

AMÉLIORATION DES MÉTHODES D'ESTIMATION D'ACCORDS ET DE TONALITÉ DEPUIS UNE REPRÉSENTATION MUSICALE SYMBOLIQUE

Thomas Rocher
LaBRI
rocher@labri.fr

Matthias Robine
LaBRI
robine@labri.fr

Pierre Hanna
LaBRI
hanna@labri.fr

RÉSUMÉ

Cet article présente de nouvelles méthodes d'estimation des accords et de la tonalité d'un morceau de musique représenté de manière symbolique. Elles reposent sur la comparaison d'un profil caractérisant la fréquence d'apparition des notes sur une période donnée avec un profil de référence. Les profils utilisés sont ici obtenus à partir d'informations tonales moins bruitées que les notes de musique. Pour construire le profil destiné à l'estimation de la tonalité, la suite des accords est utilisée. Pour le profil destiné à l'obtention des accords, les notes utilisées sont non seulement pondérées selon leur durée, mais également selon leur contribution à l'harmonie. Le nombre de notes jouées simultanément sera ainsi utilisé. Pour chacune des méthodes, une description ainsi qu'une évaluation seront présentées. Nous montrons ainsi que l'efficacité des méthodes proposées est supérieure aux méthodes existantes qui reposent sur la comparaison entre un profil d'observation et des profils de référence.

1. INTRODUCTION

L'estimation automatique d'accords et de tonalité est une problématique importante, notamment en musique occidentale où les paramètres issus de l'harmonie jouent un rôle prépondérant. Ainsi, les différentes modulations d'une pièce de musique et la suite des accords la composant sont des paramètres musicaux essentiels. Les applications de l'analyse de ces paramètres sont multiples, et s'étendent de la génération d'accompagnement automatique à l'assistance à la composition.

Nous choisissons ici de nous pencher sur les méthodes d'estimation de la tonalité d'un morceau de musique, ainsi qu'à l'analyse des accords. Nous commençons par présenter les méthodes existantes dans la section 2, avant d'exposer de nouvelles méthodes dans la section 3. Pour l'estimation de la tonalité comme pour celle des accords, une évaluation de l'efficacité est proposée, et nous comparons les résultats de nos méthodes avec ceux des méthodes existantes.

Enfin, nous présentons les perspectives ouvertes par ces travaux dans la section 4.

2. MÉTHODES EXISTANTES

Dans cette section, nous présentons les méthodes existantes d'estimation de tonalité, puis d'estimation d'accords.

2.1. Tonalité

Les travaux de Chew reposent sur un modèle de l'harmonie dans lequel les notes de musique s'enroulent le long d'une spirale en trois dimensions, et dans le même ordre que dans le cercle des quintes. Dans ce modèle, la tonalité du morceau est estimée par calcul du barycentre des notes de la pièce [2]. D'autres travaux dans l'analyse de la tonalité d'un morceau de musique utilisent un vecteur caractérisant la fréquence d'apparition des notes dans la pièce étudiée. Pour le signal audio, ces vecteurs sont généralement des chromas, qui représentent l'intensité des fréquences correspondant aux 12 demis tons ramenés sur une octave [15, 5]. Dans les études se focalisant sur la représentation symbolique de la musique, ce sont des profils de tonalités (ou *pitch profiles*) qui sont utilisés [19]. Dans ces méthodes, le vecteur utilisé est comparé à des profils de référence qui représentent l'importance de chaque note dans une tonalité donnée. Cette importance peut être considérée comme une consonance [10] ou comme une probabilité d'apparition caractéristique [20] de chaque note dans une tonalité donnée.

2.1.1. Profils vectoriels

Un des premiers algorithmes d'estimation de tonalité fut utilisé en 1990 par Krumhansl et Schmuckler [9], à la suite d'expériences psychoacoustiques [10]. Deux profils de référence (l'un majeur et l'autre mineur) ont été construits en faisant écouter à des musiciens une gamme majeure ou mineure, suivie d'une note quelconque. Il a alors été possible d'évaluer la consonance de cette dernière note selon l'harmonie de la gamme. Pour les deux modes, il a été

donc possible de créer un profil de tonalité : chacun des 12 demis tons s'est vu associer une valeur représentant la consonance par rapport au mode. Une fois le profil majeur (resp. mineur) obtenu, toutes les tonalités majeures (resp. mineures) pouvaient alors trouver un profil de référence en effectuant une rotation d'indice : sol possède la même consonance en Do Majeur que mi en La Majeur par exemple.

Une fois tous les profils de référence obtenus, ils sont comparés un à un au profil de la pièce étudiée. Le profil de la pièce fait correspondre à chacun des 12 demis tons une valeur caractérisant la fréquence d'apparition dans la pièce. Il est construit en comptant le nombre de do, de do#, de ré, etc... de la pièce, en pondérant chaque note par sa durée (une blanche se voit par exemple attribuer un coefficient 2 quand une noire a un coefficient 1). 24 corrélations sont ensuite calculées pour estimer le degré de ressemblance entre le profil de la pièce et les 24 profils de référence. Plus la corrélation entre le profil de la pièce et un profil de référence est élevée, plus ces deux profils sont similaires. Il suffit alors de considérer le profil de référence offrant la plus grande corrélation avec le profil de la pièce pour pouvoir estimer sa tonalité.

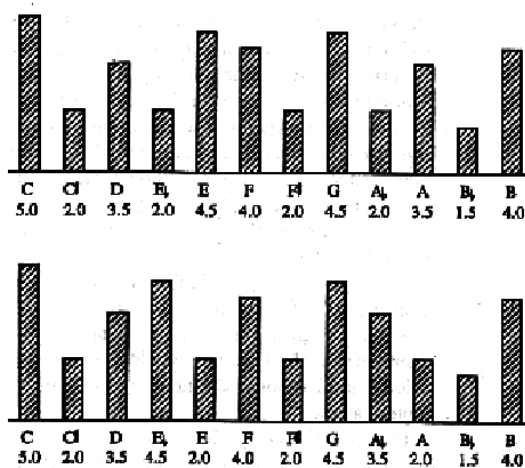


Figure 1. Les profils de référence de Temperley [19]. En haut, le profil de Do majeur et en bas, celui de Do mineur.

Temperley [19] a repris cet algorithme de référence pour y ajouter quelques modifications. Il a par exemple rehaussé l'importance de la sensible (la septième note d'une gamme) pour les profils de référence représentés sur la figure 1. Il propose le calcul de corrélation suivant, pour une tonalité donnée :

$$C_{pitch} = \sum_{i=1}^{12} V_{ref}[i].V_{piece}[i] \quad (1)$$

où V_{piece} le profil de notes de la pièce, V_{ref} le profil de référence de la tonalité, et C_{pitch} la corrélation entre

les deux profils (ici obtenue par produit scalaire).

Sur le même principe, Gomez a utilisé des chromas pour estimer la tonalité depuis un signal audio [7]. Chai s'est également inspiré de ces travaux en utilisant des profils de mode pour discerner les tonalités majeures/mineures [1].

2.1.2. Profils matriciels

En 2007, Madsen et Widmer [14] ont adopté une approche différente. En utilisant les observations de Li et Huron [13], ils ont pris en compte l'ordre des notes, en plus de leurs fréquences d'apparition. Deux distributions de notes identiques peuvent en effet correspondre à deux séquences de notes différentes, qui peuvent donc potentiellement induire deux tonalités différentes. La figure 2 montre ainsi un exemple où deux extraits musicaux constitués des mêmes notes (mais dans un ordre différent) semblent appartenir à deux tonalités différentes.



Figure 2. Deux extraits donnant le même profil de notes : un extrait de la célèbre fugue de Bach, en haut. En bas, un extrait composé des mêmes notes, mais dans un ordre différent. Ce dernier semble plutôt être en La mineur, alors que les mêmes notes que l'extrait en Ré mineur sont jouées.

Madsen et Wildmer ont donc adopté une représentation matricielle, qui prend en compte non plus le nombre d'occurrences de chaque note, mais le nombre d'occurrences d'intervalles entre les notes. On compte ainsi le nombre d'intervalle (do,ré), (do, do#), (do, ré), etc... Chacun des 144 (12x12) intervalles possibles se voit donc attribuer une probabilité d'apparition, représentée dans une matrice 12x12 (nommée *interval profile*, voir figure 3). Comme pour les profils de Krumhansl et Temperley, la matrice d'intervalles d'une pièce de musique est comparée aux matrices de référence, et un calcul de corrélation est mené pour estimer la tonalité :

$$C_{interval} = \sum_{i=1}^{12} \sum_{j=1}^{12} M_{ref}[i, j].M_{piece}[i, j] \quad (2)$$

où M_{piece} est le profil (matriciel) d'intervalles de la pièce, M_{ref} le profil de référence et $C_{interval}$ la corrélation entre les deux profils.

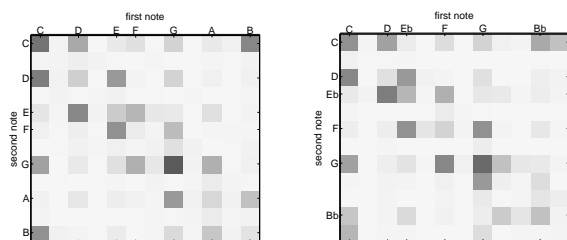


Figure 3. Les profils d’intervalles de Madsen (majeur à gauche, mineur à droite), construits à partir des chorals de Bach (communication personnelle).

2.1.3. Complémentarité des deux profils

Les résultats de l’évaluation de la méthode matricielle se sont avérés comparables aux résultats des méthodes utilisant des profils vectoriels [14]. Sur la base de données utilisée par Madsen [4], les deux méthodes obtiennent des scores avoisinant les 75 % de tonalités correctement détectées. En revanche, elles donnent des résultats différents sur différentes pièces de musique, la méthode matricielle pouvant échouer là où la méthode vectorielle donne le bon résultat et vice versa. Madsen a ainsi constaté que dans quasiment 85 % des cas, au moins une des deux méthodes donne le bon résultat. Puisque les informations issues de ces deux analyses semblent être différentes, il apparaît judicieux de tenter les combiner pour améliorer les résultats.

2.1.4. Méthode hybride

Nous avons proposé une nouvelle méthode d’évaluation de la tonalité [17]. Elle repose sur la combinaison des profils vectoriels et matriciels. Une telle combinaison peut être effectuée de plusieurs manières. Il a dans un premier temps été choisi de sommer les corrélations normalisées entre les profils de notes et d’intervalles de la pièce de musique et les profils de notes et d’intervalles de référence, afin que chaque méthode fournisse le même apport en information.

$$C_{hybrid} = f(C_{pitch}, C_{interval}) \quad (3)$$

où C_{hybrid} est la corrélation finale, C_{pitch} la corrélation entre les profils de notes, $C_{interval}$ la corrélation entre les profils d’intervalles et f une fonction de combinaison (par exemple, une somme).

2.2. Accords

Dans le domaine de la reconnaissance d’accords depuis le signal audio, de nombreux travaux utilisent des modèles de Markov cachés [8, 11]. Dans ces travaux, une phase d’apprentissage est nécessaire, afin d’entraîner le système sur une base de données annotée.

Différents paramètres comme la probabilité de transition d’un accord à un autre, ou la probabilité d’un accord suivant un descripteur (tel le chroma) sont définis lors cette phase. Ils sont ensuite utilisés lors de la phase de reconnaissance, qui utilise la programmation dynamique pour trouver le meilleur chemin dans le graphe des accords candidats. Les chromas peuvent également être utilisés à la manière d’un profil de notes, en cherchant ensuite à maximiser une corrélation avec un vecteur de référence, comme dans les travaux de Gomez [6].

Dans le domaine de la musique symbolique, on citera le modèle en spirale de Chew [2], qui en plus de la tonalité, peut estimer la suite d’accords d’un morceau en utilisant une fenêtre d’analyse. D’autres travaux utilisent un système de règles pour déterminer les accords d’après les notes de musique, comme Melisma Music Analyzer, le logiciel proposé par Temperley [19], et l’approche récemment proposée par Illescas et al. [16].

Pour l’analyse d’accord comme pour l’analyse de de tonalité, les méthodes existantes utilisent les notes de musique pour construire un profil qui sert ensuite à l’analyse. Or, les notes de musique fournissent une information tonale qui est très bruitée, en raison du nombre important de notes de passages ou d’ornement, qui sont des notes n’appartenant pas nécessairement aux accords ou à la tonalité recherchés.

3. MÉTHODES PROPOSÉES

Nous proposons ici de nouvelles méthodes d’analyse. Au lieu de prendre les notes en considération, nous utilisons d’autres informations tonales moins bruitées. Ainsi, pour l’estimation de la tonalité, les profils servant à l’analyse sont obtenus à partir des accords.

3.1. Analyse de la tonalité

Afin de débruiter l’information tonale utilisée pour construire le profil nécessaire à l’estimation de la tonalité, nous avons choisi d’utiliser les accords d’une pièce de musique plutôt que les notes la composant. Ainsi, pour chaque accord, les notes de la triade correspondante sont ajoutées au profil de la tonalité. Nous avons choisi d’attribuer à la fondamentale un poids deux fois plus important que pour la tierce et la quinte.

```

initialiser le profil de tonalite
pour toute fondamentale x des accords
    ajouter x (deux fois), tierce(x)
    et quinte(x) au profil de tonalite
fin pour tout
retourner la tonalite maximisant
la correlation avec le profil de reference

```

Le calcul du maximum de corrélation est identique à celui proposé par Temperley [19], avec les mêmes profils de référence (voir figure 1). L'intérêt de cette méthode est donc de filtrer les notes non harmoniques (ornements, notes de passages). Chaque accord ayant un rôle harmonique plus important qu'une seule note de musique, l'apport tonal peut sembler plus grand si seulement les notes issues des accords sont ajoutées au profil de tonalité.

Exemple

La figure 4 représente les premières mesures de *Little Wing*, de Jimi Hendrix. Ces mesures peuvent faire l'objet d'un profil de tonalité dit "classique", où chaque note de la partition sera comptabilisée dans le profil avec un poids dépendant de sa durée (profil du haut sur la figure 5).



Figure 4. Les quatre premières mesures de *Little Wing*, de Jimi Hendrix. Les accords sont notés au dessus de chaque mesure. Le morceau est en Mi mineur.

La méthode que nous proposons utilise un profil ne correspondant non plus à la répartition de toutes les notes de l'extrait, mais des notes des accords de l'extrait. Ainsi, cet extrait comportant 4 accords (Em, Gmaj, Amin, Em), les notes ajoutées aux profils seront les suivantes. Le poids de chaque note est mentionné entre parenthèses :

- mi(2) sol(1) si(1) (pour Emin),
- sol(2) si(1) re(1) (pour Gmaj),
- la(2) do(1) mi(1) (pour Amin),
- mi(2) sol(1) si(1) (pour Emin).

Les profils de tonalités issus de cette méthode et celui issu de la méthode classique sont différents, comme le montre la figure 5. On remarque par exemple que le mi se voit conférer une plus grande importance avec le profil issu de la méthode proposée, tout comme le sol et le si, qui sont les autres notes de la triade correspondant à Mi mineur. Le profil issu de la méthode classique marque moins cette différence entre les notes de la triade et les autres notes, et apparaît donc comme plus bruité, notamment en raison du poids important conféré au la, qui se retrouve quasiment aussi important que le sol et bien plus que le si. La figure 6 présente le résultat des corrélations entre les profils de tonalité obtenus avec les deux méthodes différentes et les profils de référence de Temperley [19].

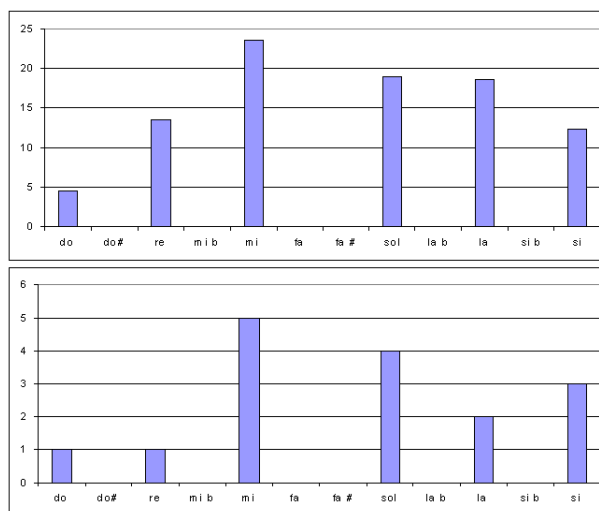


Figure 5. Profils de tonalité des premières mesures de *Little Wing* (en Mi mineur), en utilisant la méthode classique (en haut) et avec la méthode proposée (en bas). Les poids utilisés pour la méthode classique sont les suivants : 1 pour une double croche, 2 pour une croche, 4 pour une noire, ...

On remarque que le profil calculé d'après la méthode proposée présente une corrélation maximale avec le profil de référence correspondant à une tonalité de Mi mineur, qui est la tonalité exacte de l'extrait musical. En revanche, le profil classique maximise la corrélation pour un vecteur de référence correspondant à une tonalité de Sol Majeur (qui est le ton relatif de Mi mineur). On remarque également que Do Majeur présente également une corrélation supérieure à Mi mineur, et donc que la tonalité correcte ne se retrouve qu'en troisième position en utilisant le profil classique. Dans cet exemple, la prise en compte des accords pour construire le profil de tonalité d'un morceau a pour effet de filtrer le bruit harmonique causé par la prise en compte de chacune des notes de la pièce étudiée.

Cet exemple permet d'avoir un premier aperçu du fonctionnement de la méthode proposée, ainsi que de sa supposée efficacité. Dans la section suivante, une évaluation plus complète de cette méthode est présentée, ainsi que sa comparaison avec la méthode classique.

3.1.1. Évaluation

Cette section présente les morceaux utilisés pour l'évaluation, ainsi que la méthode utilisée pour calculer l'efficacité de l'estimation de tonalité. Les scores de la méthode proposée sont ensuite comparés aux scores de la méthode classique.

Bases de données

Pour estimer une tonalité à partir des accords, nous avons besoin de disposer d'une vérité terrain des accords sur un nombre suffisamment important

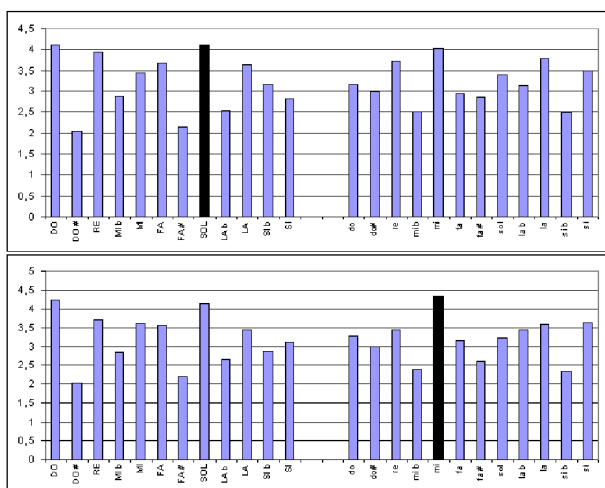


Figure 6. Résultats des corrélations (par produit scalaire) entre les profils de référence et les profils du début de la pièce *Little Wing* obtenus par la méthode classique (en haut) et la méthode proposée (en bas). Les tonalités majeures apparaissent en majuscule (à gauche) et les tonalités mineures en minuscule (à droite). La tonalité de l'extrait (Mi mineur) obtient le score maximal de corrélation avec la méthode proposée, et est donc correctement estimée. Ce n'est pas le cas pour la méthode classique, qui indique une tonalité de Sol Majeur.

de morceaux de musique. Les 12 morceaux ayant servi à la précédente évaluation n'étaient donc pas en assez grand nombre pour l'évaluation de la tonalité, puisque chaque morceau donne lieu à une seule évaluation (au lieu de l'évaluation de l'ensemble de ses accords). C'est donc pour cela que nous avons eu recours aux transcriptions de Christopher Harte¹ de l'université Queen Mary (Londres), qui propose une vérité terrain des accords de 180 titres des Beatles. Néanmoins, la vérité quant à la tonalité de chacun des morceaux n'est pas indiquée. Nous avons donc eu recours aux travaux d'Allan Pollack². Ce dernier a réalisé une étude de la structure des morceaux des Beatles, en mentionnant en particulier pour chacun la tonalité. Les morceaux sélectionnés pour l'évaluation ont donc les morceaux des Beatles dont :

- nous disposons de la transcription accord par accord par Christopher Harte,
- nous disposons d'une tonalité unique grâce aux travaux d'Allan Pollack (dans ses notes, Allan Pollack cite notamment des morceaux dont la tonalité apparaît comme ambiguë, comme par exemple *Birthday*, où deux tonalités différentes sont mentionnées),
- nous disposons d'une représentation symbolique (fichier MIDI).

Ainsi, **86 morceaux** ont finalement été retenus pour l'évaluation.

¹ contact personnel avec le chercheur

² <http://www.icce.rug.nl/soundscapes/DATABASES/AWP/awp-alphabet.shtml>

Protocole d'évaluation

Pour chaque titre de la base de données, deux tonalités sont calculées :

- la première, selon la méthode d'analyse classique : toutes les notes de la pièce de musique sont ajoutées au profil qui est ensuite comparé aux profils de référence (section 2.1.1),
- la seconde, selon la méthode d'analyse proposée : pour chaque accord de la vérité terrain, les notes de la triade correspondante sont ajoutées au profil du morceau, qui est ensuite comparé aux profils de référence. Comme pour la méthode classique, la plus grande corrélation avec le profil du morceau donne la tonalité (section 3.1).

Les deux tonalités sont ensuite comparées à la vérité terrain, en distinguant plusieurs cas correspondant à l'attribution des différents scores pour le concours MIREX [3] :

- la tonalité à évaluer est exactement la tonalité de la vérité terrain. Dans ce cas, la tonalité est déclarée *correcte*, et se voit attribuer un score de 1,
- la tonalité à évaluer est voisine de la vérité terrain (exemple, on détecte Do Majeur alors que la vérité terrain stipule Fa Majeur). Dans ce cas, la tonalité est déclarée *voisine* et se voit attribuer un score de 0.5
- la tonalité à évaluer est relative de la vérité terrain (exemple, on détecte Do Majeur alors que la vérité terrain stipule La mineur). Dans ce cas, la tonalité est déclarée *relative* et se voit attribuer un score de 0.3,
- la tonalité à évaluer présente la même tonique que la vérité terrain, mais un mode différent (exemple, on détecte Do Majeur alors que la vérité terrain stipule Do mineur). Dans ce cas, la tonalité est déclarée *parallèle* et se voit attribuer un score de 0.2,
- sinon, la tonalité est déclarée *autre*, et se voit attribuer un score de 0.

3.1.2. Résultats

Les résultats de l'évaluation de la tonalité sont présentés dans le tableau 1. Les scores sont exprimés en pourcentage du nombre de morceaux total (86). Le score MIREX, qui prend en compte les types d'erreurs rencontrés, est également mentionné.

On remarque une amélioration de plus de 10 % concernant la détection de tonalités correctes. Cette amélioration s'accompagne d'une forte diminution du nombre de détection de tonalité *autre* (plus de 8 % de moins pour notre méthode), qui implique que près de 90 % des tonalités détectées sont soit correctes, soit voisines de la tonalité correcte, ce qui peut ouvrir des perspectives quant à la correction

Méthode	Tonalité correcte	Erreurs				Score Mirex
		relatif	voisin	parallèle	autre	
depuis notes	66,28	1,16	12,79	1,16	18,6	73,26
depuis accords	77,91	0	11,62	0	10,47	83,72

Table 1. Résultats de l'analyse des tonalités. "depuis notes" désigne la méthode classique, en remplissant le profil du morceau avec les notes, alors que "depuis accords" désigne notre méthode (avec laquelle le profil est construit à partir des accords).

de ces erreurs. Enfin, le score MIREX, qui mesure l'efficacité globale, est amélioré de plus de 10 points.

3.2. Analyse des accords

Nous présentons dans cette section deux méthodes d'estimation d'accords. Plutôt que d'avoir recours aux modèles de Markov cachés, nous avons privilégié une approche similaire aux méthodes d'analyse de tonalité, avec, pour chaque accord à estimer, la comparaison entre un profil d'accord et un profil de référence. L'utilisation d'un modèle de Markov caché suppose en effet l'entraînement de ce modèle sur une base de données annotées, qui n'existe pas pour l'heure.

3.2.1. Une première méthode directe

En cherchant à adapter directement les méthodes d'estimation de tonalité à l'estimation des accords, nous avons adopté une première approche. Nous construisons pour chaque fenêtre d'analyse un profil de notes de la même manière que pour la méthode classique pour la tonalité, mais seules les notes étant jouées pendant la fenêtre d'analyse sont prises en compte (voir figure 7).

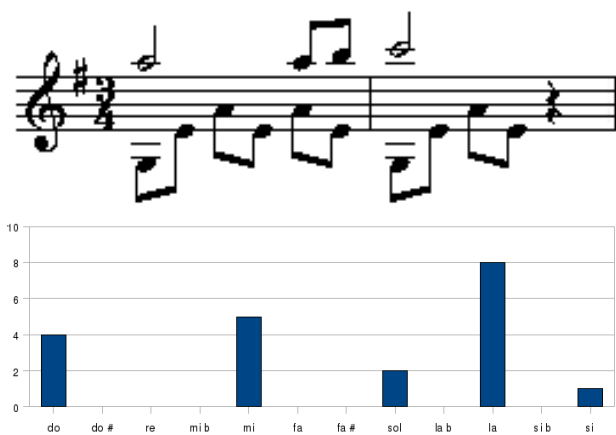


Figure 7. Sur une fenêtre d'analyse comportant les notes de la figure du haut, le profil obtenu est représenté sur la figure du bas. Le poids de chacune des notes est ici pondéré selon la longueur de ces dernières : poids 1 pour une croche, 2 pour une noire,...

Le profil de la fenêtre d'analyse est ensuite comparé aux profils de référence (voir figure 8). Pour ces

derniers, nous avons choisi d'utiliser des profils de type triade : ils contiennent une fondamentale (de poids 2), et les notes à la tierce (majeure ou mineure selon le mode) et à la quinte. Le fait qu'un petit nombre de notes est attendu dans chaque fenêtre d'analyse (en comparaison avec le nombre de notes dans toute la pièce de musique), peut justifier pourquoi seulement trois notes sont présentes dans le profil en triade. Le profil de référence ayant la plus grande corrélation avec le profil de la fenêtre d'analyse donne l'accord estimé.

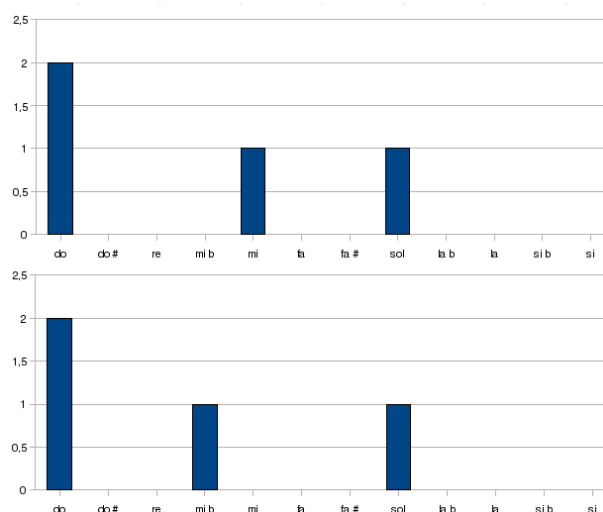


Figure 8. Profils de référence en triade de Do Majeur (en haut) et Do mineur (en bas).

3.2.2. Méthode proposée

Dans la méthode directe, le profil de chaque fenêtre d'analyse est rempli à partir des notes présentes dans la fenêtre, sans faire de distinction quant à leurs importances respectives dans l'harmonie. D'un point de vue musical, il semble néanmoins intéressant de conférer un plus grand rôle aux notes sonnantes simultanément, car elles constituent le fondement de l'aspect vertical de la musique, c'est à dire de l'harmonie. Nous considérons ainsi comme un *groupe de notes* l'ensemble des notes jouées simultanément dans une pièce de musique. Les groupes de notes sont construits grâce à une transformation des notes, appelée *transformation homorythmique*. Le principe est de mod-

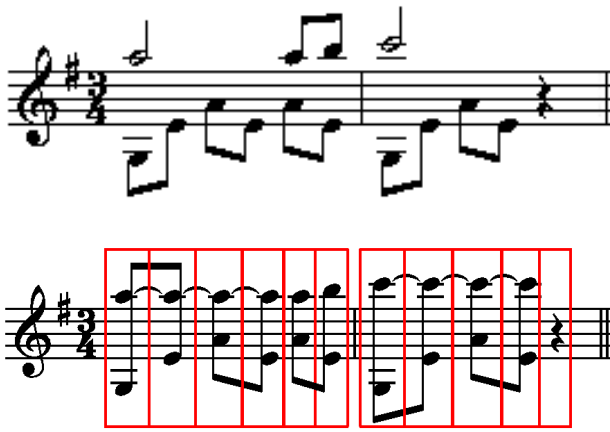


Figure 9. La figure du haut représente un extrait musical. La figure du bas représente le même extrait ayant subi une transformation homorythmique. Les groupes de notes apparaissent encadrés.

ifier la longueur de certaines notes de sorte que toutes les notes qui sont jouées en même temps débutent et terminent toutes aux même instants. On trouvera un exemple sur la figure 9.

Notons que la transformation homorythmique ne modifie pas la manière dont sonne une pièce de musique, mais seulement la manière de représenter les notes de musique la composant. De plus, nous avons, pour chaque groupe de notes, défini une fondamentale selon la définition suivante : la fondamentale d'un groupe de note A est la note x qui minimise l'empilement de tierces nécessaires (depuis x) pour rencontrer toutes notes de A .

Il a ainsi été choisi d'utiliser une approche qui tenait davantage compte de l'harmonie, en considérant les propriétés des groupes de notes de chaque fenêtre pour remplir le profil caractéristique de cette fenêtre. Voici les deux principales modifications que cette prise en compte induit :

- plus un groupe de notes est riche (plus il comporte de notes différentes), plus les notes qui le composent se trouvent renforcées harmoniquement parlant. Le poids de chacune des notes dépend donc du nombre de notes différentes appartenant au même groupe qu'elle,
- pour chaque groupe de notes, la fondamentale joue un rôle plus important sur l'harmonie que les autres notes. Le poids de cette dernière est donc augmenté lorsque le groupe de notes comporte plus d'une note.

Les autres réglages de la méthode directe sont conservés, comme le fait de donner plus de poids aux notes plus longues. Pour chaque profil, on cherche ensuite le maximum de corrélation avec des profils de référence en triade, comme dans la méthode exposée en 3.2.3.

Exemple

Considérons l'extrait musical représenté sur la figure 10, et supposons que la fenêtre d'analyse ait pour taille une mesure (soit la longueur de l'extrait).



Figure 10. Une mesure en Do Majeur.

La figure 11 représente deux profils correspondant à cet extrait :

- le premier, en haut, obtenu avec la méthode directe : chaque note de la fenêtre est ajoutée au profil de la fenêtre, avec un poids dépendant de sa longueur (1 pour une croche, 2 pour une noire et 4 pour une blanche par exemple),
- le second, en bas, obtenu avec la méthode proposée : chaque note de la fenêtre est ajoutée au profil de la fenêtre, avec un poids dépendant non seulement de sa longueur, mais également du nombre de notes jouées en même temps que cette note, et éventuellement du fait qu'elle soit ou non la fondamentale d'un groupe de notes.

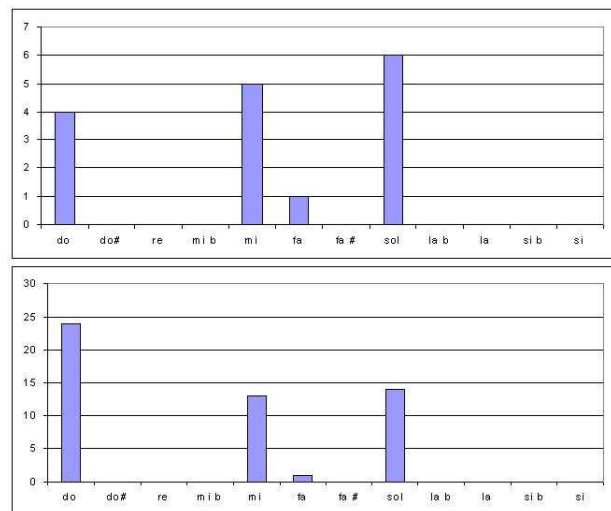


Figure 11. En haut se trouve le profil de la fenêtre obtenu avec la méthode directe, pour lequel chaque note a un poids pondéré seulement par sa durée. En bas se trouve le profil obtenu avec la prise en compte de l'harmonie, chaque note ayant un poids dépendant non seulement de la durée, mais également des notes jouées même temps. On remarque que ce deuxième profil présente une forme plus proche du profil en triade de Do Majeur (do,mi,sol) que le profil obtenu avec la méthode directe.

Ainsi, avec la méthode proposée, le do grave figurant dans l'accord de blanches au début de la mesure n'a plus un poids de 4, comme dans la méthode directe, mais un poids de $4 \times 3 \times 2 = 24$ (4 pour la durée, 3

pour le nombre de notes dans le groupe de notes, et 2 pour le fait que do est fondamentale de ce groupe de notes). Le fa entre le troisième et le quatrième temps voit son poids inchangé entre les deux méthodes (1 dans les deux cas). La méthode proposée renforce considérablement l'importance harmonique du do dans cette mesure. Sur le profil correspondant à la méthode proposée, le do a un poids plus élevé que le sol, alors qu'il apparaît moins longtemps dans la mesure (2 temps pour le do, contre 2 temps et demi pour le sol). On remarque que ce deuxième profil présente une forme plus proche du profil en triade de Do Majeur que le profil obtenu avec la méthode directe.

La figure 12 présente les résultats du calcul des corrélations pour les 12 accords majeurs et les 12 accords mineurs. Les deux profils de la mesure présentés précédemment ont été normalisés pour ce calcul, afin que chacun présente une somme de ses composantes égale à 1.

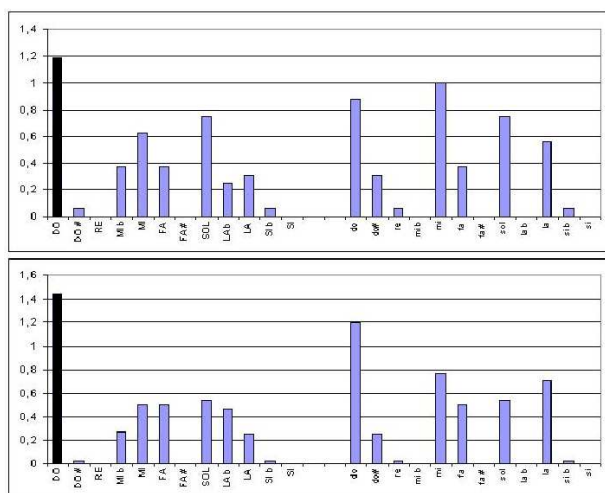


Figure 12. En haut se trouvent les scores de corrélation du profil issu de la méthode directe avec les 12 accords majeurs (en majuscule, à gauche) et les 12 accords mineurs (en minuscule, à droite). En bas sont représentées les mêmes scores pour le profil issu de la méthode proposée. Ce dernier a tendance non seulement à renforcer la corrélation correspondant à l'accord à déterminer (ici, Do Majeur), mais également à augmenter la différence le séparant des autres profils de référence, en diminuant leurs scores respectifs.

On remarque ainsi que le score de corrélation correspondant au Do Majeur est meilleur pour le profil issu de la méthode proposée (1,44 contre 1,19 pour le profil issu de la méthode directe). De plus, la méthode proposée permet d'augmenter l'écart entre le score de Do Majeur et le deuxième meilleur score : la différence avec le deuxième meilleur score (Mi mineur) est de 0,25 pour la méthode proposée, alors qu'elle est de 0,19 entre Do Majeur et Lab mineur pour la méthode directe. Enfin, on remarque que la méthode proposée permet de faire baisser les valeurs des

corrélations pour les accords éloignés de l'accord à déterminer : 6 accords ont des scores au dessus de 0,6 (dont 3 au dessus de 0,8) avec la méthode directe, alors que seulement 4 sont au dessus de 0,6 (dont 2 au dessus de 0,8) avec la méthode proposée.

Si l'étude de l'exemple précédent montre l'intérêt de la prise en compte des informations tonales issues des groupe de notes pour déterminer les accords, une évaluation de la détection d'accords doit être envisagée sur une plus grande quantité d'information. Une telle évaluation est présentée dans la section suivante.

3.2.3. Évaluation

Cette section présente les morceaux utilisés pour l'évaluation, ainsi que la méthode utilisée pour calculer l'efficacité de l'estimation d'accord. Les scores de la méthode proposée sont ensuite comparés aux scores de la méthode directe.

Bases de données

Les morceaux employés pour cette évaluation ont été choisis tels qu'ils :

- comportent un changement d'accord *au plus* toutes les mesures, ceci pour pouvoir choisir la mesure comme fenêtre d'analyse. Il est en effet nécessaire qu'un seul accord soit joué dans chaque fenêtre d'analyse : les résultats seraient en effet erronés si plus d'un accord était joué dans la même fenêtre d'analyse,
- comportent des accords dont le type est identifiable non seulement par notre méthode d'analyse, mais également par la méthode à laquelle nous souhaitons nous comparer, à savoir la méthode directe construits à partir des notes. Il a donc été choisi de restreindre les accords analysés aux accords *majeurs* et *mineurs*.

Ainsi, **12 morceaux** ont été retenus selon ces trois critères. Ces 12 morceaux représentent au total **1195 accords**.

Protocole d'évaluation

Pour chaque morceau, deux listes d'accords sont générées :

- la première, selon la méthode d'analyse directe : pour chaque fenêtre d'analyse, un profil est créé en tenant uniquement compte des notes présentes dans la fenêtre, ce profil étant ensuite comparé aux mêmes profils de référence,
- la seconde, selon la méthode d'analyse proposée : pour chaque fenêtre d'analyse, un profil est créé en tenant compte des informations des groupes de notes de la fenêtre (voir 3.2), ce profil étant

Méthode	Moyenne pondérée par titre	Moyenne pondérée par accord
depuis notes	91,04	90,28
depuis groupes	96,06	96,44

Artiste	Radiohead	The Animals	Otis Redding	The Eagles
Titre	Creep	House Of The Rising Sun	Dock Of The Bay	Hotel California
depuis notes	93,33	94,19	80,88	93,27
depuis groupes	95,56	98,26	100	99,33
Artiste	Pink Floyd	Pink Floyd	Pink Floyd	The Beatles
Titre	The Wall	Wish You Were Here	Comfortably Numb	Can't Buy Me Love
depuis notes	88,61	91,14	80	100
depuis groupes	89,87	89,87	98,5	98,75
Artiste	Barbara	The Beatles	The Beatles	The Beatles
Titre	L'Aigle Noir	Magical Mystery Tour	Come Together	Ticket To Ride
depuis notes	92	94,74	86,47	97,83
depuis groupes	100	94,74	91,26	96,52

Table 2. Résultats de l'analyse des accords. "Depuis note" désigne la méthode directe, alors que "depuis groupes" désigne la méthode proposée (avec laquelle les profils construits tiennent compte des groupes de notes). On observe une amélioration des scores avec la méthode proposée.

ensuite comparé aux profils de référence proposés par Temperley (voir 2.1.1).

L'évaluation se fait ensuite en comparant les deux suites d'accords générées à la vérité terrain. Chaque accord est considéré comme correct s'il présente la même fondamentale et le même mode que la vérité terrain.

3.2.4. Résultats

Les résultats de l'évaluation des accords sont présentés dans le tableau 2. Les scores présentés sont en pourcentage d'accords corrects par morceau. On trouvera également deux moyennes des scores, la première étant une moyenne sur les morceaux, la seconde sur les accords.

On remarque que la méthode proposée améliore significativement les résultats, en permettant une augmentation d'au moins 5 % des résultats, que ce soit en considérant la moyenne par morceau ou par accord. Il est à noter que les résultats de la méthode directe réalisant déjà un score élevé (plus de 90 %), un écart de 5 % n'est pas négligeable. On notera également que, sur une minorité de morceaux, notre méthode fait moins bien que la méthode directe (comme pour *Can't Buy Me Love* ou *Ticket To Ride* des Beatles), mais la différence est dans ce cas relativement faible (jamais plus de 2 %). En revanche, sur les titres où notre méthode fait mieux que la méthode directe, les résultats sont améliorés de manière plus significative : 18,5 % d'amélioration sur *Comfortably Numb* de Pink Floyd, et presque 20 % pour *Dock Of The Bay* d'Otis Redding, par exemple.

4. CONCLUSION ET PERSPECTIVES

Nous avons présenté dans cet article de nouvelles méthodes d'estimation de la tonalité et des accords d'un morceau de musique. Les méthodes existantes reposent pour la plupart sur le calcul d'une corrélation entre un profil représentant la fréquence d'apparition des notes sur une durée d'observation donnée, et un profil de référence. Suivant l'hypothèse que les notes de musiques seules apportent une information tonale relativement bruitée, nous avons proposé une nouvelle manière d'établir les profils d'observation.

Pour l'établissement des profils de tonalité, nous avons utilisé les notes issues des triades des accords de la pièce de musique, au lieu de l'ensemble des notes la composant. Pour la construction du profil servant à l'estimation des accords, nous avons pondéré la contribution des notes de musique non seulement selon leurs durées, mais également selon le nombre d'autres notes de musique sonnantes en même temps qu'elles. Pour l'estimation de la tonalité, les évaluations réalisées pour ces méthodes ainsi que pour les méthodes existantes montrent une amélioration de la précision pour les méthodes proposées. Concernant l'estimation des accords, nous avons montré que la méthode proposée se révélait plus efficace qu'une méthode directe adaptant le principe de certaines méthodes existantes pour l'estimation de tonalité au domaine des accords.

A la lumière de ces résultats, plusieurs perspectives peuvent être envisagées. D'autres paramètres pourraient être pris en compte dans l'obtention des pro-

files d'accord ou de tonalité, comme le rôle de la basse de chaque groupe de notes. Des techniques comme la suppression des notes de passages pourraient également être intégrées aux méthodes proposées pour diminuer le bruit harmonique [12, 18]. Il est également souhaitable d'adopter une segmentation non-uniforme et sans apprentissage pour la détection d'accord : au lieu d'estimer un accord pour chaque période d'analyse (la mesure par exemple), le début et la fin de chaque accord pourraient être détectés automatiquement. Enfin, il reste à adapter ces méthodes d'estimation à l'étude du signal audio. Une détection de fondamentale pour chaque chroma pourrait par exemple être envisagée pour donner plus de poids à cette dernière, les chromas pouvant être apparentés à des groupes de notes, car donnant une description des notes sonnantes simultanément.

5. SUPPORT

Les travaux présentés dans cet article s'inscrivent dans le cadre du projet Simbals JC07-188930, financé par l'Agence Nationale pour la Recherche. Ces recherches sont également soutenues par le Conseil Régional d'Aquitaine.

6. REFERENCES

- [1] W. Chai and B. Vercoe, "Detection of Key Change in Classical Piano Music", *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR)*, Londres, Royaume-Uni, 2007.
- [2] E. Chew, "Towards a Mathematical Model of Tonality", *PhD. Thesis, Operations Research Center, MIT, Cambridge, États-Unis*, 2000.
- [3] J. Downie, K. West, A. Ehmann and E. Vincent, "The 2005 Music Information retrieval Evaluation Exchange (MIREX'05) : Preliminary Overview", *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR)*, Londres, Royaume-Uni, 2005.
- [4] T. Eerola and P. Toivainen, "Suomen Kansan esävelmät. Finnish Folk Song Database", disponible à : <http://www.jyu.fi/musica/sks/>, 2004.
- [5] T. Fujishima, "Realtime chord recognition of musical sound : A system using Common Lisp Music", *Proceedings of the International Computer Music Conference (ICMC)*, Pékin, Chine, 1999.
- [6] E. Gómez, "Tonal Description of Music Audio Signals", *Thèse doctorale, Université de Pompeu Fabra, Barcelone, Espagne*, 2006.
- [7] E. Gómez and P. Herrera, "Estimating The Tonality Of Polyphonic Audio Files : Cognitive Versus Machine Learning Modelling Strategies" *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR)*, Barcelone, Espagne, 2004.
- [8] C. Harte, and M. Sandler, "Automatic Chord Identification Using a Quantised Chromagram", *Proceedings of the Audio Engineering Society*, Madrid, Espagne, 2005.
- [9] C. Krumhansl, "Cognitive Foundations of Musical Pitch", *Oxford University Press, New York*, 1990.
- [10] C. Krumhansl and E. Kessler, "Tracing the Dynamic Changes in Perceived Tonal Organisation in a Spatial Representation of Musical Keys", *Psychological Review*, volume 89, pages 334–368, 1982.
- [11] K. Lee and M. Stanley, "A Unified System for Chord Transcription and Key Extraction Using Hidden Markov Models", *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR)*, Vienne, Autriche, 2007.
- [12] F. Lerdahl and R. Jackendoff, "A Generative Theory of Tonal Music", *MIT Press, Cambridge*, 1985.
- [13] Y. Li and D. Huron, "Melodic Modeling : A Comparison of Scale Degree and Interval", *Proceedings of the International Computer Music Conference (ICMC)*, New Orleans, États-Unis, 2006.
- [14] S. Madsen and G. Widmer, "Key-Finding with Interval Profiles", *Proceedings of the International Computer Music Conference (ICMC)*, Copenhagen, Danemark, 2007.
- [15] K. Noland and M. Sandler, "Key Estimation Using a Hidden Markov Model", *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR)*, Victoria, Canada, 2006.
- [16] P. Illescas, D. Rizo and J. M. Iñesta, "Harmonic, Melodic, and Functional Automatic Analysis", *Proceedings of the International Computer Music Conference (ICMC)*, pages 165–168, Copenhagen, Danemark, 2007.
- [17] M. Robine, T. Rocher, and P. Hanna, "Improvements of Key-Finding Methods", *Proceedings of the International Computer Music Conference (ICMC)*, Belfast, Irlande du Nord, 2008.
- [18] H. Schenker, "Der Frei Satz", Vienna : Universal Edition, 1935. Note : publié en anglais sous le titre *Free Composition*, traduit par E. Oster, Longman, 1979
- [19] D. Temperley, "The Cognition of Basic Musical Structures", MIT Press, 1999.
- [20] D. Temperley, "A Bayesian Approach to Key-Finding", *Music and Artificial Intelligence*, pages 195–206, Berlin : Springer-Verlag, 2002.