Methods of Artificial Intelligence in Blind People Education*

Bohdan Macukow and Wladyslaw Homenda

Faculty of Mathematics and Information Science
Warsaw University of Technology
Plac Politechniki 1, 00-660 Warszawa, Poland
{macukow,homenda}@mini.pw.edu.pl

Abstract. This paper presents the idea of recognition of music symbols to help the blind people reading music scores and operating music notation. The discussion is focused on two main topics. The first topic is the concept of the computer program, which recognizes music notation and processes music information while the second is a brief presentation of music processing methods including recognition of music notation - Optical Music Recognition technology - based on artificial neural networks. The short description and comparison of effectiveness of artificial neural networks is also given.

1 Introduction

Over the past few decades computers developed enormously. Along with the progress in hardware the researchers have been working hard to bring on computers to the activities of everyday life. Starting from turning raw interfaces to user friendly communication devices new methods have been studied and developed to make the computers not only efficient but mechanistic tools (like typewriters and counting machines) but also interacting with the human partners in an intelligent way. This required, of course, the use of methods that firmly belong to the domain of Artificial Intelligence.

In this paper we attempt to study an application of methods of Artificial Intelligence in the real life computer program that is supposed to handle musical notations. The term "Artificial Intelligence", though widely used by computer researchers, has neither a common definition nor is it uniquely understood by the academic community. However, it is not our aim to provoke a discussion on what artificial intelligence is and which methods does it embed. Instead, we rather use the term in a common sense though in an intuitive way.

Computer-based music processing methods have been developing since sixties, c.f. [13] and have found their commercial applications during last two decades, c.f. [5]. Music processing embraces several types of computer programs, including MIDI sequencers, AUDIO players and notation editors. However, these

* This work is supported under State Committee for Scientific Research Grant no 3T11C00926, years 2004-2007

kinds of programs, in their basic forms, cannot be classified as employing artificial intelligence methods; they are rather simple tools like record players or music analogs of typewriters. Above and beyond implementation of simple computer programs for music processing more sophisticated methods have been developed that truly belong to the field of AI. Two most important areas of utilization of AI methods are: knowledge representation and processing focused on music information, and pattern recognition exploited in music notation recognition and recognition of music in audio form.

Music computer program for the blind is what is possible at the cutting-edge of technology. There are many commercial programs for recognition of music notation processing that have been developed during the last decade. In early nineties MIDISCAN was developed. It was then superseded by the Smart Score [17], SharpEye and PhotoScore in the forthcoming years. Several notation editors as Smart Score [17], Finale, Sibelius and many MIDI sequencers were in use for music processing. There are, however, only a few programs of music processing for the blind musicians, c.f [16, 18].

In this paper we discuss application of artificial intelligence methods in music education for the blind. The discussion is focused on two aspects of computer processing of music: information representation and processing and optical music recognition. We then present a concept of computer aimed specifically at the blind people.

2 Representation and processing of music information

Knowledge representation and processing is the most important part of any music processing system. Music itself is one of human communication languages. It has extremely high level of sophistication, has not been codified in its wholeness and is still evolving. Music notation, an emanation of music, is a description of music in graphical form. Music notation can also be seen as human communication tool, it is highly complicated, its codification does not describe the whole notation and it is still evolving (like other tools of human activities). Music processing is governed by rules that neither are unique, nor complete, nor certain. Music processing cast on music notation is characterized by the same features as music processing. All these features require carefulness in music notation representation. In fact, music notation is a two dimensional language in which the geometrical relations between its symbols if of similar importance that the symbols alone. Music notation representation format has to store data about symbols, symbols placement and also contextual information about relative placement of symbols. Incorrectly design music notation representation will result in difficulties in music processing and even may make such processing impossible.

There are two important aspects related to music representation: structure of music notation and music description. The first aspect is a derivative of music structure (title, composer, parts of music piece, instruments, voices, etc.) and of geometrical limits of paper sheets (breaking music to measures, systems, pages, removing silent instruments from systems, etc.), c.f. Figure 1.

Fig. 1. An example of music notation page

Another aspect of music representation is related to information storage about music notation symbols and their properties. Music notation has complicated structure with many implicit relations between items of music data. Music symbols vary in size, shape and are arranged in much more complex and confusing way. In fact, music notation is a two dimensional language in which the geometrical relations between its symbols if of similar importance that the symbols alone. Therefore any music notation representation has to store data about symbols, symbols placement, properties of symbols, suggestions and indications for performers, etc. It also must store contextual information about relative placement of symbols and allow for repossessing contextual information indirectly expressed by the notation.

3 Acquiring music information

Any music processing system must be supplied with music information. Manual inputs of music symbols are the easiest and typical source of music processing systems. Such inputs could be split in two categories. One category includes inputs form - roughly speaking - computer keyboard (or similar computer peripheral). Such input is usually linked to music notation editor, so it affects computer representation of music notation. Another category is related to electronic instruments. Such input usually produce MIDI commands which are captured by a computer program and collected as MIDI file representing live performance of music.

Besides manual inputs we can distinguish inputs automatically converted to human readable music formats. The two most important inputs of automatic conversion of captured information are automatic music notation recognition which is known as Optical Music Recognition technology and audio music recognition known as Digital Music Recognition technology. In this paper we discuss basics of automatic music notation recognition as a source of input information feeding music processing computer system.

3.1 Optical Music Recognition

Printed music notation is scanned to get image files in TIFF or similar format. Then, OMR technology converts music notation to the internal format of computer system of music processing. The structure of automated notation recognition process has two distinguishable stages: location of staves and other components of music notation and recognition of music symbols. The first stage is supplemented by detecting score structure, i.e. by detecting barlines and then systems and systems' structure and detecting other components of music notation like title, composer name, etc. The second stage is designed on finding placement and classifying symbols of music notation. The step of finding placement of music notation symbols, also called segmentation, must obviously precede the step of classification of music notation symbols. However, both steps segmentation and classification often interlace: finding and classifying satellite symbols often follows classification of main symbols.

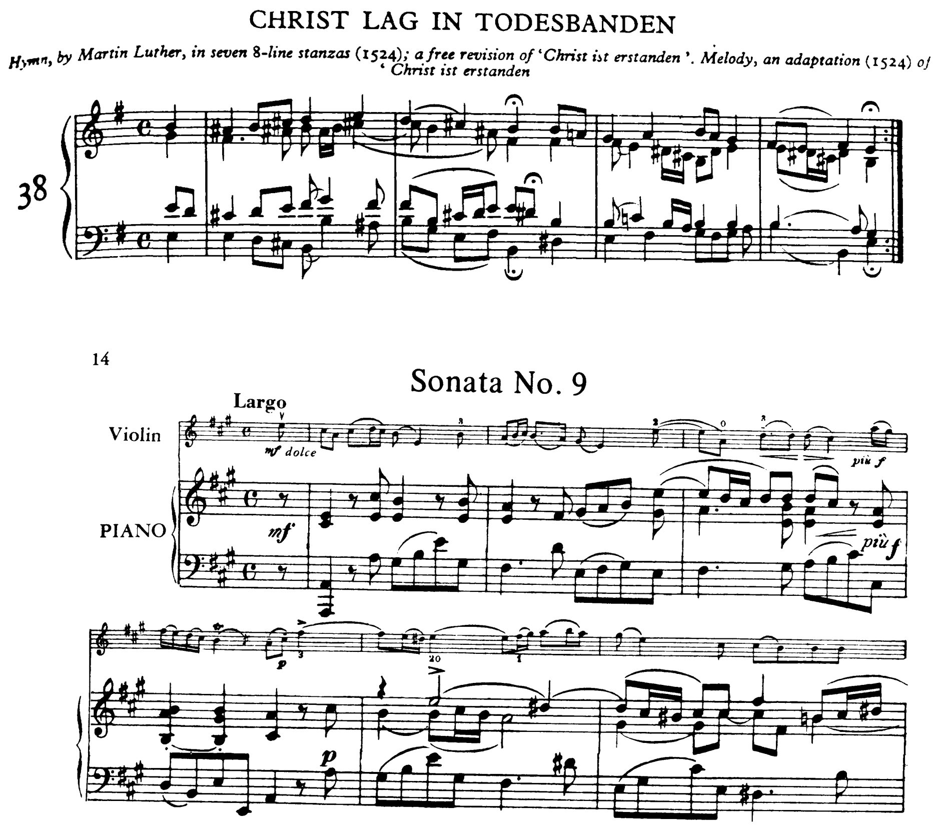

Fig. 2. Examples of real notations subjected to recognition

Staff lines and systems location

Music score is a collection of staves which are printed on sheets of paper, c.f. [6]. Staves are containers to be filled in with music symbols. Stave(s) filled in with music symbols describe a part played by a music instrument. Thus, stave assigned to one instrument is often called a part. A part of one instrument is described by one stave (flute, violin, cello, etc.) or more staves (two staves for piano, three staves for organ).

Staff lines location is the first stage of music notation recognition. Staff lines are the most characteristic elements of music notation. They seem to be easily found on a page of music notation. However, in real images staff lines are distorted raising difficulties in recognition. Scanned image of a sheet of music is often skewed, staff line thickness differs for different lines and different parts of stave, staff lines are not equidistant and are often curved, especially in both endings of the stave, staves may have different sizes, etc., c.f. [5, 6] and Figure 2.

Fig. 3. Printed symbols of music notation - distortions, variety of fonts

Having staves on page located, the task of system detection is performed. Let us recall that the term system (at a page of music notation) is used in the meaning of all staves performed simultaneously and joined together by beginning barline. Inside and ending barlines define system's structure. Thus, detection of systems and systems' structure relies on finding barlines.

Score structure analysis

Sometimes one stave includes parts of two instruments, e.g. simultaneous notation for flute and oboe or soprano and alto as well as tenor and bass. All staves, which include parts played simultaneously, are organized in systems. In real music scores systems are often irregular, parts which not play may be missing.

Each piece of music is split into measures which are rhythmic, (i.e. time) units defined by time signature. Measures are separated from each other by barlines.

The task of score structure analysis is to locate staves, group them in systems and then link respective parts in consecutive systems. Location of barlines depicts measures, their analysis split systems into group of parts and defines repetitions.

Music symbol recognition

Two important problems are raised by symbol recognition task: locating and classifying symbols. Due to irregular structure of music notation, the task of finding symbol placement decides about final symbol recognition result. Symbol classification could not give good results if symbol localization is not well done. Thus, both tasks are equally important in recognizing of music symbols.

Since no universal music font exits, c.f. Figures 1 and 2, symbols of one class may have different forms. Also size of individual symbols does not keep fixed proportions. Even the same symbols may have different sizes in one score. Besides usual noise (printing defects, careless scanning) extra noise is generated by staff and ledger lines, densely packed symbols, conflicting placement of other symbols, etc.

A wide range of methods are applied in music symbol recognition: neural networks, statistical pattern recognition, clustering, classification trees, etc., c.f. [1, 2, 4, 9, 11]. Classifiers are usually applied to a set of features representing processed symbols, c.f. [9]. In next section we present application of neural networks as example classifier.

3.2 Neural networks as symbol classifier

Having understood the computational principles of massively parallel interconnected simple neural processors, we may put them to good use in the design of practical systems. But neurocomputing architectures are successfully applicable to many reallife problems. The single or multilayer fully connected feedforward or feedback networks can be used for character recognition, c.f. [8].

Experimental tests were targeted on classification of quarter, eight and sixteen rests, sharps, flats and naturals, c.f. Figure 3 for examples music symbols. To reduce dimensionality of the problem, the images were transformed to a space of 35 features. The method applied in feature construction was the simplest one, i.e. they were created by hand based on understanding of the problem being tackled. The list of features included the following parameters computed for bounding box of a symbol and for four quarters of bounding box spawned by symmetry axes of the bounding box:

- mean value of vertical projection,
- slope angle of a line approximating vertical projection,
- slope angle of a line approximating histogram of vertical projection;
- general horizontal moment m_{10},
- general vertical moment m_{01},
- general mixed moment m_{11}.

Table 1. Chromatic symbols recognition rate of selected classifiers

flats	sharps	naturals	Classifier	
99.11%	97.98%	98.81%	backpropagation	5 - 3 - 2
92.30%	86.81%	91.74%	counterpropagation	15 - 8 - 1
96.52%	93.33%	89.11%	counterpropagation	25 - 8 - 1

The following classifiers were utilized: backpropagation perceptron, feedforward counterpropagation maximum input network and feedforward counterpropagation closest weights network. An architecture of neural network is denoted by a triple input - hidden - output which identifies the numbers of neurons in input, hidden and output layers, respectively, and does not include bias inputs in input and hidden layers.

Table 1 presents results for three symbols on music notation: flats, sharps and naturals, c.f. [9]. Classifier applied: backpropagation perceptron, feed-forward counterpropagation maximum input network and feedforward counterpropagation closest weights network. An architecture of neural network is denoted by a triple *input - hidden - output* which identifies the numbers of neurons in input, hidden and output layers, respectively, and does not include bias inputs in input and hidden layers.

4 Braille Score - bringing research to practice

Braille Score is a computer program to process music information. Braille Score is an integrated music processing computer program directed to a broad range of people. It is the part of the project under development in Warsaw University of Technology. The program together with a man creates an integrated system. It has special features allowing its usage by blind people. It is intended as a tool supporting blind people in dealing with music. Its important application could be placed in music education of blind students. Braille Score directly applies methods of artificial intelligence in practice. Its main modules deal with music information acquisition, storage and processing, communication with both blind people and good eyesight users. Main functions of Braille Score are:

- creating scores from scratch,
- capturing existing music printings and converting them to electronic version,
- converting music between different formats of music representation,
- processing music, e.g. transposing musing to different keys, extracting parts from given score, creating a score from given parts,
- preparing teaching materials,
- creating and storing own compositions and instrumentation.

4.1 Representing and processing of music information

A software for music processing is usually built around the model of an electronic format of music representation. Such a format stores music in a computer memory, processing it, exchange music data between different music equipment but also present music in a form of music notation. But from the other side the proper recognition of music notation is still under development. Information acquired at the stage of pattern recognition has to be stored in the form allowing for its further usage. This stage is based on the methods of knowledge representation and processing, c.f. [3, 5, 15]. The following topics could exemplify music knowledge storage and processing:

- designing a format of music representation,
- recognizing context relations: inheriting accidentals, linking articulation and ornamentation symbols to notes, linking lyrics to notes,
- structuring recognized music symbols, grouping notes into chords, grouping chords into beamed sequences, grouping accidentals into key signatures,
- identifications of rhythmic groupings,
- identifications of voices,
- converting music between different formats of music representation.

4.2 Acquiring music information

Braille Score is capable to acquire music information from several sources. Its main and distinguishable input source is printed music notation, which is subjected to automatic recognition of the structure and symbols. Only limited set of music notation symbols is intended to be recognized and process in Braille Score at the current version. The set of recognized symbols includes notes, chords, rests, accidentals, clefs, bar lines, key signatures, time signatures, change of key and time signature. Assuming future development of Braille Score, BSF include wider set of symbols including rhythmic, articulation and ornamentation figures and other symbols.

Braille Score can also read music information represented in MIDI, NIFF, MusicXML and Braille Music formats. Conversely, Braille Score can also output music information to the same sources. This way it is able to exchange music information with broad range of music software.

4.3 User interface extensions for blind people

Braille Score is addressed to blind people. Its user interface extensions allow blind user to master the program and to perform operations on music information. The most important feature of Braille Score is its ability to read, edit and print music information in Braille format. Blind user is provided the following elements of interface: Braille Notation editor, keyboard as input tool, sound communicator.

Blind people do not use pointing devices. In consequence, all input functions usually performed with mouse must be mapped to computer keyboard. Massive communication with usage of keyboard requires careful design of interface mapping to keyboard, c.f. [12].

Blind user usually do not know printed music notation. Their perception of music notation is based on Braille music notation format, c.f. [10] presented at Braille display or punched sheet of paper. In such circumstances music information editing must be done on Braille music notation format. Since typical Braille display is only used as output device, such editing is usually done with keyboard as input device. In Braille Score Braille representation of music is online converted to internal representation and displayed in the form of music notation in usual form. This transparency will allow for controlling correctness and consistency of Braille representation, c.f. [12].

Sound information is of height importance for blind user of computer program. Wide spectrum of visual information displayed on display screen for user with good eyesight could be replaced by sound information. Braille Score provides sound information of two types. The first type of sound information collaborates with screen readers, computer programs dedicated to blind people which could read contents of display screen and communicate it to user in the form of synthesized speech. This type of communication is supported by contemporary programming environments. Braille Score uses tools provided by Microsoft .NET programming environment. The second type of sound information is based on own Braille Score tools. Braille Score has embedded mechanism of sound announcements based on own library of recorded utterances.

5 Conclusions

In this paper we describe a concept of Braille Score the specialized computer program which should help blind people to deal with music and music notation. The use of artificial intelligence tolls like neural networks can improve the program part devoted for the recognition of the music symbols. The first results with Braille Score show its to be a practical and useful tool.

References

1. Bainbridge D., Bell T., The challenge of optical music recognition, Computers and the Humanities 35 (2001) 95-121.
2. Carter N. P., Bacon R. A., Automatic Recognition of Printed Music pp. 456-465 in: Structured Document Analysis, Analysis, H.S.Baird, H.Bunke, K.Yamamoto (Eds), Springer-Verlag, 1992.
3. Dannenberg R., Music Representation Issues, Techniques, and Systems, Computer Music Journal (17:3) (1993) 20-30.
4. Fujinaga I., Adaptive optical music recognition, 16th Inter. Congress of the Inter. Musicological Society, Oxford, Oxford University Press, 2001.
5. Homenda W., Automatic recognition of printed music and its conversion into playable music data, Control and Cybernetics, vol. 25, no. 2 (1996) pp. 353-367.
6. Homenda W., Granular Computing as an Abstraction of Data Aggregation - a View on Optical Music Recognition, Archives of Control Sciences, Vol. 12 (XLVIII) (2002) No. 4, pp 433-455.
7. Homenda W., Optical Music Recognition: the Case Study of Pattern Recognition, in: Computer Recognition Systems, Kurzyski et al (Eds.), pp. 835-842, Springer Verlag, 2005.
8. Homenda W., Luckner M., Automatic Recognition of Music Notation Using Neural Networks, Proc. of the International Conference On Artificial Intelligence and Systems, Div-nomorskoye, Russia, September 3-10, 2004.
9. Homenda W., Mossakowski K., Music Symbol Recognition: Neural Networks vs. Statistical Methods, EUROFUSE Workshop On Data And Knowledge Engineering, Warsaw, Poland, September 22 - 25, 2004, pp. 265-271
10. Krolick B., How to Read Braille Music, 2nd Edition, Opus Technologies, 1998.
11. McPherson J. R., Introducing feedback into an optical music recognition system, Third Internat. Conf. on Music Information Retrieval, Paris, France, 2002,.
12. Moniuszko T., Design and implementation of music processing computer program for blind people (in Polish), Master Thesis, Warsaw University of Technology, Warsaw, 2006.
13. Pruslin D. H., Automatic Recognition of Sheet Music, PhD Thesis, Massachusetts Institute of Technology, 1966.
14. MIDI 1.0, Detailed Specification, Document version 4.1.1, February 1990.
15. http://www.musique.umontreal.ca/personnel/Belkin/NIFF.doc.html 1990.
16. http://www.dancingdots.com/goodfeel.htm
17. http://www.musitek.com/
18. http://members.optusnet.com.au/∼terryk/toccata.htm