

IMUTUS – an interactive system for learning to play a musical instrument

Erwin Schoonderwaldt, Kjetil Falkenberg Hansen & Anders Askenfelt

Dept. of Speech, Music and Hearing, KTH, Stockholm, Sweden

Key words: *music education, interactive learning, feedback, musical performance, performance skills*

Abstract:

IMUTUS (Interactive Music Tuition System) is a EU project that aims to develop a practising environment for the recorder, combining new technologies and new approaches for music learning. Automatic analysis and evaluation of student performances play a central role in the student-system interaction. The performance evaluation module identifies typical performance errors, and provides feedback that relates to performance skills, helping the student to improve. The performance evaluation process is based on the knowledge and experience of recorder teachers, obtained via questionnaires, interviews and structured evaluations of recorded student performances. Another important feature of the performance evaluation is that it can be guided by teachers writing the content for IMUTUS by means of annotations.

1 Introduction

Computers were introduced in music education more than 20 years ago, and have become more and more common. The computer has probably been most successfully applied for training fundamentals of music, such as interval recognition and rhythmic and melodic dictation and, where the response of the student can be easily assessed. The obvious advantages of the computer here are its availability and replacement of the piano. Other areas of music education, in which the computer is applied are music performance skills, music analysis and music composition [1].

The IMUTUS system is mainly aimed at training musical performance skills on the recorder, but also facilitates training of fundamentals of music and simple composition. In this paper the discussion will be focused on systems for training performance skills in general, and the specific approach of IMUTUS.

The basic facilities for every practising environment for training of musical performance skills are: a database of lessons (content), display of the score on the screen (note reading), and the possibility to listen to a (synthesized) performance of the score. In the simplest implementation, the computer could be used to present the material to the student. That does not fundamentally differ from a songbook with an audio cassette, apart from interactive features, such as showing the fingering by clicking on a note.

More advanced practising environments offer a performance analysis as well, enabling the computer to “listen” to the student’s performance. This opens the way for assessing the performance and providing feedback to the student, which adds another dimension to the interactivity. However, there are some fundamental problems associated with analysis and assessment of musical performances. The first problem is that the computer has to “know” what notes the student played and how these relate to the music score. This requires

segmentation of the recorded audio into recognized notes and a mapping of the notes in the score on the performance.

The second problem is directly related to the expressive nature of musical performance. Musical performance is characterized by “expressive” deviations in timing, dynamics and pitch compared to the score. A straight performance of the score will lead to so called *mechanical* or *dead-pan* performance.

There is only a limited number of existing systems for practising musical instruments with “listening” facilities. A good example is Musicalis, a commercially available program for training performance skills on a variety of instruments [2]. The problem of knowing the score position is solved by letting the student synchronize her playing with an accompaniment or a metronome, presented over headphones. A cursor shows the supposed score position. The feedback is presented in the form of a transcription of the student’s performance, indicating notes with wrong pitch or duration with red color.

Probably the most accomplished practising system developed so far is Piano Tutor [3], which uses a score follower for tracking the student’s position in the score. This gives the student more freedom in playing. The feedback is presented to the student by voice messages and a marker in the score, indicating the position of the performance error in question. Piano Tutor is limited to the use of a MIDI keyboard as instrument; it does not contain the audio recognition techniques, necessary for converting the sound of acoustical instruments.

2 The IMUTUS system

The aim of the IMUTUS project is to develop a practising environment for the recorder by integrating new and state-of-the-art technology into a platform, supporting a new approach towards music learning. The target group is recorder students 9-14 years old, including beginners and up to medium level. The most efficient use of IMUTUS would be as a “assistant teacher” during the practising sessions at home, as a complement to weekly traditional lessons with a teacher.

The core of the IMUTUS system consists of a “virtual teacher” and a score viewer [4]. During a student performance, the virtual teacher (in the role of a listener) keeps track of the student’s position in the score and turns the page if needed. After the performance the virtual teacher (in the role of an analyst) provides an evaluation of the performance on request of the student. The feedback is presented in the in the score, where mistakes and errors (one at a time) is highlighted and clarified with a text message. Additional system components helping the student are a tuner, a metronome and a virtual-reality fingering viewer.

The design of the curriculum shares the idea of the [3]. By completing lessons, the student acquires skills, which are collected in a “student profile.” This profile is compared with the prerequisites of the next lessons the student can choose from.

The system also facilitates a platform for distance learning. The student can download lessons from a central server and send her performance to a (distance) teacher. The teacher can add new lessons to the server using the content authoring tools.

3 Automatic performance evaluation in IMUTUS

Automatic performance evaluation plays a key role in the student-system interaction. The purpose is not only to inform the student about her achievements, but also to make the practising sessions more efficient by providing structured feedback immediately after each performance. The performance evaluation module (PEM) identifies typical errors or mistakes in the student performance and gives constructive feedback on the three top-ranked errors as output. In this section it will be explained how performance evaluation is done in IMUTUS, and how it relates to teaching practise.

3.1 Basic performance skills

The performance evaluation module in IMUTUS is based on a model of novice student performance skills, covering different aspects of playing the recorder. Errors or mistakes in a student performance can be classified according to a set of basic performance skill categories. In this approach, performance errors are closely related to performance skills.

The basic performance skill categories (BPSC) and their relative importance were determined at the initial phase of the project as part of the user requirements. This was done with questionnaires and interviews with recorder teachers. In total 40 music teachers from France, Italy and Sweden responded to the questionnaires. Additional interviews were held with 5 Swedish recorder teacher.

The results are summarized in Table 1. A set of 9 basic performance skill categories was considered to capture the essential characteristics of the students' performances, as judged by recorder teachers. The teachers were also asked to rank these categories in order of importance. The average was used to determine the ranking as shown in Table 1. It can also be noted that basic performance skill categories refer both to aspects of instrument control (IC), which are specific for the recorder, and more general instrument-independent aspects of musical performance (MP).

The ranking is also reflected in the development of a student. For beginning students, who do not master the instrument yet, the main focus is on aspects of instrument control, while more advanced students spend more attention to aspects of musical performance.

Table 1: Basic performance skill categories, average ranking, and performance aspects. The ranking refers to the relative importance of the errors/mistakes during the first four terms of playing the recorder. Aspects of instrument control as well as musical performance are represented. Instrument control is considered most important to develop in the early stages of learning.

No	Basic performance skill category (BPSC)	Average ranking	Instrument control(IC) / Musical performance (MP)
1	Airflow	1.7	IC
2	Fingering	1.7	IC
3	Rhythmic performance	2.0	MP
4	Attack	2.0	IC
5	Melodic performance	4.2	MP
6	Tempo	5.0	MP
7	Intonation	5.3	IC
8	Phrasing	6.0	MP
9	Articulation	7.5	IC/MP

3.2 Typical performance errors

More specific knowledge of typical performance errors was obtained in an empirical way. For this purpose recordings were made of recorder students with varying level of progress. A representative selection of 8 recordings was made and 5 Swedish recorder teachers were asked to provide in a structured analysis of these student performances. Except the recordings, the teachers also received the score. The structured analysis consisted amongst others of:

- The most important errors,
- The feedback they would give to the student,
- A grading of the performance on a scale from 1 to 5.

The performance errors, reported by the teachers were classified according to the basic performance skill categories, resulting in a detailed overview of typical student errors. This overview formed an important basis for the development of error recognition algorithms in PEM.

The collected teacher evaluations were entirely based on auditory information. This implies that in principle all the information needed for the evaluation, provided that the score is known, is coded in the audio recording of the student performance. Performance errors can be recognized by *symptoms*, present in the acoustical signal, combined with comparison of the performance with the score.

A major advantage of this approach is that identified performance errors represent a higher level interpretation of symptoms. Recognition of typical performance errors provides the key to more adequate feedback.

The fact that only sound is considered means also that the evaluation is limited to audible performance errors. For example posture, which is a very important aspect for the beginning student cannot be considered by PEM.

3.3 Design of the Performance Evaluation Module

This strategy of finding and classifying performance errors is reflected in the design of the performance evaluation module. Before PEM is called, the recorded audio is analysed, resulting in signals of fundamental frequency and amplitude, as well as a list of recognised notes, with pitch, onset times and durations. This note list is matched to the score the student was supposed to play. This forms the main input of PEM.

PEM scans this extracted performance information for typical performance errors and sorts them in order of priority according to the basic performance skill categories. The most important errors are selected for feedback to the student. This will be explained in more detail in section 3.4.

PEM does not only look for errors, but also tries to identify well performed aspects. For example, if no error was found in a certain basic performance skill category, a positive comment, referring to this skill is generated.

3.4 Feedback

Immediately after the performance, before the feedback is displayed, a *grading* is given. This gives a rough estimation of how well the student performed, and could be useful as a momentary stimulus or reward to the student. A set of twinkling recorders appear on the screen, with 3 recorders being the top grade. The grading is based on a weighted sum of the number of errors and positive comments.

The detected errors in a performance may be numerous. For pedagogical purposes the mistakes/errors need to be prioritised and a small selection (typically 3) to be made before presenting the feedback to the student. As only a small number of mistakes/errors will be reported after each performance, the priority will be critical for the pedagogical value of the evaluation. In the current implementation, the priority is mainly made according to the basic performance skill categories (see Table 1). There are a few exceptions, in which errors of a lower category have a high priority. For example, bad tuning of the recorder has a high priority, even if it falls under the *Intonation* category.

Except for reporting errors, it is also important to inform the student about the strong points of her performance. Rewarding feedback is important to keep the student motivated. For this reason, PEM selects a positive comment about every second performance. Positive comments are ordered in the same way as performance errors, but have priority over higher ranked errors in this case.

The feedback is communicated to the student via a GUI. In Figure 1 it is shown what this might look like. The student is guided through the feedback process following the priority of

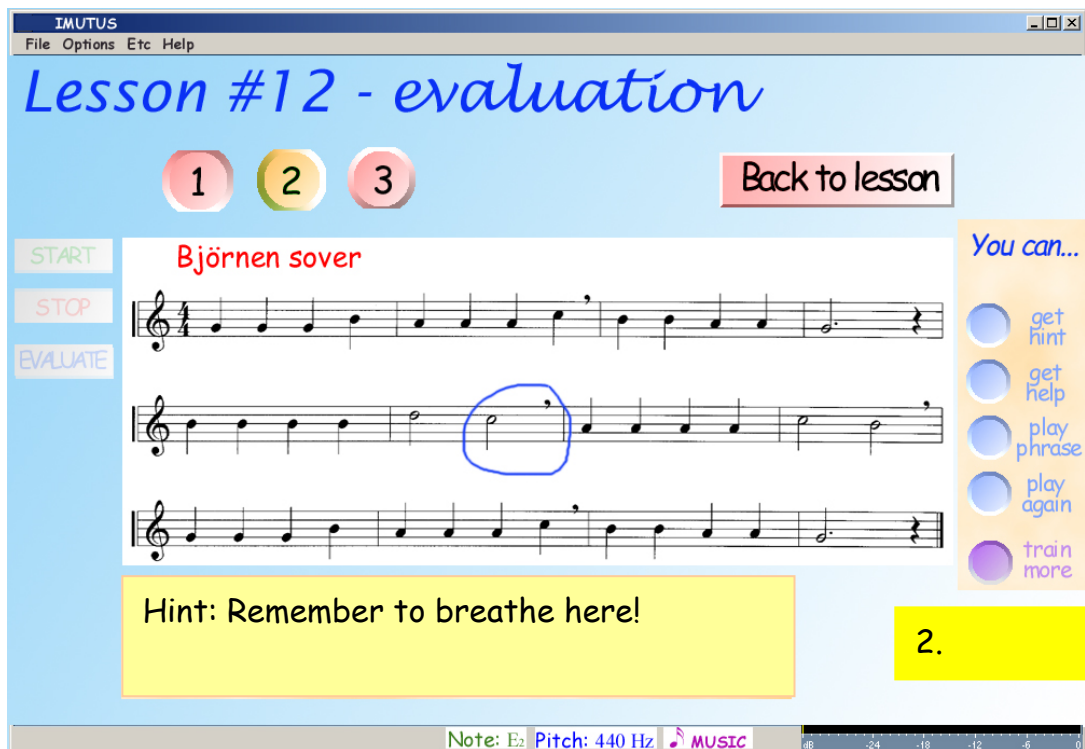


Figure 1: Screen view example. The student has pressed error button 2 and the position in the score and the descriptive message “Phrasing!” are displayed. The student has then asked for “Get hint” and a feedback message is displayed.

the feedback. The location of each mistake/error in the score and an initial, short descriptive message will be displayed when pressing one of the three error buttons.

There are two levels of feedback, which can be shown on request by the student. The first, light level is ‘Get hint’ and the second, deeper level is ‘Get help’. Each typical performance error has a predefined set of feedback messages. Some of the messages are available in different wordings, for avoiding boring repetitions of the same message.

3.5 Support for error detection by teachers

Teachers writing the content for IMUTUS have also the opportunity to influence the evaluation by PEM. This can be done in two ways. The teacher can add *score annotations*, marking difficult spots in the score where an error of a certain type is likely to occur. Such annotations can concern accidentals, rhythmic patterns, difficult finger transitions and so forth. By means of score annotations, knowledge of the teacher is made explicit to PEM. A nice side effect is that it enables PEM to provide well founded positive feedback when the student has played a difficult passage correctly.

Another way to tailor PEM towards the needs of a specific exercise is through the use of *projections*. In a projection, only one or a few performance aspects are considered, for instance rhythm only, and as a result the feedback will be more focused.

4 Discussion & conclusions

Currently, the components of the IMUTUS system are not yet fully integrated. The first prototype is scheduled to be available in the near future and will be used for usability tests and validation. However, preliminary tests of a stand-alone version of PEM show promising results: the evaluations of the selected student recordings show a high degree of agreement with the teacher evaluations. This indicates that the approach of error recognition and classification might be a good basis for automatic performance evaluation.

Acknowledgements

This work was supported by the European Community under the Information Society Technology (IST) RTD programme, contract IST-2001-32270

References:

- [1] Brandão, Pain & Wiggins, *Computers in Music Education*. Proceedings of the AISB'99 Symposium on Musical Creativity, 1999
- [2] Musicalis (software for music tuition): <http://www.musicalis.fr/v1/>
- [3] R.B. Dannenberg, M. Sanchez, A. Joseph, R. Joseph, R. Saul & P. Capell, *Results from the Piano Tutor Project*. Proceedings of the Fourth Biennial Arts and Technology Symposium, Connecticut College, (March 1993), pp. 143-150
- [4] D. Fober, S. Letz, Y. Orlarey, A. Askenfelt, K. Falkenberg Hansen & E. Schoonderwaldt, *IMUTUS – an interactive music tuition system*. Proceedings of Sound and Music Computing 2004, IRCAM, France.

Author(s):

Erwin Schoonderwaldt. M.Sc.
Kjetil Falkenberg Hansen
Anders Askenfelt, Dr.

Royal Institute of Technology (KTH)
Dept. of Speech, Music and Hearing
Lindstedsvägen 24
S-100 44, Stockholm, Sweden

{schoondw, hansen, andersa}@speech.kth.se