

Rapport de CSI

Table des matières :

1. [Introduction](#)
2. [Corrélation entre descripteurs audio et paramètres de jeu](#)
 1. [Introduction](#)
 2. [Méthodologie](#)
 1. [Données](#)
 2. [Descripteurs audio](#)
 3. [Test statistique](#)
 3. [Résultats](#)
 1. [Résultats de la selection des meilleurs descripteurs](#)
 4. [Discussion et Conclusion](#)
 3. [Identification individuelle de violon par leur son](#)
 1. [Introduction](#)
 2. [Méthodologie](#)
 1. [Base de données](#)
 2. [Descripteurs audio](#)
 3. [Résultats](#)
 4. [Discussion et Conclusion](#)
 4. [Expérience au CNSMDP](#)
 1. [Introduction](#)
 2. [Méthodologie](#)
 1. [Protocole expérimental](#)
 2. [Enregistrements](#)
 3. [Traitement des enregistrements](#)
 4. [Traitement des notes](#)
 3. [Résultats](#)
 1. [Enregistrements](#)
 2. [Notes](#)
 4. [Discussion et Conclusion](#)
 5. [Conclusion](#)
 6. [Bibliographie](#)
 7. [Annexes](#)
 1. [Segmentation automatique](#)
 1. [Enveloppe](#)
 2. [Seuillage](#)
 3. [Prolongement et contraintes](#)

Introduction

L'identité sonore d'un violon — si tant est qu'elle existe — résulte-t-elle principalement des caractéristiques de l'instrument ou bien est-elle influencée par la façon de jouer de l'instrumentiste ?

Il est possible de mesurer nombre de caractéristiques propres à un violon : son âge, sa géométrie, les propriétés des matériaux utilisés dans sa construction... Parmi elles, les réponses en fréquence du corps du violon et du chevalet ont déjà été utilisées pour juger de la qualité d'un instrument [1] voire pour en simuler le son (réf : Claudia ?). Par ailleurs, de nombreuses études se sont intéressées à la perception des violonistes, notamment au vocabulaire employé afin de noter différents instruments.

Il reste cependant difficile de corrélérer les mesures acoustiques (sans violonistes) avec les qualités perçues par les instrumentistes.

Se posent alors les questions suivantes : un violon a-t-il une identité sonore propre ? Quelle est l'influence des violonistes sur le son produit ? Est-il possible de dissocier le son d'un violon du son produit par un·e violoniste ?

La suite de ce rapport est structurée de la manière suivante, plus ou moins fidèle à la chronologie de ma première année : tout d'abord, j'identifierai des corrélations entre des paramètres de jeu du violon et certains descripteurs audio. Ensuite, je me concentrerai sur le problème de l'identification de violon par leur son. Enfin, je présenterai une expérience en cours au CNSMDP visant à mettre en évidence l'influence du jeu sur un instrument.

Corrélation entre descripteurs audio et paramètres de jeu

Introduction

Certains paramètres de jeu du violon ont une influence flagrante sur le son produit. Par exemple, une note jouée *ponticello* (avec un archet proche du chevalet) n'aura pas la même sonorité que cette même note jouée *sul tasto* (avec un archet sur la touche).

Dans (réf : The Violinist's Sound Palette: Spectral Centroid, Pitch Flattening and Anomalous Low Frequencies), les auteurs s'intéressent à l'influence de trois paramètres (vitesse d'archet, distance archet-chevalet et force d'archet) sur le son produit à l'aide d'un monocorde et d'une machine à frotter.

Cependant, ces conditions expérimentales semblent assez artificielles par rapport à la réalité, où l'instrumentiste est moins précis·e et répétable qu'une machine et où l'instrument est plus complexe qu'un monocorde.

Est-il possible de vérifier les corrélations observées dans [réf] dans des conditions réelles de jeu ?

Je commencerai par présenter la méthodologie utilisée, avant de détailler les résultats obtenus. Enfin, je concluerai par une discussion des principaux enseignements de cette étude.

Méthodologie

Données

Afin de vérifier les résultats présentés dans [2], j'ai réalisé des enregistrements sur plusieurs violons en faisant varier les mêmes paramètres à savoir la vitesse d'archet, la force appliquée, et la distance au chevalet. Dans le but de correspondre autant que possibles aux conditions réelles de jeu, je n'ai pas mesuré les niveaux de ces différents paramètres mais me suis donné des consignes "musicales". Ainsi, une vitesse d'archet lente, avec une force appliquée forte et une distance au chevalet faible se traduit en une note longue, jouée *forte* et *ponticello*. Chaque paramètre pouvait prendre trois valeurs :

- Vitesse d'archet \in [lente, moyenne, rapide]
- Force appliquée \in [piano, mezzo, forte]
- Distance au chevalet \in [ponticello, ordinario, sul tasto]

12 enregistrements de ces 27 combinaisons possibles ont été réalisés sur une note tenue (plus précisément, une corde de ré à vide). Ces enregistrements ont été réalisés avec une fréquence d'échantillonnage de 44.1Hz, dans une pièce peu réverbérante, et sur plusieurs jours afin d'introduire une certaine variabilité dûe au violoniste.

Descripteurs audio

Table 1. Liste des descripteurs utilisés

Type	Nom	Description	Formule	Utilisé dans
Temporelle	Attack	Durée de l'attaque d'une note	$t_{\text{attack end}} - t_{\text{attack start}}$	
Temporelle	Attack slope	Durée de l'attaque d'une note	$t_{\text{attack end}} - t_{\text{attack start}}$	
Temporelle	Decrease slope	Durée de l'attaque d'une note	$t_{\text{attack end}} - t_{\text{attack start}}$	
Temporelle	Log-attack	Logarithme de la durée de l'attaque d'une note	$\log(t_{\text{attack end}} - t_{\text{attack start}})$	
Temporelle	Duration	Durée de la note	$t_{\text{end}} - t_{\text{start}}$	
Temporelle	RMS	Quantifie le volume d'un signal sonore	$RMS[m] := \frac{1}{W} \sum_{k=mW}^{(m+1)W} x[k]^2$	
Temporelle	ZCR	Nombre de fois que le signal passe par 0. Quantifie l'harmonicité.	$ZCR[m] := \frac{1}{W} \sum_{k=mW}^{(m+1)W} 1_{x[k] < 0}$	
Temporelle	Temporal Centroid	Représente le centre de gravité temporel d'une note	$\mu := \frac{\sum_{n=n_1}^{n_2} t[n] RMS[n]}{\sum_{n=n_1}^{n_2} RMS[n]}$	
Spectral	Spectral Centroid	Représente le centre de gravité du spectre	$\mu_1[n] := \frac{\sum_{k=k_1}^{k_2} f[k] S[k, n]}{\sum_{k=k_1}^{k_2} S[k, n]}$	
Spectral	Spectral Spread	Représente la déviation autour du Spectral Centroid	$\mu_2[n] := \frac{\sum_{k=k_1}^{k_2} (f[k] - \mu_1[n])^2 S[k, n]}{\sum_{k=k_1}^{k_2} S[k, n]}$	
Spectral	Spectral Skewness	Quantifie l'asymétrie du spectre autour du Spectral Centroid	$\mu_3[n] := \frac{\sum_{k=k_1}^{k_2} (f[k] - \mu_1[n])^3 S[k, n]}{\mu_2^3 \sum_{k=k_1}^{k_2} S[k, n]}$	
Spectral	Spectral Spread	Quantifie le fait que le spectre soit pointu autour du Spectral Centroid	$\mu_4[n] := \frac{\sum_{k=k_1}^{k_2} (f[k] - \mu_1[n])^4 S[k, n]}{\mu_2^4 \sum_{k=k_1}^{k_2} S[k, n]}$	

Spectral	Spectral Slope	Pente de l'amplitude du spectre vue comme fonction de $f[k]$	$\alpha := \frac{\sum_k (f[k] - \bar{f})(S[k, n] - \bar{S}[n])}{\sum_k (f[k] - \bar{f})^2}$
Spectral	Spectral Decrease	Représente la décroissance du spectre	$SD := \frac{\sum_{k=k_1+1}^{k_2} \frac{S[k, n] - S[k_1, n]}{k - k_1}}{\sum_{k=k_1}^{k_2} S[k, n]}$
Spectral	Spectral Rolloff	Bande de fréquence contenant $\alpha\%$ de l'énergie du spectre	$f[k_{RO}] \text{ tel que } \sum_{k=k_1}^{k_{RO}} S[k, n] = \alpha \sum_{k=k_1}^{k_2} S[k, n]$
Spectral	Spectral Variation	Mesure la variation entre deux trames successives d'un spectre	$SV := 1 - \frac{\sum_k S[k, n - 1]S[k, n]}{\sqrt{\sum_k S[k, n - 1]^2} \sqrt{\sum_k S[k, n]^2}}$
Spectral	Spectral Flux	Quantifie le fait que le spectre soit pointu autour du Spectral Centroid	$\mu_4[n] := \frac{\sum_{k=k_1}^{k_2} (f[k] - \mu_1[n])^4 S[k, n]}{\mu_2^4 \sum_{k=k_1}^{k_2} S[k, n]}$
Spectral	Spectral Flatness	Quantifie le fait que le spectre soit pointu autour du Spectral Centroid	$\mu_4[n] := \frac{\sum_{k=k_1}^{k_2} (f[k] - \mu_1[n])^4 S[k, n]}{\mu_2^4 \sum_{k=k_1}^{k_2} S[k, n]}$
Spectral	Spectral Crest	Quantifie le fait que le spectre soit pointu autour du Spectral Centroid	$\mu_4[n] := \frac{\sum_{k=k_1}^{k_2} (f[k] - \mu_1[n])^4 S[k, n]}{\mu_2^4 \sum_{k=k_1}^{k_2} S[k, n]}$

Test statistique

Afin de savoir quels descripteurs décrivent le mieux les changements de paramètres de jeu, un algorithme de sélection des meilleurs descripteurs est appliqué. Cet algorithme, issu de la bibliothèque *scikit-learn* de Python, calcule un test d'analyse de variance (ANOVA) permettant de savoir quels descripteurs permettent de séparer au mieux différents niveaux d'un paramètre de jeu. Pour chaque paramètre de jeu, les meilleurs descripteurs sont ainsi obtenus.

Résultats

Résultats de la sélection des meilleurs descripteurs

Les figures , et montrent respectivement les meilleures corrélations pour les paramètres de vitesse d'archet, de force / nuance et de distance au chevalet.

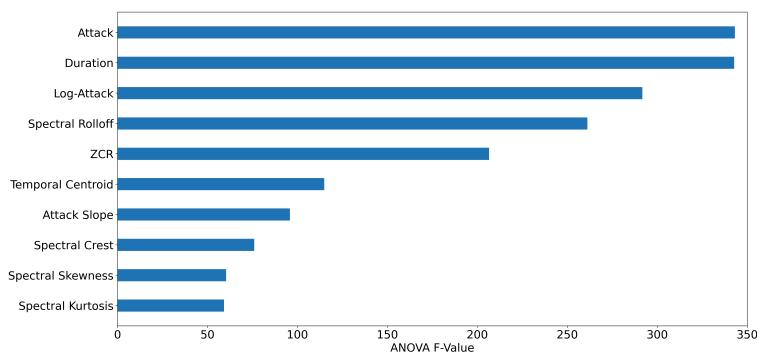


Figure 1. Corrélation des descripteurs avec la vitesse d'archet

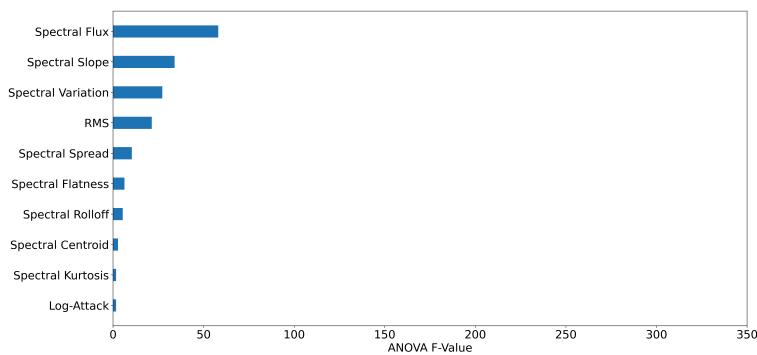


Figure 2. Corrélation des descripteurs avec la force / nuance

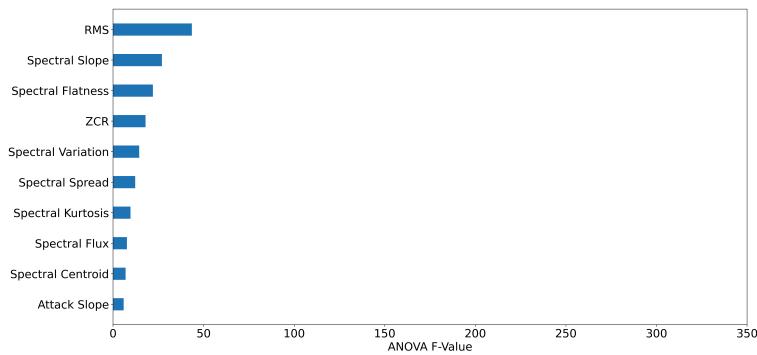


Figure 3. Corrélation des descripteurs avec la distance au chevalet

D'une manière générale, on remarque que les résultats du test statistique sont plus élevés quand il s'agit du paramètre  que pour les deux autres. On peut donc en conclure qu'il semble plus simple de distinguer plusieurs niveaux de vitesse d'archet que plusieurs niveaux des deux autres paramètres.

Les trois descripteurs les plus corrélés à la vitesse d'archet sont la durée de l'attaque, la durée de la note ainsi que le logarithme de la durée de l'attaque. Le fait que la durée de la note soit corrélée avec la vitesse de l'archet est ici trivial, puisque les notes enregistrées étaient plus courtes avec une vitesse d'archet élevée. Une majorité des descripteurs les plus corrélés sont des descripteurs issus de l'enveloppe du signal, et non pas de son spectre.

On remarque cependant que pour les paramètres de jeu

◇◇◇◇◇◇◇◇, les descripteurs les plus corrélés semblent être de nature spectrale. Ceci peut s'expliquer par le fait que des phénomènes tels que le *flattening*, qui fait baisser la fréquence de la note lorsque l'on joue sur la touche en appliquant beaucoup de force. Cela rejoue en partie les conclusions de [3].

Discussion et Conclusion

La corrélation entre certains descripteurs audio et les principaux paramètres contrôle de l'archet (vitesse, force, distance au chevalet) ont été étudiés ici à l'aide d'enregistrements réalisés dans des conditions réalistes de jeu. Les données ont été récoltées en subdivisant ces trois paramètres en trois niveaux d'intention musicale : *lent*, *moyen*, *rapide* pour la vitesse d'archet, *piano*, *mezzo*, *forte* pour la force, et *ponticello*, *ordinario* et *sul tasto* pour la distance au chevalet.

Les résultats montrent que dans ces conditions, les descripteurs classiques sont plus corrélés à la vitesse d'archet qu'aux autres paramètres. Plus précisément, ce sont des descripteurs issus de l'enveloppe du signal qui semblent être le plus corrélés à ce paramètre. Les deux autres (force et distance au chevalet) semblent se traduire par des variations d'ordre spectral (fréquence fondamentale, contenu harmonique, ...).

Afin de vraiment pouvoir comparer les résultats de cette analyse avec les résultats de [2], il serait intéressant de limiter le calcul des descripteurs à la partie stable de la note : cela enlèverait la plupart des descripteurs travaillant sur l'enveloppe du signal, mais rendrait peut-être l'analyse moins sujette aux variations en début et en fin de note.

Identification individuelle de violon par leur son

Introduction

L'identification d'instruments de musique est une tâche commune en *Music Information Retrieval* (MIR). Elle consiste à retrouver l'ensemble des instruments présents dans un enregistrement. C'est un problème qui a été longuement étudié, et, pour des enregistrements monophoniques (constitués d'un seul instrument), les modèles actuels permettent d'obtenir une précision de l'ordre de 100%.

Cependant, peu d'articles se sont intéressés à l'identification d'instruments au sein d'une même famille, comme par exemple l'identification de violons à partir de leur enregistrement.

Le but de cette étude est de passer en revue et de comparer différentes façons d'aborder ce problème, depuis la collection et l'exploration des données jusqu'aux différents algorithmes permettant de faire cette classification. Des conseils sur la manière de réaliser les enregistrements seront donnés et une version à temps long des MFCCs sera introduite.

Cet article est structuré comme suit : la section présente la méthodologie générale, décrivant la collecte des données, l'extraction des descripteurs audio, leur exploration et enfin la classification à l'aide de méthodes d'apprentissage automatique.

tique. Les résultats sont discutés dans la section . Enfin, des conclusions sont tirées dans la section , qui esquisse également d'éventuels développements futurs.

Méthodologie

Base de données

Dans le cadre du projet Bilbao, treize violons ont été fabriqués afin d'établir une relation entre leurs caractéristiques matérielles et géométriques et leur qualité tonale [4]. Ces violons ont été joués en 2019 par vingt-trois violonistes professionnels, chacun ayant enregistré une gamme sur chaque violon ainsi qu'un court extrait musical sur un violon de leur choix. Les enregistrements ont été réalisés dans des conditions identiques, dans une grande salle de répétition du conservatoire de Bilbao, en maintenant une distance constante entre le musicien et le microphone. Notre ensemble de données se compose ainsi de 13×23 gammes, plus 1×23 extraits musicaux.

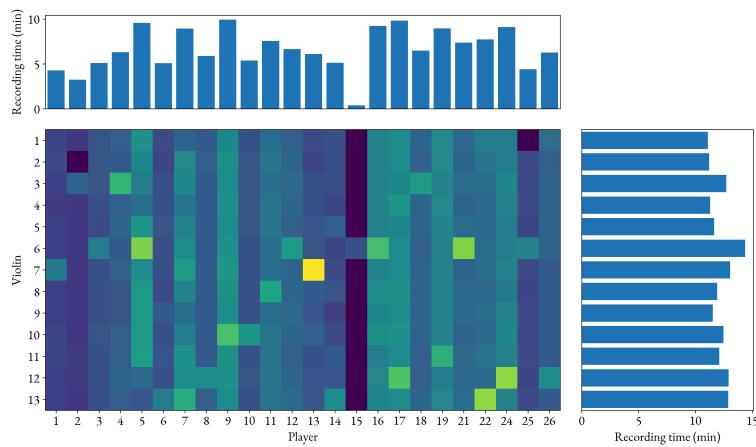


Figure 4. Temps d'enregistrement disponible par violoniste et par violon.

Six de ces treize violons (violons numéros 1, 4, 5, 9, 11 et 13) ont été apportés au workshop de Villfavard en 2024 et ont été enregistrés à nouveau. Ils ont été joués librement par quatre nouveaux musiciens dans une petite salle, dans des conditions assez différentes de celles des enregistrements de 2019.

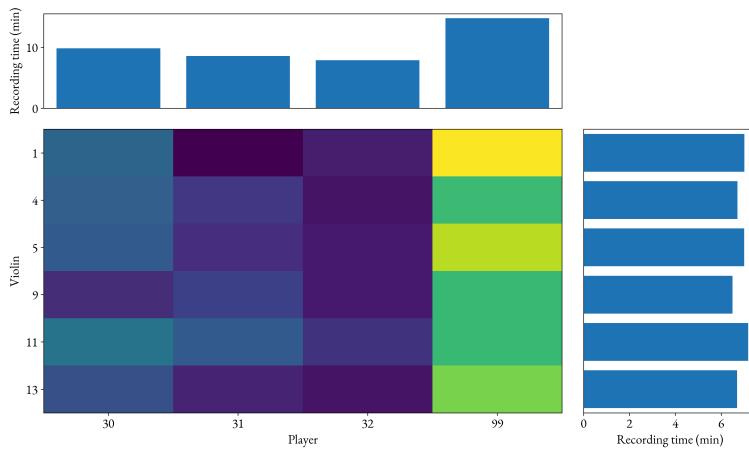


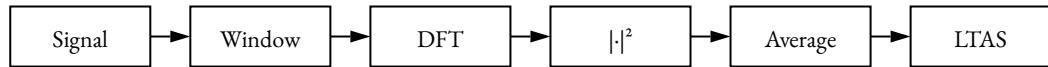
Figure 5. Temps d'enregistrement disponible par violoniste et par violon.

Descripteurs audio

Les descripteurs suivants ont été comparés lors de la classification :

Long Time Average Spectra (LTAS)

Le *Long-Time Average Spectra* (LTAS) d'un enregistrement est obtenu en divisant le signal d'entrée en segments superposés, puis en calculant la DFT (transformée de Fourier discrète) fenêtrée de chaque segment, et enfin en moyennant la puissance de ces DFTs.

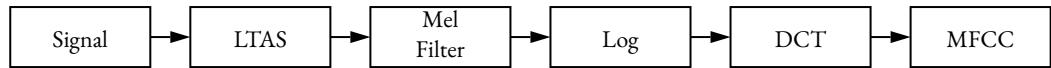


Le LTAS a été utilisé dans [5] afin de comparer la qualité tonale des violons. Plus précisément, le son des violons italiens anciens (Stradivari/Guarneri) et des violons modernes a été comparé. L'auteur conclut que des différences entre ces deux groupes peuvent être mises en évidence en utilisant le LTAS.

Mel-Frequency Cepstral Coefficients (MFCC)

Les MFCCs sont obtenus en calculant en projetant les fréquences d'un spectres sur une échelle Mel (une échelle perceptuelle constituée de hauteurs jugées équidistantes par des auditeur·ices), en prenant le logarithme, puis en calculant la DCT (Transformée en Cosinus Discrète) du résultat.

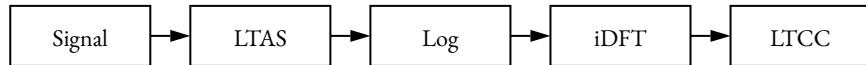
Comme ici on s'intéresse à des caractéristiques ayant une signification à long terme (on cherche une signature sonore d'un violon), un LTAS a été utilisé comme base du calcul des MFCCs. Il s'agit alors d'une version à temps long des MFCCs classiques.



Les MFCCs sont des descripteurs qui ont été largement utilisés pour la reconnaissance automatique des locuteurs et pour la classification des instruments.

Long-Term Cepstral Coefficients (LTCC)

Les LTCCs ont été introduits dans [6] pour l'identification individuelle d'instruments. Leur calcul est similaire à celui des MFCCs, sauf que la projection sur l'échelle Mel n'est pas appliquée et que l'étape finale est constituée d'une transformée de Fourier discrète inverse (iDFT).

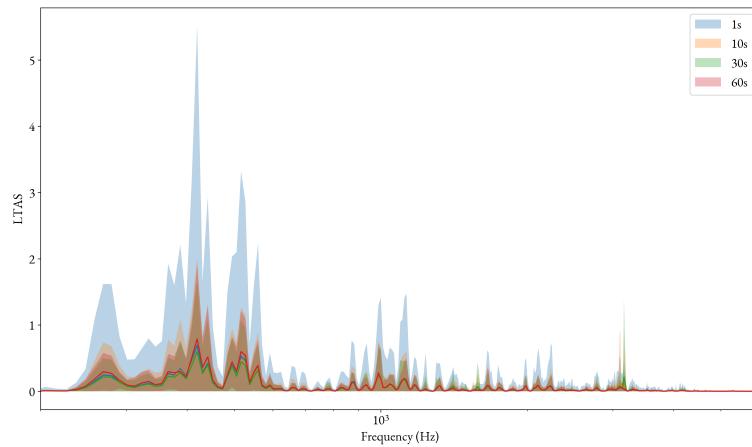


Exploration des données

Variabilité des LTAS

Un paramètre majeur intervenant lors du calcul du LTAS est la durée de l'extrait sur lequel calculer ce spectre moyen. La figure ci-dessous montre la variabilité de ce calcul pour quatre durées d'extraits pour un même violon :

Figure 9. Moyennes (traits pleins) et écarts-types (fond transparent) des LTAS d'un même violon en fonction de la durée de l'extrait choisi



Nous remarquons que la variance des LTAS des extraits d'une seconde est beaucoup plus forte que pour les autres durées. De plus, il ne semble pas y avoir d'énormes différences de variances entre les durées au delà de 10 secondes. Ceci indique que la variabilité des LTAS atteint une certaine stabilité dès que nous considérons des extraits d'une durée de quelques dizaines de secondes. Par la suite, les LTAS seront donc calculés en considérant des extraits de cet ordre de grandeur.

Sélection des descripteurs

Il est possible de lancer des algorithmes de sélection de descripteurs afin de voir quels coefficients discriminent le mieux les violons.

Figure 10. LTAS des violons (au-dessus) et meilleurs coefficients sélectionnés par l'algorithme.

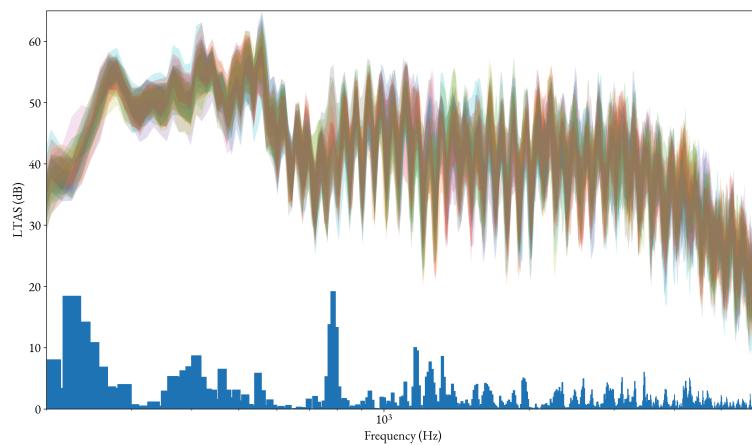


Figure 11. MFCCs des violons (au-dessus) et meilleurs coefficients sélectionnés par l'algorithme.

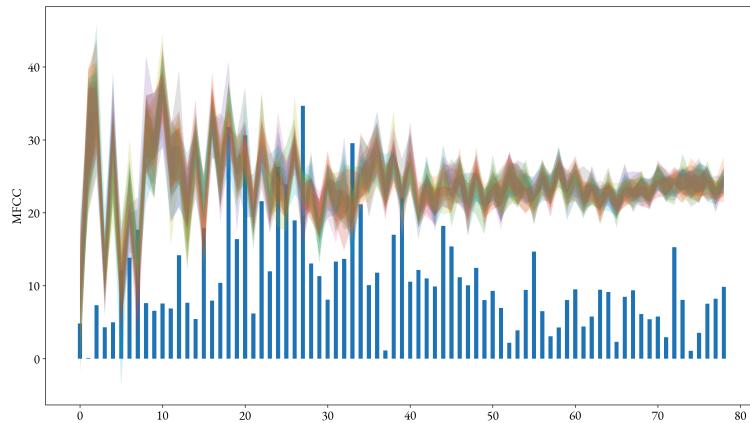
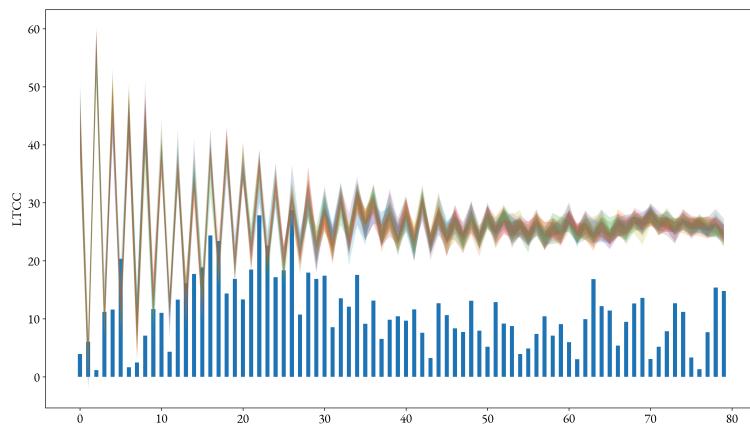


Figure 12. LTCCs des violons (au-dessus) et meilleurs coefficients sélectionnés par l'algorithme.



Certains coefficients semblent plus discriminants que d'autres quant à la reconnaissance de violons à partir de leur enregistrement. En particulier, sur la figure relative aux LTAS, il est possible de voir quels modes du violon sont les plus discriminants. Nous remarquons que le mode A0 (autour de 300Hz) ne l'est pas beaucoup, tandis qu'entre les modes B1- et B1+ (les deux pics autour de 500Hz), le premier semble avoir un meilleur pouvoir de différenciation.

Classification

Nous comparons les résultats de trois algorithmes de classification parmi les plus populaires : les K-Plus Proches Voisins, les Machines à Vecteurs de Support et le Perceptron Multicouche. Ces trois algorithmes utilisent différentes stratégies d'apprentissage et donneront ainsi des résultats différents sur nos données.

K-Nearest Neighbours

Les K-Plus Proches Voisins est une méthode qui trouve les points d'entraînement les plus proches d'un nouveau point de test et prédit sa classe à partir de ceux-ci. Étant donné que la prédiction est effectuée directement à partir des données

d'entraînement, cette méthode est non paramétrique (ou non généralisante), ce qui est un avantage lorsque la frontière de décision est irrégulière.

Support Vector Machines

Les Machines à Vecteurs de Support sont une méthode d'apprentissage supervisé utilisée pour la classification. Elles fonctionnent en trouvant un hyperplan optimal qui maximise la distance entre chaque classe dans les données d'entraînement. Cet algorithme est coûteux en termes de calcul, mais possède généralement de bonnes propriétés de généralisation.

Multilayer Perceptron

Le perceptron multicouche est une méthode d'apprentissage supervisé qui apprend une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^o$ à l'aide de données d'entraînement. Pour ce faire, il utilise des couches de neurones comme illustré dans la Figure \ref{f}. Chaque neurone transforme le résultat de la couche précédente en formant une sommation linéaire pondérée $w_0 + w_1x_1 + \dots + w_nx_n$, suivie d'une fonction d'activation non linéaire.

Cette méthode a montré son efficacité dans divers domaines, car les fonctions d'activation non linéaires peuvent modéliser des relations complexes dans les données, tandis que les couches cachées peuvent apprendre des représentations hiérarchiques à partir des données d'entrée. Cependant, les perceptrons multicouches sont sujets au surapprentissage, en particulier sur de petits ensembles de données.

Résultats

Dans cette section, nous présentons les résultats de la tâche de classification réalisée sur différents sous-ensembles du jeu de données. Nous comparons les résultats de chaque descripteur associé à chaque algorithme de classification. Les résultats présentés ci-dessous sont ceux obtenus après l'ajustement des hyperparamètres, qui peuvent être des hyperparamètres dans le calcul des descripteurs (fréquence d'échantillonnage, taille des fenêtres de pondération, etc.) ou des hyperparamètres de l'algorithme de classification (comme k pour les K-Plus Proches Voisins, par exemple).

Table 2. Résultats en utilisant les gammes comme données d'entraînement et les extraits musicaux comme données de test

Méthode	Descripteur	Entraînement	Test
K-NN	LTAS	100%	37%
	LTCC	100%	100%
	MFCC	100%	100%
SVM	LTAS	100%	37%
	LTCC	100%	100%
	MFCC	100%	100%
MLP	LTAS	100%	37%
	LTCC	100%	100%
	MFCC	100%	100%

Table 3. Résultats en utilisant les gammes comme données d'entraînement et les extraits musicaux comme données de test

Méthode	Descripteur	Entraînement	Test
K-NN	LTAS	100%	38%
	LTCC	100%	76%
	MFCC	100%	83%

Discussion et Conclusion

Expérience au CNSMDP

Introduction

Il est souvent avancé que le son d'un instrument évolue au fil des années. Cela peut être dû à de nombreuses causes physiques : évolution de la structure et des matériaux, modifications des réglages de l'instrument, etc. Des études, passées [] ou en cours [], s'intéressent à la quantification de ces changements de sonorité via des mesures acoustiques. Les luthiers et les violonistes avancent aussi que la sonorité d'un instrument évolue au fur et mesure qu'il est joué. Ainsi, un violon neuf sera dit "vert" ou "fermé".

Cependant, à ce jour, aucune étude n'a permis d'affirmer que de tels changements de sonorité pouvaient être dûs au jeu des violonistes.

Nous pouvons alors nous poser les questions suivantes : l'interprète est-il aussi à l'origine de l'évolution de l'instrument ? Dans quelle mesure laisse-t-il son empreinte sur le son de celui-ci ? Est-ce que cette évolution n'est pas que ressentie par l'interprète qui a progressivement modifié son jeu pour tirer le meilleur de son instrument ?

Pour répondre à ces questions, une expérience est en cours au sein du Conservatoire National Supérieur de Danse et de Musique de Paris, sous la coordination de Julien Dubois, responsable du parc instrumental, et avec le concours de l'enseignante, concertiste-interprète et cheffe d'orchestre Stéphanie Marie Degand. Pendant trois mois à partir de septembre, Stéphanie Marie Degand va jouer un instrument du parc, essayant de lui donner un son différent et d'y laisser son empreinte sonore. L'objectif de cette expérience est de vérifier s'il y a eu d'éventuelles évolutions d'un point de vue acoustique et/ou s'il y a eu des évolutions dans la façon de jouer de l'interprète.

Nous commencerons par présenter la méthodologie utilisée, avant de détailler les résultats (partiels) obtenus. Enfin, nous conclurons par une discussion des principaux enseignements de cette étude.

Méthodologie

Protocole expérimental

L'expérience fait intervenir trois violons : un test (Klimke, 2012) et deux de contrôle (Levaggi, 2012 et Stopanni, 2018). Elle fait aussi intervenir une quinzaine de violonistes de très bon niveau : une violoniste test (Stéphanie Marie Degand), et une quatorzaine de violonistes de contrôle. Au cours de l'expérience, le violon test va être joué par Stéphanie Marie Degand dans le but de l'ouvrir.

L'expérience se divise en deux sessions : une première ayant eu lieu en septembre, une seconde qui aura lieu quand le violon test sera considéré ouvert par Stéphanie Marie Degand.

Avant le début de l'expérience, les violons ont été réglés suivant les consignes de Stéphanie Marie. Ils ne seront pas réglés par la suite.

Pour chaque session, voici un tableau récapitulant les actions demandées aux participant-es :

Condition	Déroulé de chaque session	Durée (min)
Aveugle	Accueil, explications	3
	Période de familiarisation avec les trois instruments. Notation (sur 10) de chaque instrument suivant les trois critères suivants : puissance, facilité de jeu et beauté du timbre.	15
	Enregistrements (3 gammes + morceaux du répertoire)	20
Non-aveugle	Période de familiarisation avec les deux instruments (cf. ci-dessous). Notation suivant les mêmes critères.	10
	Enregistrements (3 gammes + morceaux du répertoire)	10
Discussions et conclusion		5

Lors de la partie en aveugle, les participant·es se voient présenter les trois violons successivement. En non aveugle, nous ne leur présentons plus que deux violons : le Klimke et le Stopanni (à noter que nous ne précisons pas qu'il s'agit des mêmes violons que ceux ayant été joués en aveugle : la plupart des participants pensent jouer 5 instruments en tout). Cette partie de l'expérience doit servir à vérifier si l'aspect visuel des instruments influent la perception qu'en ont les violonistes. En effet, le Klimke est un violon avec un vernis plein, très brillant, et d'aspect très "neuf". En revanche, le Stopanni a un aspect imitant celui d'un vieil instrument.

Enregistrements

Les enregistrements ont été réalisés au CNSMDP, sur le plateau 1. C'est une pièce relativement grande et sèche, équipée de tout le matériel nécessaire pour enregistrer. Les captations ont été faites à l'aide d'un couple de micros DPA-4010, qui ont une réponse en fréquence quasiment plate, placés à une cinquantaine de centimètres des violonistes. Les données ont été collectées avec une fréquence d'échantillonnage de 48kHz.

Traitement des enregistrements

Les enregistrements seront traités grâce aux méthodes présentées dans la partie sur l'identification individuelle de violon. En effet, ce travail a permis de faire émerger des descripteurs permettant de discriminer le son produit par plusieurs violons. À l'issu de l'expérience, il sera donc possible de quantifier à quel point ces descripteurs appliqués aux enregistrements du violon test ont évolué entre les deux sessions. Il sera également possible de vérifier que ces même descripteurs n'ont pas significativement changé pour les violons de contrôle entre les deux sessions.

Nous espérons aussi que ces descripteurs soient suffisamment représentatifs du son d'un violon pour pouvoir calculer une distance entre le son de chaque violon de l'expérience et le son du violon de chaque participant·e.

Traitement des notes

Les notes recueillies seront traités par un test statistique (ANOVA) afin de savoir quels facteurs (violon, violoniste, condition et session) les affectent significativement.

Résultats

Enregistrements

Travail en cours... Peut-être sera-t-il prêt d'ici le CSI ? Seul le temps nous le dira.

Notes

Voici les résultats donnés par le test statistique sur les notes récoltées lors de l'expérience :

	F Value	Num DF	Den DF	Pr > F
Violon	0.009	1	8	0.925
Critère	3.426	2	16	0.058
Condition	0.153	1	8	0.706
Session	0.837	1	8	0.387
Violon:Critère	1.962	2	16	0.173
Violon:Condition	0.484	1	8	0.506
Critère:Condition	3.404	2	16	0.059
Violon:Session	0.484	1	8	0.506
Critère:Session	0.985	2	16	0.395
Condition:Session	0.155	1	8	0.704
Violon:Critère:Condition	2.566	2	16	0.108
Violon:Critère:Session	0.691	2	16	0.515
Violon:Condition:Session	1.111	1	8	0.323
Critère:Condition:Session	0.308	2	16	0.739
Violon:Critère:Condition:Session	4.951	2	16	0.021

Aucun facteur ne semble significativement affecter les notes. Que le violon n'ait pas un effet significatif peut s'expliquer de plusieurs raisons : d'une part, les violons utilisés sont tous trois de bonne facture et ne présentent pas de défaut flagrant. Deuxièmement, les critères et surtout l'échelle utilisée pour les notés ne sont peut-être pas suffisamment précis pour les discriminer. En effet, lors de l'expérience, nous nous sommes rendus compte que la plupart des candidat·es ne donnaient que très rarement des notes en dehors de l'intervalle [6; 9]. Peut-être faudra-t-il forcer les participant·es à utiliser une plage de notes plus grande.

Par ailleurs, il ne semble pas y avoir un effet significatif de la condition (aveugle / non-aveugle) dans laquelle les violons ont été notés. En l'état, nous ne rejetons donc pas l'hypothèse nulle selon laquelle l'aspect visuel n'a pas d'influence sur les notes attribuées par les violonistes (ce qui ne signifie pas que cette hypothèse soit vraie, juste que nous n'avons pas suffisamment de preuves).

Discussion et Conclusion

Le but de cette étude est de savoir à quel point les violonistes peuvent laisser leur empreinte sonore sur leur instrument. Un objectif secondaire, est de savoir à quel point l'aspect visuel des instruments peut modifier la perception des musicien·nes.

Les données permettant de répondre à la première question sont encore en cours de collecte : après une première session d'enregistrement, le violon *Klimke* va être joué pendant plusieurs mois par Stéphanie Marie Degand dans le but de changer sa sonorité. La deuxième session d'enregistrement aura lieu courant janvier : il sera alors possible de quantifier cette évolution à l'aide des descripteurs audio permettant d'identifier les violons. Pour ce qui est de la deuxième question, nous ne pouvons pas affirmer que l'aspect visuel a eu un impact sur les notes qu'ont donné les participant·es.

Une idée serait alors de modifier légèrement l'objectif secondaire : est-ce que l'histoire d'un violon influe la perception que les violonistes en ont ? Nous pourrions par exemple dire aux participant·es que tel violon a été joué par tel soliste. Cela aura-t-il une influence sur les notes qui lui sont données ?

Conclusion

Voici la conclusion de ce rapport qui, je l'espère, n'a pas été trop désagréable à lire. Les deux questions auxquelles je me suis principalement attaquées sont :

1. Les violons ont-ils une identité sonore propre, dissociable du violoniste qui les joue ?
2. Dans quelle mesure les violonistes laissent leur empreinte sonore sur leur instrument ?

La première partie, correspondant au début de ma thèse, a permis de montrer qu'il est possible de remonter à certains

paramètres de jeu via des descripteurs audio classiquement utilisés dans la littérature, et ce, même dans des conditions de jeu réalistes. La deuxième partie s'est attaquée à la définition d'une signature sonore des violons. Pour ce faire, des descripteurs audio à temps long ont été comparés, et deux d'entre eux (LTCCs et MFCCs) permettent d'obtenir des résultats encourageants. Enfin, l'expérience en cours au CNSMDP va permettre de savoir dans quelle mesure une violoniste peut laisser son empreinte sonore sur un violon en le jouant.

Une des principales pistes d'ouverture identifiées serait de réussir à relier les descripteurs utilisés pour l'identification à la perception humaine. En effet, les distances calculées entre les descripteurs ne correspondent pas forcément aux distances relevées lors de questionnaires soumis à un groupe de violonistes et de luthiers.

Bibliographie

1. R. Malvermi, S. Gonzalez, F. Antonacci, A. Sarti, R. Corradi, *A Statistical Approach to Violin Evaluation*. Applied Sciences, 2022.
2. F. Setragno, M. Zanoni, A. Sarti, F. Antonacci, *Feature-based characterization of violin timbre*. 2017 25th European Signal Processing Conference (EUSIPCO), 2017.
3. E. Schoonderwaldt, *The Violinist's Sound Palette: Spectral Centroid, Pitch Flattening and Anomalous Low Frequencies*. Acta Acustica united with Acustica, 2009.
4. C. Fritz, V. Salvador, G. Stoppani, *The Bilbao project: searching for relationships between sound and playing properties of violins with their construction parameters.* , 2021.
5. A. Buen, *COMPARING THE SOUND OF GOLDEN AGE AND MODERN VIOLINS: LONG-TIME-AVERAGE SPECTRA.* , 2005.
6. E. Lukasik, *Long Term Cepstral Coefficients for Violin Identification*. Journal of The Audio Engineering Society, 2010.

Annexes

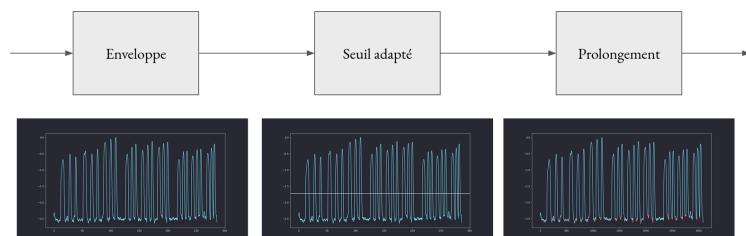
Segmentation automatique

Le but de cette annexe est de présenter un algorithme simple (mais en pratique assez efficace) permettant de découper automatiquement un enregistrement contenant plusieurs extraits musicaux. Ceci est utile lors du traitement d'enregistrements issus d'une expérience où pour des raisons de temps il n'est pas possible d'arrêter et de relancer sans arrêt les enregistrements. Les données prennent alors la forme de plusieurs dizaines de minutes d'enregistrement, parmi lesquelles seuls les extraits contenant du violon nous intéressent.

L'idée est tout simplement de modéliser un extrait joué au violon comme un signal ayant un volume fort et maintenu dans le temps. L'algorithme se découpe alors en trois étapes :

1. le calcul de l'enveloppe du signal (modélisant le volume)
2. le calcul d'un seuil adapté (permettant de différencier les moments de silence / conversation des moments où un violon est joué)
3. le prolongement de ce seuillage afin d'obtenir des extraits suffisamment bien découpés

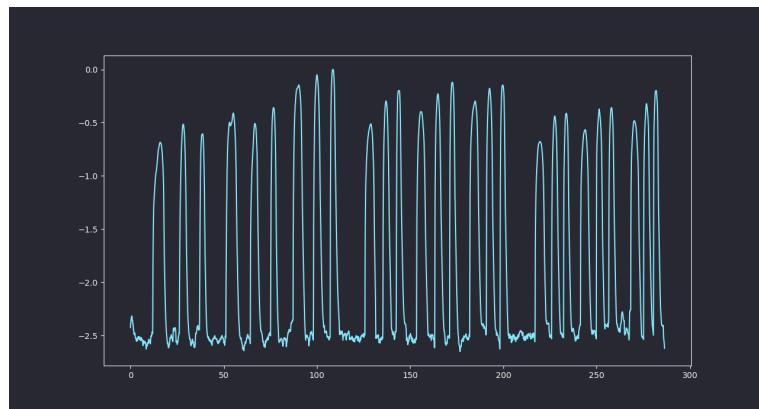
Figure 13.



Enveloppe

L'enveloppe est calculé via le calcul de la RMS du signal, avec une fenêtre assez large afin de lisser le résultat

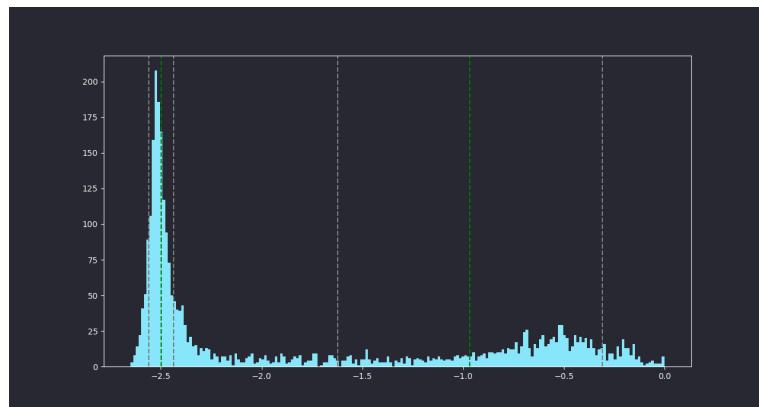
Figure 14.



Seuillage

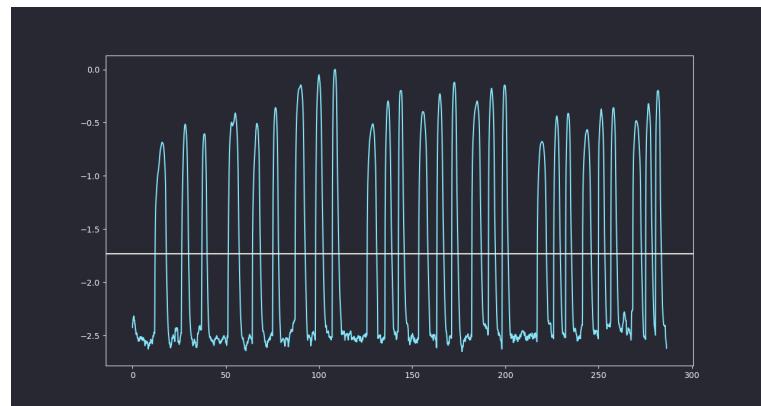
Afin de trouver automatiquement un seuil permettant de faire la différence entre un bruit de fond et un son de violon, l'histogramme des valeurs de la RMS du signal est analysé. Nous faisons ensuite l'hypothèse qu'il peut être décrit comme une somme de deux gaussiennes : une correspondant au bruit de fond (prenant des valeurs d'énergie moins élevées), une autre correspondant au son du violon (valeurs plus élevées).

Figure 15.



Le seuil est alors choisi comme la moyenne entre les moyennes de ces gaussiennes.

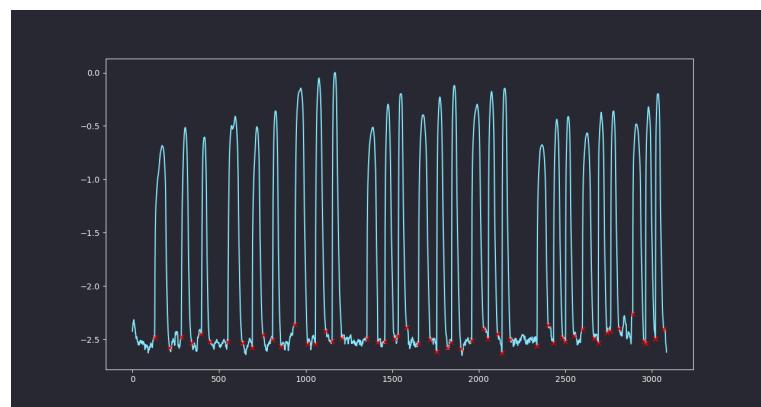
Figure 16.



Prolongement et contraintes

Enfin, afin d'éviter que les extraits ne soient découpés trop abruptement, un prolongement est effectué à gauche et à droite de chaque extrait (tant que la dérivée de change pas de signe).

Figure 17.



Les extraits trop courts sont éliminés, et deux extraits séparés par un temps inférieur à un certain paramètre sont fusionnés.