Automatic Generation of Piano CG Animation: Real-time Rendering of Human Skin Appearance and Humanizing Playing Motion

Sho Muramatsu*
Kensuke Tobitani*
Junya Hirata*
Hiroyoshi Miwa*
Noriko Nagata*

We are promoting research into the analysis and CG expression of realistic and natural piano-playing. This paper proposes a system that automatically generates an interactive animation of a pianist playing the piano from a musical score in real-time. In the first step, a generation of realistic piano performance animation from a musical score, this system automatically defines the fingering, optimizes the playing motion, and generates bony framework animation by sending time-series MIDI note number data to the system. Furthermore, this system can generate a variety of playing motion styles, such as amateur and professional, by setting the stiffness of each joint as an optimization cost. Setting the stiffness is based on knowledge that pianists with superior skill reduce their finger muscle load during keystrokes. In the second step, a generation of realistic piano performance animation from a musical score, we utilize pre-computed textures as a means to implement a realistic real-time skin shade. For directional lighting, we pre-compute textures corresponding to the direction of each light source in mapping them to 3d modeling in real-time. Finally, we generated realistic piano performance animation from a musical score by the system implementing automatic generation of the piano-playing motion and pre-computed GI textures.

Keywords: Real-time, Animation, Global illumination

1. Introduction

Technologies generating piano performance in the form of CG animation are eagerly anticipated in various fields, such as content production and music education. However, much of the past research were from musical and kinematic viewpoint, structure of hand and hand motion (1)~(3). Research that has focused on the piano performance is scant.

We could have been promoting research into the analysis and CG expression (4) of realistic and natural piano playing. Our goal is to generate "the virtual pianist system" that is 3D CG character. When a user enters the musical score data into this system, it plays the piano automatically.

There are two technical issues to achieve this goal. The first issue is to improve the texture of human skin at rendering. The second issue is to implement piano performance like a human.

In rendering, we developed global illumination (GI) and sub-surface scattering (SSS) (5) to express the skin texture. It is necessary to use a GPU (graphics processing unit) to implement high-quality and high-speed rendering for real-time processing. In our previous generation method of piano performance, we measured human motion using a motion capture system. This method had drawbacks in price, time, and data capacity. We have proposed a method for calculating the fingering of piano music from data. Some problems are human likeness of appearance and motion and piano performance differences in piano skill.

* Kwansei Gakuin University 2-1 Gakuen, Sanda, Hyogo 669-1337, Japan
sho.m@kwansei.ac.jp

2. Overview of piano CG animation

We convert a musical scores into MIDI as input data for the virtual pianist system. Alternatively, we could use piano performance data acquired by a motion capture system. There are advantages to data input from each method. The advantage of the automatic generation of piano performance by using the score is that it does not require a pianist and expensive measuring equipment. On the other hand, the advantage of the motion capture system is that it can better reproduce human movement.

In this study, we employed the method of automatic generation of piano performance using the music score, and determined the piano fingering using obtained MIDI data. Next, we automatically generated piano performances by the optimization method. Last, we provided the simulated piano performance to a hand model, and generated piano CG animation. Figure 1 shows our method of generating piano CG animation. In the procedure, we propose new methods to generate pre-

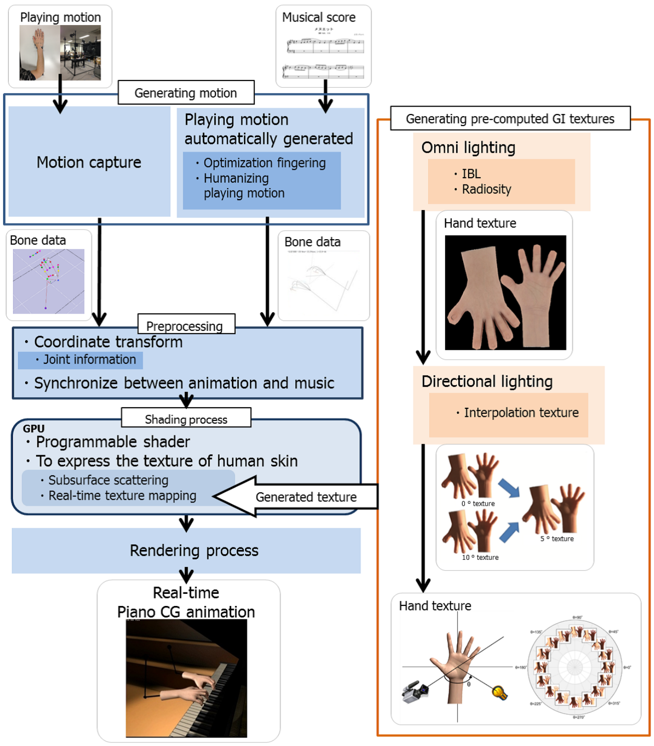

Fig. 1. Overview of piano CG animation.

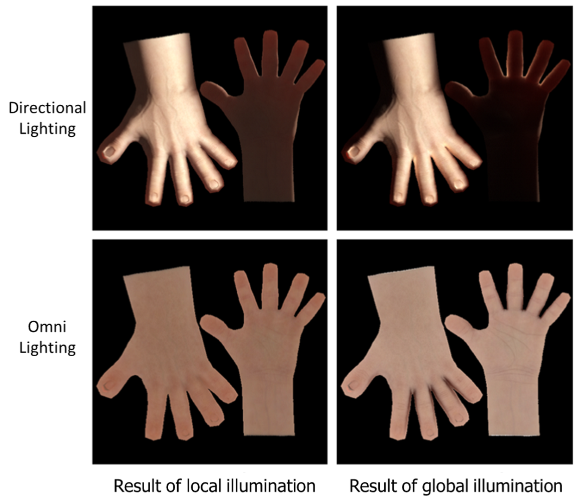

Fig. 2. Obtained pre-computed GI textures.

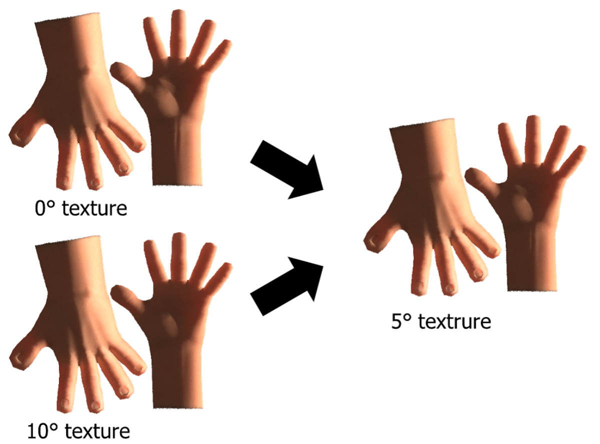

Fig. 3. Overview of interpolation texture.

computed GI textures and to humanize generated piano playing motion. These methods are described in Section 3 and 4.

3. Real-time Rendering of Human Skin Appearance

Implementing real-time rendering of a high quality model places a high load on the GPU and CPU. Therefore, we gave real-time technology priority, like a previous study [6]. On the other hand, research of texture is not enough. We propose a pre-computed GI texture to express the texture of human hand skin for a model quality improvement in real-time, and to express comparison of piano performance difference. without applying a high load to map the hand model in real-time.

3.1 Generation of pre-computed GI texture

This method performs SSS simulation and pre-computes lighting to generate high-quality hand model skin textures.

3.1.1 Pre-computed lighting By using omni-directional lighting, it is possible to remove the shadow of the complex surface of a hand model; therefore, it is possible to generate a high-quality texture without shadows. We utilized the radiosity method to calculate GI. First we used the image based lighting (IBL) [7] to unify the strength of the lighting. Then, we used directional lighting on the hand model that used textures generated by omni directional lighting. We applied lighting to the hand model to 360 degrees from 1 degree in the horizontal direction to generate the textures.

3.1.2 Data volume reduction through interpolated texture Obtaining the texture that can correspond directional lighting requires at least pre-computed textures. To optimize the CG interactivity, we seek to minimize the amount of data processed in real-time. This method allows us to generate an interpolated texture from the two pre-generated textures, and thus reduce the amount of data. Figure 2 shows obtained pre-computed GI textures using omni directional lighting and directional lighting on local illumination and global illumination.

Interpolated textures are obtained by the alpha blending the two textures. For example, if the sampling was carried out at intervals of 10 degrees, blending the textures obtained from the lighting environment at 0 degrees and 10 degrees, it is possible to obtain artificial textures from 5-degree lighting. In this case, compared to an interval of 1 degree, the size of the texture file becomes one-tenth. Figures 3 shows an overview of the interpolated textures.

3.1.3 Making the model correspond to the light source color In our previous pre-computed GI in real-time, there was an issue that changes in the light source color were not reflected on the model. By changing the ratio of the light source color, the original

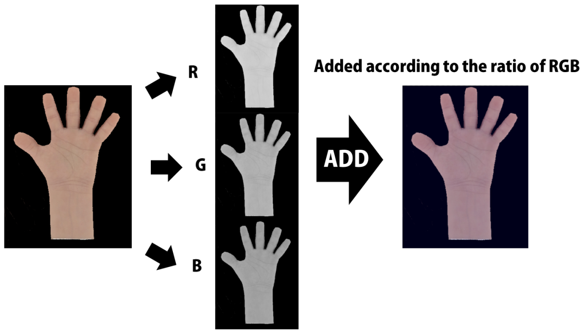

Fig. 4. Changing the RGB ratio of the textures..

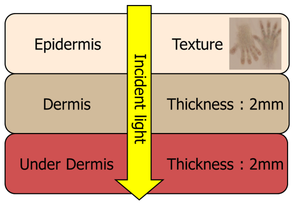

Fig. 5. Conceptual diagram of SSS shader.

texture's color can be changed. If there is a change in the light source color, the system must obtain the ratio of its RGB, and match the RGB ratio of the textures to it; therefore, it is possible to correspond to real-time changes in the light source color. Figure 4 shows how we changed the RGB ratio of the textures.

3.2 Subsurface Scattering We adopted a subsurface scattering (SSS) technique to express the most realistic skin texture possible. This technique is based on a physical phenomenon that the incident light reflect multiple times at a translucent object ,then the light is emitted from a different point. This technique is utilized to express an object with a sense of transparency, such as human skin. As shown in Figure 5, we used a simulation of SSS to consider a complex incident through the Radiosity algorithm; therefore, making it possible to obtain an accurate shading.

3.3 Real-time mapping Here we sought to utilize the pre-computed GI texture generated in Section 3.1 to respond to changes in the light environment in real-time rendering. In this method, we compute the angle from the model location and light source location, and utilize the textures that match the model; therefore, allowing us to implement real-time texture mapping by using pre-computed GI textures. Figure 6 shows the results of real-time rendering with the pre-computed GI texture.

4. Automatic generation of piano performance using a musical score

4.1 Previous work of automatic generation of piano performance We are now developing a system that automatically generates an animation of a pianist playing a piano from a musical score. This system is mainly composed of the function determining an optimal piano fingering and the function determining the trajectories of all control points in the bony frameworks of hands based on the fingering. It is necessary to design efficient algorithms to realize these functions.

First, we deal with an optimal fingering method. For the sake of simplicity, we restrict our attention to the problem of determining an optimal fingering of a right hand for a musical score that contains no chords (its notes are played one at a time) or rests (silences). We number the fingers from one to five, starting with the thumb and ending with the pinkie, and number all notes. We refer to motion $(j, \tilde{j}; i, \tilde{i})$ as the transition from the situation (i, j) that finger i is on note j to the situation $(\tilde{i}, \tilde{j})$ that finger $\tilde{i}$ is on note $\tilde{j}$. As the structure of bones and tendon strength of a hand determine the difficulty of a motion, we define the difficulty of motion $(j, \tilde{j}; i, \tilde{i})$ as a positive real value $c(j, \tilde{j}; i, \tilde{i})$; the lower the value of a motion is, the easier one can play the motion. For the purposes of this research, we determined these values based on our own piano skills.

For a musical score, that is, a sequence of notes *sp* (p = 0, 1, 2, ..., n), we define the optimal fingering for the score as the sequence of the motions to minimize the sum of the costs of all the motions contained in the score. By using a dynamic programming method, we can easily solve this optimization problem. In general, we can use another appropriate objective function as the cost of a fingering. We evaluated this method by comparing the fingering generated by it with the fingering of a person with experience. As a result, the difference in costs between these fingerings was small.

Next, we deal with a trajectory control method. We assume that the control points move so as to reduce the load of fingering to a hand. We refer to a position (i, j) as the set of coordinates of all control points for the situation (i, j). For motion $(j, \tilde{j}; i, \tilde{i})$, we determine a position $(\tilde{i}, \tilde{j})$ to minimize the weighted square sum of the difference between the position (i, j) and the position $(\tilde{i}, \tilde{j})$, and the difference between the position $(\tilde{i}, \tilde{j})$ and the prescribed natural position of the situation $(\tilde{i}, \tilde{j})$. As it is difficult to get the global optimum solution of this problem, we derived a solution through a heuristic algorithm. Then we generate all key frames by interpolating linearly between positions (i, j) and $(\tilde{i}, \tilde{j})$. Thus, the trajectories of the control points are determined.

In the subsection, we describe the method that applies humanization processing to automatically generated piano performance.

4.2 Humanization processing Humanization processing is to change monotonous motion to human-like movement. In the robotics field, robots that mimic the human skeletal muscle have been developed to perform human-like behavior. Moreover, in the CG field, the importance of human facial expression is recognized [8] [9]. On the basis of research for the purpose of piano skills improvement, we implement the humanization process to piano performance.

For the control of human likeness of motion, we set parameters that are the cost of each joint and the default position of hand model.

Fig. 6. Result of real-time rendering.

4.3 Playing motion differences in piano skills By comparing the muscle activity of professional and amateur pianists, Furuya et al. identified muscles and joints used in piano performance (10). In particular, they showed that the professional pianist uses a proximal rather than a distal joint to reduce the amount of peripheral muscle load. Based on this information, we implement in the piano CG animation the difference in proficiency by setting the parameters of each joint.

5. Results and Discussion

5.1 Rendering result Table 1 shows our research rendering environment. Figures 6 and 7 show our rendering images.

By using the pre-computed GI textures, it became possible to express more realistic skin texture than existing piano CG animation, and to keep around 60 fps at real-time rendering. Moreover, if a user changes the hand model, it is possible to regenerate the skin texture by recalculating the lighting and SSS simulation though the existing skin textures corresponds only to the specific model. Therefore, our skin textures have a high versatility. Furthermore, this approach can be applied to various CG objects.

As shown in Figure 8, we able to generate piano performance with high human likeness of appearance and motion automatically. Moreover, it is possible to express amateur and expert performances by the humanizing process. By comparing between a good and bad performance, it is easy to understand the difference between them; therefore, the system implemented our methods can help to teach piano performance. Moreover, students can see the piano performance from various angles because the animation can change the position of the camera.

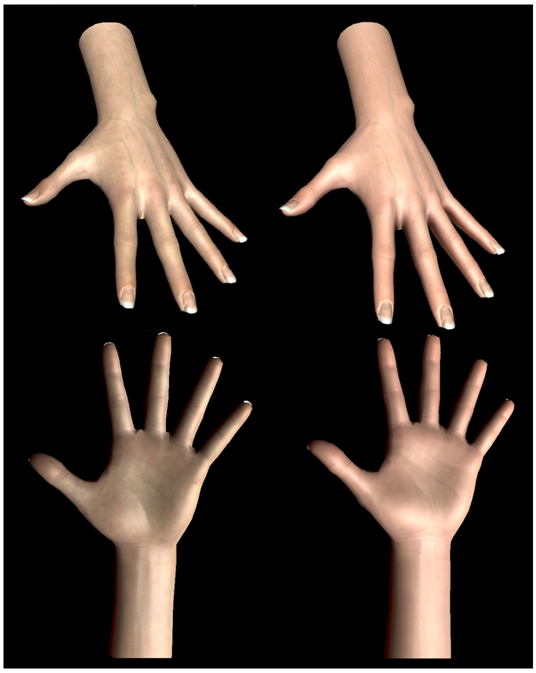

Fig. 7. Comparison of original textures and pre–computed GI textures.

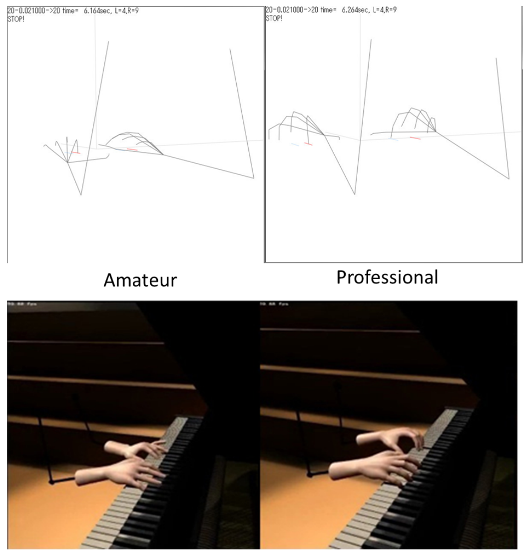

Fig. 8. Upper figure: comparison of amateur and professional bone frame animation. Lower figure: comparison of amateur and professional 3DCG animation.

Table 1. Rendering environment.

CPU	Intel(R) Core(TM) i7 CPU 2.80GHz
GPU	NVIDIA GeForce GTX 480
OS	Microsoft Windows 7 Professional
Memory	8GB RAM
DirextX version	DirectX 11.0
Resolution	1440 × 900

6. Conclusion

In this research, we significantly improved our piano CG animation's skin texture by implementing pre-computed GI textures. Moreover, we were able to transform a mechanical piano performance into something more like a human piano performance and to express piano performance differences in skill by humanization processing. Future work will include the following: comparing the motion by motion capture and by our automatic generation of piano performance from a musical score; using our pre-computed GI texture with various CG objects, and modeling other body parts for the virtual pianist.

References

(1) J. Lin, Y. Wu and T. S. Huang: "Modeling the constraints of human hand motion. Human Motion", Proceedings. Workshop on,vol. 2, no. 1, pp. 52-77 (2000).

(2) K. Funahashi, T. Yasuda, S. Yokoi and J. Toriwaki: "A Model for Cooperative Manipulation of Virtual Objects with Both Hands", Trans. IPSJ, vol. 36, no. 5, pp. 1334-1342 (1998).

(3) G. Elkoura and K. Singh: "Handrix: Animating the Human Hand",ACM SIGGRAPH2003. 110-119 (2009).

(4) N. Kugimoto, K. Takai, R. Miyazono, K. Omori, T. Fujimura, S. Furuya, H. Katayose, H. Miwa, and N.Nagata: "CGAnimation for Piano Perfor- mance", ACM SIGGRAPH2009, Aug. 3-7 (2009)

(5) M. Hasan, M. Fuchs, W. Matusik, H. Pster and S. Rusinkiewicz: "Physical reproduction ofmaterials with specied subsurface scattering",ACM SIG- GRAPH 2010 papers, SIGGRAPH '10, New York, NY, USA, ACM, pp. 61:1-61:10 (2012)

(6) M. Oshita and K. Suzuki: "Artist-Oriented Real- Time Skin Deformation Using Dynamic Patterns", Arts and Technology, pp 88-96 (2009)

(7) P. Debevec: "Image-Based Lighting", IEEE Comput. Graph. Appl., Volume 22 Issue 2, Page 26-34 (2002)

(8) S.Villagrasa and A. S. Sanchez: "FACe! 3D Facial Animation System based on FACS", IV Iberoamer- ican Symposium in Computer Graphics - SIACG 2009, pages 203-209 (2009)

(9) R. Bowden, P. Kaewtrakulpong, and M. Lewin:"Jeremiah : the face of computer vision", SMARTGRAPH '02, pages 124-128 (2002)

(10) S. Furuya, T. Goda, H. Katayose, H. Miwa and N. Nagata: "Distinct interjoint coordination during fast alternate keystrokes in pianists with superior skill", Frontiers in Human Neuroscience, 5:50 (2011)