Chapter V
Optical Music Recognition: Architecture and Algorithms

Pierfrancesco Bellini
University of Florence, Italy

Ivan Bruno
University of Florence, Italy

Paolo Nesi
University of Florence, Italy

ABSTRACT

Optical music recognition is a key problem for coding western music sheets in the digital world. This problem has been addressed in several manners, obtaining suitable results only when simple music constructs are processed. To this end, several different strategies have been followed to pass from the simple music sheet image to a complete and consistent representation of music notation symbols (symbolic music notation or representation). Typically, image processing, pattern recognition, and symbolic reconstruction are the technologies that have to be considered and applied in several manners; the architecture of the so called OMR (optical music recognition) systems. In this chapter, the O^3MR (object oriented optical music recognition) system is presented. It allows producing from the image of a music sheet the symbolic representation and saving it in XML format (WEDELMUSIC XML and MUSICXML). The algorithms used in this process are those of the image processing, image segmentation, neural network pattern recognition, and symbolic reconstruction and reasoning. Most of the solutions can be applied in other fields of image understanding. The development of the O^3MR solution with all its algorithms has been partially supported by the European Commission in the IMUTUS Research and Development project, while the related music notation editor has been partially funded by the research and development WEDELMUSIC project of the European Commission.
INTRODUCTION

Systems for music score recognition are traditionally called OMR (optical music recognition) systems. This term is tightly linked to OCR (optical character recognition) systems that are used for reading textual documents. Strictly speaking, OCR refers to systems based on the segmentation and recognition of single characters. OMR systems are used for many applications of education, cultural heritage, and publishing. The typical application of OMR is for accelerating the conversion of image music sheets to a symbolic music representation that can be manipulated to create new and revised music editions. Also, educational applications use the OMR systems for the same purpose (IMUTUS, 2004), generating, in this way, a customized version of music exercises. A different usage is for the extraction of incipit or full description for the image score retrieval (Byrd, 2001).

Typically, OCR techniques cannot be used in music score recognition since music notation presents a bidimensional structure. In a staff, the horizontal position denotes different duration for notes and the vertical position defines the height of the note (Roth, 1994). Several symbols are placed along these two directions. OMR is quite a complex problem, since several composite symbols are typically arranged around the note head. Despite the various research systems for OMR (Baimbridge, 1996, 2003; Byrd, 2001, 2006; Carter, 1989, 1994; Cooper, Ng, & Boyle, 1997; Coüasnon, 1995; Fujinaga, 1988, 1996; Kobayakawa, 1993; McPherson, 2002, Modayur, 1996; Ng, & Boyle, 1994, 1996; Prerau, 1970; Selfridge-Field, 1993; Tojo, 1982) as well as commercially available products (MIDISCAN, PIANOSCAN, NOTESCAN in Nightingale, SightReader in FINALE, PhotoScore in Sibelius, etc.), none of them is fully satisfactory in terms of precision and reliability. They provide a real efficiency, close to 90%, only if quite regular music sheets are processed and the estimation is not always objective (Bellini, 2007). This datum justifies the current research work focused on building reliable OMR systems and tools. OMR systems can be classified on the basis of the granulation chosen to recognize the music score’s symbols. There are two main approaches to define basic symbols. They can be considered as (1) the connected components remaining after staff lines removal (chord, beam with notes, etc.), or (2) the elementary graphic symbols, such as note heads, rests, hooks, dots, that can be composed to build music notation (Bellini, 1999; Blostein, 1992; Heussenstamm, 1987; Ross, 1970). With the first approach, symbols can be easily isolated out of the music sheet (segmented), and yet, the number of different symbols remains very high. The second approach has to cope with a huge number of different symbols obtained from the composition of basic symbols. This leads to an explosion of complexity for the recognition tool. A compromise is necessary between complexity and the system’s capabilities.

The architecture of an OMR system, and the definition of basic symbols to be recognized, are related to the methods considered for symbol extraction/segmentation and recognition. Generally, the OMR process can be divided into four main phases: (1) the segmentation, to detect and extract basic symbols from the music sheet image, (2) the recognition of basic symbols from the segmented image of the music sheet, (3) the reconstruction of music information, to build the logic description of music notation, and finally (4) the building of the music notation model for representing music notation as a symbolic description of the initial music sheet. In this chapter, the architecture, algorithms, methods, and results related to the O’MR system (object oriented optical music recognition) developed at the DSI University of Florence are discussed. The development of the O’MR solution with all its algorithms has been partially supported by the European Commission, in the IMUTUS IST FP5 Research and Development project (IMUTUS, 2004), while the related