



RECONNAISSANCE DES PARTITIONS MUSICALES

R. RANDRIAMAHEFA , J-P COCQUEREZ* ,
F. PEPIN* , S. PHILIPP***

**** Institut National des Jeunes Aveugles, 56 Bd des Invalides 75007 PARIS**

*** E. N. S. E. A., Allée des Chênes Pourpres, 95014 CERGY-PONTOISE CEDEX**

RÉSUMÉ

Dans cet article, nous discutons des problèmes d'interprétation de partitions musicales. Ces problèmes concernent la segmentation (détection et effacement des lignes de portée), la représentation sous forme d'un graphe valué du squelette polygonalisé de l'objet et la reconnaissance proprement dite. Les attributs de ce graphe sont les paramètres géométriques et topologiques des segments associés aux noeuds. L'interprétation de l'image se fait d'une manière hiérarchique. Elle utilise une coopération de plusieurs méthodes: méthode structurelle utilisant le graphe et méthodes utilisant l'image complète de la portée ou de l'objet.

ABSTRACT

In this paper, we describe the different problems of printed music recognition. These problems are the segmentation (detection and elimination of staff lines), the attributed graph representation of the polygonalized skeleton and the recognition proper. The graph attributes are the geometrical and topological characteristics of segments corresponding to the nodes. A hierarchical approach is used in the recognition of music symbols. Different methods are applied: structural method using graph and others methods using the complete image of staff or symbols.

1- INTRODUCTION

La recherche dans le domaine de la lecture optique des partitions musicales commençait vers la fin des années 60 et continue actuellement dans plusieurs pays [1, 2, 3, 4, 5, 6, 7]. A cause de la complexité du problème, certains chercheurs travaillent sur des partitions simples (partition non surchargée, portée à simple voix etc...); les résultats de la reconnaissance sont donc encore limités (certaines catégories de symboles ne sont pas reconnues). Cette réalité nous a poussés à développer un système abordant réellement les problèmes (type de partitions élargi: portée à plusieurs voix, portée avec des accords, symboles collés, etc...).

Ce travail rentre dans le cadre d'un projet de transcription automatique des partitions en Braille musical. Cette chaîne de transcription comprend: la scannérisation, la reconnaissance des partitions musicales, la correction à l'aide d'un éditeur, la traduction en Braille musical et l'impression en Braille.

Comme tout système de reconnaissance, la reconnaissance des partitions musicales comprend la segmentation (détection et effacement des lignes de portée), l'extraction des attributs des objets isolés et le système de décision. A ces trois étapes sont ajoutés l'interprétation globale des résultats et le vérificateur syntaxique utilisant la grammaire musicale.

Dans cet article, nous commençons par exposer les

problèmes de détection et d'effacement des lignes de portée ainsi que les solutions proposées, puis nous discutons de la représentation de l'objet sous forme d'un graphe valué et enfin de la coopération de plusieurs méthodes de reconnaissance.

2- SEGMENTATION

Les partitions sont numérisées par un scanner MICROTEK MS300A avec le logiciel EYESTAR. La résolution choisie est de 300 dpi (118 points par cm). La luminosité et le contraste sont choisis en fonction de la qualité de l'image.

Une partition musicale comprend des portées, chaque portée étant composée de cinq lignes sur lesquelles ou entre lesquelles sont écrits les symboles musicaux. Afin d'isoler ces symboles, il est nécessaire de détecter ces lignes et de les éliminer. La détection de celles-ci nous permet aussi de repérer la position des différents symboles (ex: hauteur des notes) et de trouver la valeur de l'interligne d'une portée (distance moyenne entre deux lignes voisines). Cette valeur de l'interligne sera utilisée pour fixer la valeur des différents seuils.



-2-1- Détection des lignes de portée

Effectuer une projection de l'image sur l'axe vertical est la méthode la plus simple pour détecter ces lignes. Pour des lignes parfaitement horizontales, celles-ci sont représentées par des pics de la projection. Cette méthode ne convient pas pour des partitions réelles avec des lignes rarement horizontales et présentant des faibles courbures, ou certaines parties des lignes sont absentes. Pour ce genre de situation, la projection ne met pas en évidence les pics représentant les lignes de portée.

Ces problèmes nous ont poussés à développer une autre méthode capable de traiter des partitions avec des lignes dégradées, inclinées ou présentant de faibles courbures. Cette méthode est composée de plusieurs étapes:

- la projection de l'image sur l'axe horizontal,
- le filtrage de la projection,
- la recherche des zones de minima locaux de la projection,
- la projection de ces zones sur l'axe vertical,
- la liaison des différents pics appartenant à une même ligne.

Après filtrage, la projection est composée de plusieurs zones de différentes hauteurs et de différentes largeurs. Les zones de minima locaux correspondent aux zones où il n'y a que les lignes de portée. Ces zones sont obtenues en effectuant un double seuillage par rapport à la hauteur suivi d'un seuillage par rapport à la largeur.

Une fois ces zones de minima locaux déterminées, on fixe la largeur de chaque zone à une valeur plus faible afin que les pics obtenus lors de leur projection sur l'axe vertical représentent bien les points des lignes de portée. Puis on effectue la projection de chaque zone sur l'axe vertical pour mettre en évidence une série de pics et on valide les pics suivant leur hauteur et leur largeur.

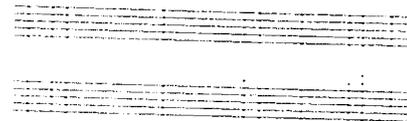
A ce stade, le problème est de combiner les pics constituant la même ligne. Deux méthodes ont été testées pour effectuer cette liaison: la première, utilisant la transformée de HOUGH, a été abandonnée car le résultat était peu satisfaisant. La deuxième solution qui a été retenue utilise la méthode de régression linéaire, en faisant en sorte de tolérer les faibles courbures des lignes de portée. Des tests de validation sont ensuite effectués pour éviter la détection des fausses lignes. Ces tests sont faits sur le nombre de pics appartenant à la ligne, sur la longueur et la pente de la ligne. Cette technique de détection est robuste car même si ces lignes sont discontinues (perte d'informations à la scannérisation), présentent de faibles courbures ou sont fortement inclinées, elles sont toujours détectées (Fig-1a, 1b, 1c).



(a) Partie d'un ensemble de portée



(b) Projection sur l'axe vertical des différentes zones



(c) Lignes détectées

Fig-1 Détection des lignes de portée

2-2- Effacement des lignes de portée

Une fois détectées, l'effacement de ces lignes est effectué en se basant sur leur épaisseur à chaque point. Si cette épaisseur est inférieure à un certain seuil, les pixels de la ligne à cet abscisse sont éliminés sinon ils sont conservés (Fig-2a, 2b). Cette méthode exige une bonne précision des équations de ces lignes. Sa simplicité et sa rapidité compensent le défaut de l'élimination des parties de tangence entraînant une division d'objets. Ce problème sera pris en compte pendant la reconnaissance.



(a) Image avant effacement des lignes



(b) Lignes effacées

Fig-2 Elimination des lignes de portée

3- GRAPHE VALUE

Après l'effacement des lignes de portée, les symboles musicaux sont isolés. Une représentation minimale de chaque objet appelée graphe valué est ensuite construite. Ce graphe nous permet d'appliquer une méthode structurale pour la reconnaissance.

La construction du graphe comprend:

- la squelettisation et la polygonalisation,
- l'attribution à chaque segment de la distance minimale de ses points au contour,
- la transformation du graphe.

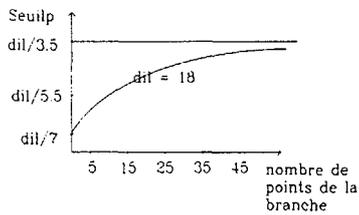
3-1- Squelettisation et Polygonalisation

La squelettisation a été effectuée par l'algorithme de Zhang et Suen amélioré par Holt [8]. Cet algorithme a été choisi pour son insensibilité aux variations brusques du contour.

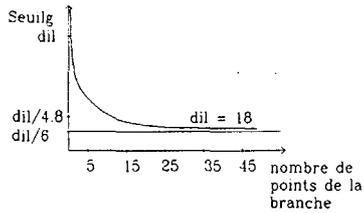
La polygonalisation consiste à représenter chaque branche du squelette par un ou plusieurs segments. Une branche est définie entre deux points particuliers (points pendants ou points multiples). La polygonalisation comprend:

- la recherche des points particuliers,
- le suivi des branches entre deux points particuliers,
- la polygonalisation des branches par l'algorithme de la corde.

Le seuil utilisé par cet algorithme est une fonction exponentielle de l'interligne et du nombre de points de la branche. Cette fonction est différente pour les objets de petite et de grande taille afin de garder la morphologie de l'objet initial (Fig-3).



(a) Pour les objets de petite taille



(b) Pour les objets de grande taille

Fig-3 Courbes des seuils

Après polygonalisation du squelette, nous attribuons à chaque segment la valeur de sa distance au contour.

3-2- Distance du segment au contour

Ce paramètre de distance est utilisé pour classer les segments en deux catégories. Les segments appartenant aux têtes de note (ellipse pleine) ont une distance assez importante, par contre les autres segments ont une distance assez faible.

Ce paramètre est calculé à partir des distances attribuées à chaque point du segment. Pendant la squelettisation, une distance au contour est attribuée à chaque point du squelette (égale au rang d'itérations à partir de laquelle le pixel est gardé).

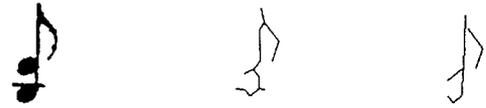
Nous avons constaté que la variation de la courbe des distances est importante pour les segments appartenant aux têtes de notes et constante pour les autres. Notre but est donc d'attribuer la valeur maximale des distances de ses points aux segments de tête de note et la valeur minimale aux autres. La moyenne des pentes des cordes entre une extrémité et un point courant est utilisée pour évaluer cette variation.

À la fin de la polygonalisation, certaines configurations du graphe sont transformées: certains segments sont éliminés ou certains segments voisins sont fusionnés afin d'obtenir une

structure simple (chaque queue, chaque tête et chaque barre de note sont composées d'un seul segment). Ces transformations utilisent les attributs géométriques et topologiques du graphe de l'objet. Quelques exemples sont présentés sur les figures suivantes:



(a) Représentation de la queue par un seul segment



(b) Représentation d'une note par un seul segment

Fig-4 Transformation du graphe

À la fin de ces transformations, nous obtenons un graphe valué dont les noeuds correspondent aux segments et dont les arcs représentent les connexions entre segments (Fig-5). À chaque noeud sont associées la valeur de la distance du segment au contour et les coordonnées des extrémités du segment.

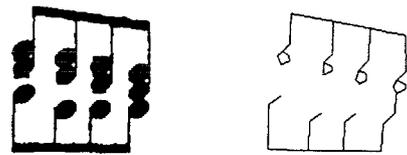


Fig-5 Graphe d'un objet

4- RECONNAISSANCE

Une première classification est effectuée en considérant le paramètre "distance au contour du segment". Dans le graphe des objets ayant une ellipse pleine il y a au moins un noeud dont ce paramètre dépasse un seuil proportionnel à l'interligne. Cette première classification nous permet de distinguer les objets ayant une ellipse pleine (noire, croche, demi-croche, etc...) des autres (blanche, dièse, bécasse, clef, etc...).

4-1- Reconnaissance des symboles ayant une ellipse pleine

La reconnaissance d'une note (ellipse pleine) se résume ainsi:

- détection de la note,
- détermination du sens de la queue,
- recherche de la queue,
- détermination du nombre de crochets,
- détermination de la hauteur de la note et détermination de la portée.

La détection des notes est différente pour une tête de note simple et pour un groupe de têtes (accord).

Pour le premier cas, la tête de note est composée d'un seul segment et le graphe suffit pour la reconnaître. Par contre, pour le second, l'ensemble des têtes de notes est représenté soit



par un circuit soit par plusieurs noeuds. La reconnaissance utilisant uniquement le graphe ne suffit pas d'où l'application d'une méthode de corrélation. Un masque d'ellipse pleine est balayé dans la région de l'objet définie par les groupes de têtes de notes. Des nuages de points apparaissent et ceux-ci représentent les têtes de note.

La position de la note par rapport au segment vertical voisin détermine le sens de la queue. La queue de la note est composée d'un ou de plusieurs segments verticaux ayant une faible distance au contour. Les crochets sont des segments horizontaux de faible distance au contour, voisins de la queue et situés au bout de la queue. La position de la note est déterminée: pour une note simple, par le point du segment possédant la valeur maximale de distance au contour; pour un groupe de têtes de notes, par le centre de gravité du nuage de points. Avec cette valeur, on détermine la hauteur de la note et la portée à laquelle elle appartient.

4-2- Reconnaissance des blanches, des barres de mesure, des barres de reprise, des barres terminales et des liaisons

Pendant l'effacement des lignes de portée, certaines parties des têtes de blanches sont éliminées. Ainsi, les têtes de note ne figurent plus dans le squelette. La reconnaissance des blanches ou des groupes de blanches se résume ainsi:

- détection de la queue de note,
- corrélation de l'image avec un masque d'ellipse creuse,
- recherche des crochets,
- détermination de la hauteur de la note et détermination

de la portée.

La queue de la note est un segment suffisamment long et d'une faible distance au contour. Comme pour les notes ayant des ellipses pleines, les crochets sont des voisins de la queue et la position de la note est déterminée par le centre de gravité du nuage de points.

La reconnaissance des barres de mesure, des barres de reprise et des barres terminales est effectuée à partir du graphe. Ces barres sont composées de segments verticaux. L'interprétation de l'ensemble du résultat détermine la nature exacte de la barre.

Une liaison est composée d'un ou plusieurs segments longs et presque horizontaux. Le graphe est suffisant pour l'identifier.

5- CONCLUSION

Les équations des lignes de portée déterminée par la technique utilisée sont précises. La valeur de l'interligne utilisée pendant tout le processus, le résultat de l'effacement des lignes de portée sont étroitement liés à cette précision. Le taux de reconnaissance obtenu sur quelques tests est voisin de 100%. Le résultat de reconnaissance est lié à la qualité de l'image (objets divisés ou collés).

Références

- [1] H.Kato and S.Inokuchi, "A Recognition System for Printed Piano Music Using Musical Knowledge and Constraints," Proceedings of the International Association for Pattern Recognition workshop on Syntactic and Structural Pattern Recognition, Murray Hill, New Jersey USA, June 13-15, 1990.
- [2] P.Martin and C.Bellissant, "Low-Level Analysis of Music Drawings Images," Proceedings of the First International Conference on Document Analysis and Recognition, Saint-Malo, France 1991.
- [3] N.P.Carter and R.A.Bacon, "Automatic Recognition of Printed Music," Proceedings of the International Association for Pattern Recognition workshop on Syntactic and Structural Pattern Recognition, Murray Hill, New Jersey USA, June 13-15, 1990.
- [4] T.Matsushima, T.Harada, I.Sonomoto, K.Kanamori, A.Uesugi, Y. Nimura, S.Hashimoto and S.Ohteru, "Automated Recognition System for Musical Score - The Vision System Wabot-2," Bulletin of Science and Engineering Research Laboratory, Waseda University, 112, September 1985, pp. 25-52.
- [5] A.T.Clarke, B.M.Brown and M.P.Thorne, "Using a Micro to Automate Data Acquisition in Music Publishing," Microprocessing and Microprogramming 24 (1988) 549-554.
- [6] I.Fujinaga, B.Pennycook and B.Alphonse and K.Hogan, "Optical Music Recognition: Progress Report," Proceedings of the I.C.M.C pp.66-73 (1991?).
- [7] D. Prerau, "Computer Pattern Recognition of Printed Music," Proceedings of the Fall Joint Computer Conference 1971, Montvale, NJ 39, pp. 153-162, AFIPS Press.
- [8] C.M.Holt, A.Stewart, M.Clint & R.H.Perrott, "An improved parallel thinning algorithm," Communications of the ACM, Vol 30, N° 2, Fev. 1987.