Received April 6, 2014; Accepted June 15, 2014

International Journal of Asia Digital Art & Design

Gocen: Appropriating Simplified Handwritten Notation as a Musical Interface

Tetsuaki Baba
Tokyo Metropolitan University
baba@tmu.ac.jp

Yuya Kikukawa
Tokyo Metropolitan University
kikukawayuya@gmail.com

Kumiko Kushiyama
Tokyo Metropolitan University
kushi@tmu.ac.jp

Abstract
Not only in childhood but also adulthood, we need some training to read music scores, which sometimes make music hard to learn and enjoy. In this article, we shall propose the system that enables users to play their handwritten musical notations by our musical interface. Since 1960's, Optical Music Recognition (OMR) has become matured in the field of printed score. In recent, some products were released on a market that uses OMR for music composition and playing. However, few researches on handwritten notations have been done, as well as on interactive system for OMR. We combined notating with performing in order to make the music more intuitive for users and give aids to learn music to users. We set 6 design criteria for evaluation of our musical interface. Through users tests and user observations, we shall report our system has high accurate image recognition for handwritten notation and is success to help users to create music along with our design criteria.

Keywords: handwritten notation, optical music recognition, optical character recognition, interaction design

1 Introduction

We are interested in developing tools that support the enjoyable education and performing process for music. Since so many digital musical instruments have been developed, the interface of electronic musical instrument has been evolved new types, such as gestural, wearable, computer vision, etc.

From the imitation of the classic musical instrument to a totally new one, these kinds of digital instruments provide new possibility for performance and education of music. To learn and performing some piece of music, we generally use staff based score that consists of five horizontal lines and notes. The staff originated from musically annotated text, though the Gregorian Chants around the 12th to 13th centuries. It is a basic and important literacy that not only professional musicians but also beginners can use staff based score for learning, playing and composing. Professional musicians have literacy to read and write the staff. But other than the classical music, some users and semiprofessional users often cannot read and the write score.

Not only in childhood but also adulthood, we need some training to read music scores, which sometimes make music hard to learn and enjoy. In this article, we shall propose the system that enables users to play their handwritten musical notations by our musical interface. By combining notation with performance, our system makes music more intuitive and accessible to enjoy and learning music. Concretely, users can play by scanning their handwritten notation with a kind of scanning device. Users can do intuitive interaction: play, re-play and play backwards. Furthermore, users can make some articulations, such as pitch bending and a vibrato by moving scanning device during playing.

Figure 1 shows sketch image of our system. The user writes simple handwritten score on a paper, and then makes sound by tracing with the device. As mentioned next section, it is difficult to recognize and process handwritten score as an interactive system for the contemporary research field of Optical Music Recognition.

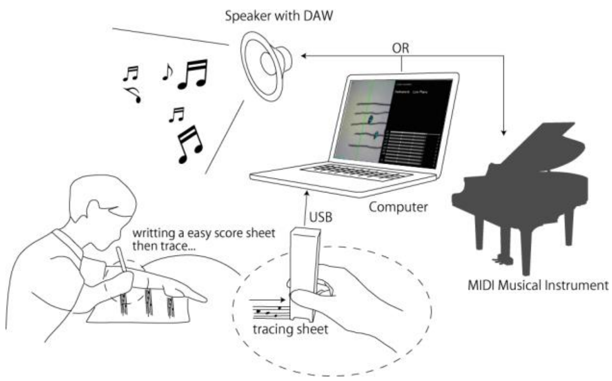


Figure 1 Abstract of our system

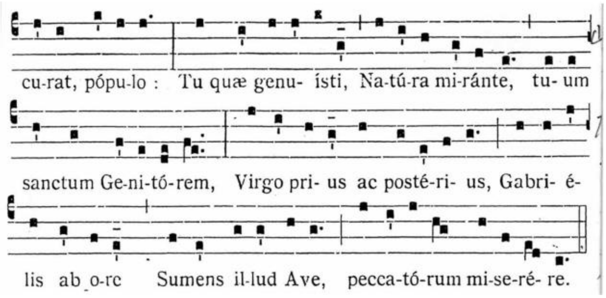


Figure 2 A sample score of Gregorian chant

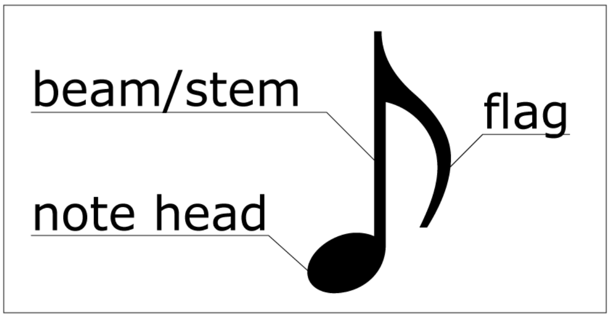


Figure 3 Composition of each part of musical note

In this study, we resolve handwritten recognition accuracy and speed of processing by simplify scores that consists of note head and five line and some accidentals. This kind of score is similar to Gregorian chant. Figure 2 shows sample score of Gregorian chant. Then we design interface and interaction for the system. Note head is a part of note that is shown figure 3. We call this filled part "note head".

2 Related Work

2.1 Optical Music Recognition

Optical musical recognition for handwritten can be roughly classified into two kinds of input methods: online and offline input. Online input is a method that users can input stroke information via pointing devices: touch-pen, mouse, etc. On the other hand, offline input is a method to scan optical information via kind of paper.

In the case of online input, There are several researches for applying stroke information to character recognition system [2,8], Especially Forsbergs [8] is well known as the simplified stroke based interactive system. Whereas, We uses offline input for optical music recognition (OMR).
OMR for offline printed musical scores has been matured and active research field. Since Kassler [14] in 1963, Pruslin [6] in 1966, Prerau [22] in 1970 have published papers on OMR. Up to the present time, OMR research field have been still active. In Japan, Since 1980, Miyao[19], Ohteru[20], Matsushima [18] have reported on OMR. Although many researchers are trying to enhance recognition accuracy, it is still difficult to recognize all kinds of music scores perfectly because of diversity of musical notation and symbols. On the other hand, some techniques of OMR has been applied to commercial applications and OMR libraries for developers: Audiveris [3], OpenOMR [5] and Gamera Framework [7] has been distributed.

As an example for interactive OMR, Kawai has released OMR smartphone application [15], which enables users to play music from musical scores easily. Kawai has also released "PDF Musician" [16] that users can play musical scores by touching and tracing. It provides users with familiar and intuitive OMR interactive system. This system uses text PDF that score information is implemented in. It could not realize musical scores precisely and instantaneously unless users prepare a PDF file that includes text PDF. Yamamoto and et al. [25] proposed musical interface system named "on-Note". They use physical marker less musical scores to play music intuitively. The notes on the score are captured by a camera and are processed by the system that retrieves the music from a score database. In addition, the system can do a real-time recognition of the paper's position and the rotation. Thus, by physically moving and connecting the musical scores we can play music intuitively.

These kinds of researches are similar to ours on the view that users can directly use musical scores for playing music. But our methods are different from the view that users can use own handwritten score and interactive techniques for playing.

"The Music Wand" by Hoerter [12] is most similar to ours. A user can play music by using printed scores for playing with hand-held device. They reported recognition of note position is about 70 % and accidentals is less than note position. In our first prototype, we have used hand-held scanning device to read music score and its recognition accuracy is same as Hoerters. Furthermore, Their system needs a light box to get binalized image and a kind of ruler to support uses interaction, which puts an obstacle for usability.

Many researches are focused on printed scores, but some researches about OMR focused on handwritten score has been also reported since 1970s. Bulis showed that computerized recognition of musical notes can be accomplished by a relatively of comparing horizontal and vertical histograms of symbols. Yadid-Pcht and et al. reported recognition of handwritten musical notes, based on neural network model. And they showed average 80 % recognition rate for note recognition. In the case of interactive system for OMR, we have to develop a solution against recognition rate and responsibility.

2.2 Tangible UI & Media Art

In the field of media art or tangible UI, many researchers reported kinds of new musical instruments. Especially several researches or works are focused on enhancing comprehensibility for music performing and learning as a motif of music score.

Bottelo proposed Tangible user interface for music learning to help children learning the music theory. He developed the software system that enables users to create music by putting objects labeled optical marker on the table [4]. A user can not only put objects but also rotate or extend objects to make change pitch or length of note.

Noteput by Jonas Friedemann Heuer [11] is an interactive music table based on five line staff, which combines all three senses of hearing sight and touch to make learning the classical notation of music for children and pupils easier and more interesting. This kind of timeline based musical interfaces have already well known by Golan Levin [17], Toshio Iwai [23] and et al.

Drawn by Liberman [27] presents a whimsical scenario in which painted ink forms appear to come to life, rising off the page and interacting with the very hands that drew them. Inspired by early filmic lightning sketches, in which stopmotion animation techniques were used to create the illusion of drawings escaping the page, drawn presents a modern update: custom-developed software alters a video signal in real time, creating a seamless, organic and even magical world of spontaneous and improvised performance of hand and ink. Calligraphy is one of art form and it is the supreme art form in China. So we could not separate writing and paper. This kind of primitive and intrinsic interaction between human and paper is important for our creative activity.
Tsandilas and et al. [26] focus on the creative use of paper in the music composition process and proposed Musink to provide composers with a smooth transition between paper drawings and OpenMusic as a flexible music composition tool by using the Anoto pen. In another project, they insisted that composers can use paper effectively on the first stage of composing [9].
Our goal is to develop music performing and learning system that do not use mouse, keyboard and touch panel on a computer by using handheld scanning device and designing user interaction between users and paper.

3 Gocen

3.1 Design criteria

Poupyrev and et al. [22] mentioned design issues for musical controller for good design of musical interface on workshops, chi2001. Therefore we also same criteria for evaluating our interface. Furthermore we adds one more criteria "uniqueness" to theirs, which means single interface must control all interactions. We aim to develop a system that do not require the legacy and generic interfaces such as pc keyboard, mouse, touch pen and touch pad.

- Usability and comprehensibility
- Expressiveness
- Sensitivity and sophistication
- Aesthetics
- Hedonics
- Uniqueness

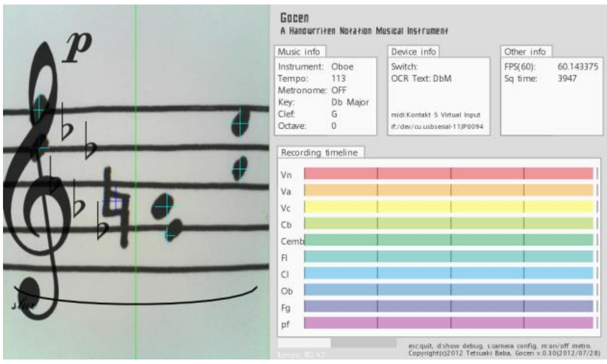


Figure 4 A screenshot of our software. Left captured image shows camera image on which is overlaid musical information. Right panel shows text information, timeline and more.

3.2 System abstract

Our system mainly consists of a scan device, computer and sound module. A user can play simple music by tracing notes with the scan device as shown in figure 1.

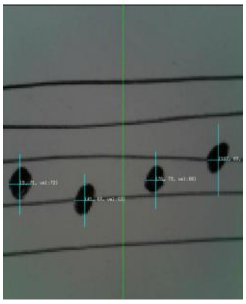


Figure 5 A screenshot showing recognized notes

The computer processes captured images by using OpenCV

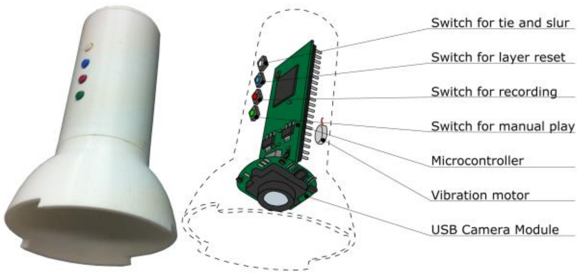
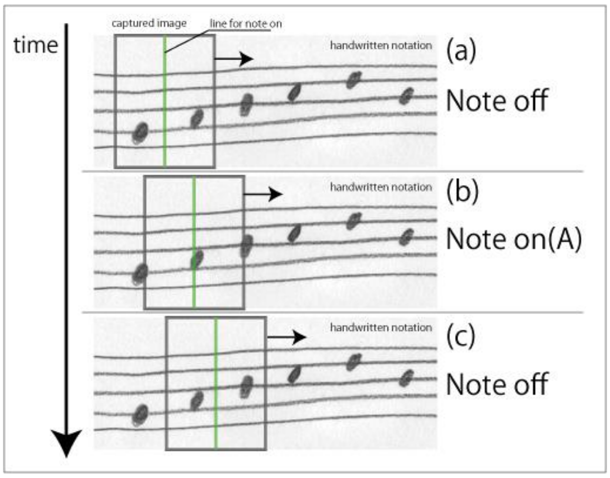


Figure 6 Gocen device consisting of USB camera, microcontroller, switches and vibration motor

and our algorithm at 30fps, then outputs sounds according to the data from the notations.

We do not need any special materials other than this system. A user can use normal white paper and his/her own pen. Our device is built with a USB camera, microcontroller, and vibration motor (see figure 5). The vibration motor is used for tactile feedback while the user is playing.



Our software is made on some libraries: openFrameworks[21], OpenCV, OpenGL, portmidi, Ocrad[1]. The system can select

Figure 7 Flow to make a note on/off

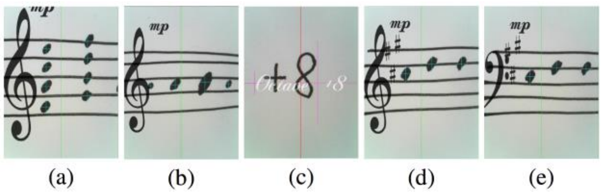


Figure 8 Example images of interaction. (a)Chord, (b)velocity, (c)octave, (d)key in A Major, (e)bass clef

Figure 9 Example of score of "Twinkle, twinkle, little star" that is shown to users.

Table 1 Result of the user test for pitch recognition.

trial	participants	total	correct(Ave.)	accuracy
1	21	42	40.9	97.4 %
2	9	42	42	100 %

any midi devices. On this stage, we use Kontakt Player[13] for MIDI device. Figure 4 shows screenshot of our software.

Our system can be roughly classified into two kinds of processing. One is OMR processing which is mainly control performing. Another is OCR processing which is a mainly control setting, such as key, key range, instruments and etc. We uses 3rd party library for OCR function. Therefore we do not refer to recognition accuracy about it.

3.3 Interaction

Our system is not only an OMR system but also performance system. We developed several musical interactions for this interface.

Note on/off

First of all, a user write 5-line staff and note head on a paper. Then a user holds the handheld device and put it on the paper. A scanned image will be shown on the screen. Figure 6 is a screen shot of a image that shows each position of note head as cross hairs. The vertical size of a cross hairs shows velocity of the note. Figure 4 shows a screenshot of application on PC screen. A user can confirm some information: selected instrument, options, timeline for recording and so on. While pushing green button on the device, a user can play each note head by moving left/right to right/left over a note head. Figure 7 shows flow to make a note. To place green bar on the screen upon a note, a user can make a sound.

As a result, these kinds of operation enable uses to ignore the information of musical note length. No matter how a note flag shows sound length. A user can move the scan position with the handheld device by his/her hand manually. That's why we simplified the staff-based musical notation.

Chord

Our system can recognize a chord up to 32 notes at the same time. Figure 8(a) shows example notes of a chord.

Velocity

It is poplar to use kind of symbols: such as pp, mp, mf, ff and etc. for instructing velocity of a note. For example, a user can change velocity into mezzo piano to put the device on a word mp that a user wrote on a paper. Whereas, a user can also control velocity with the vertical size of a note. Larger note makes louder sound. Smaller note makes small sound. Figure 8(b) shows that our system changes velocity depending on the vertical size of a note.

Pitch Bending

In the case of kind of string instrument, such as violin, musician expresses articulations: vibrato, turn and etc. techniques. In our system, a user can use pitch bending by moving the device up or down after making a note on. The volume of pitch bending is defined as $|Y_1 - Y_2|$. Y_1 is first vertical position of a note on. Y_2 is the vertical position of the phonetic note. In the case of $(Y_1 - Y_2 < 0)$, the pitch is increased.

Tonal Range

The default tonal range is set from C4 to A5. It can be extended by writing an additional line or octave mark. In our system, a user can select the tonal range to read some characters such as "+15", "+8", "0", "-8", "-15" with OCR function (see Figure 8(c)).

key

The default key is C major. To change the key of score, a user put the device on handwritten characters, "AM", "Am", "BM", and etc.

Key Transition

A user can make key transition by operating note on/off after read with OCR function.

Changing a instrument
A user can change the instrument of sound by covering a text he/she wrote to indicate the name of instrument, such as pf(piano), bs(bass), gt(guitar), dr(drums), etc., with the device, by means of OCR.

Changing the kind of clef
A user can change the kind of clef to read “C.F” or “C.G” with the device. Figure 8(e) shows a status on clef F or bass clef.

Loop sequencer
A user can record sound events into timeline, while pressing recording button and make a loop like a sequencer. Each recorded note will be set in the quantized timeline.

Figure 10 Example of scores that users wrote.

Table 2 Result of the user test for accidental recognition.

	1st trial			2nd trial		
user	**sharp**	**flat**	**natural**	**sharp**	**flat**	**natural**
1	73 %	77 %	90 %	97 %	93 %	100 %
2	70 %	90 %	100 %	87 %	90 %	100 %
3	90 %	30 %	100 %	100 %	97 %	93 %
4	87 %	10 %	100 %	100 %	37 %	97 %

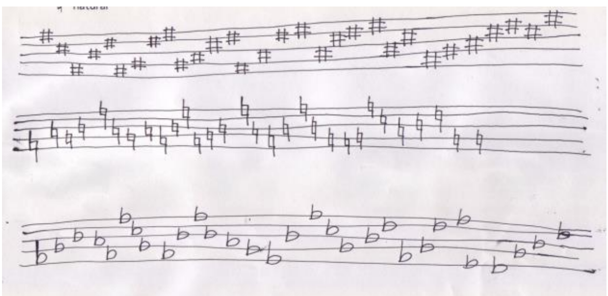


Figure 11 Example of scores that users wrote.

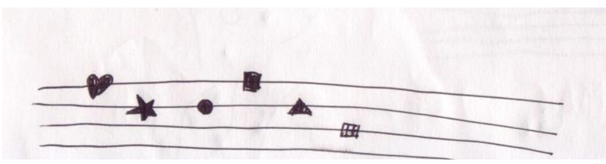


Figure 12 A user wrote no musical symbols

4 User Study

4.1 User test for recognition accuracy of pitch

We find how users can play a correct pitch of note they want by simple user test. We recruited 19 participants (9 male) from our university students who has experience of notation: “almost everyday”: 0, “often”: 3, “rarely”: 8, “almost nothing”:8, Age average is 21.9.

Before the experiment, we distributed a experiment paper and instruct about our system. After that, we let participants write a simple piece (twinkle twinkle little star) and play it. An example score had been already printed on experiment paper (see figure 10). It is made on DTP software because an example score might affect users' writing. Below is the procedure of experiment.

1. Instruction (5min)
2. The user writes and plays their handwritten notation.
3. If user finds some mistake about a pitch, the user modifies their score and plays it again.
4. If user finds some mistake, goes back to 3.
5. Free description questionnaire

We provided two kinds of pen. One is 0.7mm oil base ballpoint pen. Another is 0.5mm oil base ballpoint pen. Participants can chose the pen that they want. We also provided participants with general plain paper. All number of notes of the score is 42 which is shown figure 9.

10 users could play all notes correctly on the first trial. After modification, all other 9 users could play all notes correctly. Table 2 shows the result. Average recognition rate on the first trial is 97.4 % and 100 % on the second trial. Figure 11 shows sample sheet that users wrote on their paper. Users answered almost affirmative response about our system such as fun, interesting, this is an epochmaking, better to commercialize, great, If I had this system in childhood, I could practice piano harder, and etc. On the other hand, we got dissatisfaction response such as, it need precision for interaction, it makes better if I could use colorful pens, difficult at first.

4.2 User test for recognition accuracy of accidentals

We developed a contour-based CV algorithm with Support Vector Machine. Fig.13 shows an example flow of a sharp accidental symbol from an original image. We use edge detection features for machine learning algorithm.
We recruited 4 participants (2 male) from our university students. Before the experiment, we distributed a experiment paper and instruct about our system. After that, we let participants write 30 symbols for each symbols (sharp, flat, natural). An example score had been already printed on experiment paper (see figure 9). It is made on DTP software because an example score might affect users' writing. Below is the procedure of experiment.

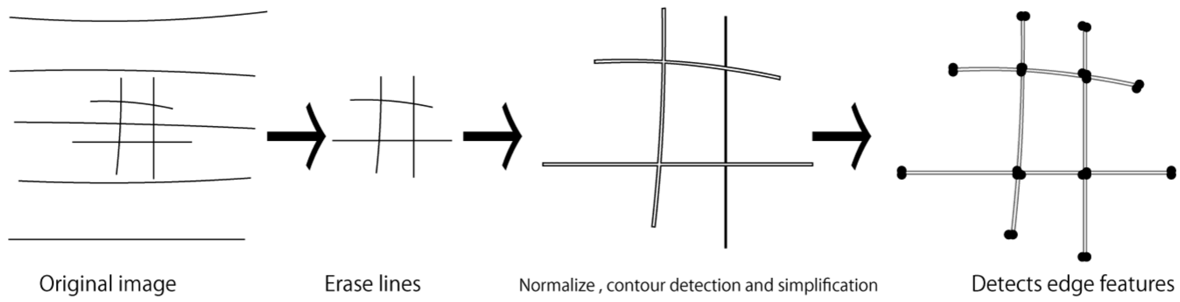


Figure 13 Our accidental symbol detection algorithm.

1. Instruction (5min)
2. 1st trial: The user writes 30 handwritten notations for each accidental symbol.
3. An experimenter gives him/her some advices about how to write recognizable.
4. 2nd trial: The user writes 30 handwritten notations for each accidental symbol.

On 1st trial, we showed participants printed accidental symbols. Figure 11 shows sample sheet that a participant wrote.

We had several exhibitions on domestic events and a international conference. Through exhibitions we have observing users carefully. Score notes that users wrote has 90 pages and over 300 users have experienced our system. A user draw symbols that is not musical symbol for his creative expression. Figure 12 is the image the user drew. We did not mention about other kind of symbols. This user showed this image to us, then suggested to implement to recognize these symbols. Many users spent more time for writing notations than OCR. Although we prepared many OCR settings, many users enjoyed writing notation and playing.

5 Discussions

Through the user study, we found many users can use easily our system whatever they has literacy to read/write music scores. On the user test, we found expected high recognition accuracy of pitch and accidentals, because of simplification of handwritten notation.

5.1 Validity of Design Criteria

We shall review the evaluation of our design criteria on the basis of the results of the user test and observation.

Usability and comprehensibility

Many users could play and enjoy with our system as soon as we told them how to play, because our common mental image of music notation help us to understand our system easily. Through exhibitions, we found that users who are around 6 to 60 years old enjoyed our system. Concerning with usability, especially design of our prototype device should be taken consideration. Several users could not get know how to hold our device, because of its no intuitive design. And we found tangled cables of the device interfered user interactions.

Expressiveness

As we mentioned before, our system has 10 interactions for playing. By using these, it enables users to play many kind of musical piece. Moreover we showed a possibility to extend our system by ensemble. But we have to verify the responsibility of our system carefully on future work. Musical score has more expressiveness than our system. For example, our system could not recognize two columns of 5-line staff. We have to continue develop the system constantly, but our system has achieved to a certain degree for expressiveness.

Sensitivity and sophistication

As a result that we developed specific OMR processing procedure, we implemented fast interaction with 30 fps. During the user test, one participant told us it is surprising that this system processes faster than he expected. At a recognition accuracy point of view, we showed our system has enough precision as interactive system.

Aesthetics

We have to continue to consider design about the shape of prototype carefully in future. We found some users often play music that they create with our system, it is caused that our system could provide users with creativity about music. Figure xx also shows that our system is success for providing users with contingency that is important human activity for art creation. In near future, we are going to create sophisticated performance to show how our system can make a attractive performing.

Hedonics

We could observe some children or users who can read/write score plays our system 30 minutes and more. We got comments about tactile feedback such as it feels good, feels like I playing

Uniqueness

We designed that a user can use all interactions with only one device. A result that we could not get any unsatisfied comment about from users shows a success for offering uniqueness of a device.

6 Conclusions

There are many researches about OMR. But few researches on interactive handwritten notation have been done. We proposed the system that enables a user to play their handwritten notation on a paper with our device by scanning. Our system can recognize handwritten note head and accidentals with high accuracy by simplifying musical notations. We could implemented not only accurate recognition system but also fast processing (30fps).The FPS except for capturing image is about 150 [fps]. It is possible to improve operability by improvement of computer and camera specification.

References

[1] Ocrad - the gnu ocr -. www.gnu.org/software/ocrad/ (last accessed, Jun, 2012).

[2] Anstice, J., Bell, T., Cockburn, A., and Setchell, M. *The design of a pen-based musical input system.* In Proceedings of the 6th Australian Conference on Computer-Human Interaction (OZCHI '96), OZCHI '96, IEEE Computer Society (Washington, DC, USA, 1996), 260–.

[3] Bitteur, H. Audiveris: *Open music scanner*, 2000 – 2012. http://audiveris.kenai.com/(last accessed, Jun, 2012).

[4] Bottello, L. *Tangible user interface for music learning*, 2010. http://cloverleafdesign.weebly.com/(last accessed, Jun, 2012).

[5] Desaedeleer, A. Openomr. http://sourceforge.net/projects/openomr/(last accessed, Jun, 2012).

[6] D.H.PRUSLIN. *Automatic Recognition of Sheet Music*. PhD thesis, Doctoral Thesis MIT Cambridge MA January 1967, 1967.

[7] Droettboom, M., Fujinaga, I., MacMillan, K., Chouhury, G. S., DiLauro, T., Patton, M., and Anderson, T. *Using the gamera framework for the recognition of cultural heritage materials.* In Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries, JCDL '02, ACM (New York, NY, USA, 2002), 11–17.

[8] Forsberg, A., Dieterich, M., and Zeleznik, R. *The music notepad*. In Proceedings of the 11th annual ACM symposium on User interface software and technology, UIST '98, ACM (New York, NY, USA, 1998), 203–210.

[9] Garcia, J., Tsandilas, T., Agon, C., and Mackay, W. *Interactive paper substrates to support musical creation*. In Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems, CHI '12, ACM (New York, NY, USA, 2012), 1825–1828.

[10] George, S. E. *Clustering on-line dynamically constructed handwritten music notation with the self-organising feature map*. In Proceedings of the 16th international conference on Developments in applied artificial intelligence, IEA/AIE'2003, Springer Springer Verlag Inc (2003), 93–103.

[11] Heuer, J. F. *noteput is an interactive music table with tangible notes, that combines all three senses of hearing, sight and touch to make learning the classical notation of music for children and pupils more easy and interesting*. http://www.jonasheuer.de/index.php/noteput/(last accessed, Jun, 2012).

[12] Hoerter, N., and Chatt, T. *The music wand: Real-time optical scanning of sheet music*, 2008. http://people.ece.cornell.edu/land/courses/ece4760/FinalProjects/s2008/tjc42_nah35/tjc42_nah35/index.html(last accessed, Jun, 2012).

[13] Instruments, N. Kontakt player. http://www.native-instruments.com/(last accessed,Jun, 2012).

[14] Kassler, M., and of Music, P. U. D. *An Essay Toward Specification of a Music-reading Machine...* Princeton University, Department of Music, 1963.

[15] Kawai Musical Instruments Mfg. Co., L. Gakufu camera, 2011.

[16] Kawai Musical Instruments Mfg. Co., L. Pdf musician, 2012.

[17] Levin, G., and Lieberman, Z. *Sounds from shapes: audiovisual performance with hand silhouette contours in the manual input sessions.* In Proceedings of the 2005 conference on New interfaces for musical expression, NIME '05, National University of Singapore (Singapore, Singapore, 2005), 115–120.

[18] Matsushima, T., Kanamori, K., and Ohteru, S. *Automated recognition system of musical score (vision system of the wabot-2)*. Journal of the Robotics Society of Japan 3, 4 (1985), 354–361.

[19] Miyao, H. *Symbol recognition for printed piano scores based on the musical knowledge* (in japanese). The transactions of the Institute of Electronics 75, 11 (1992), 1848–1855.

[20] Ohtheru, S., Matsushima, T., and Kanamori, K. *Automatic recognition of printed music*(in japanese). journal of Acoustic Society of Japan 41, 6 (1985-06-01), 412–415.

[21] openFrameworks community. openframeworks. www.openframeworks.cc (last accessed, Jun, 2012).

[22] Poupyrev, I., Lyons, M. J., Fels, S., and Blaine (Bean), T. *New interfaces for musical expression*. In CHI '01 extended abstracts on Human factors in computing systems, CHI EA '01, ACM (New York, NY, USA, 2001), 491–492.

[23] Prerau, D. S. *Computer pattern recognition of printed music*. In Proceedings of the November 16-18, 1971, fall joint computer conference, AFIPS '71 (Fall), ACM (New York, NY, USA, 1971), 153–162.

[24] Toshio, I. Chess on the table, 1997.

[25] Tsandilas, T., Letondal, C., and Mackay, W. E. *Musink: composing music through augmented drawing.* In Proceedings of the 27th international conference on Human factors in computing systems, CHI '09, ACM (New York, NY, USA, 2009), 819–828.

[26] Yamamoto, Y., Uchiyama, H., and Kakehi, Y. *onnote: playing printed music scores as a musical instrument*. In Proceedings of the 24th annual ACM symposium on User interface software and technology, UIST '11, ACM (New York, NY, USA, 2011), 413–422.

[27] Zachary,L. Drawn. http://thesystemis.com/projects/drawn/ (last accessed, Jun, 2012).