



HAL
open science

La compression des signaux, résultat d'une approche par objectif, dans un contexte de ressources limitées -Scilab, LAME, WAV, MP3

Pierre Maréchal, Hugues Duflo

► To cite this version:

Pierre Maréchal, Hugues Duflo. La compression des signaux, résultat d'une approche par objectif, dans un contexte de ressources limitées -Scilab, LAME, WAV, MP3. Congrès National de la Recherche des IUT (CNRIUT), Jun 2023, Saint-Pierre de la Réunion, France. hal-04220317

HAL Id: hal-04220317

<https://hal.science/hal-04220317>

Submitted on 27 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La compression des signaux, résultat d'une approche par objectif, dans un contexte de ressources limitées – Scilab, LAME, WAV, MP3

Pierre Maréchal

pierre.marechal@univ-lehavre.fr

Hugues Duflo

hugues.duflo@univ-lehavre.fr

IUT Le Havre, Université Le Havre Normandie
Laboratoire Ondes et Milieux Complexes, LOMC UMR 6294 CNRS

Thèmes – Informatique - Électronique - Mathématiques - Communication - Signal

Résumé – *Les programmes mis en œuvre ici visent à soulever certains détails des signaux que nous utilisons au quotidien. La qualité des signaux que nous recevons via nos terminaux résultent d'un choix a priori par des experts compétents en la matière, mais dont l'utilisation est trop souvent relayée à des utilisateurs pressés par la simplicité et l'efficacité. Que ces préoccupations soient légitimes n'empêche pas, qui plus est dans une formation universitaire de s'intéresser au codage de l'information et d'en analyser les mécanismes et les limites. Les sons ont cette qualité de faire appel à nos sens, et nos perceptions de soulever la question de la qualité d'un signal. Parmi les outils à disposition, les logiciels libres permettent de faire des choix de compression audio (<https://lame.sourceforge.io/index.php>), et de constituer des éléments de décision et d'analyse quantitative (<https://www.scilab.org/>). La question de la compression est très souvent relayée à l'arrière plan bien que centrale dans nos usages, d'où l'intérêt de sensibiliser et former les étudiants à ces questions auxquelles ils sont implicitement confrontés, que ce soit dans leur vie personnelle, leur vie d'étudiant et leur vie professionnelle. Les signaux en question sont compressés pour des raisons historiques de limites de mémoire et de débit. S'intéresser au format MP3 (<https://en.wikipedia.org/wiki/MP3>), c'est s'intéresser aussi à l'histoire logicielle, celle des formats de fichiers et des motivations de diffusion (<https://museumofportablesound.com/mp325-timeline/>). Bien que de nombreuses limites matérielles aient été dépassées, le compromis compression/qualité demeure une préoccupation majeure pour les hébergeurs et diffuseurs de contenus. Les vertus d'un codage et d'une compression optimale constituent donc un sujet majeur pour la formation au choix de compromis techniques, économiques, mais aussi éthique au sens de l'amélioration continue de l'utilisation des ressources à disposition.*

Mots-Clés – Logiciel libre, compression audio, qualité du signal.

Introduction

Le traitement de données numériques constitue un enjeu de premier plan pour les étudiants, en particulier ceux nés après 2000. La multiplication des systèmes d'exploitation grand public, dont l'ergonomie a permis une large diffusion s'est faite au détriment de la maîtrise des codages associés. Que ce soit pour les formats vidéo, image ou audio, la multiplicité de l'offre par des codages propriétaires viennent confisquer la question de l'accessibilité aux ressources logicielles et aux choix attenants. La question des logiciels libres est souvent éludée par la culture d'un pragmatisme débridé. Et pourtant, la complémentarité de l'efficacité et la culture du partage des codes source n'est plus à démontrer [1], tant au niveau technique qu'économique, au sens d'outils communs à des nombreux acteurs. Pour preuve, les rachats de Android par Google en 2005 et GitHub par Microsoft en 2018, respectivement pour 5,4 et 7,5 milliards de dollars ont bousculé le monde de l'open source. Sur d'autres plans, de nombreux logiciels et formats libres démontrent que les usages, la mutualisation des besoins et des moyens produisent des alternatives ouvertes [2] au bénéfice de tous. Au-delà de l'aspect informatique, c'est le choix du logiciel libre et ouvert qui est illustré *via* la problématique de la compression des signaux audio et l'analyse de la quantification d'erreur du format MP3 avec les logiciels libres LAME pour la compression MP3 et Scilab pour l'analyse des signaux WAV.

1 Contexte et motivation

Bien qu'incontournable, la maîtrise de la numérisation des données et des données numériques n'est pas chose évidente, et est trop souvent contournée par des béquilles logicielles. Derrière l'ergonomie apparente se cache la délégation totale de problématiques complexes. En 1840, le code Morse fait référence avec un codage compressé permettant une transmission de l'information via un codage numérique. Cette approche empirique marque le début d'un essor qui ne cessera plus. La formalisation du codage de l'information numérique par Shannon en 1948 pose la question de l'optimisation du codage [3]. Par la suite les algorithmes de codage de Huffman [4], puis LZW (pour Lempel, Ziv et Welch) [5] permettront de compresser des données de tous types. Pour les images, viendront les algorithmes spécifiques au codage d'images dont les standards JPG (destructif) [ITU-T81, 1992] et PNG (non destructif) [ISO/CEI 15948:2004]. Les signaux audio verront apparaître les formats (destructifs) de compression FLAC (2000), MP3 (1992), OGG (2000). Enfin, pour les données vidéo, les formats (destructifs) DivX (Rota, 1999), Xvid (GNU GPL, 2001), H.264/MP4 (ISO/CEI 14496-10:2003).

Les formats de données visuelles, tout comme celles audio tirent avantage des limites de détection, traitement, et analyse de nos sens. La qualité des signaux obtenus peut s'évaluer *via* des *benchmarks*, mais aussi des panels d'utilisateurs qui évaluent en double aveugle la qualité des compressions réalisées. Dans le cas du travail présenté ici, on s'intéresse à la quantification numérique de l'erreur relative moyenne entre une source audio WAV et sa version compressée par une méthode destructive MP3.

2 Compression audio

La compression audio telle que réalisée par le format MP3 comporte différentes étapes, dont une compression destructive dite "psycho-acoustique" tirant avantage du masquage audio permettant d'effacer les données relatives à des sons rendus inaudibles par d'autres sur une fenêtre temporelle donnée. Cette approche a permis de remplir le cahier des charges initiales qui consistait à opérer une compression d'un taux supérieur à 10 relativement au standard WAV [6]. L'objectif s'est trouvé atteint avec le MP3 *via* le débit de données (*bitrate*) fixé en standard à 128 kbps (kilobits par seconde) contre le standard WAV avec une fréquence d'échantillonnage à 44,1 kHz, une quantification à 16 bits par échantillon, et 2 canaux pour la stéréo, soit $44100 \text{ ech/s} \times 16 \text{ bits/ech/canal} \times 2 \text{ canaux}$, *i.e.* un *bitrate* à 1440 kbps pour le WAV en stéréo.

2.1 Méthode

Un jeu de fichiers audio de référence est choisi puis les erreurs relatives de compression sont comparées pour ces différentes sources [7]. Dans l'objectif d'évaluer la performance de la compression destructive MP3, des signaux d'instruments connus pour leur richesse harmonique ont été choisis, avec un jeu de cinq fichiers source au format WAV. Pour l'analyse d'erreur relative, les fichiers ont été tronqués à 3 s soit $N=132\ 300$ échantillons. La méthodologie adoptée a donc été de compresser ces fichiers avec un *bitrate* constant fixé de 32 à 320 kbps, puis de les comparer au fichier source. Le logiciel libre LAME a été utilisé pour effectuer les compressions et conversions WAV à MP3, et réciproquement, puis le traitement des signaux temporels a été programmé avec Scilab.

2.2 Quantification de l'erreur

Afin de quantifier les modifications apportées par la compression MP3, un estimateur doit être choisi. Par souci de simplicité, l'erreur relative $\varepsilon(k)$ est évaluée pour chaque échantillon k (dont la source est non nulle) entre le fichier WAV source et le fichier compressé MP3 :

$$\varepsilon(k) = \frac{s_{\text{WAV}}(k) - s_{\text{MP3}}(k)}{s_{\text{WAV}}(k)} \quad (1)$$

Dans l'objectif de comparer les erreurs relatives obtenues pour chaque échantillon (eq. (1)) pour chaque fichier et pour chaque *bitrate*, l'erreur relative moyenne est calculée :

$$\varepsilon_{moy} = \frac{1}{N} \sum_{k=1}^N \left| \frac{s_{WAV}(k) - s_{MP3}(k)}{s_{WAV}(k)} \right| \quad (2)$$

3 Résultats

Pour illustration, l'erreur relative (eq. (1)) est évaluée à partir d'un fichier source au format WAV sur une compression MP3 à 128 kbps (Fig. 1).

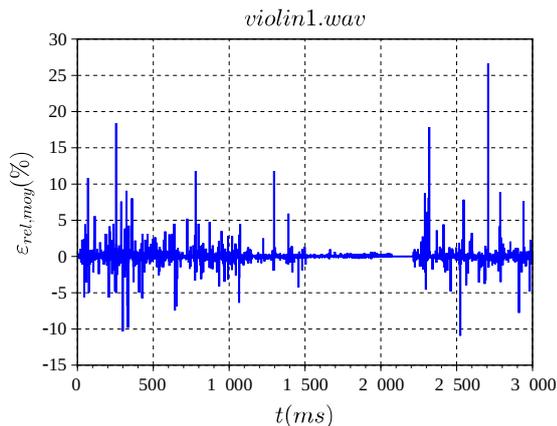


Figure 1 – Erreur relative MP3 vs WAV relevée pour un *bitrate* fixe de 128 kbps.

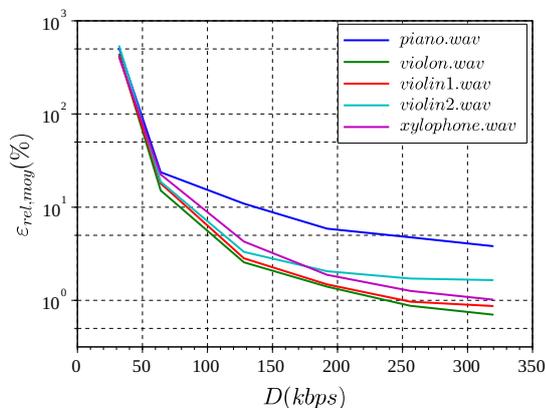


Figure 2 – Erreur relative moyenne liée au *bitrate* fixe de compression MP3.

Les erreurs relatives moyennes ont été calculées sur cinq fichiers et les résultats obtenus sont illustrés (Table 1) pour différentes compressions, dont un *bitrate* minimal à 32 kbps, standard à 128 kbps, et maximal à 320 kbps :

Table 1 – Erreur relative moyenne ε_{moy} (%)

D (kbps)	piano	violon	violin1	violin2	xylo
32	513	445	424	545	411
128	10,9	2,56	2,83	3,31	4,27
320	3,81	0,70	0,87	1,65	1,02

4 Discussion

Au-delà du résultat attendu de la réduction de l'erreur relative moyenne qui décroît avec le taux de compression (Fig. 2), les résultats diffèrent d'un fichier à l'autre. Dans le cas du piano, la dégradation liée à la compression MP3 reste conséquente, et même avec un *bitrate* de 320 kbps, on atteint l'erreur relative moyenne de 3,81 %. Pour cet instrument, des contributions harmoniques significatives ont été tronquées. Cette dégradation est nettement visible *via* une analyse temps-fréquence réalisée avec Scilab et fait bien ressortir le filtrage, inhérent à la méthode de compression MP3.

5 Conclusion

La programmation mise en œuvre *via* des logiciels libres tels que LAME et Scilab ont permis de traiter et d'analyser des signaux compressés au format audio MP3. La comparaison des résultats obtenus a mis en évidence les aspects destructifs liés à la compression MP3, en fonction du *bitrate*, mais aussi une quantification de l'erreur relative moyenne produite pour cinq échantillons de fichiers. Les outils développés gagneraient à être diffusés et repris pour des analyses étendues à de nouvelles études de cas.

Références

- [1] Williams, S., Stallman, R. and Masutti, C.: Richard Stallman et la révolution du logiciel libre. Framasoft; 2010.
- [2] Giesler, M.: "Conflict and Compromise: Drama in Marketplace Evolution", *Journal of Consumer Research*, vol.34, p.739-753, 2008, (doi.org/10.1086/522098).
- [3] Shannon, C.: "A mathematical theory of communication", *Bell System Technical Journal*, vol.27, p.379-423, 1948, (doi.org/10.1002/j.1538-7305.1948.tb01338.x).
- [4] Huffman, D.A.: "A method for the construction of minimum-redundancy codes", *Proc. IRE*, vol.40, p.1098-1101, 1952, (doi.org/10.1109/JRPROC.1952.273898).
- [5] Welch, T.: "A technique for high-performance data compression", *Computer*, vol.17, p.8-19, 1984, (doi.org/10.1109/MC.1984.1659158).
- [6] Le Guyader, A., Philippe, P. and Rault, J.B.: "A review of speech and audio coding standards (ITU-T, ETSI and ISO/MPEG)", *Annales Des Télécommunications*, vol.55, p.425-441, 2000, (doi.org/10.1007/BF02995200).
- [7] Papadakis, N.M., Aroni, I. and Stavroulakis, G.E.: "Effectiveness of MP3 coding depends on the music genre: Evaluation using semantics", *Acoustics*, vol.4, p.704-719, 2022, (doi.org/10.3390/acoustics4030042).