Computerized Recognition of Hand-Written Musical Notes

Alez Bulis, Roy Almog, Moti Gerner & Uri Shimsony
Laboratory of Computer Music Engineering
Department of Electrical Engineering
Technion, Israel Institute of Technology,
Haifa, Israel,

Introduction

This work describes a software unit called CNR (Computerized Note Recognition) whose function is to interpret and recognize handwritten musical notes. The handwritten notes are read from paper by an optical scanner and the image of the page is stored in memory as a bit-map file. This file serves as an input to the CNR software. Note recognition is done by the histogram method: each symbol is specified in a library of symbols by prerecorded horizontal and vertical histograms. Histograms of an unknown symbol are compared with the library histograms and the symbol with the best fit to the unknown symbol is chosen. A horizontal histogram describes the sum of on-pixels for each column of the symbol. Similarly, a vertical histogram describes the sum of on-pixels for each row of the symbol, as seen in the Figure.

![Figure. Horizontal and vertical histograms of a half-note symbol](image)

The output of CNR is a musical alpha-numeric code, describing the pitches of notes, durations, etc. This code, used in our Lab and called Music Script, can be easily converted to MIDI code, or can be played directly by a sound generator. The output can also be used for quality printing of musical notes.

There are several methods to identify musical notes, however the advantages of the histogram method of identification are: simple identification algorithm, fast recognition process, simple prerecorded symbol histograms and easily constructed library of prerecorded symbol histograms. The histogram method of identifying symbols sometimes fails, especially when
the symbols are too complicated and the histograms cannot characterize them well. However, in the case of musical notes the number of symbols is not too great, the symbols are simple enough and they do not change much in size, so that the histogram method works very well. As in other pattern recognition cases, one cannot expect always perfect identification and some human help is needed to correct machine errors and put the final touch.

The Recognition Process

The recognition process consists of two main stages. The first one is the staff search stage in which all the staffs in a page are identified. The second one is the symbol identification stage in which every musical symbol is identified and recognized one after the other, in the order they appear on the staff.

There is an option of manual pixel or line editing of the image. This option is to be used in situations where noise and smudges appear in the image. Early human intervention can aid the identification which follows.

The staff search stage
A preliminary operation of rotation is often required to align the lines of the page image parallel to the lines of the display, because usually the page in the scanner is not aligned perfectly. A special check decides if rotation is required and performs the rotation. Care must be taken that in the rotation extreme parts of the image are not lost.

To search for the staffs, the program scans the image of the page horizontally from top to bottom. A line is found if the number of on-pixels is greater than a certain threshold value, usually taken as 60%. This threshold can be changed by the operator of the software, from 6% to 100%, as required. When the image of the page is "noisy" or "dirty" this threshold is lowered, otherwise the lines cannot be detected well.

A staff line should be only one pixel wide, so that heavier horizontal lines are not interpreted as staff lines. Sometimes an unwanted bottom horizontal line appears in the image due to the edge of the scanned paper. Therefore, a single bottom line is discarded by default, unless the operator chooses to look at it first.

To define a staff, five horizontal lines having a specified spacing between them should be found.

The gaps found between the staffs should be within predefined boundaries which may also be changed by the operator.

Once the whole staff setup of the page is found, corrections are done to straighten staff lines. This is done by moving up or down by few pixels parts of the staff with the symbols written on them.

The whole process of staff search can be automatic or it can be executed step by step. In the latter case the operator can follow the progress of the process and interfere if something goes wrong.

The symbol identification stage
There are two possible modes of symbol identification, automatic and manual. In the automatic mode the program chooses the best fit between library histograms and the histograms of an unknown symbol to be identified. In the manual mode the operator is presented on
the display with the best ten histogram fits, in order of fit quality, and he has to make the choice.

The symbol to be identified must conform to size tolerances, otherwise it is not considered, as a candidate for identification. The default tolerances are usually between 50% and 150%.

Before a symbol is subjected to the identification process it is cleaned from the staff lines that cross it. Its histograms are then formed and compared with the histogram collection in the library. The vertical position on the staff is of course retained before the cleaning of the symbol so that the pitch can be determined later.

The program provides a special menu option for creating and expanding the library of symbol histograms. Every time that the user discovers that a certain symbol is consistently not identified well, he can redefine the histograms in the library. He can also add new symbols to the library. For best results the library of histograms should be constructed specially for each handwriting. Library histograms are also defined from symbols which were cleaned from horizontal lines crossing it.

The identification decision is based on the calculation of a difference function which sums the squares of differences between the library histograms and the unknown symbol histograms. Usually equal weights are assigned to the horizontal and vertical histograms. Sometimes however it is better to give the vertical difference a higher weight than the horizontal difference, because details of the musical symbols are usually more pronounced in the vertical direction. Weights of between 50%/50% and 70%/30% were found to work well. During the identification process the operator may choose to look at the difference function and judge himself the quality of identification.

The present version of CNR works on any compatible PC equipped with a scanner. Because of memory shortage, the image is loaded from the hard disk part by part, as required. The whole operation is automatic and transparent to the user, using a virtual memory mode built into CNR.

The scope of the present version of CNR was limited to symbols which are not connected graphically to each other, in order to simplify things. A later version will overcome this limitation and enable to identify also chord notes and notes connected by eighth bar, sixteenth bar, etc.

Conclusion

The CNR project shows that computerized recognition of musical notes can be accomplished by a relatively simple method of comparing horizontal and vertical histograms of symbols. Although the identification of notes is not always perfect, it is highly efficient and can save much time in the process of inputting musical notes, even printed notes, into computer.