

Extracción de líneas melódicas a partir de imágenes de partituras musicales

Ángel Sánchez, Juan J. Pantrigo, and José Ignacio Pérez

Universidad Rey Juan Carlos, Madrid, Spain
{angel.sanchez, juanjose.pantrigo, joseignacio.perez}@urjc.es

Resumen El objetivo de este trabajo ha sido el desarrollo de una herramienta automática para la segmentación de partituras musicales no manuscritas, que permite extraer pentagramas, compases y duración y tono de las notas incluidas en ella. Una vez procesada la imagen, un segundo objetivo es obtener líneas melódicas y, en ellas, ciertos patrones musicales proporcionados por el usuario. El método elegido para el análisis se ha derivado de diferentes propuestas encontrados en el estado del arte de los sistemas OMR (*Optical Music Recognition*) combinadas con aportaciones propias de los autores. Los experimentos se han realizado usando diversas piezas de música de contrapunto (invenciones) de J. S. Bach.

1. Introducción

La música computacional (*computer music*) es un área interdisciplinar de creciente interés. Su ámbito de aplicación engloba tareas tan diversas como reconocimiento de estilos musicales [2], clasificación de piezas por estilos musicales [1], reconocimiento de notas en partituras musicales [3], búsqueda de similitudes [4], estudios relacionados con el ritmo [5], generación automática de piezas musicales [6] o reconocimiento de elementos en partituras musicales (OMR) [7]. La mayor parte de los sistemas que utilizan una representación simbólica de la imagen escaneada de una partitura musical coinciden en una serie de pasos comunes. Éstos comprenden: detección de las líneas del pentagrama, segmentación de elementos musicales y análisis de la imagen segmentada. Estas son las etapas típicas de un sistema de tratamiento digital de imágenes. La mayoría de los autores considera la detección de las líneas de los pentagramas como el paso preliminar. En primer lugar, se analizan los ángulos de inclinación de los pentagramas para rotar las partituras en caso de que sea necesario. Un gran número de métodos se opta por identificar las líneas, empleando la transformada de Hough en la mayoría de los casos (con el consiguiente consumo de recursos debido a la complejidad de dicho algoritmo función), para obtener las ecuaciones de las rectas de la imagen, y posteriormente eliminarlas. Otro método consiste en el análisis de la proyección horizontal de la imagen. El histograma resultante presenta picos debidos a la concentración de píxeles que conforman las líneas del pentagrama.

Una vez eliminadas las líneas de los pentagramas, en la partitura permanecen símbolos aislados que han de ser identificados. Estos símbolos pueden ser notas musicales, caracteres (generalmente numéricos y para cuyo reconocimiento se puede hacer uso de un motor OCR, *Optical Character Recognition*), texto o elementos ornamentales de la partitura. Puesto que, generalmente, sólo interesa la información musical contenida en el pentagrama, se procura eliminar los caracteres y otros elementos. Es en esta etapa donde más diversidad de métodos existe. La mayor parte de éstos optan por realizar una búsqueda de las componentes conexas para segmentar y buscar patrones en partituras musicales [15], usando principalmente grafos de adyacencias de líneas (LAG). Una vez elaborado el grafo se procede al estudio de cada una de las componentes conexas para caracterizarlas. Otra alternativa es el uso de plantillas, que es el método propuesto en este trabajo, debido a su mayor simplicidad y los buenos resultados que proporciona. En algunos casos, se emplean conjuntamente ambos métodos. El formato de las partituras se tiene en cuenta para la identificación de la información. Hay algunas características comunes en toda partitura, como por ejemplo, la ubicación de las claves, que siempre aparecen en el lado izquierdo del primer compás de un pentagrama. El uso de bases del conocimiento musical simplifica el proceso de segmentación. Finalmente, una vez identificados los elementos que componen la partitura, la siguiente etapa consiste en asignar una notación determinada que permita establecer la relación entre la partitura original y la representación simbólica de la misma.

2. Metodología propuesta

En esta sección se describe la metodología desarrollada para la segmentación y análisis de partituras musicales. En ella se pueden distinguir cuatro etapas fundamentales: preproceso de la imagen, identificación de los compases, búsqueda de notas y determinación de su tono y duración.

2.1. Preproceso de la imagen

En esta etapa se acometen dos tareas fundamentalmente: binarización de la imagen de la partitura y eliminación de información innecesaria para su posterior análisis. Para convertir una imagen en niveles de gris (256 niveles de gris) en una imagen binaria, se calcula el umbral a partir del cuál un valor de gris será considerado como blanco o negro en la representación binaria. El cálculo de dicho umbral se realiza empleando el método de Otsu [8], que elige un umbral para reducir al mínimo la variación de los rangos en los que un pixel será considerado como negro o blanco en la imagen binaria resultante. Una vez calculado dicho umbral y transformada la imagen original a su representación binaria, se elimina la información innecesaria que aparece en la partitura. Se considera información irrelevante el número de página y el título de la pieza musical. Para localizar estas regiones de la imagen se ha optado por caracterizar la morfología del histograma de la proyección horizontal de toda la partitura.

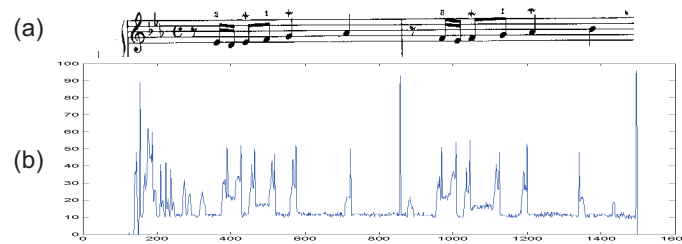


Figura 1. Etapa de identificación de los compases: (a) imagen original del pentagrama y (b) histograma de proyección vertical.

2.2. Identificación de los compases

Esto se consigue mediante sucesivos refinamientos en el análisis de la imagen binaria resultante del preproceso. En primer lugar, se detectan los sistemas es decir, grupos de pares de pentagramas, puesto que se analizan piezas escritas para piano. Para ello, nos basamos nuevamente en la estructura del histograma de la proyección horizontal de la partitura. Una vez obtenidas las coordenadas de los sistemas, se prosigue en la tarea de detección de los pentagramas, utilizando nuevamente el histograma horizontal de la partitura. Existe una pequeña zona de separación entre un pentagrama y el siguiente de cada par considerado. Basándonos en dicha separación, pueden establecerse las dimensiones de la ventana que delimita a un determinado pentagrama.

Por último, se procede a la extracción de los compases de cada pentagrama. En una partitura, los compases vienen separados unos de otros mediante unas líneas verticales, que producen un pico en un histograma vertical del pentagrama. En la Figura 1, es posible identificar las barras de separación de los compases de un pentagrama a través de los picos que producen en el histograma.

2.3. Búsqueda de notas

Esta fase es la más compleja del proceso, debido a la exhaustividad de las exploraciones que se realizan sobre cada uno de los compases en la búsqueda de las notas musicales. El procedimiento desarrollado traza, en primer lugar, una serie de “pistas virtuales” en el compás, que recorren horizontalmente las líneas y los espacios del pentagrama. Sobre dichas pistas, se desplaza posteriormente una máscara de la cabeza de una nota musical. El uso de dichas pistas se justifica por los siguientes motivos:

1. Eficiencia: en lugar de recorrer una a una las posiciones de la matriz que contiene la imagen correspondiente al compás, sólo se buscarán las notas en unas determinadas filas en las que la probabilidad de hallar una nota es mayor, puesto que la disposición de las notas en un compás está limitada.



Figura 2. Etapa de identificación de las notas: (a) imagen original del pentagrama, (b) Posición de las pistas virtuales, y (c) resultado de la detección de notas (se marca coloreada la cabeza de las notas).

2. Efectividad: el uso de las pistas virtuales permite asignar el tono de una nota de modo sencillo.

Para elegir estas pistas, se utiliza el histograma de la proyección horizontal del compás de la Figura 2(a). A continuación, se trazan las pistas virtuales. Además se estiman unas líneas intermedias, que recorren los espacios. El resultado se muestra en la Figura 2(b). La siguiente tarea consiste en recorrer cada una de las pistas con la máscara, cuyo tamaño se determina experimentalmente. Durante este desplazamiento, se compara la región de la imagen de tamaño igual a la máscara elegida que rodea al punto en el que se desea saber si hay una nota, y la máscara de la cabeza de una nota musical. Se estudia el índice de solapamiento entre ambas regiones, y si supera un cierto umbral, se considerará candidata a nota musical.

2.4. Determinación del tono y duración de las notas

La siguiente tarea consiste en asignar un valor que identifique el tono y la duración de la nota. Para asignar estos valores se ha elegido la nomenclatura propuesta en [9]. En esta referencia, se proponen diferentes notaciones, de las cuales se han adoptado las notaciones implícita (*splitted implicit notation*) y diferencial (*splitted differential notation*). En la notación implícita se representa el valor de la escala de tonos mediante un valor numérico. La diferencia entre un tono y el siguiente se traduce en un incremento numérico de dos unidades, a excepción de los pasos de *Mi* a *Fa* y de *Si* a *Do*, entre los que sólo existe un semitono de diferencia, lo cual implica un incremento de una unidad.

La notación diferencial se basa en la notación implícita. Consiste en calcular la diferencia entre el tono de una nota y de la siguiente. De este modo se consigue una abstracción de la tonalidad, preservando el patrón melódico. Para representar la duración de una nota, se codifica mediante una letra: *n* para las negras, *c* para las corcheas, *s* para las semicorcheas, *f* para las fusas, etc.

3. Resultados experimentales

Los algoritmos de tratamiento de partituras musicales se han implementado en Matlab, como un conjunto de *scripts*. En este apartado se presentan los resultados de identificación de las partituras analizadas. Estas partituras forman

Cuadro 1. Evaluación del rendimiento del OMR desarrollado.

	Número	Porcentaje
Total Notas	1786	100 %
Falsos Positivos	63	3.53 %
Falsos Negativos	13	0.73 %
Errores en el Tono	61	3.42 %
Errores en la Duración	524	29.34 %

parte del repertorio de invenciones de J.S. Bach. Se trata de piezas polifónicas de música de contrapunto. En la tabla 1 se muestra el número de notas contenidas en cada compás y el número y tipo de los errores. Los errores considerados son: (i) falsos positivos (notas que no aparecen en la partitura y que el sistema ha detectado), (ii) falsos negativos (notas que aparecen en la partitura y que el sistema no ha detectado), (iii) errores cometidos en la determinación del tono de la nota y (iv) errores cometidos en la determinación de la duración de la nota.

Por lo tanto, se detectan correctamente más del 95 % de todas las notas contenidas en las partituras analizadas. Del total de notas halladas, menos del 4 % de las veces se comenten errores en el cálculo del tono de dichas notas. Los peores resultados se obtienen en la determinación de la duración de las notas, donde el porcentaje de error se eleva al 29 %.



Figura 3. detección del patrón “2,-2,-2” en una partitura.

Como característica adicional, se ha desarrollado una herramienta de búsqueda de patrones en la partitura. Esta herramienta puede ser de utilidad para el análisis del estilo de una composición. En la Figura 3 se muestra una partitura y el resultado de la identificación del patrón “2,-2,-2” (en notación diferencial) en recuadros.

4. Conclusiones

En todo sistema OMR el índice de fiabilidad es muy importante. En este trabajo, nos hemos centrado en la determinación precisa del tono de las notas contenidas en una partitura de polifonía barroca. Como resultado, se ha desarrollado una aplicación que, pese a distar de las funcionalidades que ofrecen sistemas OMR comerciales, proporciona unos resultados aceptables (índice de fiabilidad del 95%), si bien es cierto que se omite parte de la información contemplada por los sistemas comerciales como por ejemplo, silencios y alteraciones. En este sentido, un trabajo futuro que se plantea está orientado a describir fielmente estas otras características relevantes de una partitura, así como la mejora de la tasa de errores correspondientes a la duración de las notas. Otro trabajo planteado consiste en la adaptación de los algoritmos propuestos al análisis de música armónica.

Referencias

1. W. Chai, B. Vercoe: Folk Music Classification using Hidden Markov Models. Proc. of the International Conference on Artificial Intelligence (2001)
2. Buzzanca, G.: A Rule-Based Expert System for Musical Style Recognition. Proceedings of the 1st Int. Conf. “Understanding and Creating Music” (2001)
3. Hori, T., Wada, S., Howzan T., Kung, S.Y.: Automatic music score recognition/play system based on decision based neural network. IEEE 3rd Workshop on Multimedia Signal Processing (1999) 183–184
4. Madsen, S. T., Widmer, G.: Evolutionary Search for Musical Parallelism. LNCS 3449 (2005) 488–497
5. Brown, A. R.: Exploring Rhythmic Automata. LNCS 3449 (2005) 488–497
6. Gartland-Jones, A., Copley, P.: The Suitability of Genetic Algorithms for Musical Composition. Contemporary Music Review, **22**:3 (2003) 43-55
7. Bainbridge, D.; Bell, T.C.: Dealing with superimposed objects in optical music recognition. Sixth International Conference on Image Processing and Its Applications, vol 2 (1997) 756–760
8. Otsu, N.: A Threshold Selection Method from Gray-Level Histograms, IEEE Transactions on Systems, Man, and Cybernetics, **9**:1 (1979) 62–66
9. Cruz-Alcázar, P.P., Vidal-Ruiz, E.: Modeling Musical Style using Grammatical Inference Techniques: a Tool for Classifying and Generating Melodies. Proceedings of the WEDELMUSIC’03 IEEE (2003) 77–84