

ANALYSE DE REPRÉSENTATIONS SPATIALES DE LA MUSIQUE PAR DES OPÉRATEURS SIMPLES DE MORPHOLOGIE MATHÉMATIQUE

Paul Lascabettes^{1,2,3}

1. École Normale Supérieure

paul.lascabettes@
ens-paris-saclay.fr

Isabelle Bloch²

2. LTCI, Télécom Paris
Institut Polytechnique de Paris

isabelle.bloch@
telecom-paris.fr

Carlos Agon³

3. STMS
Sorbonne Université, IRCAM

agon@ircam.fr

RÉSUMÉ

La morphologie mathématique, développée au cours des années 1960 à l'École des Mines de Paris, est une théorie qui, dans sa branche déterministe, repose sur des formulations mathématiques essentiellement algébriques, donnant des opérateurs non linéaires. Ses développements en analyse d'images reposent sur des opérateurs permettant d'étudier les propriétés des objets (forme, topologie, apparence) et de leur voisinage. Ses applications vont des pré-traitements à l'analyse (segmentation, reconnaissance d'objets) et à l'interprétation de scènes.

Aujourd'hui, il n'existe que peu d'applications directes à la musique. Ainsi, cet article a-t-il pour but de proposer des axes d'application de la morphologie mathématique à la musique. En représentant la musique sous forme d'une image du type *Piano Roll*, les opérateurs morphologiques permettent d'extraire, d'enrichir ou de transformer les accords, d'isoler la mélodie, ou encore de détecter certains motifs musicaux. La pertinence des différentes transformations morphologiques sera étudiée et discutée dans cet article dans le but de mieux comprendre comment étendre les domaines d'application de la morphologie mathématique aux représentations spatiales de la musique.

1. INTRODUCTION

Développée dans les années 1960 par Georges Matheron et Jean Serra à l'École des Mines de Paris, la morphologie mathématique est un axe de recherche qui suscite de plus en plus d'intérêt dans la communauté scientifique internationale. Elle repose sur des notions issues de la théorie des ensembles, de la topologie, de la théorie des treillis ou encore des probabilités. Ici on ne considère que le cadre déterministe de la morphologie mathématique, et principalement des opérations croissantes. Le cadre le plus classique est alors celui des treillis complets, qui permet d'utiliser la morphologie mathématique dans un grand nombre de structures algébriques munies d'une relation d'ordre partiel définissant un treillis, comme des ensembles, des fonctions, des graphes ou hypergraphes, ou encore des propositions logiques.

Principalement appliquée en analyse d'images, la morphologie mathématique permet de modifier la forme des

objets, leur taille et leurs propriétés topologiques grâce à des transformations non linéaires et non inversibles. Le principe de base des opérations les plus classiques de la morphologie mathématique est d'analyser des formes ou des structures dans les images à l'aide d'un élément structurant (relation binaire, le plus souvent de voisinage) et de ses relations avec ces objets ou structures.

Il existe très peu d'applications directes de cette théorie à l'analyse musicale. Karvonen et coll. [8, 9] ont utilisé les outils morphologiques pour détecter des motifs presque identiques à un motif type dans une base de données. Pour cela, des opérations morphologiques sont appliquées à une représentation de type *Piano Roll*. Agon et coll. [1] ont appliqué la morphologie en travaillant avec des treillis de concepts formels construits sur des intervalles musicaux, tandis que Lascabettes [10] a proposé une méthode de détection de la structure musicale en utilisant les opérateurs morphologiques pour filtrer des matrices d'auto-similarité. Cet article propose une nouvelle méthode d'application de la morphologie mathématique à la musique en travaillant à partir des opérateurs de base de la morphologie mathématique sur des fichiers MIDI. Chaque opérateur sera étudié et illustré musicalement dans cet article afin d'en comprendre les caractéristiques.

Cet article est organisé comme suit. En section 2, les principes fondamentaux de la morphologie mathématique sont rappelés. La section 3 détaille comment nous allons représenter la musique pour pouvoir y appliquer les outils morphologiques. La section 4 présente les résultats qu'il est possible d'obtenir en appliquant les opérateurs morphologiques de base sur un accord. La section 5 reprend ces idées en travaillant sur une section rythmique. La section 6 permet de comprendre la pertinence de certaines opérations morphologiques pour extraire des données d'un fichier MIDI. Enfin, une discussion accompagnée de travaux futurs, proposée en section 7, conclut l'article.

2. RAPPELS DE MORPHOLOGIE MATHÉMATIQUE

Nous rappelons ici brièvement les concepts de la morphologie mathématique qui seront utiles dans la suite. Plus de détails peuvent être trouvés par exemple dans [2, 5,

6, 12, 13, 14, 15]. La majorité des transformations morphologiques déterministes découlent de deux opérations de base qui sont la *dilatation* et l'*érosion*. De manière générale, dans un treillis complet, une dilatation est définie comme un opérateur qui commute avec le supremum et l'érosion comme un opérateur qui commute avec l'infimum (au sens de l'ordre partiel défini sur le treillis). Ces définitions abstraites sont souvent spécifiées sous une forme concrète impliquant un élément structurant, représentant une relation binaire (ou de voisinage dans une interprétation spatiale). Nous nous plaçons pour cette présentation dans le cas simple du treillis des parties d'un ensemble E , muni de la relation d'inclusion : $(\mathcal{P}(E), \subseteq)$. Tout d'abord, il est nécessaire de réintroduire quelques notions élémentaires dans le cas où l'espace E est \mathbb{R}^n ou \mathbb{Z}^n . Soit $X \subseteq E$, on rappelle que :

- son *complémentaire* est $X^c = \{x \in E \mid x \notin X\}$;
- son *translaté* par $t \in E$ est $X_t = \{x + t \mid x \in X\}$;
- son *symétrique* est $\check{X} = \{-x \mid x \in X\}$.

Soit $S \in \mathcal{P}(E)$ un *élément structurant*, la *dilatation* δ_S et l'*érosion* ε_S par S sont définies par :

$$\delta_S : \mathcal{P}(E) \longrightarrow \mathcal{P}(E) \quad (1)$$

$$X \longmapsto X \oplus S$$

$$\varepsilon_S : \mathcal{P}(E) \longrightarrow \mathcal{P}(E) \quad (2)$$

$$X \longmapsto X \ominus S$$

où \oplus et \ominus sont respectivement l'*addition* et la *soustraction de Minkowski* introduits en 1903 par Minkowski [11] et en 1950 par Hadwiger [4], et définis pour tout $X, S \in \mathcal{P}(E)$ par :

$$\begin{aligned} X \oplus S &= \bigcup_{s \in S} X_s = \{x + s \mid x \in X, s \in S\} \\ &= \{t \in E \mid \check{S}_t \cap X \neq \emptyset\} \\ X \ominus S &= \bigcap_{s \in S} X_{-s} = \{t \in E \mid S_t \subseteq X\} \end{aligned} \quad (3)$$

De plus, ces deux opérations sont *duales par complémentation*, c'est-à-dire :

$$X \oplus S = (X^c \ominus \check{S})^c \quad X \ominus S = (X^c \oplus \check{S})^c \quad (4)$$

Pour illustrer ces concepts, un ensemble X en gris et un élément structurant S en rouge sont représentés sur la figure 1.

Il est important de remarquer que l'élément structurant est défini avec une *origine* O qui a un impact direct sur l'addition et la soustraction de Minkowski, et donc modifie le résultat de la dilatation et de l'érosion. Par défaut, l'origine est localisée au centre mais dans notre exemple,



Figure 1. Exemple d'un ensemble X en gris et d'un élément structurant S en rouge.

l'origine se situe dans le coin supérieur du triangle. En règle générale, si l'origine est à l'intérieur de l'élément structurant, la dilatation est *extensive* ($X \subseteq \delta_S(X)$ si et seulement si $O \in S$), tandis que l'érosion est *anti-extensive* ($\varepsilon_S(X) \subseteq X$ si et seulement si $O \in S$). De plus les deux opérations sont croissantes et vérifient la propriété d'adjonction : $\forall X, Y, \delta_S(X) \subseteq Y \iff X \subseteq \varepsilon_S(Y)$. D'autres propriétés peuvent être trouvées dans les ouvrages cités.

Pour avoir une intuition de la forme obtenue après une dilatation, il faut imaginer que l'origine du triangle va se déplacer le long de X . La dilatation sera constituée de l'ensemble des points que balaie le triangle. La figure 2a illustre la transformation effectuée par une dilatation où la frontière de l'ensemble dilaté est de couleur noire. De la même façon, pour visualiser l'érosion, il est possible d'utiliser la dualité par complémentation. Ainsi, il faut imaginer de faire une dilatation du complémentaire de X par le symétrique de S . L'origine de l'élément structurant glisse le long de la frontière de X mais l'élément structurant est inversé et la frontière de l'ensemble érodé sera l'ensemble des points où le triangle ne peut aller. La frontière du résultat de l'érosion est indiquée en noir sur la figure 2b. Dans les deux cas, la forme de X a changé et les transformations ne sont pas seulement des homothéties.

Pour résumer, une dilatation remplit les trous plus petits que l'élément structurant (au sens de l'inclusion), rassemble les composantes qui sont proches topologiquement et élargit les formes. En revanche, une érosion élimine les composantes connexes plus petites que l'élément structurant, élargit les trous et réduit la taille des objets. Ce sont des opérations non réversibles et duales l'une de l'autre. Ainsi, la figure 3 représente ce lien entre les deux principaux opérateurs sur un graphe.

Les deux autres opérations fondamentales résultent de la composition des fonctions précédentes. En effet, l'*ouverture* γ_S est la composition d'une érosion et d'une dilatation et la *fermeture* ϕ_S est une dilatation suivie d'une érosion.

$$\gamma_S : \mathcal{P}(E) \longrightarrow \mathcal{P}(E) \quad (5)$$

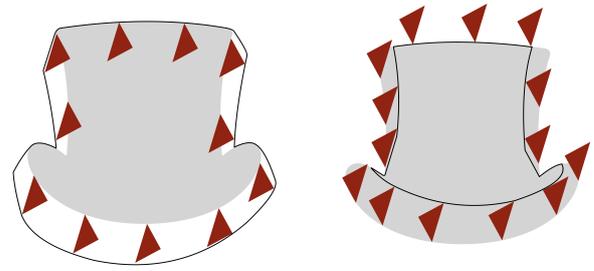
$$X \longmapsto X \circ S$$

$$\phi_S : \mathcal{P}(E) \longrightarrow \mathcal{P}(E) \quad (6)$$

$$X \longmapsto X \bullet S$$

où pour tout $X, S \in \mathcal{P}(E)$:

$$X \circ S = (X \ominus S) \oplus S \quad X \bullet S = (X \oplus S) \ominus S \quad (7)$$



(a) Dilatation de X par S

(b) Érosion de X par S

Figure 2. Dilatation et érosion de X par S .

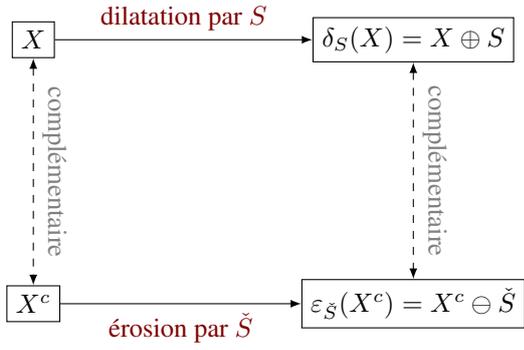


Figure 3. Graphe illustrant les relations entre la dilatation et l'érosion.

De plus, ces opérations sont *duales* :

$$X \circ S = (X^c \bullet \check{S})^c \quad X \bullet S = (X^c \circ \check{S})^c \quad (8)$$

L'ouverture de X est l'ensemble des points que balaie S lorsque S est inclus dans X . Par dualité, la fermeture revient à ajouter des points à X où \check{S} ne peut pas aller en restant dans le complémentaire de X . Il est important de noter que l'origine de l'élément structurant S n'a plus d'importance pour ces deux opérations, seule la forme de l'élément structurant est importante. La frontière du résultat de ces opérations est représentée en noir sur la figure 4 en utilisant les mêmes ensembles X et S qu'auparavant.

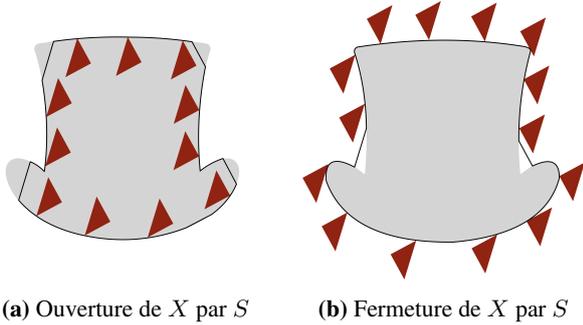


Figure 4. Ouverture et fermeture de X par S .

Une ouverture lisse les formes, élimine les composantes connexes plus petites que l'élément structurant, ne préserve pas (toujours) la topologie et réduit l'objet, γ_S est *anti-extensif* ($\forall X, \gamma_S(X) \subseteq X$), alors qu'une fermeture remplit les trous plus petits que l'élément structurant, ne conserve pas (toujours) la topologie, soude les formes rapprochées et agrandit l'objet, ϕ_S est *extensif* ($\forall X, X \subseteq \phi_S(X)$). Ainsi, les inclusions suivantes restent vérifiées :

$$\forall X, S \in \mathcal{P}(E), \quad \gamma_S(X) \subseteq X \subseteq \phi_S(X) \quad (9)$$

De plus, ces deux opérations sont croissantes (comme compositions d'opérations croissantes), et idempotentes, c'est-à-dire :

$$\gamma_S \circ \gamma_S = \gamma_S \quad \phi_S \circ \phi_S = \phi_S \quad (10)$$

En composant successivement l'ouverture et la fermeture, on peut obtenir quatre nouveaux filtres : ouverture

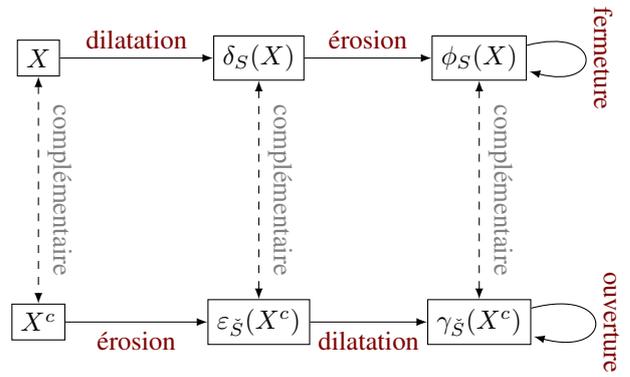


Figure 5. Graphe illustrant les relations entre les opérations de base de la morphologie.

suivie d'une fermeture, fermeture suivie d'une ouverture, ouverture suivie d'une fermeture et d'une ouverture, fermeture suivie d'une ouverture et d'une fermeture. Toutes ces opérations sont idempotentes, et aucun autre opérateur ne peut être obtenu par une composition supplémentaire comme l'a montré Serra en 1988 [15]. Pour finir, la figure 5 résume sur un graphe le lien entre les quatre opérations que nous venons de définir et qui forment la base de la morphologie mathématique.

3. MORPHOLOGIE MATHÉMATIQUE ET REPRÉSENTATIONS SPATIALES DE LA MUSIQUE

La présentation précédente des opérateurs de base de la morphologie mathématique s'applique directement à des images binaires. L'extension à des images à niveaux de gris et en couleurs repose sur des principes similaires, mais dans des treillis différents (fonctions scalaires ou vectorielles, munies d'une relation d'ordre partiel). Pour appliquer la morphologie mathématique à la musique, nous allons donc suivre cette logique, et commencer par utiliser les concepts de morphologie sur les images binaires. Pour ce faire, il est nécessaire de trouver une bonne représentation spatiale de la musique, c'est-à-dire de réduire un morceau de musique à une image binaire sans perdre trop d'informations. Il existe de nombreuses façons de le faire. Par exemple une représentation temps-fréquence d'un signal musical peut être interprété comme une image. La trajectoire d'une pièce de musique dans le cercle chromatique ou dans un *Tonnetz* est un autre exemple.

Une autre idée intuitive, que nous allons explorer ici, consiste à représenter le morceau de musique comme un *Piano Roll* particulier. En effet, dans ce *Piano Roll*, on ne tient pas compte de la durée des notes. Pour traiter les données plus facilement, seule l'apparition d'une note sera importante. De plus, pour obtenir une image binaire, les coefficients doivent être 0 ou 1. La valeur 1 signifiera qu'une note est jouée, alors que 0 est similaire à l'absence de note. Par conséquent, les notes sont toutes d'intensité constante. Cette première représentation permet de comprendre comment appliquer les outils de la morphologie

mathématique à la musique. Un des problèmes fondamentaux de la communauté MIR – *Music Information Research community* – est de trouver une bonne représentation de la musique pour appliquer les théories qui proviennent de l'analyse d'images comme la morphologie mathématique. Le *Piano Roll* présenté en figure 8 est une première idée, mais les formes musicales qui apparaissent ne sont pas compactes et ne sont pas connexes. Comme les opérateurs de morphologie fonctionnent très bien sur les images qui contiennent des formes connexes, il faut s'efforcer de modifier les opérateurs et/ou de modifier la représentation musicale. Les sections suivantes se concentrent sur la visualisation classique de la musique sous forme de *Piano Roll* où il est possible d'adapter les opérateurs issus de la morphologie mathématique en particulier avec des choix judicieux des éléments structurants. Nous présenterons des exemples simples, qui pourraient bien sûr être traités par d'autres approches, mais qui permettent d'illustrer le potentiel de la nouvelle approche proposée ici.

4. APPLICATION DES PRINCIPAUX OPÉRATEURS MORPHOLOGIQUES À UN ACCORD MAJEUR

Avant d'appliquer la morphologie mathématique sur des pièces musicales, il est nécessaire d'étudier quelles sont les transformations qu'il est possible d'obtenir en travaillant sur un accord. Dans ce cas, la dimension temporelle n'est pas prise en compte, ce qui réduit directement les éventuelles applications. Néanmoins, il est possible d'obtenir certains résultats des quatre opérations de base de la morphologie. Pour cela, l'accord le plus populaire de la musique occidentale, c'est-à-dire l'accord de *C majeur* sera étudié. Cet accord, composé de trois notes et représenté en figure 6, sera vu comme une matrice colonne $(0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1)^T$.



Figure 6. Représentation de l'accord de *C majeur*.

Tout d'abord, lorsque l'élément structurant est composé de 0 et d'un seul 1, la *dilatation* est équivalente à une translation de l'accord. Par conséquent, l'accord sera plus aigu ou plus grave, mais gardera les mêmes propriétés (nombre de notes et intervalles entre ces notes). Cette opération est très utilisée par les musiciens et apparaît dans de nombreux morceaux populaires, car elle induit un effet de surprise et de renouveau. La table 1 illustre cette idée avec

Dilatation		
Élément structurant	Résultat de l'opération	Représentation dans le cercle chromatique

Table 1. La dilatation vue comme une translation de l'accord initial, ou une rotation dans le cercle chromatique.

Érosion		
Élément structurant	Résultat de l'opération	Représentation dans le cercle chromatique

Table 2. L'érosion peut permettre de détecter la fondamentale de l'accord initial.

un élément structurant très simple : $(1, 0, 0)^T$, où l'origine se situe au centre. La conséquence d'une dilatation par cet élément structurant est une translation d'un demi-ton, le résultat est un accord de *C# majeur*. La transformation est aussi représentée dans le cercle chromatique où les notes en gris correspondent à l'accord initial et celles en rouge à l'accord obtenu après transformation. Notons que la translation dans l'espace du *Piano Roll* correspond à une rotation sur le cercle chromatique.

Par suite, l'érosion permet de simplifier l'accord pour garder uniquement certaines propriétés, par exemple sa fondamentale. En effet, en utilisant $(1, 0, 0, 0, 1)^T$ comme élément structurant, où l'origine est située sur le dernier coefficient, l'érosion va détecter uniquement la fondamentale de l'intervalle de tierce majeure. Le résultat est illustré dans la table 2. Dans ce cas, le résultat d'une érosion sera $(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1)^T$ ce qui correspond à la note *C* qui est la fondamentale de l'accord. Plus généralement,

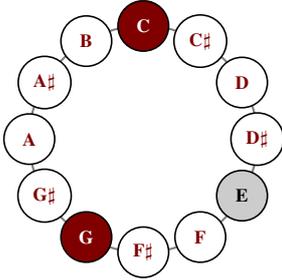
Ouverture														
Élément structurant	Résultat de l'opération	Représentation dans le cercle chromatique												
	<table border="1"> <tr><td>B</td></tr> <tr><td>A#</td></tr> <tr><td>A</td></tr> <tr><td>G#</td></tr> <tr><td>G</td></tr> <tr><td>F#</td></tr> <tr><td>F</td></tr> <tr><td>E</td></tr> <tr><td>D#</td></tr> <tr><td>D</td></tr> <tr><td>C#</td></tr> <tr><td>C</td></tr> </table>	B	A#	A	G#	G	F#	F	E	D#	D	C#	C	
B														
A#														
A														
G#														
G														
F#														
F														
E														
D#														
D														
C#														
C														

Table 3. L'ouverture sélectionne un motif musical particulier de l'accord initial, par exemple en gardant uniquement l'intervalle de quinte.

en sélectionnant un intervalle comme élément structurant, l'érosion permettra d'extraire une seule note par rapport à cet intervalle. Cette note correspond à la position de l'origine dans l'élément structurant. Dans l'exemple précédent, l'érosion détecte la fondamentale, mais il est possible de choisir de détecter n'importe quelle note en adaptant la position de l'origine.

L'opération morphologique qui a le plus de sens dans ce contexte musical est sans doute l'*ouverture*, pouvant extraire des motifs musicaux particuliers. La recherche de motifs musicaux est un domaine de recherche très actif, comme on peut le voir avec [3] ou [7], et cette opération pourrait sans doute permettre de grandes avancées. Par exemple, en choisissant $(1, 0, 0, 0, 0, 0, 1)^T$ comme élément structurant, qui est un intervalle de quinte, l'ouverture va sélectionner les intervalles de quinte de l'accord initial. La table 3 permet d'illustrer le résultat de cette opération. Seul ce qui est similaire à l'élément structurant sera présent dans l'ouverture. Par conséquent, l'accord de *C majeur* sera transformé en un accord de quinte *C/G*, autrement dit, la tierce majeure a été enlevée de l'accord. En appliquant cette opération à une pièce musicale, il sera possible d'extraire directement tous les intervalles de quintes de la pièce ce qui est pertinent pour permettre de simplifier et d'épurer le morceau.

Pour finir, la *fermeture* est la transformation morphologique la moins intuitive. Puisque $\forall(X, S), X \subseteq \phi_S(X)$ (extensivité de la fermeture), cette opération va enrichir l'accord initial. En choisissant $(1, 0, 1, 0, 1)^T$ comme élément structurant, la fermeture va ajouter un *D* à l'accord de *C majeur*, ce qui est illustré en table 4. Néanmoins, il n'est pas possible d'ajouter n'importe quelle note à l'accord de base, ce qui rend cette opération plus compliquée à utiliser en pratique. En effet, dans l'exemple d'un accord de *C majeur* il n'est pas possible d'ajouter un *B* pour former un accord de *C majeur 7* qui est très utilisé par les musiciens. Il sera possible d'obtenir un accord de *C majeur 7* par d'autres opérations morphologiques (en parti-

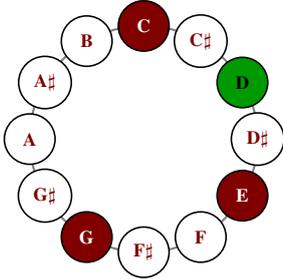
Fermeture														
Élément structurant	Résultat de l'opération	Représentation dans le cercle chromatique												
	<table border="1"> <tr><td>B</td></tr> <tr><td>A#</td></tr> <tr><td>A</td></tr> <tr><td>G#</td></tr> <tr><td>G</td></tr> <tr><td>F#</td></tr> <tr><td>F</td></tr> <tr><td>E</td></tr> <tr><td>D#</td></tr> <tr><td>D</td></tr> <tr><td>C#</td></tr> <tr><td>C</td></tr> </table>	B	A#	A	G#	G	F#	F	E	D#	D	C#	C	
B														
A#														
A														
G#														
G														
F#														
F														
E														
D#														
D														
C#														
C														

Table 4. La fermeture peut permettre l'ajout d'une note à l'accord initial, par exemple ajouter un *D*.

culier avec une transformation en *tout ou rien* [14]¹).

5. APPLICATION DES PRINCIPAUX OPÉRATEURS MORPHOLOGIQUES À UNE SECTION RYTHMIQUE

La section 4 a permis de clarifier et de synthétiser les applications possibles de chacune des quatre opérations de base de la morphologie mathématique sur un accord. Puisque l'accord est représenté verticalement, il est possible d'appliquer les mêmes opérations *horizontalement*, c'est-à-dire sur une échelle purement rythmique et non mélodique. En représentant l'accord précédent horizontalement, on obtient une section rythmique qui est présentée en figure 7. Ce rythme de longueur 12 est composé de trois impulsions aux temps 1, 5 et 8 ce qui est représenté par la matrice $(1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0)$.



Figure 7. Représentation d'une section rythmique.

De la même façon que précédemment, la *dilatation* va effectuer une translation horizontale de la section rythmique en travaillant avec un élément structurant composée de 0 et d'un seul 1. Par exemple, une dilatation par $(0, 0, 1)$, où l'origine est au centre, va effectuer une translation de la section rythmique d'une case vers la droite. Il est aussi possible d'utiliser la dilatation pour générer et complexifier le rythme de base, ce qui est plus compliqué à faire verticalement à cause de problèmes de tonalités. C'est pourquoi la dilatation est bien plus intéressante d'un point de vue rythmique que mélodique. En effet, la dilatation par un élément structurant composé de quelques 1 va engendrer une grande complexité sur un rythme simple. Ensuite, l'*érosion* va simplifier le rythme pour garder une seule impulsion s'il y a des intervalles rythmiques similaires à ceux de l'élément structurant dans la section ryth-

¹ Cette transformation a été utilisée dans [9] pour rechercher un motif donné dans des représentations de type *Piano Roll*.

mique de base. L'*ouverture* va extraire les rythmes de la section rythmique qui sont identiques à l'élément structurant à translation horizontale près. Le résultat d'une ouverture sera une simplification de la section rythmique. Une fois de plus, le résultat est similaire à ceux de la section précédente mais en travaillant horizontalement. Enfin, la *fermeture* complexifie le rythme en ajoutant des notes aux rythmes similaires à l'élément structurant, mais comme expliqué dans la section précédente, elle reste compliquée à utiliser car il n'est pas possible d'ajouter n'importe quel temps par rapport à un rythme donné.

6. APPLICATIONS À L'EXTRACTION DE DONNÉES D'UN FICHIER MIDI GRÂCE AUX OPÉRATEURS MORPHOLOGIQUES

Après avoir présenté comment appliquer les principaux opérateurs morphologiques à un accord en section 4 et à un rythme en section 5, nous allons nous focaliser sur l'étude d'un fichier MIDI à partir de la morphologie mathématique. La dilatation et la fermeture permettent de modifier et d'enrichir le morceau, alors que l'érosion et l'ouverture agissent comme un filtre pour extraire certaines données du morceau. Par conséquent, ce sont ces deux dernières opérations qui seront pertinentes dans l'étude d'un fichier MIDI. Pour illustrer ces concepts, l'introduction de *Hey Jude* des Beatles sera analysée.

Dans le but de permettre une compréhension plus claire, ce n'est pas la version officielle mais une adaptation au piano qui sera étudiée et qui est représentée en figure 8. La mélodie est présente en haut du *Piano Roll* car elle est jouée dans les aigus, tandis que les basses sont situées bien en-dessous. Les accords qui sont joués pendant le morceau sont au niveau de la partie centrale du *Piano Roll*.

Ces mesures sont construites sur des accords majeurs, il est donc naturel de choisir un accord majeur comme élément structurant pour pouvoir en extraire la partie centrale. Comme un accord majeur est composé de trois notes : la fondamentale, la tierce majeure et la quinte juste, il peut être représenté de trois manières différentes si l'ordre est maintenu, comme illustré en figure 9.

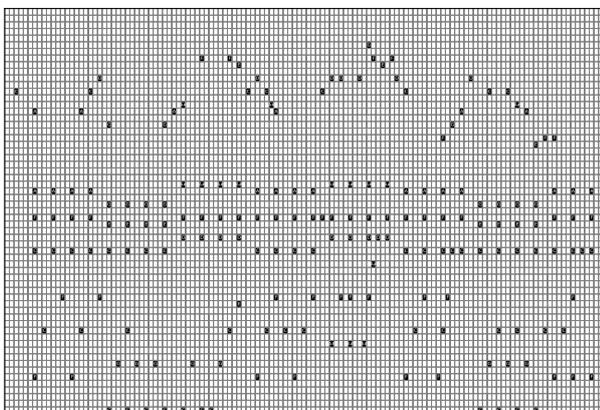


Figure 8. Introduction de *Hey Jude* représentée sous forme de *Piano Roll*.

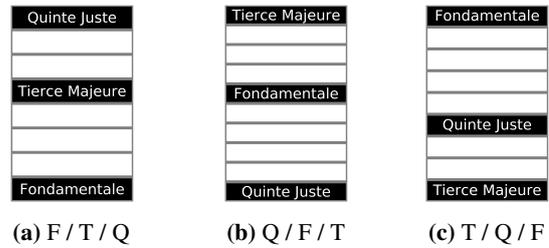


Figure 9. Trois représentations différentes d'un accord majeur, considérées comme trois éléments structurants.

Par suite, il est nécessaire de choisir plusieurs éléments structurants et donc de généraliser certains concepts morphologiques. Une érosion $\varepsilon_{S_1, \dots, S_n}$ et une ouverture γ_{S_1, \dots, S_n} de X avec plusieurs éléments structurants S_1, \dots, S_n sont définies respectivement comme :

$$\varepsilon_{S_1, \dots, S_n}(X) = \bigcup_{i \in \{1, \dots, n\}} \varepsilon_{S_i}(X) \quad (11)$$

$$\gamma_{S_1, \dots, S_n}(X) = \bigcup_{i \in \{1, \dots, n\}} \gamma_{S_i}(X) \quad (12)$$

Notons qu'une réunion d'ouvertures est aussi une ouverture dans un sens algébrique, c'est-à-dire une opération *croissante*, *anti-extensive* et *idempotente*. Pour cela, nous allons choisir les trois représentations d'un accord majeur de la figure 9 :

- $(1, 0, 0, 1, 0, 0, 0, 1)^T$ correspond à la permutation *fondamentale / tierce / quinte*, où la fondamentale est la note la plus grave,
- $(1, 0, 0, 0, 1, 0, 0, 0, 1)^T$ correspond à *quinte / fondamentale / tierce*,
- et pour finir, $(1, 0, 0, 0, 0, 1, 0, 0, 1)^T$ correspond à un ordre *tierce / quinte / fondamentale*.

L'*ouverture* par ces trois éléments structurants procède à l'extraction des accords du morceau. Le résultat de cette opération est présenté en figure 10. Ainsi, la mélodie et les basses ont été supprimées après cette transformation.

Une fois de plus, les éléments structurants sont choisis pour que toutes les notes de l'accord soient jouées en même temps pour plus de clarté. Néanmoins, il est tout à fait possible d'imaginer des éléments structurants

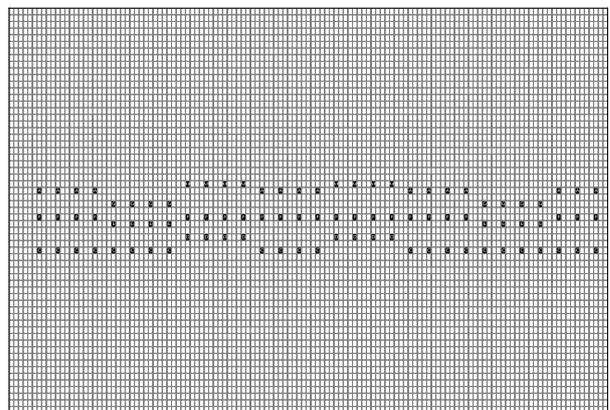


Figure 10. Ouverture de l'introduction de *Hey Jude* par des accords majeurs.

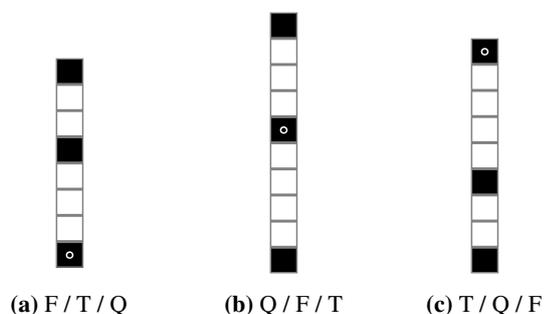


Figure 11. Indication de la position de l'origine sur les trois éléments structurants.

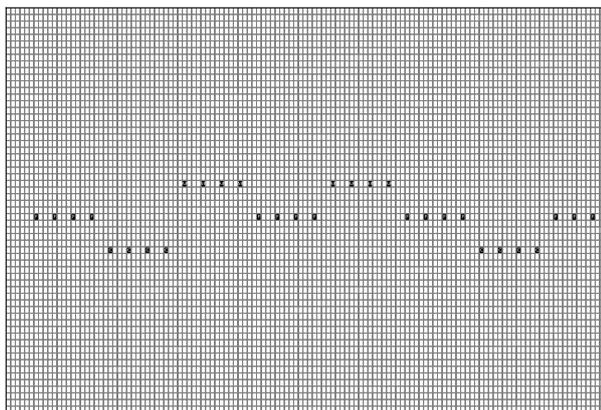


Figure 12. Érosion de l'introduction de *Hey Jude* pour détecter la fondamentale de chaque accord.

où l'accord est étalé dans le temps comme des arpèges. En reprenant ces trois éléments structurants mais en ajoutant une origine située sur la fondamentale, l'érosion va permettre d'extraire uniquement la fondamentale de chaque accord. La figure 11 indique la position de l'origine sur les trois éléments structurants.

Le résultat de l'érosion est illustré en figure 12. Cela permet de conclure que l'introduction de cette adaptation de *Hey Jude* est composée des accords suivants, chacun répété 4 fois : $F-C-B\flat-F-B\flat-F-C-F$.

Ainsi, ces deux exemples permettent de comprendre l'intérêt des outils morphologiques pour extraire des données et analyser un fichier MIDI. Pour aller plus loin, il est aussi possible de combiner certaines opérations morphologique pour obtenir d'autres résultats. Par exemple, il est possible de retirer les accords du morceau de base dans le but d'extraire la mélodie. Cela peut être obtenu avec l'opération morphologique de *chapeau haut-de-forme* qui consiste en une soustraction de l'ouverture à l'ensemble de départ.

7. CONCLUSION ET TRAVAUX FUTURS

Nous avons proposé dans cet article de nouvelles méthodes d'analyse musicale grâce aux opérateurs de morphologie mathématique. Les quatre opérations principales de la morphologie ont été étudiées et la pertinence de chacune a été exposée. Ces exemples simples initient un nouvel axe de recherche, original, associant des formulations

mathématiques (essentiellement algébriques) à des structures musicales.

Puisque cette étude se limite aux outils de la morphologie mathématique appliquée aux images binaires, un travail similaire est envisageable en s'intéressant aux images en niveaux de gris puis aux images en couleurs. En effet, la vélocité, qui correspond à l'intensité d'une note, est un paramètre à rajouter qui nous amènerait à utiliser des fonctions, interprétées comme des images non binaires. Une note avec une forte vélocité correspondrait à une valeur proche de 1 tandis qu'une faible vélocité se traduirait par une valeur presque nulle. De plus, il serait intéressant de comparer les résultats obtenus sur des fichiers MIDI et MP3 en travaillant avec un spectrogramme.

Un autre axe de développement de cet article est la recherche d'éléments structurants caractérisant un corpus donné, par exemple déterminer les éléments structurants non-triviaux qui maximisent le nombre de notes après une ouverture. La connaissance de ces éléments permettrait de comprendre la caractéristique de certains morceaux et d'obtenir une approche supplémentaire à la détection de motifs musicaux.

8. RÉFÉRENCES

- [1] Agon, C., Andreatta, M., Atif, J., Bloch, I., Mascara Relano, P. « Musical Descriptions based on Formal Concept Analysis and Mathematical Morphology », Proceedings of the International Conference on Conceptual Structures, Edinburgh, Royaume-Uni, 2018.
- [2] Bloch, I., Heijmans, H., Ronse, C. « Mathematical Morphology », *Handbook of Spatial Logics*, Aiello, M., Pratt-Hartman, I., van Benthem, J. (dir.), Springer, Cham, 2007, chap. 13, pp. 857-947
- [3] Conklin, D. « Discovery of distinctive patterns in music », *Intelligent Data Analysis* 14 (2010), p. 547-554.
- [4] Hadwiger, H. « Minkowskische Addition und Subtraktion beliebiger Punktmengen und die Theoreme von Erhard Schmidt », *Mathematische Zeitschrift* 53 (1950), p. 210-218.
- [5] Heijmans, H. *Morphological Image Operators*, Academic Press, Boston, 1994.
- [6] Heijmans, H., and Ronse, C. « The Algebraic Basis of Mathematical Morphology – Part I: Dilations and Erosions », *Computer Vision, Graphics and Image Processing* 50/3 (1990), p. 245-295.
- [7] Hsu, J.-L., Chen, A. « Efficient Repeating Pattern Finding in Music Databases », Proceedings of the International Conference on Information and Knowledge Management, Bethesda, États-Unis pages 281-288, 1998.
- [8] Karvonen, M., Lemström, K. « Using Mathematical Morphology for Geometric Music Retrieval », Proceedings of the International Workshop on Machine Learning and Music, Helsinki, Finland, 2008.

- [9] Karvonen, M., Laitinen, M., Lemström, K., Vikman, J. « Error-tolerant content-based music-retrieval with mathematical morphology », Proceedings of the International Symposium on Computer Music Modeling and Retrieval, Málaga, Espagne, 2010.
- [10] Lascabettes, P. « Mathematical Morphology Applied to Music », mémoire de Master, sous la dir. de I. Bloch, C. Agon, École Normale Supérieure, Paris–Saclay, 2019.
- [11] Minkowski, H. « Volumen und Oberfläche », *Mathematische Annalen* 57/4 (1903), p. 447-495.
- [12] Najman, L., Talbot, H. (dir.) *Mathematical morphology : from theory to applications*, ISTE-Wiley, 2010.
- [13] Ronse, C., Heijmans, H. « The Algebraic Basis of Mathematical Morphology – Part II: Openings and Closings », *Computer Vision, Graphics and Image Processing* 54/1 (1991), p. 74-97.
- [14] Serra, J. *Image Analysis and Mathematical Morphology*, Academic Press, London, 1982.
- [15] Serra, J. (dir.) *Image Analysis and Mathematical Morphology, Part II : Theoretical Advances*, Academic Press, London, 1988.

Texte édité par Corentin Guichaoua.