MEASUREMENT AND REPRODUCTION ACCURACY OF COMPUTER-CONTROLLED GRAND PIANOS

Werner Goebl[1,2] *& Roberto Bresin*[2]

[1] Austrian Research Institute for Artificial Intelligence (ÖFAI), Vienna

[2] Department of Speech, Music, and Hearing
Royal Institute of Technology (KTH), Stockholm

werner.goebl@oefai.at, roberto@speech.kth.se

ABSTRACT

The recording and reproducing capabilities of a Yamaha Disklavier grand piano and a Bösendorfer SE290 computer-controlled grand piano were tested, with the goal of examining their reliability for performance research. An experimental setup consisting of accelerometers and a calibrated microphone was used to capture key and hammer movements, as well as the acoustic signal. Five selected keys were played by pianists with two types of touch (*staccato and legato*). Timing and dynamic differences between the original performance, the corresponding MIDI file recorded by the computer-controlled pianos, and its reproduction were analysed. The two devices performed quite differently with respect to timing and dynamic accuracy. The Disklavier's onset capturing was slightly more precise (±12 ms) than its reproduction (from −20 to +30 ms). The Bösendorfer performed generally better, but its timing accuracy was slightly less precise for recording (−9 to 3 ms) than for reproduction (±2 ms). Both devices exhibited a systematic (linear) error in recording over time. In the dynamic dimension, the Bösendorfer showed higher consistency over the whole dynamic range, while the Disklavier performed well only in a wide middle range. Neither device was able to capture or reproduce different types of touch.

1. INTRODUCTION

Current research in expressive music performance mainly deals with piano interpretation because obtaining expressive data from a piano performance is easier than from other instruments. Pianists are able to control only a few parameters on their instruments. These are the tone[1] onsets and offsets, the intensity (measured as the final hammer velocity), and the movements of the two pedals[2]. Computer-controlled grand pianos are a practical device to record and to measure these expressive parameters and simultaneously provide a natural and familiar setting for pianists in a recording situation. Two systems are most commonly used in performance research: the Yamaha Disklavier and the Bösendorfer SE system.

The measurement results for the Yamaha Disklavier were already reported elsewhere [1] and are compared here with data from a Bösendorfer SE290 computer-controlled grand piano.[3]

[1] The onset of a sounding tone is very often called 'note onset', because of the MIDI world's terminology. In this paper, the terms 'tone' and 'note' are used synonymously, since we are not talking about musical notation.

[2] The middle or *sostenuto* pedal only prolongs certain tones and is not counted as an individual expressive parameter.

[3] For further introductory references, refer also to [1].

2. METHOD

The two computer-controlled grand pianos were the same as in [2]. For the experimental setup, equipment, calibration, and procedure, please also refer to [2].

In addition to the readings reported in [2], the **MIDI note onset time**, and the **MIDI velocity value** were taken from the MIDI file or the corresponding internal file format of the Bösendorfer. The onset differences between the original recording and the MIDI file, and those between the original recording and its reproduction were calculated[4]. Since the three measurements (original recording, MIDI file, and reproduction) were not synchronised in time by the measurement procedure, their first attacks were defined as being simultaneous. Care was taken that the first tones always were loud attacks in order to minimise synchronisation error, since timing error was smaller the faster the attack was. If there was a soft attack at the beginning, the files were synchronised by the first occurring louder attack (hammer velocity over 1 m/s).

3. RESULTS & DISCUSSION

3.1. Timing accuracy

In Figure 1, the note onset delays of the MIDI file in comparison to the original recording are plotted against the recorded time separately for the two pianos. It is evident that both MIDI files show a constantly increasing anticipation over time. This timing error in the MIDI file was larger for the SE system than the Disklavier. The origin of this systematic timing error is yet unknown, but it is likely that the internal counters of the systems (in the case of the SE system, it is a personal computer) did not operate in exactly the desired frequency, probably due to a rounding error.

To illustrate the recording accuracy without this systematic error, the residual timing error (the differences between the fitted lines and the data) is plotted in Figure 2 separately for the two pianos against recorded MIDI velocity.[5] The variance was larger for the Disklavier than the SE system (Yamaha mean: 1.4 ms, standard deviation (SD): 3.8 ms; Bösendorfer mean: 0.2 ms, SD: 2.1 ms), but for both pianos, the residual timing error bore a trend with respect to the loudness of the recorded tones. The Disklavier tended to record softer tones later than louder ones, the SE showed the opposite trend, but to a smaller extent and with much less variation (Figure 2).

[4] $Delay_{MIDI} = t_{MIDI} - t_{orig}$; $Delay_{repro} = t_{repro} - t_{orig}$.

[5] In an earlier conference contribution, a different normalisation method was applied on the same data of the Disklavier [1].

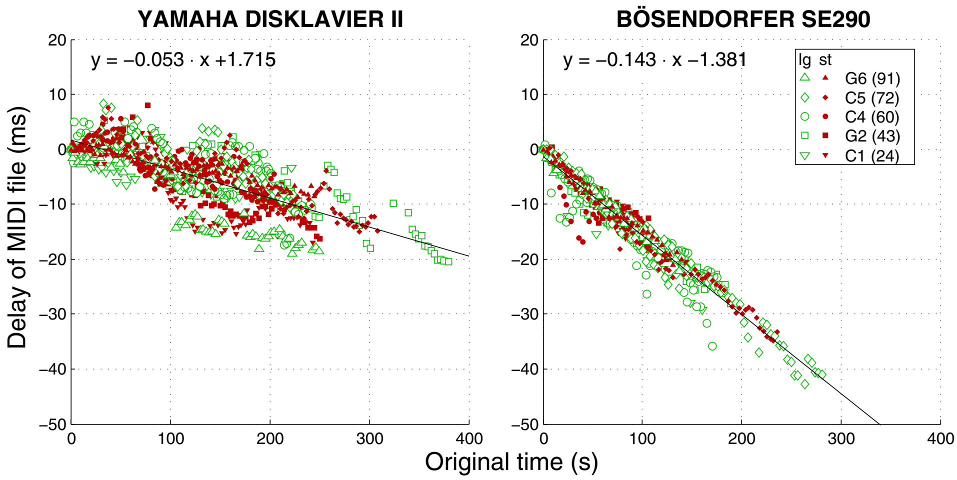

Figure 1: Timing delays (ms) as a function of recorded time (s) between the original recording and the MIDI file as recorded by the computer-controlled grand pianos for two types of touch: *legato* ('lg') and *staccato* ('st'). Negative values indicate that an onset in the MIDI file was earlier than in the original recording. The straight lines are linear fits of the whole data.

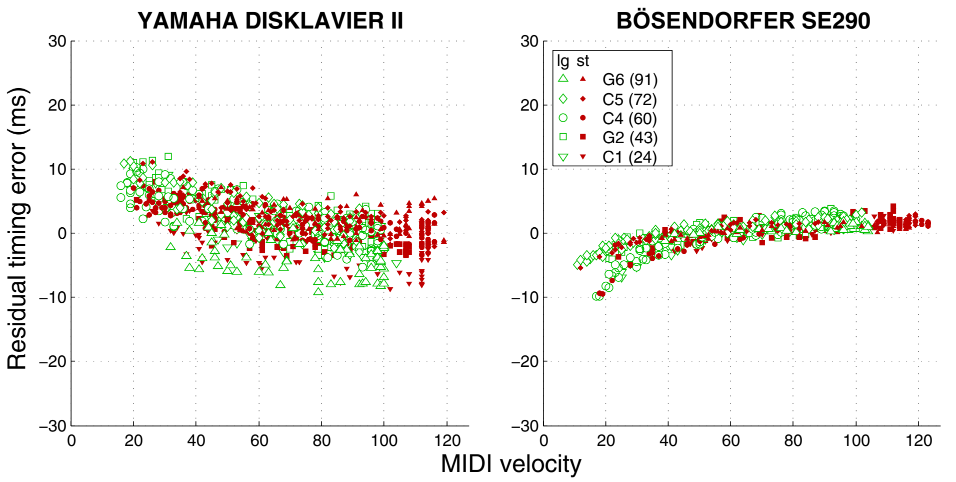

Figure 2: The residual timing error (ms) between the MIDI file and the original recording as a function of MIDI velocity, as recorded by the computer-controlled pianos. Again, negative values indicate onsets too early in the MIDI data, in comparison to the original file.

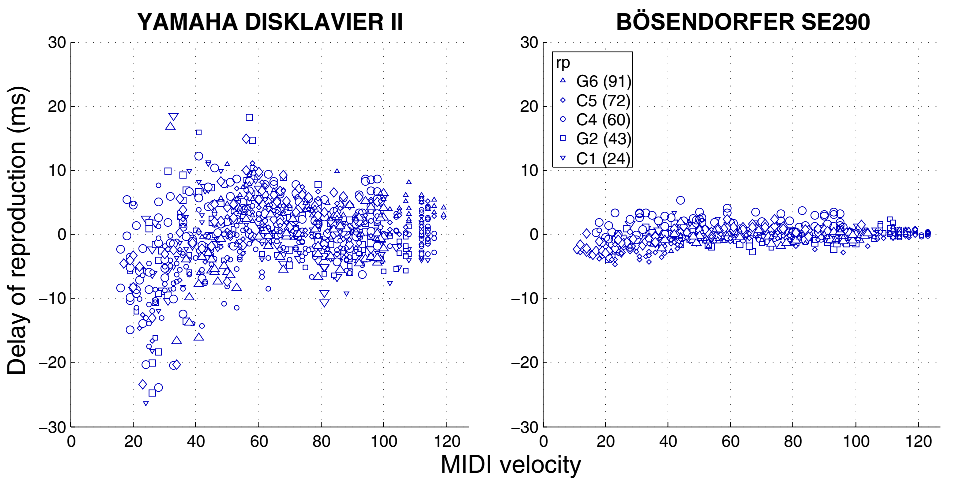

Figure 3: Timing delays (ms) between the original and its reproduction by the computer-controlled piano. (No systematic trend had to be removed.)

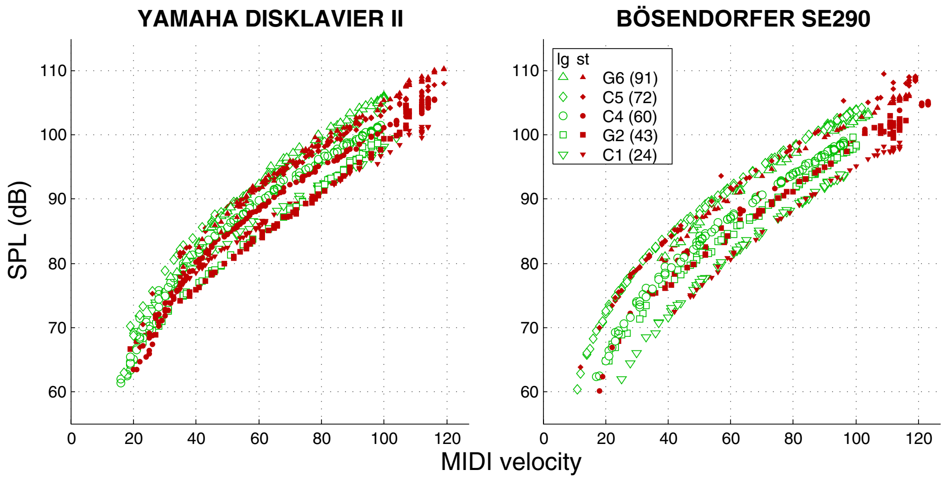

Figure 4: Peak sound pressure level (dB) against MIDI velocity as recorded by the two computer-controlled pianos. Different markers denote the five tones, different grey scales *legato* ('lg') and *staccato* ('st') touch.

The timing delays between the original recording and its reproduction are plotted in Figure 3 separately for the two pianos. There was no systematic timing error to remove in the data; the error in recording was evidently cancelled out by the same error in reproduction. The difference between the two systems becomes most evident in this display. While the reproduced onsets of the Disklavier can differ as much as +20 and -28 ms (mean: -0.3 ms, SD: 5.5 ms) from the actual played onset, the largest timing error of the SE system rarely exceeded ± 3 ms with a tendency of soft notes coming up to 5 ms too soon (mean: -0.1 ms, SD: 1.3 ms). Interestingly, the recording accuracy of the SE system was lower than its reproduction accuracy. Obviously, its internal calibration function aimed successfully to absolute precise reproducing capabilities. It could also be that the SE takes the first trip point (5 mm before the strings) as being the note onset, but calibrates itself correspondingly to overcome this conceptual mistake. However, this assumption was contradicted by information obtained by the SE's developer, W. Stahnke [3].

3.2. Dynamic accuracy

The second of the investigated parameters was dynamics which was measured in terms of the speed of the hammer hitting the strings (m/s) or peak sound pressure level (dB). Due to the limited space, we report here on the dynamic recording and reproduction capabilities only in terms of peak sound pressure level (dB–SPL). In Figure 4, dB–SPL is displayed against MIDI velocity units as recorded by the reproducing systems. On both instruments, different pitches exhibited different curves. The higher the pitch, the louder the radiated sound at the same MIDI velocity.

Peak dB–SPL as measured from the reproductions by the systems is plotted against dB–SPL as measured from the pianists' original recordings in Figure 5. It becomes evident that the Disklavier's solenoids were not able to reproduce very strong tones above a certain intensity. This varied slightly between keys, e.g. the G6 (with less hammer mass than hammers at a lower pitch) could be reproduced properly up to 107 dB, whereas a C1 (with a comparatively heavy hammer) only up to 93 dB. On the SE system, this ceiling effect could not be observed and there was no obvious effect of pitch as for the Disklavier. The superior reproduction of very loud tones for the Bösendorfer was due to its stronger solenoids. However, this does not explain why soft tones were reproduced considerably louder by the Disklavier, but very consistently by the Bösendorfer. Even silent notes were reproduced by the Bösendorfer, while this was not the case for the Disklavier.

4. GENERAL DISCUSSION

In this study, we measured the recording and reproducing accuracy of two computer-controlled grand pianos (Yamaha Disklavier, Bösendorfer SE) with an accelerometer setting in order to determine their precision for piano performance research. Both devices showed a systematic timing error over time which was most likely due to a rounding error in the system clock (the internal hardware at the Disklavier, a common personal computer at the SE). This linear error removed, the Bösendorfer had a smaller (residual) timing error than the Disklavier, but both exhibited a certain trend with respect to the loudness of the tones. The Disklavier tended to record soft tones too late whereas the SE had the tendency to record soft tones too early. But within these tendencies, the SE was more consistent. During reproduction, the superior performance of the Bösendorfer became more evident: the timing error was smaller than during recording whereas the Disklavier increased in variance in comparison to its recording.

The important point for performance research is the recording accuracy of these systems. Apart from the systematic error that only marginally affects the measured tempo value (0.0053% or 0.014%, resp.), the residual timing error (Fig. 2) was considerably large for the Disklavier and smaller for the Bösendorfer. The measurement precision could be improved by calculating out these trends using polynomial curve approximations.

To examine reproducing accuracy in the loudness dimension, we used peak sound pressure level as a measure. Here, the SE system revealed a much more precise reproducing behaviour over the whole dynamic range than the Disklavier. For the latter, the dynamic extremes flattened out, soft tones were played back too loudly and very loud tones too softly. The Disklavier's poor reproduction of loud tones was due to its smaller solenoids. Its reproduction of soft notes was limited, because the tested Disklavier prevented very soft tones from being silently reproduced with a minimum velocity matrix, adjustable by the internal control unit. It was also due to this function that the Disklavier was not able to reproduce silent notes at all, a crucial feature especially for music of the 20th century. The Bösendorfer exhibited linear reproducing behaviour over the whole dynamic range (from 60 to 110 dB SPL).

As another, and indeed very important criterion of recording and reproducing capability, we did not investigate the two pedals.[6]

[6]We are talking only of the right and the left pedal of grand pianos, since the middle pedal—the *sostenuto* pedal—only varies the tone length of certain keys depressed during its use, which is recorded and reproduced

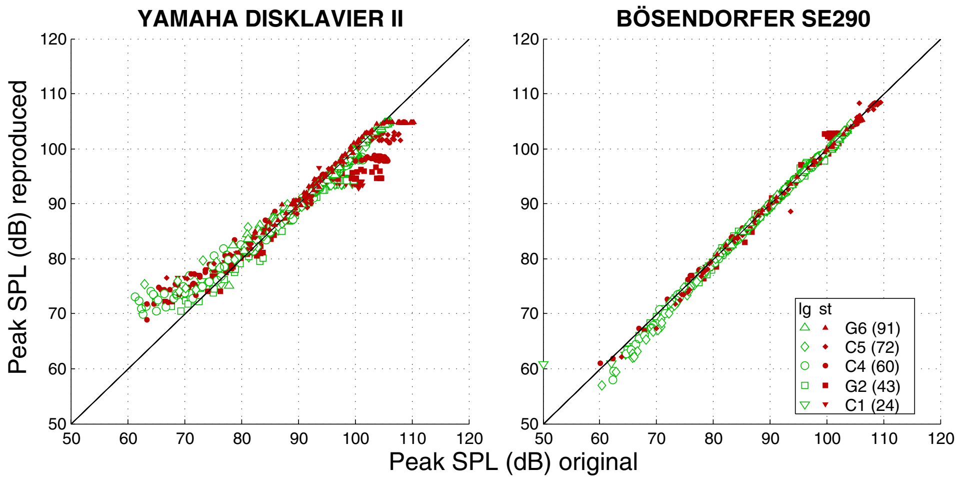

Figure 5: Peak sound pressure level (dB) as measured in the tones performed by the pianists (x axes) and reproduced by the computer-controlled pianos (y axes).

Both the Disklavier and the SE system are based on the same underlying principle. That is, to measure and reproduce movement of the piano action (and the pedals), in particular the final speed of the hammer before touching the strings. This principle is fundamentally different to what a performing artist does when playing expressively. The artist controls his/her finger and arm movements in order to reproduce a certain mental image of the sound by continuously listening to the resulting sound and by feeling the haptosensory feedback of the keys [4]. In this way, the performer is able to react to differences in the action, the voicing, the tuning, and the room acoustics, just to mention a few variables that have a certain influence on the radiated sound. On the other hand, a reproducing piano aims to reproduce a certain final hammer velocity independently of whether room acoustics, tuning, or voicing changed since the recording or not. Even if the reproduction takes place on the same piano and immediately after the recording, the tuning might not be the same anymore and the mechanical reproduction, as good as it might be, does not result in an identical sounding performance as the pianist played it before. This obvious limitation of such devices becomes most evident when a file is played from a different piano in a different room. Especially, if the damping point (the point of the right pedal where it starts to prevent the strings from freely oscillating) is a different one on another piano, the reproduction could sound too 'wet' or too 'dry'. One possible solution to this problem could be a reproducing device with 'ears', in other words, the piano should be able to control its acoustical output via a feedback loop through a built-in microphone. If put into a different room, the device could check the room acoustics, its pedal settings, and its current tuning and voicing before the play-back starts, much the same as a pianist warming up before a concert. Such a system would require a representation of loudness or timbre other than MIDI velocity, indicating at what relative dynamics a certain note was intended to sound in a pianist's performance.

The Disklavier measured in this study is certainly not the top model produced of the Yamaha corporation. Since then, Yamaha issued the Mark III series and the high end series, called 'Pro' (e.g., the special 'Pro2000 Disklavier'). The latter series uses an extended MIDI format (with a velocity representation using more than 7 bits), and additional measures like key release velocity to reproduce the way the pianist released a particular key. It can be expected that these newer devices perform significantly better than the tested Mark II grand piano. Since these more sophisticated devices were not available for the authors or were too far away from the accelerometer equipment, which was too costly to transport, this has to remain subject for future investigations.

by simply holding down the corresponding keys the same time this pedal was depressed.

5. ACKNOWLEDGEMENTS

This study was supported by the European Union (Marie Curie Fellowship, HPMT-GH-00-00119-02, the *Sounding Object* project (SOb), IST-2000-25287, *http://www.soundobject.org*, and the MOSART IHP network, HPRN-CT-2000-00115) and the START program from the Austrian Federal Ministry for Education, Science, and Culture (Grant No. Y99–INF). The Austrian Research Institute for Artificial Intelligence acknowledges basic financial support from the Austrian Federal Ministry for Education, Science, and Culture. We are indepted to Alf Gabrielsson and to the Bösendorfer company for providing the two instruments, and to Tore Persson and Friedrich Lachnit for maintaining and servicing them. Specials thanks are due to Anders Askenfelt, Simon Dixon, Alexander Galembo, and Erik Jansson.

6. REFERENCES

[1] Werner Goebl and Roberto Bresin, "Are computer-controlled pianos a reliable tool in music performance research? Recording and reproduction precision of a Yamaha Disklavier grand piano," in *Workshop on current research directions in computer music, November 15–17, 2001*, Claudia Lomeli Buyoli and Ramon Loureiro, Eds., p. 45–50. Audiovisual Institute, Pompeu Fabra University, Barcelona, Spain, 2001.

[2] Werner Goebl, Roberto Bresin, and Alexander Galembo, "The piano action as the performer's interface: Timing properties, dynamic behaviour, and the performer's possibilities," in *this volume*. 2003.

[3] Werner Goebl, "Melody lead in piano performance: Expressive device or artifact?," *Journal of the Acoustical Society of America*, vol. 110, no. 1, pp. 563–572, 2001.

[4] Alexander Galembo, "Perception of musical instrument by performer and listener (with application to the piano)," in *Proceedings of the International workshop on Human Supervision and Control in Engineering and Music, September 21–24, 2001*, p. 257–266. University of Kassel, Kassel, Germany, 2001, *http://www.engineeringandmusic.de/*.