

# Rapport de l'analyse d'un sujet spécifique de l'intelligence artificielle

Wenhao LUO

November 24, 2022

code permanent : **LUOW09129805**

# 1 Introduction

Ce document est le résumé sur une analyse d'un sujet spécifique de l'intelligence artificielle, demandé par le cours d'intelligence artificielle de l'UQAC (8INF878). Le sujet choisi pour cette analyse est "reconnaissance optique de musique", cette analyse est basée originalement sur un état de l'art écrit[2]. Au cours de la lecture, on trouve que ce sujet est très vaste, donc après une introduction générale, on se concentre sur une problématique spécifique, à l'aide d'un autre article[1].

## 2 Présentation du sujet

La reconnaissance optique de musique, ou *optical music recognition* en anglais, est une recherche sur les approches informatiques pour la lecture des documents musicaux (partitions). C'est un sujet qui est recherché depuis plus que 50 ans. Elle est considérée originalement comme une extension de la reconnaissance optique de texte ; pourtant l'article étudié nous rappelle que une partition porte les informations de 2D, qui est plus compliqué par rapport un document textuel, qui porte les informations de 1D.

Les recherches étudient principalement avec les notations *Common Western Music Notation*, qui sont les notations utilisées couramment. Un pipeline de traitement a été établi basé sur ces notations avec l'ordre suivante :

- prétraitement d'image ;
- détection des symboles musicaux ;
- reconstruction des notations ;
- encodage et archivage.

On se concentre sur un des traitements dans la partie détection. C'est la détection de *staff*<sup>1</sup>. Ce sont les lignes horizontales dans des partitions qui permettent de fixer la position des autres symboles. L'article [1] présente une solution nouvelle, nommée *stable paths*.

## 3 Analyse de l'article

### 3.1 Présentation sur les algorithmes précédents

Cet article commence par une introduction, dans laquelle il a donné un état de l'art bref, ensuite une présentation sur les algorithmes utilisés pour cette problématique. Les solutions comme :

- utiliser une projection horizontale avec une recherche de minima locale ;

---

<sup>1</sup>Une traduction en français doit être "la portée", mais je ne suis pas sûr.

Les solutions précédentes sont mentionnées comme des références, et elles sont également utilisées pour montrer la différence entre la solution proposée dans cet article.

- *mathematical morphology algorithms* ;
- *rule-based classification*.

## 3.2 Définitions conventionnelles

### 3.2.1 Graphe dans une partition

Cet article a commencé par proposer son algorithme par définir un graphe dans une partition musical. On considère que une partition scannée est un ensemble des pixels, et on traite les pixels comme des nœuds d'un graphe. Ensuite les arêtes sont définies : ce sont les connexions entre des nœuds et ses voisins. En fonction de la position relative des deux nœuds connectés, chaque arête est affectée une valeur. C'est le coût, ou la distance entre un pixel/nœud et un autre.

### 3.2.2 Définition d'un chemin

Ensuite, un chemin est défini comme une suite des nœuds connectés, et la distance de ce chemin égale à la somme de la valeur affectée sur chacune des arêtes entre les nœuds voisins.

Les chercheurs ont ensuite élargi la définition de distance, et ils ont l'appliquée entre deux ensembles des pixels. On peut donc définir la distance entre deux régions sur une partition.

## 3.3 Problématique proposée

Sous l'hypothèse que une ligne du *staff* doit être continue et connecté entre la partie gauche et la partie droite d'une partition avec un ligne quasiment droite, le problème donc se transforme en cherchant des chemins plus courts dans un graphe. Un chemin est "stable" entre deux régions s'il est le chemin plus court entre chaque élément d'une région et l'autre. L'enjeu de cet algorithme est donc retrouver toutes les chemins stables de ces deux parties. Selon le modèle simplifié que cet article a fourni, toutes les chemins plus courts sont donc le *staff*.

Ici, cet élargissement de la définition est quelque chose typique dans l'analyse mathématique. Je n'explique pas détaillément car il semble assez évident.

## 3.4 Algorithme proposé pour le calcul

Le algorithme présenté dans l'article est le suivant<sup>2</sup> :

Preprocessing:

```
compute staffspaceheight and stafflineheight
compute weights of the graph
```

Main Cycle:

```
compute stable paths
validate paths with blackness and shape
erase valid paths from image
```

---

<sup>2</sup>C'est la description origine dans cet article.

```
    add valid paths to list of stafflines
    end of cycle if no valid path was found
Postprocessing:
    uncross stafflines
    organize stafflines in staves
    smooth and trim stafflines
```

### 3.5 Résultat présenté

Cet article a montré également le résultat en appliquant cet algorithme proposé. Selon les data, il atteint un taux de false positive de 1.3% et un taux de false négative à 1.4%.

## 4 Conclusion

Dans ce rapport, j'ai présenté un domaine de recherche lié avec l'intelligence artificielle ; c'est la reconnaissance optique de musique. J'ai étudié un état de l'art ainsi un algorithme spécifique de cette problématique en lisant deux articles[1][2]. Ensuite j'ai présenté l'algorithme utilisé dans l'article [1].

Personnellement je trouve que le domaine de recherche est très intéressant, et j'apprécie l'algorithme proposé par cet article, qui est concis et efficace ; la créativité de traduire le problème d'un domaine appliqué vers un autre plus théorique est très impressionnante.

## References

- [1] Jaime dos Santos Cardoso, Artur Capela, Ana Rebelo, Carlos Guedes, and Joaquim Pinto da Costa. Staff detection with stable paths. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(6):1134–1139, 2009.
- [2] Elona Shatri and György Fazekas. Optical music recognition: State of the art and major challenges. *arXiv preprint arXiv:2006.07885*, 2020.