire des sons plus forts. Cela permet au compresseur d'avoir moins d'amplitude d'information à prendre en compte et donc potentiellement d'utiliser moins de bits pour coder cette information.

Avec un analyseur de niveau sonore (sonomètre), il est possible de voir à quel point ce traitement transforme radicalement les sons. En écoutant « Le Boléro » de Maurice Ravel dans un enregistrement sans compression, les courbes qui défilent sur l'écran de l'appareil sont d'une grande amplitude. Il y a des silences au début, des notes à peine audibles à 40 dB et puis viennent à la fin les éclats à plus de 85 dB. Donc l'oreille n'est pas tout de temps sollicitée. Il n'y a rien de tout ça sur le même morceau compressé : les courbes démarrent d'emblée à 70-75 dB et se maintiennent à ce niveau pendant toute la durée de l'enregistrement. La dynamique, c'est-à-dire la différence entre un niveau sonore fort et un niveau faible, a disparu. L'oreille est donc plus souvent sollicitée et se fatigue donc plus. Vous ajoutez à cela le fait d'écouter avec un casque et, suivant le temps d'écoute, vous avez un cocktail plus ou moins d'étonnant !

1.4 Lecteurs Audio

La plupart des lecteurs (players) audio logiciels lisent la quasi-totalité des formats. Certains, comme Windows Media Player, chargent les codecs dynamiquement pour de nouveau codage. Il n'en va pas de même des lecteurs intégrés à du matériel dit de salon ou vos téléphones mobiles. Les Lecteurs typiques et les extensions des fichiers de base associés sont :

- RealOne Player .ra
- Windows Media Player .wma
- Winamp .mp3, .ogg
- ou un lecteur générique comme VLC qui permet de lire à la fois l'audio et la vidéo.

Vous pourrez à l'occasion charger le lecteur logiciel qui vous conviendra le mieux. Nous verrons par la suite que ses capacités à lire plusieurs formats vidéo et de streaming sont tout aussi déterminantes.

1.5 Dispositif de restitution (CD audio)

Pour graver votre propre CD audio, vous partez d'une liste de fichiers. Vous pourrez ainsi écouter ce CD sur n'importe quel lecteur ou platine de CD.

Les données gravées sur un CD sont des données non compressées (format .raw, semblable au format .wav), avec une fréquence d'échantillonnage de 44100 Hz. Nombreux sont les logiciels de gravure qui peuvent vous être fournis avec un graveur de CD, comme Néro, BurnAtOnce, Express Burn, ...

TD séance n° 13

Multimédia Son

2 Exercices

Durant les exercices qui suivent, nous allons tenter de comprendre par la pratique le codage du son, le recodage et la compression dans différents formats ainsi que les traitements que l'on peut effectuer une fois le son numérisé.

2.1 Installation des logiciels

Nous utiliserons le magnétophone de Windows pour effectuer les compressions/recodages (ex. réglages de l'échantillonnage, compression, ...) et Audacity pour étudier/manipuler le signal audio (amplitude, spectre en fréquence,...).

2.1.1 Audacity

Audacity est un logiciel libre pour la manipulation de données audio numériques. Audacity permet d'enregistrer du son numérique par le biais d'une entrée audio sur une carte son. Il permet d'appliquer divers filtres et effets : pitch, tempo, réduction de bruit, égaliseur, filtres de Fourier, augmentation de fréquences précises, compression, amplification, normalisation, écho, phaser, wahwah, inversion...

Pour installer Audacity, vous pouvez utiliser le lien :

<http://audacity.sourceforge.net/download/windows>

2.1.2 Magnétophone

Pour ouvrir les fichiers audio, nous utiliserons une application fournie sur Windows : le magnétophone de Windows (dans Programmes / Accessoires / Divertissement, ...).

Pour ceux qui travaillent sous Windows 7, il vous faudra réinstaller le logiciel `sndrec32.exe` s'il n'est pas présent dans `%SystemRoot%\system32\`.

2.2 Recodage

2.2.1 Audio non compressée

Pendant que vous téléchargez ou installez les logiciels nécessaires à ce TD, nous allons commencer par un petit calcul théorique.

TD séance n° 13

Multimédia Son

3 Exercices

Exercice n°1:

Quel sera la taille d'un fichier audio en stéréo, d'une durée d'une heure, encodé à 44 100Hz (qualité CD audio) en 16 bits ? A quoi cette quantité trouvée correspond-t-elle ?

Nous travaillerons maintenant sur des exemples avec les `music1.wav` et `music2.wav`.

Exercice n°2:

Quelles sont ses caractéristiques de numérisation de ces deux sons : durée, nombre de canaux, fréquence d'échantillonnage, nombre de bits de quantification ?

Comment expliquez-vous la différence de taille des deux fichiers pour à peine 1 seconde de différence ?

Nous allons maintenant faire quelques manipulations qui vont nous permettre d'illustrer les variations de qualité de l'objet multimédia audio que nous pourrions obtenir selon les propriétés des dispositifs d'acquisition et des codages de base.

Il s'agit, par exemple, de comparer le format 8 bits et 16 bits en utilisant le fichier `music1.wav`. D'origine la bande est enregistrée en 16 bits. Pour la convertir en 8 bits, vous pouvez utiliser le magnétophone de Windows (Fichier / Propriétés / Convertir...). A partir du fichier `music1.wav`, passez-le en 8 bits, et sauvegardez le sous `music1_8bits.wav`.

Exercice n°3:

Écoutez les deux fichiers obtenus. Que se passe-t-il ? Quelles sont les tailles des fichiers `music1.wav` et `music1_8bits.wav` ? Pourquoi ?

Continuons en comparant des sons de plus faible intensité. Avez Audacity, reprenez le fichier `music1.wav`, vous l'atténuez de 30dB (Effets / Amplification / -30dB). Enregistrez le sous `music1_attenué.wav`. Puis convertissez-le au format 8 bits, et sauvegardez le sous `music1_8bits_attenué.wav`.

Exercice n°4:

Écoutez les deux fichiers obtenus. Pourquoi êtes-vous obligés d'augmenter le volume de vos enceintes ?

Vous pouvez constater de plus une perte de qualité par rapport aux sons non atténués après avoir compensé avec votre volume sonore. En fait dans ce cas le volume de la musique a la même amplitude que le bruit.

Vous allez reprendre le fichier d'origine `music1.wav`. Vous allez convertir le fichier d'origine en qualité radio stéréo (`music1_radio.wav`) et en qualité téléphonique (`music1_tel.wav`).

Exercice n°5:

Pour chacun des deux formats : que pensez-vous de la qualité du signal ? Que pouvez vous en conclure ?

3.1.1 Audio compressée

Vous pouvez grâce au menu « Fichier » exporter le son à éditer, ou une partie du son seulement, dans trois formats différents :

- le format « WAVEform audio format » (WAV)
- le format « MPEG-1/2 Audio Layer 3 » (MP3)
- le format « Ogg » (OGG)
- ...

Après avoir quitté Audacity, installez `FFmpeg_vo.6.2_for_Audacity_on_Windows` et `Lame_v3.99.3_for_Windows`.

TD séance n° 13

Multimédia Son

Nous allons maintenant pouvoir comparer les codecs audio WAV, MP3, OGG Vorbis, WMA Audio V2 et GSM.

Exercice n°6:

Compresserez le fichier `music.wav` dans chacun des formats de compression ci-dessus. Puis dans un tableau, vous donnerez vos conclusions en ce qui concerne la qualité des sons compressés obtenus et la taille des fichiers.

Exercice n°7:

Que pouvez-vous conclure sur les performances de chacun de ces codecs audio ?

3.2 Edition et manipulation audio numérique

Audacity permet aussi de faire des manipulations sur le son numérique. Le logiciel possède des fonctions avancées pour ceux qui sont habitués à l'édition audio : montage, mixage, conception, échantillonnage, enregistrement, sound design, mastering ... D'autres logiciels gratuits pour l'édition audio existent comme le logiciel Wavosaur (<http://fr.wavosaur.com/>).

Exercice n°8:

Donnez le nom des opérations à effectuer si l'on désire :

- Supprimer les silences de début et de fin
- Atténuer le son en début et en fin de lecture
- Générer un morceau de musique en lecture inverse (les plus mélomanes reconnaîtront une manipulation à l'origine du « backmasking », notamment utilisé par un groupe de musique célèbre, lequel ? wikipedia est mon ami ...)

Le reste est une affaire de curiosité et de temps ...