

**Laurent Pugin**

**Lecture et traitement informatique de typographies musicales anciennes  
Un logiciel de reconnaissance de partitions par modèles de Markov cachés**

Thèse de Doctorat en Musicologie, Université de Genève, Faculté des Lettres, 2006

Composition du jury

Professeur Etienne DARBELLAY, directeur de thèse

Professeur Brenno BOCCADORO, président du jury

Professeur Eric WEHRLI, doyen de la Faculté des Lettres

Professeure Eleanor SELFRIDGE-FIELD (Stanford University, USA)

Professeur Ichiro FUJINAGA (McGill University, Montreal, Canada)

Professeur Christian PELLEGRINI (Faculté des Sciences, Laboratoire d'Intelligence Artificielle)

**Description du travail**

Ce travail de thèse traite de la lecture et du traitement informatique d'imprimés musicaux en typographie des 16<sup>ème</sup> et 17<sup>ème</sup> siècles. A la base de cette recherche sont différents problèmes de philologie musicale liés aux sources typographiées de cette époque, et l'objectif principal de ce travail était d'expérimenter une solution qui permette à terme d'assister par ordinateur le traitement de ces sources. Cette thèse traite donc d'un problème de reconnaissance optique de documents, avec à la fois un objectif pratique précis et un type de document particulier. Elle constitue par essence un travail interdisciplinaire avec d'un côté des questions d'ordre musicologique pouvant être liées à la notation musicale, à la philologie musicale ou encore à l'histoire de l'imprimerie, et d'un autre côté, des questions liées à la reconnaissance optique de documents, celles-ci pouvant concerner le traitement d'image, l'intelligence artificielle ou la représentation digitale d'information musicale.

Cette thèse couvre un champ de recherche nouveau en reconnaissance optique de musique. Les caractéristiques particulières que les documents considérés présentent tant au niveau du contenu que du support ont exigé une approche différente de celle habituellement utilisée. Une des difficultés du problème été d'obtenir une vision globale de ces caractéristiques de manière à pouvoir définir une solution idoine. Cette tâche a pu être accomplie en réunissant un corpus d'imprimés recouvrant uniformément la période la plus importante de la typographie musicale (env. 1550 à 1700). Au total, 180 pages de musique ont été sélectionnées garantissant une répartition et une distribution représentatives et homogènes des imprimeurs, des répertoires (profane ou religieux et vocal ou instrumental), des systèmes de notation (parties séparées, partitions, etc.) ainsi que des formats d'impression. Ce corpus a permis de mettre en perspectives différents problèmes liés à l'évolution de la notation musicale en relation avec les contraintes imposées par la typographie.

Concernant le support, différents éléments ont du être pris en considération. La qualité des imprimés est très variable, notamment à cause de l'usure des fontes utilisées, des irrégularités d'impression ou de la qualité médiocre du papier. Ces imprimés sont très souvent fortement dégradés, avec de nombreuses transparences, des taches, voire des déchirures ou des trous. L'accès aux sources est limité dans la mesure où elles sont disponibles uniquement sous forme de microfilms ou d'images déjà numérisées, ce qui limite considérablement les possibilités de contrôler la phase d'acquisition. Pour ces différentes raisons, la mise en place d'une phase de pré-traitement spécifique a été nécessaire.

Une des originalités majeures de cette thèse est l'utilisation de *Hidden Markov Models* (HMM) pour la reconnaissance du texte musical proprement dit. Les HMMs constituent une technique très performante pour modéliser des séquences continues et ont été utilisés jusque là avec beaucoup de succès en bio-informatique et pour la reconnaissance de la parole ou de l'écriture manuscrite, mais jamais en reconnaissance optique de

musique. Cette thèse a permis de mettre au point une approche originale où les HMMs constituent un modèle typographique de la fonte utilisée, avec une correspondance directe entre les symboles à reconnaître et les caractères typographiques de la fonte. Grâce à l'utilisation des HMMs, il a été possible de mettre au point une solution qui permette une segmentation implicite de la portée sans avoir au préalable à supprimer les lignes de portée. Avec un classifieur comme un réseau de neurones, cette tâche est indispensable et reste toujours un point délicat dans les systèmes de reconnaissance de musique habituels. Elle l'eût été encore plus avec les imprimés anciens que nous considérons ici. Les lignes de portées sont traitées implicitement dans l'analyse du signal de même que les courbures que peuvent présenter les portées. Cette solution se distingue clairement de celle utilisée dans tous les travaux de reconnaissance optique de musique effectués jusqu'à ce jour.

Finalement, les taux de reconnaissance ont pu être augmentés de façon significative grâce à l'invention d'un modèle musical stochastique spécifique. Ce modèle par classe est inspiré des modèles *n-grams* utilisés dans la reconnaissance du langage (parole ou texte). Une attention toute particulière dans la définition des classes a été donnée à la distinction entre les éléments de nature typographique, de nature notationnelle et de nature musicale. Quatre sous-modèles distincts ont été définis : le modèles des symboles, le modèles des durées, le modèles des hauteurs et un modèle particulier inter-symboles, le modèle des intervalles. Cette partie de la recherche ouvre de nouvelles perspectives pour l'intégration de connaissance de plus haut niveau en reconnaissance optique de musique.

En plus des recherches historique et théorique que constitue cette thèse dans le domaine, elle a aussi donné lieu à un logiciel complet, Aruspix (<http://www.aruspix.net>), développé en lien avec la recherche. Ce logiciel a permis de tester les différents algorithmes et de générer les données pour l'entraînement des modèles typographiques (HMMs) et la réalisation des modèles musicaux. Il inclut le pré-traitement, la pré-classification et la reconnaissance mais aussi un éditeur complet permettant de corriger les erreurs dans les résultats. D'une certaine manière, ce logiciel est le résultat de la thèse et en même temps l'outil qui a permis de perfectionner la recherche.