

Optical Music Sheet Segmentation

P. Bellini, I. Bruno, P. Nesi

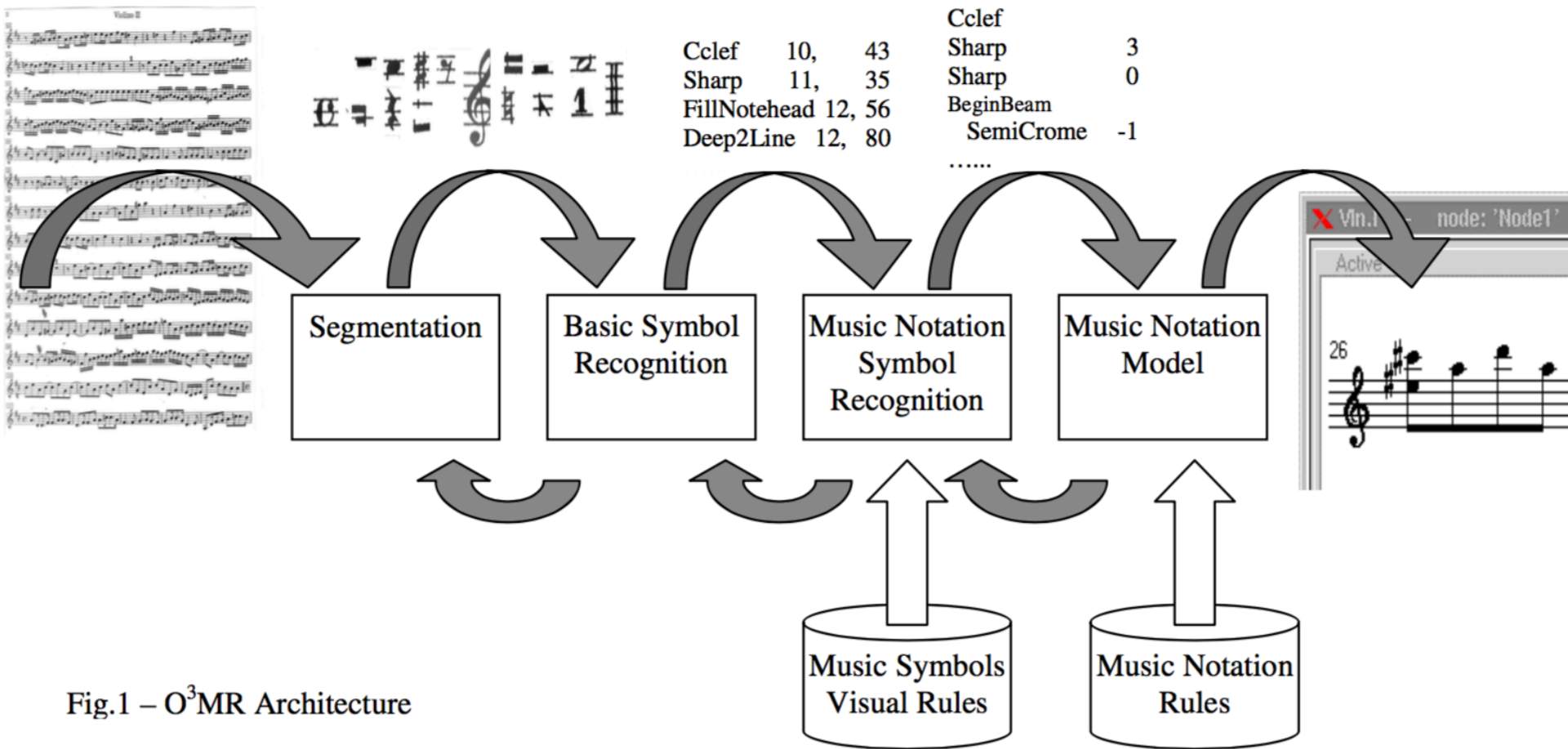


Fig.1 – O³MR Architecture

Referencial Teórico:

- Histograma
- Filtros passa alta
- Segmentação

Introdução

- Esse artigo apresentará o sistema O³MR (Object Oriented Optical Music Recognition)
- O método apresentado no artigo é baseado na adoção de projeções para a extração de símbolos básicos que constituem o elemento gráfico da nota musical.

- Os sistemas de reconhecimento de partituras musicais, comumente chamados de OMR(Optical Music Recognition), estão estritamente ligados com sistemas de reconhecimento de caracteres (OCR - Optical Character Recognition).
- Técnicas de OCR não podem ser usadas em reconhecimento de notas musicais pois essas são apresentadas numa forma bidimensional.

- O artigo em questão propõe o reconhecimento de notas musicais a partir da decomposição de símbolos compostos em básicos. A partir da segmentação da partitura, podemos reconhecer esses símbolos básicos.

Arquitetura

- A arquitetura O³MRé baseado em 4 componentes principais:

Segmentação:

- A partitura é processada com o a finalidade de extrair os símbolos básicos e suas posições. Do ponto de vista do autor, os símbolos básicos são símbolos elementares de construção das notas musicais.

- **Reconhecimento dos símbolos básicos:**
- Esse modulo reconhecerá os símbolos básicos usando uma rede neural. Como entrada ela receberá os segmentos de imagens normalizados, então uma rede neural fará o reconhecimento. O retorno desse módulo é simbólico, para cada símbolo reconhecido é retornado as coordenadas do segmento de imagens e um valor de confiança de reconhecimento é gerado.

- **Reconhecimento da notação dos símbolos musicais:**

- Símbolos menores são mapeados para símbolos mais complexos. O critério de decisão é baseado no contexto de reconhecimento: reconhecimento dos símbolos básicos e suas posições em relação as linhas horizontais, valor de confiança de reconhecimento etc.

Foi formalizado o **conjunto de regras visuais para os símbolos musicais**. De acordo com esse processo, para cada símbolo básico um conjunto de símbolos elementares é atribuído.

- **Modelo de refinamento da notação musical:**

- Uma vez que os símbolos básicos eles são serão usados como peças para construir os símbolos mais complexos, usando as Regras de Notação Musicais.

Modulo da Segmentação

- A segmentação será um processo complexo devido ao fato da quantidade de notas musicais e as variadas combinações complexas que elas podem ter.
- A partitura será recursivamente decomposta em blocos menores definido por um conjunto de linhas horizontais e verticais que permitirão o isolamento e extração dos símbolos básicos.

- Level 0:
- A partitura é segmentada para extrair subimagens que incluem cada staff. Os dois objetivos desse primeiro level são a calibração dos segmentos processados e a detecção dos segmentos da imagem que contenham as linhas horizontais.
- - Calibração dos segmentos processados:
É realizado para obter parâmetros como a grossura n das linhas horizontais e a distância d entre essas linhas.

- Para estimar esses valores, a partitura é processada coluna por coluna para gerar dois histogramas. No eixo X temos o número de pixels e no Y o valor do número de ocorrências.
- Para controlar a variância e o ruído são usado os intervalos entre o valor máximo e o mínimo dessas ocorrências

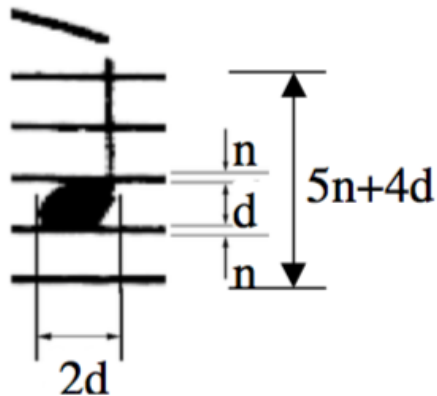


Fig.5 – Tuning

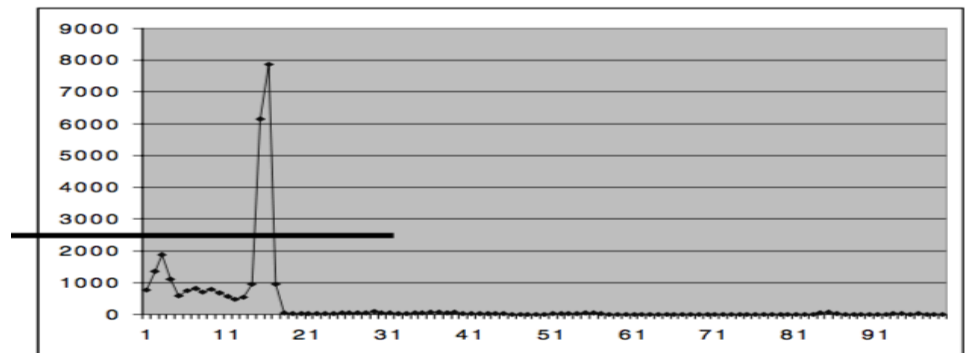


Fig. 6a – Distance between two staff lines, d

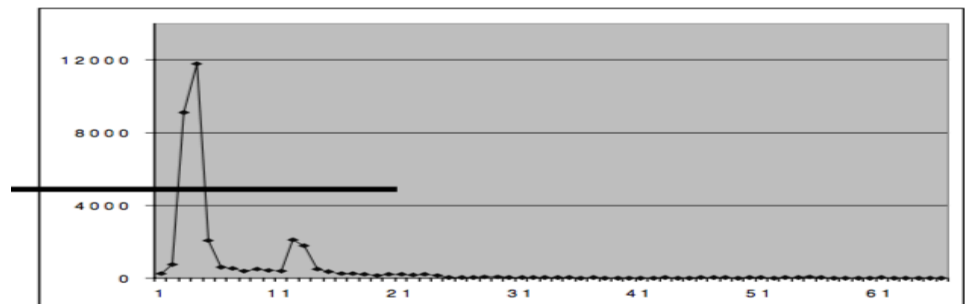


Fig. 6b – Thickness of staff line, n

- - Detecção dos segmentos da imagem que contenham as linhas horizontais:

O objetivo dessa fase é identificar a área retangular que as linhas horizontais estão localizadas, aplicando uma projeção em Y a uma porção da imagem vai apresentar um padrão regular da estrutura onde outras projeções terão um padrão variável .

Level 1

- O level 1 funciona com os segmentos de imagens produzidos no level 0.
- O objetivo desse level é a extração dos segmentos verticais da imagens que contém símbolos musicais.
- Nesse level será produzido as sub segmentos de imagens em 3 fases de detecção.

- Grupos e detecção dos símbolos símbolos isolados:
 - Nesse estágio de detecção é considerado uma janela que percorre os retangulos de imagens para detectar grupos de figuras e símbolos individuais. Ao final desse processo, é produzido o valor binário da função F.



Fig.9 – Symbol segment detection, function F

- Detecção e classificação da cabeça de nota:
 - O objetivo dessa fase é particionar os segmentos de imagens classificados com $F=1$ para procurar as cabeças de notas.
 - Uma janela que percorre a imagem detecta as cabeças de nota, que pode ser detectada pelo cálculo da projeção em Y .
 - A largura da nota (H) pode ser detectada com base no intervalo $[2n_2, 2n_2+d_1]$.

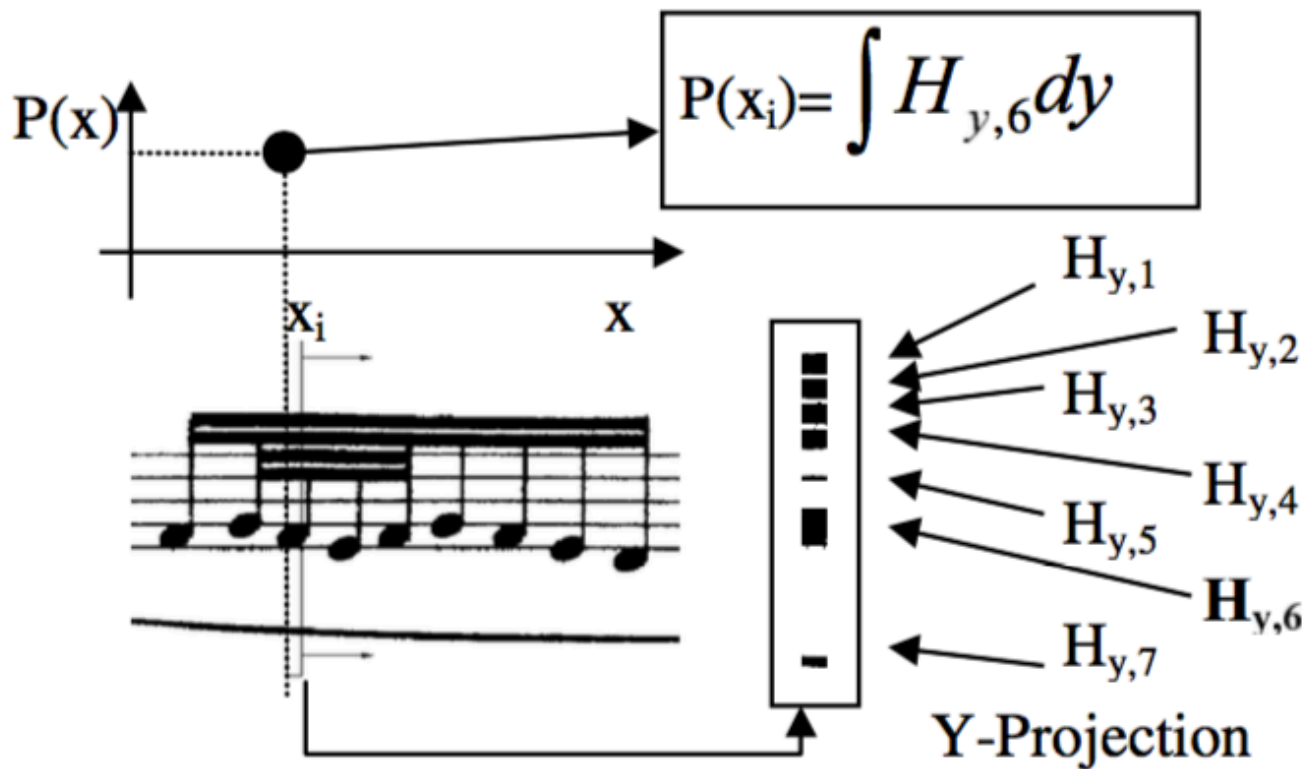


Fig.10 – X-Projection building

Level 2

- Nessa ultima fase da segmentação as imagens vindas do processo anterior serão decompostas em conjuntos de simbolos básicos.
- Nesse estágio dois métodos diferentes serão adotados para os segmentos de imagens. Os que contém cabeça de notas e os que nao contém.

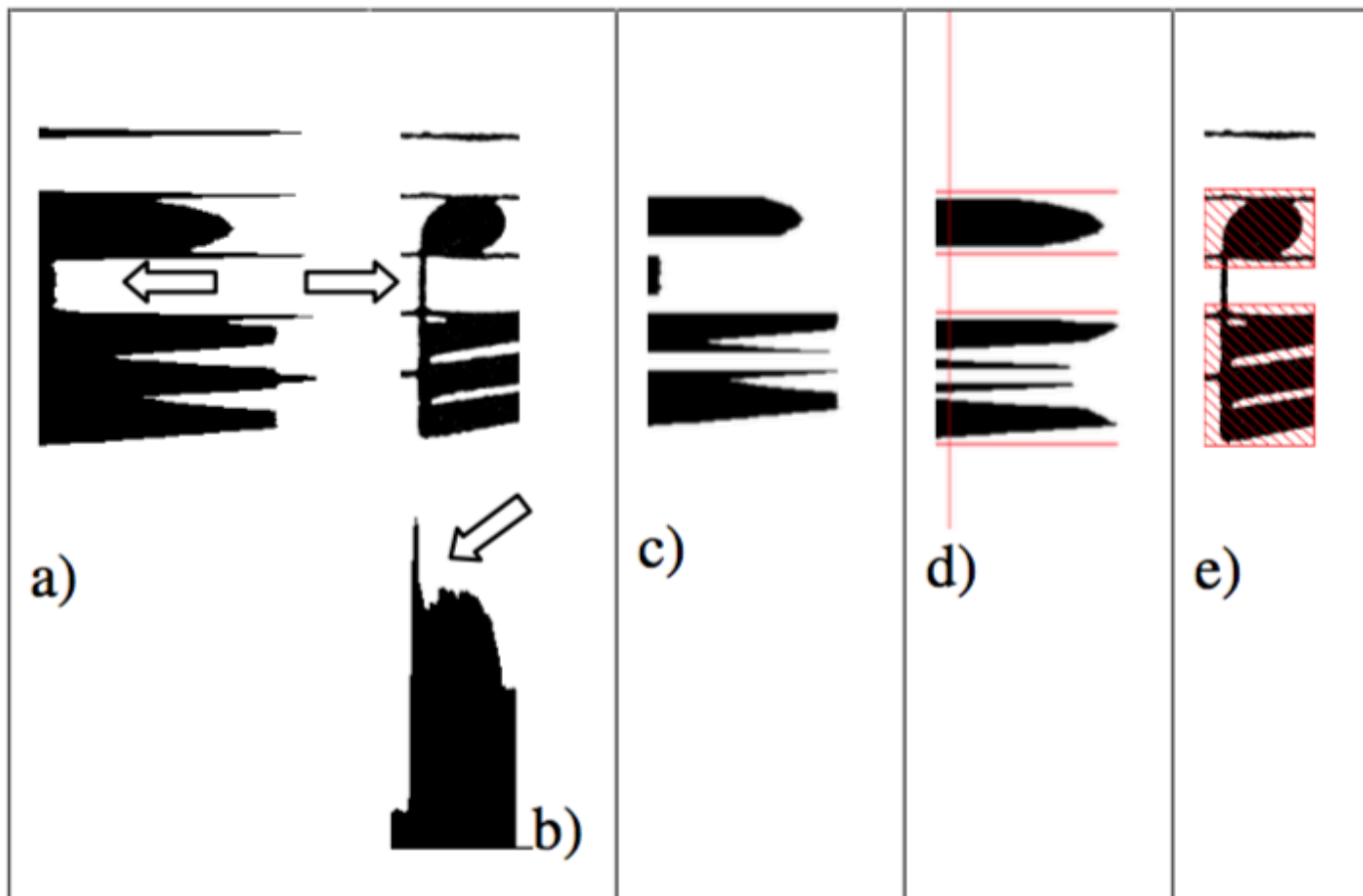


Fig. 13 – a) Y-projection of 32-th with beams. Arrows indicate the offset due to the stem; b) X-projection: the arrow indicates the spike due to the stem; c) Y-projection of the image without staff lines; d) Y-projection after staff lines removal, filtering and extraction points computation; e) identification of basic symbol after extraction point computation

Resultados

- Num teste feito com imagens adquiridas a 300dpi a fase de calibração retornará com os parametros: $[n1, n2] = [2, 4]$ e $[d1, d2] = [17, 18]$ pixels.



Fig. 14 – Music score, Image segments slicing

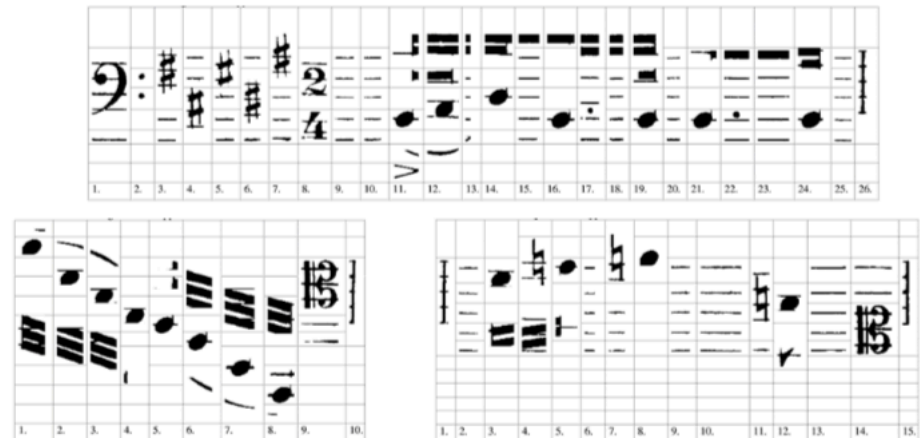


Fig. 15 – Decomposition in basic symbols