



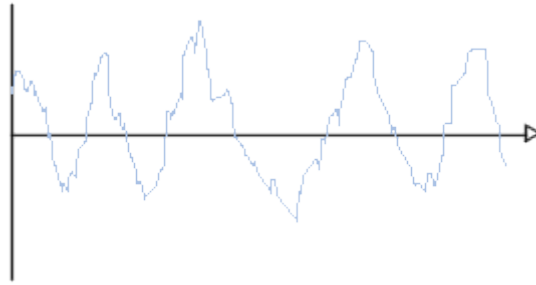
Le traitement du son

Tome 1 principes et définitions

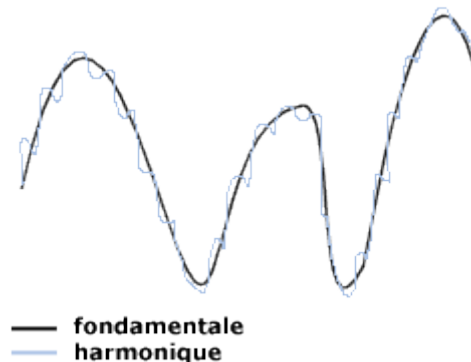
1. Les sons.....	1
1.1. Echantillonnage du son.....	1
1.2. Représentation informatique du son.....	2
1.3. Mémoire requise pour stocker un son.....	3
2. Présentation du format MP3.....	3
2.1. Le contexte.....	3
2.2. Les procédés.....	3
2.2.1. L'effet de masque.....	3
2.2.2. Le réservoir de bytes.....	4
2.3. Le joint stéréo.....	4
2.4. Le code Huffman.....	4
2.5. Les résultats.....	4
2.6. La législation.....	4
2.7. Son multi canal.....	5
2.8. Configuration spatiale.....	5
2.9. Configuration 5.1.....	5
2.10. Configuration 6.1.....	6
2.11. Configuration 7.1.....	7
2.12. Dolby Surround.....	7
2.13. Dolby Surround Pro Logic.....	7
2.14. Dolby Pro Logic II.....	7
2.15. Dolby Pro Logic IIX.....	8
3. Les types de formats musicaux.....	8
3.1. Les formats audio sans compression.....	8
3.2. Les formats audio compressés sans perte.....	9
3.3. Les formats audio compressés avec perte.....	9
4. Calcul du débit et du poids d'un fichier audio.....	12

1. Les sons.

Les ondes sonores 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 de la forme suivante :



Cette représentation d'un son est appelée **spectre de modulation d'amplitude** (modulation de l'amplitude d'un son en fonction du temps). Le **sonogramme** représente par contre la variation des fréquences sonores en fonction du temps. On peut remarquer qu'un sonogramme présente une fréquence fondamentale, à laquelle se superposent des fréquences plus élevées, appelées harmoniques.



C'est ce qui permet d'arriver à distinguer plusieurs sources sonores : les sons graves auront des fréquences basses, et les sons aigus des fréquences élevées

1.1. Echantillonnage du son

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**. 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)**. 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

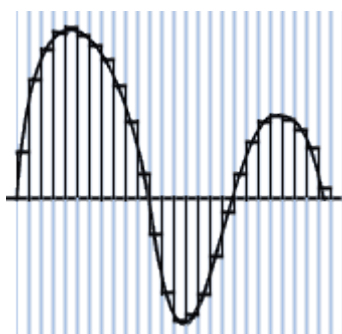
La valeur du taux d'échantillonnage, pour un CD audio par exemple, n'est pas arbitraire, elle découle en réalité du théorème de Shannon. La fréquence d'échantillonnage doit être suffisamment grande, afin de préserver la forme du signal. Le théorème de Nyquist - Shannon stipule que la fréquence d'échantillonnage doit être égale ou supérieure au double de la

fréquence maximale contenue dans ce 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. Il existe un certain nombre de fréquences d'échantillonnage normalisées :

- **32 kHz** : pour la radio FM en numérique (bande passante limitée à 15 kHz)
- **44.1 kHz** : pour l'audio professionnelle et les compact-disques
- **48 kHz** : pour les enregistreurs numériques multipistes professionnels et l'enregistrement grand public (DAT, MiniDisc...)

1.2. Représentation informatique 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'ordinateur travaille avec des bits, 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.

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.

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, et quatre à de la quadriphonie)

1.3. Mémoire requise pour stocker un son

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 :

Taux d'échantillonnage x Nombre de bits

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

Taux d'échantillonnage x Nombre de bits x 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...).

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

Taux d'échantillonnage x Nombre de bits x Nombre de secondes x Nombre de voies

2. Présentation du format MP3

Le **format MP3** (« *MPEG-1 Audio layer 3* ») est un format de compression de données audio par destruction de données, développé par l'organisation de standardisation internationale (ISO - International Standard Organization). Ce format permet de compresser à un taux de 1:12 les formats audio habituels (WAV ou CD audio).

Il permet de faire tenir l'équivalent en fichiers de douze albums de musique sur un seul CD-ROM. De plus, le format mp3 n'altère que faiblement le son pour l'oreille humaine.

2.1. Le contexte

En fait la compression MPEG layer 3 consiste à retirer des données audio les fréquences inaudibles pour l'auditeur moyen dans des conditions habituelles d'écoute. La compression vise donc à analyser les composantes spectrométriques d'un signal audio, et de leur appliquer un modèle psychoacoustique pour ne conserver que les sons « audibles ». L'oreille humaine est capable de discerner, en moyenne, des sons entre 0.02 kHz et 20 kHz, sachant que sa sensibilité est maximale pour des fréquences entre 2 et 5 kHz (la voix humaine est entre 0.5 et 2 kHz), suivant une courbe donnée par la loi de Fletcher et Munson.

La compression consiste à déterminer les sons que nous n'entendons pas et à les supprimer, il s'agit donc d'une compression destructive, c'est-à-dire avec une perte d'information.

2.2. Les procédés

2.2.1. L'effet de masque

Gabriel Bouvigne explique : « Lorsque vous regardez le soleil et qu'un oiseau passe dans son axe, vous ne le voyez pas car la lumière provenant du soleil est trop importante. En acoustique, c'est similaire. Lorsqu'il y a des sons de fort volume sonore, vous n'entendez pas les sons faibles. Prenez l'exemple d'un orgue : lorsqu'un organiste ne joue pas, vous entendez le souffle dans les tuyaux, et quand il joue, vous ne l'entendez plus car il est masqué.

Il n'est donc pas essentiel d'enregistrer tous les sons, c'est la première propriété utilisée par le format MP3 pour gagner de la place. »

2.2.2. Le réservoir de bytes

Souvent, certains passages d'une musique ne peuvent pas être encodés sans altérer la qualité. Le mp3 utilise donc un petit réservoir de bytes qui agit en utilisant des passages qui peuvent être encodés à un taux inférieur au reste des données.

2.3. Le joint stéréo

Dans beaucoup de chaînes hi-fi, il y a un boomer unique (qui produit les basses). Cependant on n'a pas l'impression que le son vient de ce boomer mais plutôt des haut-parleurs satellites. En effet, en dessous d'une fréquence donnée l'oreille humaine est incapable de localiser l'origine du son. Le format mp3 peut exploiter (en option) cette astuce en utilisant la méthode du *joint stéréo*. C'est-à-dire que certaines fréquences sont enregistrées en mono mais elles sont accompagnées d'informations complémentaires afin de restituer un minimum d'effet spatial.

2.4. Le code Huffman

La technique de l'algorithme Huffman est un algorithme de codage (et non de compression), qui agit à la fin de la compression, en créant des codes de longueurs variables sur un grand nombre de bits. Les codes ont l'avantage d'avoir un préfixe unique, ils peuvent toutefois être décodés correctement malgré leur longueur variable, et rapidement grâce à une correspondance de tables. Ce type d'encodage permet de gagner en moyenne un peu moins de 20% d'espace.

Lorsque les sons sont « purs » (lorsqu'il n'y a pas de masquage) l'algorithme Huffman est très efficace car le son digitalisé contient de nombreux sons redondants.

2.5. Les résultats

Ainsi, une minute d'un CD-audio (à une fréquence de 44.1 kHz, 16 bits, stéréo) ne prendra qu'un seul Mo. Une chanson fait donc en moyenne 3 ou 4 Mo, ce qui rend son téléchargement possible même avec un modem.

Bande passante	Mode	Débit	Qualité	Compression
11.025	Mono	8 kbps	Très mauvaise	200:1
22.050	Stéréo	64 kbps	Mauvaise	25:1
44.100	Stéréo	96 kbps	Acceptable	16:1
44.100	Stéréo	128 kbps	Bonne	12:1
44.100	Stéréo	196 kbps	Très bonne	12:1

Le décodage d'un fichier MP3 est relativement peu coûteux en utilisation du processeur, tandis que l'encodage est plus gourmand.

2.6. La législation

Le format MP3 n'est pas illégal, car il représente uniquement une façon de compresser des données numériques. Par contre son utilisation peut l'être. Lors de l'utilisation de fichiers Mp3, veillez à respecter les droits d'auteur : vous pouvez faire une copie (de sauvegarde) d'une chanson dont vous possédez l'original, mais vous ne pouvez pas télécharger ou archiver une musique d'un artiste dont les droits d'utilisation ne sont pas libérés. Il est ainsi fort peu probable que la chanson que vous rêvez de télécharger (chanson diffusée à la radio, ...) puisse légalement l'être. Il existe des plateformes de téléchargement légal, permettant d'acquérir des MP3 moyennant quelques euros. Attention toutefois, les MP3 ainsi téléchargés peuvent être protégés par des DRM (Digital Right Management), c'est-à-dire un dispositif permettant

d'assurer aux ayants-droits que le MP3 ne sera pas copié sur un autre support, limitant ainsi son utilisation et pouvant empêcher sa lecture sur certains appareils.

2.7. Son multi canal

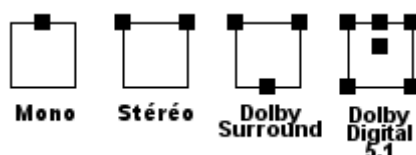
Le terme « **son multicanal** » (parfois « multi-voies ») désigne l'utilisation de plusieurs pistes audio en vue de la restitution sur un système comportant plusieurs enceintes.

Il existe une terminologie associée, constituée de deux chiffres séparés par un point (2.1, 5.1, 6.1, 7.1, etc.), permettant de classifier le type de configuration spatiale des enceintes en fonction du nombre de pistes audio utilisées.

Le premier chiffre indique le nombre de canaux principaux destinés chacun à être restitués sur une enceinte, tandis que le second désigne la présence d'effets basse fréquence (en anglais *Low Frequency Effect*, notés *LFE*) destinés à être restitué sur un caisson de basses. Ainsi, 1.0 correspond à un son audio mono (sous-entendu monocanal) et 2.0 correspond à une source sonore stéréo.

2.8. Configuration spatiale

Selon le nombre de canaux audio, il existe une configuration spatiale des enceintes permettant d'optimiser l'effet recherché. Ainsi, des icônes spécifiques présents sur les supports multicanaux permettent de symboliser le nombre de canaux et la répartition des enceintes dans l'espace grâce à de petits carrés noirs (représentant chaque canal) répartis sur un carré représentant la pièce :



2.9. Configuration 5.1

La configuration spatiale des enceintes d'un système 5.1 est d'une importance primordiale car elle conditionne directement la qualité d'écoute et le réalisme des effets sonores.

Il existe un certain nombre de règles à respecter afin de positionner au mieux chaque enceinte :

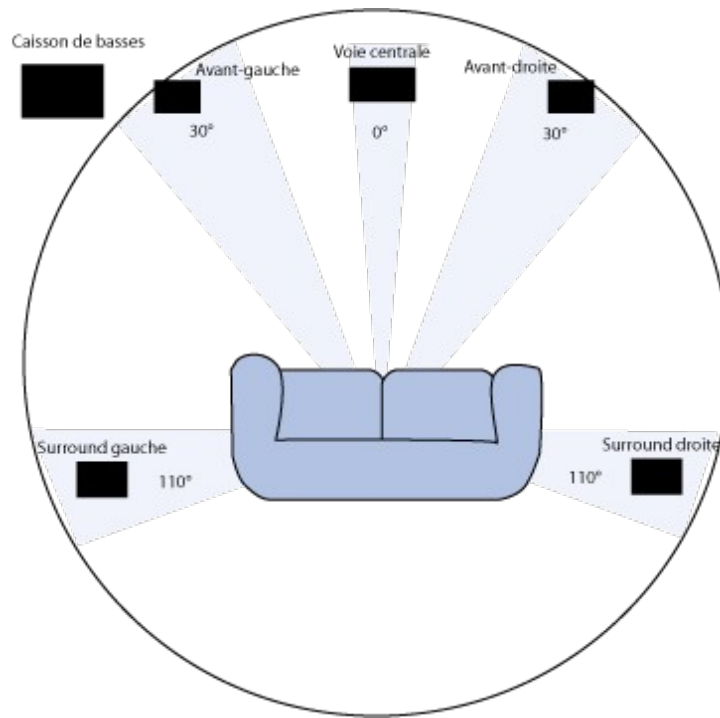
Les enceintes frontales doivent idéalement être placées à la hauteur d'écoute de l'auditeur assis. Les satellites arrière (surround) doivent être positionnés légèrement au-dessus de cette ligne d'écoute.

1- Les enceintes frontales gauche et droite doivent être disposées de part et d'autre du téléviseur en veillant à respecter le côté de chacune d'entre-elles. Dans la pratique elles devront chacune former un angle de 25° à 45° avec l'auditeur.

2- L'enceinte centrale doit être placée directement au-dessus ou en dessous de la TV étant donné qu'elle sert principalement aux dialogues des acteurs principaux.

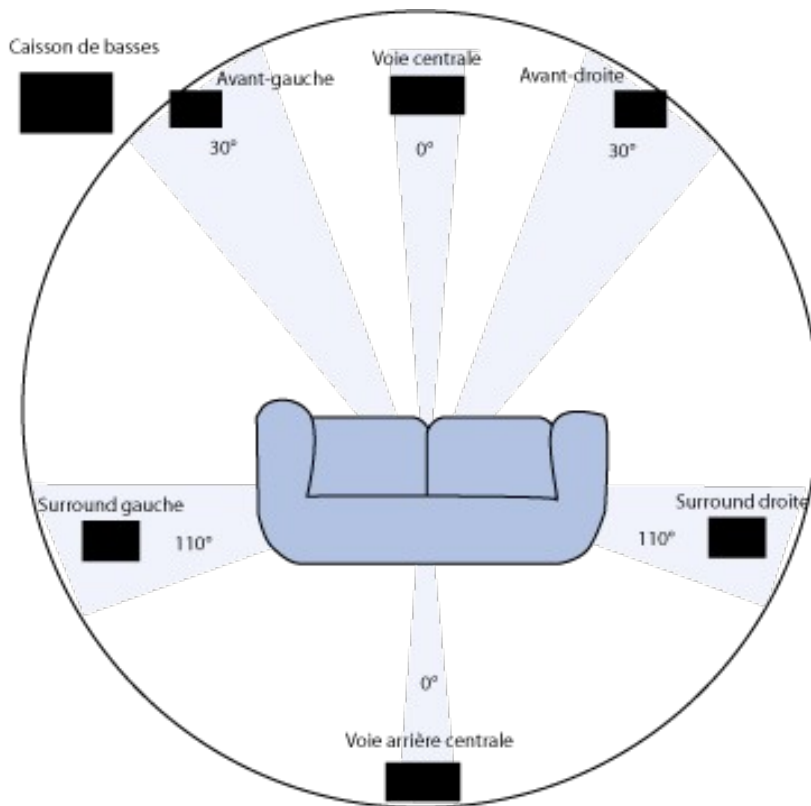
3- Le caisson de basse peut être disposé n'importe où dans la pièce, de préférence posé sur le sol pour mieux transmettre les vibrations. Le mieux est de tester différentes positions dans la pièce.

4- La position optimale pour les satellites arrière est en retrait par rapport à l'auditeur de manière à former un angle de 90° à 110° avec l'auditeur.



2.10. Configuration 6.1

La configuration 6.1 est similaire à une configuration 5.1, si ce n'est qu'elle ajoute une voie arrière centrale afin de compenser le « trou » entre les deux enceintes arrières.



2.11. Configuration 7.1

Une configuration 7.1 propose de combler le « trou » entre les deux enceintes arrières non pas par une enceinte mais par deux enceintes.

2.12. Dolby Surround

Le format **Dolby Surround**, mis au point en 1982 par la société *Dolby Labs*, est le premier système de codage de son numérique à ajouter un canal supplémentaire aux deux canaux du format stéréo. Ce canal supplémentaire, appelé « surround » permet d'ajouter une dimension supplémentaire au son (le terme anglais « *to surround* » signifie « envelopper »).

Les 3 canaux codés selon le format **Dolby Surround** sont transportées via les deux canaux stéréo traditionnels afin d'offrir une compatibilité avec le matériel existant. Ainsi, un décodeur est nécessaire pour pouvoir exploiter le 3ème canal mixé discrètement avec les deux premiers. Le canal surround possède une bande passante maximale de 7 kHz et un temps de retard d'environ 20 ms afin d'améliorer le réalisme des effets.

Les équipements possédant un décodeur *Dolby Surround* possèdent habituellement le logo suivant :



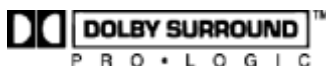
2.13. Dolby Surround Pro Logic

Le format **Dolby Surround Pro Logic** (appelé généralement *Dolby Pro Logic*), apparu en 1987, est une amélioration du format Dolby Surround ajoutant un quatrième canal aux 3 canaux du Dolby Surround, permettant de constituer une voie centrale destinée à reproduire la voix des acteurs.

Ainsi un système Dolby Pro Logic comporte les éléments suivants :

- Deux enceintes latérales en front
- Une enceinte centrale
- Deux enceintes arrières surround (en mono)

Il apporte également des améliorations en terme de qualité de restitution sonore et de réglages. Les équipements capables de décoder une source audio codées avec le procédé *Dolby Pro Logic* arborent généralement le logo suivant :



2.14. Dolby Pro Logic II

Le **Dolby Pro Logic II**, créé en août 2000, permet une reconstitution artificielle d'un environnement sonore 5.1 grâce à des traitements informatiques appliqués à une source stéréo (2.0) ou Dolby Surround (3.0/4.0/4.1).

Les équipements capables de décoder une source audio codées avec le procédé *Dolby Pro Logic II* arborent généralement le logo suivant :



Les systèmes Dolby Surround et Dolby Pro Logic sont désormais dépassés depuis l'émergence du Dolby Digital et du DTS. Ils restent néanmoins utiles pour permettent de reconstituer un son 3D à partir d'une source stéréophonique.

2.15. Dolby Pro Logic IIx

Le standard **Dolby Pro Logic IIx**, élaboré en 2003, permet de reconstituer artificiellement un environnement sonore 6.1 ou 7.1 à partir d'une source stéréophonique. Il propose plusieurs ambiances sonores selon le type d'usage que l'on souhaite réaliser :

- *Movie* pour la lecture de films,
- *Music* pour la lecture de CD audio,
- *Game* pour les jeux vidéos.

Les équipements capables de décoder une source audio codées avec le procédé *Dolby Pro Logic IIx* arborent généralement le logo suivant :



3. Les types de formats musicaux

Il est important de faire le distinguo entre un format de fichier et un codec. Un codec effectue le codage et le décodage des données brutes audio tandis que les données elles-mêmes sont stockées dans un fichier avec un format spécifique fichier audio. Certains formats de fichier audio peuvent contenir des données obtenues avec des codecs différents.

3.1. Les formats audio sans compression

Il existe un format audio non compressé, PCM, qui est généralement stocké sous forme de WAV sur Windows ou sous Aiff sur Mac OS. WAV et AIFF sont des formats de fichiers flexibles conçus pour stocker plus ou moins n'importe quelle combinaison de taux d'échantillonnage ou de bitrates (bit rate). Ce sont les formats de fichier appropriés pour le stockage et la réalisation d'enregistrements originaux.

RAW

RAW (*Real Audio Wrapper*) est un format audio utilisé pour représenter les données de son en modulation d'impulsion codée sans en-tête ni métadonnées.

WAV

Le format WAV (ou WAVE), (WAVEform audio format) est une extension de fichiers audio, il s'agit d'un conteneur capable de recevoir des formats variés. Il est basé sur le format de fichier RIFF, lequel est semblable au format IFF.

Mono ou stéréo, il a été mis au point par Microsoft et IBM.

Le suffixe des fichiers créés est **.wav**

BWF

Le BWF (Broadcast Wave Format) est un format audio standard créé par l'European Broadcasting Union en tant que successeur du WAV. Le BWF permet de stocker des métadonnées dans le fichier. Voir European Broadcasting Union : Spécification du Broadcast Wave Format (EBU Technical document 3285, juillet 1997). Il s'agit du format d'enregistrement usuel utilisé dans de nombreuses stations de travail audio professionnel de la télévision et du cinéma. Les Fichiers BWF incluent une référence standardisée Timestamp qui permet et facilite la synchronisation avec un élément d'image distincte. Stand-alone, basé sur

des fichiers, multi-enregistreurs de Sound Devices, Zaxcom, HHB USA, (**en**) en:Fostex, et Aaton tous utilisent BWF comme leur format préféré.

AIFF

L'AIFF est un format de stockage de sons sur les ordinateurs d'Apple. C'est l'équivalent du format WAV dans le monde Windows.

Les résolutions 8, 16, 20, 24 et 32 bits (à virgule flottante) sont acceptées.

Le suffixe des fichiers créés est « . **aif** »

Une variante l'AIFF-C permet de compresser la taille jusqu'à 6x.

CAF

Le CAF (*Core audio format*) a été développé par Apple pour s'affranchir des limitations de conteneur audio plus ancien comme le AIFF ou le WAV.

Il est compatible avec le système Mac OS X d'Apple depuis la version 10.3 et est lisible par Quicktime 7.

CDA

Le CDA (*Compact Disc Audio*), est un format Microsoft spécifique à Windows, des pistes des CD audio, telles qu'elles apparaissent lorsqu'elles sont insérées dans le lecteur CD-ROM. Les CD audio du commerce répondent à la norme professionnelle "Red Book". La technique d'échantillonnage du son utilisée pour les disques compacts est la modulation d'impulsion codée (en anglais PCM, pour *Pulse Coded Modulation*).

Le suffixe des fichiers créés est . **cda**

3.2. Les formats audio compressés sans perte

La compression sans perte (*lossless*) signifie qu'on utilise un algorithme tel qu'on peut toujours retrouver les données d'origine. Dans l'absolu, il existe toujours un fichier d'origine tel que l'algorithme ne ferait pas gagner d'espace disque.

Typiquement, la compression sans perte permet de diviser la taille des fichiers par deux ou trois. Elle est relativement peu utilisée, car ce gain est très faible en comparaison de ceux permis par la compression avec perte (ce qui est un gros handicap pour les échanges de fichiers), et assez gourmande en temps de calcul. Aucun standard n'a donc suffisamment convaincu pour devenir universellement lisible.

ATRAC

L'ATRAC (*Adaptive Transform Acoustic Coding*) est une technique de compression audio avec et sans pertes développée par Sony en 1992. Ce format a subi plusieurs évolutions : ATRAC3, ATRAC3plus (familièrement écrit ATRAC3+) et ATRAC Advanced Lossless se sont succédé respectivement en 1999, 2002 et 2006.

FLAC

Le format flac (*Free Lossless Audio Codec*), est un format libre de compression audio sans perte. Maintenu par la fondation Xiph.org, il est apprécié pour conserver la qualité des fichiers sonores originaux en alternative aux formats de compression avec perte type mp3.

3.3. Les formats audio compressés avec perte

La compression audio avec perte (*lossy*) se base sur des algorithmes spécialisés pour déterminer quelles transformations simplifient la représentation du son tout en étant perçue quasiment de la même manière par l'oreille humaine. Elle diminue la taille du fichier en éliminant les nuances perçues comme les moins utiles. L'élimination est définitive, créer un fichier dans un format de haute qualité à partir d'un fichier compressé avec perte ne sert strictement à rien.

Le format le plus connu est le MPEG-1/2 Audio Layer 3, dont le suffixe est .mp3. Ce format propose une qualité sonore très correcte pour un débit de 128 Kbits/s. C'est ce format qui a été massivement utilisé pour transférer les musiques via internet dès la fin des années 1990. Rapidement, des baladeurs avec une mémoire réenregistrable et capables de lire directement ce format sont apparus.

Dans la décennie 2000, de nouveaux formats ont été proposés. Vu les progrès des algorithmes, ils surpassent largement le MP3 en terme de qualité à débit égal, et peuvent atteindre des qualités supérieures. De plus, certains sont moins contraignants que le MP3 quant aux droits d'utilisation (le Ogg est un format libre). Mais le MP3 reste le plus utilisé, car l'arrivée en continu de nouveaux formats, apportant un avantage assez faible par rapport aux précédents, ne permet pas de mettre en place un standard meilleur que le MP3 et lisible par tous les baladeurs. En particulier, le fait que les iPod d'Apple ne lisent que le MP3 et l'AAC freine assez fortement les initiatives dans ce domaine.

Pour un même format de compression, il n'y a pas de manière unique de coder, car chaque algorithme cherche la meilleure manière de représenter le son d'origine suivant le langage de compression. En particulier, les codecs de MP3 ont réalisés des progrès très importants depuis le début de l'utilisation de ce format.

Elle permet typiquement un gain d'un facteur 10 de taille du fichier. Cela a rendu possible non seulement le stockage d'un temps d'écoute formidable sur les supports informatiques, mais aussi leur échange par internet, souvent illégalement.

MP3

Star incontestée des formats audio de diffusion, MP3 (*MPEG-1 Layer III*), est l'abréviation de MPEG-1/2 Audio Layer 3. Cet algorithme de compression prend naissance en 1987. L'ISO en fera un standard dans les années 92-93. La couche (*Layer*) III est la couche la plus complexe. Elle est dédiée à des applications nécessitant des débits faibles (128 Kbits/s) d'où une adhésion très rapide du monde Internet à ce format de compression. Les taux de compression (*ratio*) sont d'ordinaire de 1 pour 10 (1:10) (1:4 à 1:12). Très rapide à l'encodage. Des royalties importantes sont à payer pour exploiter la licence MP3. Utiliser l'encodeur MP3 LAME dernière version, encodé à 130 Kbit/s (V5) permet d'obtenir une *qualité comparable* au AAC (Advanced Audio Coding) encodé à 48 kbit/s. ⁶

Le suffixe des fichiers créés est. **mp3**

Type de compression : constant ou variable (VBR)

mp3PRO

Le format mp3PRO, fruit de la collaboration entre Thomson Multimédia et l'Institut Fraunhofer, combine l'algorithme MP3 et un système améliorant la qualité des fichiers comprimés appelé **(en)** SBR pour *Spectral Bandwidth Replication*.

Ce format a été publié à la fin de 2001 ; un fichier MP3pro 64 Kbit/s a une qualité équivalente à un MP3 à 128 Kbit/s.

Le suffixe des fichiers créés est «. **mp3** »

Ogg Vorbis

Le format Ogg Vorbis est un format libre, fruit de la fondation Xiph.org. Vorbis se différencie des MP3, WMA et autre AAC par son algorithme. Il segmente les sources audio en paquets successifs, l'algorithme de compression agissant dans un premier temps sur chaque paquet indépendamment des autres. Cela lui permet d'avoir très peu de faiblesses sur certaines fréquences et de conserver la même qualité quel que soit le type de musique.

Le suffixe des fichiers créés est «. **ogg** » ou parfois «. **oga** ».

Par abus de langage, on appelle 'fichier Ogg' des fichiers musicaux compressés par l'algorithme Vorbis. Ceci peut être particulièrement dérangent à l'ère des baladeurs numériques supportant audio et vidéo. En fait, Ogg est un conteneur qui peut contenir des pistes sonores (Vorbis), audio sans perte (FLAC), audio parlées (Speex) et vidéo (Theora). Un

'fichier Ogg' peut donc contenir l'un ou l'autre (ou une combinaison) de pistes. Pour être plus clair, nous devrions parler de fichier Ogg Vorbis lorsque nous mentionnons un fichier .ogg qui ne contient qu'une piste sonore au format Vorbis.

VQF ou TwinVQ

Le format TwinVQ (*Transform-domain Weighted Interleave Vector Quantization*), a été développé par NTT Cyber Space Laboratories et soutenu par Yamaha. Dans le même esprit que le MP3, il comprime encore plus et avec une meilleure qualité. On regrettera une durée d'encodage un peu trop longue, près de 10 fois plus lente que le MP3. De plus, arrivé bien plus tard, et distribué sous une licence très restrictive, il a eu peu d'adeptes et est plus ou moins abandonné.

Le suffixe des fichiers créés est **.vqf**, **.vql** ou **.vqe**

WMA

Le format WMA (*Windows Media Audio*), créé par Microsoft à partir des recommandations MPEG-4 en 1999, est utilisé par le logiciel Windows Media Player. Ce format est lié à une gestion pointue des droits d'auteurs (Gestion numérique des droits, en anglais *Digital Right Management* ou DRM) qui permet de définir par exemple une durée de vie limitée pour les fichiers ou d'interdire les possibilités de gravure.

Il existe plusieurs versions du codec (wma7.1, wma9, wma pro).

Le suffixe des fichiers créés est « **wma** »

AU

Le format AU est assez bien répandu grâce à Unix et Linux. La fréquence d'échantillonnage est comprise entre 1 kHz et 200 kHz. Mais les applications de rendu audio ne lisent principalement que trois fréquences d'échantillonnage : 8012.821 (codec entré), 22050 et 44100 hertz.

Le suffixe des fichiers créés est **.au**

Les résolutions 8, 16, 20, 24 et 32 bits (flottant) sont acceptées.

ASF

ASF *Advanced Streaming Format*, est un format conteneur de Microsoft servant au streaming audio et vidéo.

AA

(en) AA *Audible*, est un format utilisé par Apple pour les *audio-books*.

AAC ou MPEG-2 AAC

L'AAC (Advanced Audio Coding), est une extension du MPEG-2 et a été amélioré en MPEG-4, MPEG-4 Version 2 et MPEG-4 Version 3. Il a été reconnu fin avril 1997.

Le suffixe des fichiers créés est « **aac** », « **.mp4** », « **.m4a** »

Apple et l'AAC

Apple a choisi l'ACC comme codec privilégié, on le retrouve dans son iPod et son logiciel iTunes. Pour la vente musicale en ligne iTunes Music Store, la norme AAC ne proposant pas de système de gestion des droits numériques (DRMs), Apple a développé son propre système, appelé FairPlay. Il est lisible sur Mac OS et Windows, uniquement avec le logiciel iTunes. Le fait que l'AAC soit le seul format de compression à pertes plus performant que le MP3 qui soit supporté par les iPod a fortement contribué à sa popularité. Insuffisamment cependant pour qu'il s'impose comme successeur du MP3, d'autres formats cités faisant largement jeu égal en performances.

4. Calcul du débit et du poids d'un fichier audio

Calcul du débit :

Débit = fréquence d'échantillonnage x quantification x nb de piste = Kbps

Exemple avec un CD audio : débit = 44.1 KHz x 16 bits x 2 = 1411 kbps

Calcul du poids

Pour un format à débit constant :

Taille (Ko) = temps (s) x débit (Kbit/s) / 8

Exemple : pour un fichier MP3 encodé à 192 Kbit/s d'une durée de 3 minutes

Taille = 3 x 60 x 192 / 8 = 4 320 Ko

Pour un format à débit variable :

Taille moyenne(Ko) = temps (s) x débit moyen (Kbit/s) / 8

Taille maximale(Ko) = temps (s) x débit maximum (Kbit/s) / 8