Computer Animation of Pianist's Hand

Junhwan Kim

Ecole Polytechnique

junkim@poly.polytechnique.fr

Abstract

In this paper we present an algorithm for the animation of the pianist's hand and the determination of piano fingering based on the animation. The animation shows feasible movement of hands and fingers. In general, the result of the fingering determination accords with the written fingerings. A brief description of the system and the results are presented.

Keywords: hand, animation, piano, fingering

1. Introduction

The grasping motion has attracted many researchers in the human hand animation field [Rijpkema91,Mas94,Lee95,Douville96]. In this paper, we will show the animation of the human hand that plays the piano. It is automatic in the sense that it requires only the score, not the fingerings for the score, which will be determined by search method. Thus it can be used as another approach other than the rule-based determination of piano fingering [Sayegh89,Parncutt97,Viana98].

For the given score to play, the search engine generates a candidate fingering, which is taken over to the evaluation module. The evaluation module takes the piano model, the hand model, the score, and the fingering and evaluates the feasibility of that fingering. The evaluation module consists of three parts. First, the states of each finger are determined, second, the tip position of the fingers in non-idle state, or fingers that are engaged in touch, are determined. Finally, the position/orientation of wrist and the joint angles of idle fingers are calculated. If the fingering is feasible, the animation module takes it to show the playing. If not, the output value of the evaluation module is used in the search engine to generate another fingering to be evaluated. This iteration continues until either it finds a feasible fingering or it exhausts all the candidate fingerings. The animation module is virtually same as the evaluation module except the fact that it displays the results of the evaluation.

2. The searching strategy for fingering

The input of the search algorithm is a score without fingering. In a score, the attributes for each key touch include start time, end time, intensity, and liaison with the next one. The aim of the search is to find the fingering which will generate feasible movement of hands and fingers to play the score in the sense that the accumulated sum of the evaluation function in the course of playing is sufficiently low, if not lowest. Our strategy is the best-first search [Winston84] as in [Huang96]. Each node in the search tree represents a specific fingering of the same length as the depth of the node. The choice of the node to expand for further search depends on the evaluation function which assesses the promise of the node, which is defined in our search as

$$f(node) = f(node)\big|_{\text{up to present}} + f(node)\big|_{\text{future}}, \text{ where}$$

$$f(node)\big|_{\text{up to present}} = \int_{\text{start}}^{\text{present}} Evaluation(fingering, t)\,dt$$

$$f(node)\big|_{\text{future}} = \left|\text{key touch remaied}\right| \times \overline{\int_{1\,\text{key touch}} Evaluation(fingering, t)\,dt}$$

The $Evaluation(fingering, t)$ indicates the appropriateness of the hand for the given fingering at time t, which will be explained in section 5.1. The estimated average of the sum of the evaluation function for one key touch, or $\overline{\int_{1\,\text{key touch}} Evaluation(fingering, t)\,dt}$, is predefined.

3. Finger state and non-idle finger tip position determination

This module is responsible for the determination of the state of each finger and the tip position of the fingers in the non-idle state for each time for given score, fingering, hand model and piano model. A score contains start time, end time, intensity, and

Published by the Eurographics Association ISSN 1017-4656

liaison for each key touch. The *liaison* of a key touch is a measure of smoothness of the key touch and the next one. The liaison value is set as 1.0 and 0.0 for *slur* and *staccato* respectively. Otherwise, it is set to 0.5. A key touch is modeled as described in Figure 1. The height of the fingertip is plotted against time axis. For a given touch, start time of each state is determined by the heuristic as illustrated in Table 1.

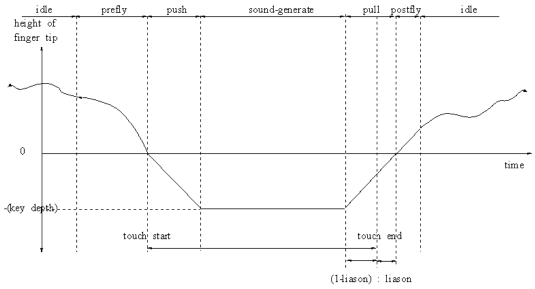

Figure 1: Height of the finger tip during a key touch.

state	start time	movement
prefly	$\left\| t^{\text{current}} - t^{\text{touch start}} \right\| \times v_{\text{tip position}} \leq \left\| \mathbf{X}^{\text{current}}_{\text{tip position}} - \mathbf{X}^{\text{touch start}}_{\text{tip position}} \right\|$	parabola
push	$t^{\text{touch start}}$	straight down
sound-generate	$t^{\text{touch start}} + \Delta t^{\text{push}}$	stop
pull	$t^{\text{touch end}} - (1 - liaison) \times \Delta t^{\text{push}}$	straight up
postfly	$t^{\text{touch end}} + liaison \times \Delta t^{\text{pull}}$	straight up
idle	$t^{\text{touch end}} + liaison \times \Delta t^{\text{pull}} + \Delta t^{\text{postfly}}$	not defined in this part

Table 1: Start time and movement of finger tip in each state

The finger in idle state is triggered to prefly state if $\left| t^{\text{current}} - t^{\text{touch start}} \right| \times v_{\text{tip position}} \leq \left| \mathbf{X}^{\text{current}}_{\text{tip position}} - \mathbf{X}^{\text{touch start}}_{\text{tip position}} \right|$ is satisfied, in other words, there remains just enough time to get to the desired key. In prefly state, the finger tip assumes a parabolic trajectory to the key to touch, where the curvature of the parabolic is determined by whether the finger tip crosses other fingers or not. The end time of prefly state is to be start time of the given key touch, which is also start time of push state. The duration of push state is given by

$$\Delta t^{\text{push}} = (\text{const.}) / \text{intensity}$$

, which implies more intense touch requires faster push. This heuristic will be improved when dynamics is considered instead of kinematics. The star time of pull state depends on the liaison of the key touch and the next one, which is modeled as

$$t^{\text{pull start}} = t^{\text{touch end}} - (1 - liaison) \times \Delta t^{\text{push}}$$

In other words, the higher liaison between the current key touch and the next one, in musical term *slur*, urges shorter interval between them. The duration of pull state, or Δt^{pull} is set by

$$\Delta t^{\text{pull}} = \Delta t^{\text{push}}$$

We assume that there is no slip. Thus in push state, the fingertip goes straight down, in contrast in pull state, the fingertip goes straight up. In sound-generate state, the fingertip does not move. At the end of push state, there comes postfly state whose duration is predefined to be rather short. The fingertip goes straight up in postfly state. When the postfly state ends, the finger returns to idle state.

4. Determination of position/orientation of wrist and joint angles of idle fingers

The tip position of non-idle fingers is determined in the module explained in the pervious section. Therefore the next step is to determine the position/orientation of wrist and the joint angles of idle fingers. This is a redundant articulated structure [Schilling90]. We solve the problem by the optimization process.

The evaluation function to be minimized depends on fingering comfort, wrist comfort, quality of sound to be produced, and collision avoidance. The constraint is the positions of the tip position of non-idle fingers. In case that the evaluation function is continuous, the position/orientation of wrist and the joint angles of idle fingers are continuous if the tip positions of the non-idle fingers are continuous. This fact supports the feasibility of the procedure of local search using the previous value as the initial value to determine the current value of the position/orientation of wrist and the joint angles of idle fingers.

4.1 The evaluation function

4.1.1 Finger comfort

At least from the point of view of fingering comfort, the number of crosses between fingers must be minimized [Neuhaus71]. Cross between thumb and other fingers is given moderate penalty whereas cross between 2-5 fingers is highly penalized although there are some advocates for it [Neuhaus71]. The penalty for cross is linearly dependent on the horizontal displacement of fingertips in case of cross to ensure continuity of evaluation function. For anatomical reason, the week fingers, or pinky/ring are given some penalty for their use [Parncutt97].

4.1.2 Wrist comfort

The orientation of wrist must be as perpendicular as possible according to the pedagogy of piano [Kentner76]. This can be expressed as:

$$f(\mathbf{S}_{\mathbf{W}}) = \sin^2 \phi_{\mathbf{W}}$$

, where $\phi_{\mathbf{W}}$ denotes the azimuth angle of the wrist.

4.1.3 Sound quality

We adopted simple heuristic that the quality of sound is the function of the orientation of the fingertip that is in touch with key. The orientation must be as perpendicular as possible for 2 reason. First, it is recommended by the pedagogy of piano [Kentner76]. Second, not to slide on the key, the fingertip orientation must be inside the frictional cone [Kijima96]. Thus the corresponding term can be expressed as

$$f(\Theta_{\mathbf{C}}) = \sum_{f \text{ is in touch state}} w_f \times \sin^2 \phi_f$$

, where touch state means push, sound-generate, and pull state, and w_f is the weight of the significance of the f th finger. We assign the w_f's in such a way that w_1 and w_5 are relatively low compared to w_3, w_4, and w_5.

4.1.4 Collision avoidance

There are 2 possible collisions. The collisions between fingers and those between keyboard and a finger. In either case, we can employ the potential method [Latombe91] to prevent the collision. It gives penalty for idle finger that is near the keyboard according to the tip position. We considered the collision between only the tip positions of the fingers not every joint for the efficiency. But the joint limit constraint guarantees that there can hardly happen that only the joints collide when there tip positions do not collide, which is supported by experiment.

Since all the functions related to each criterion are continuous, the evaluation function that is the weighted sum of those functions is also continuous. This is a sufficient feature to guarantee the continuity of the state of the wrist as explained in the next section.

4.2 Procedure

One of the necessary conditions for the acceptable animation is that for the given continuous tip positions of the non-idle fingers, all the DOFs must be also continuous. We will show that for the continuous evaluation functions as in 5.1, this feature is guaranteed.

Let $\mathbf{X_C}$ be the positions of non-idle fingers, $\Theta_{\mathbf{C}}$ and Θ_R be the joint angles of non-idle and idle fingers respectively, and $\mathbf{S}_W$ the position/orientation of wrist. For given tip positions of non-idle fingers or $\mathbf{X_C}$, we will find the position/orientation of wrist or $\mathbf{S}_W$, and the angles of fingers $(\Theta_{\mathbf{C}}, \Theta_R)$ such that the evaluation function $G(\Theta_{\mathbf{C}}, \Theta_R, \mathbf{S}_W)$ is minimized. If $\mathbf{S}_W$ is given, the joint angles of non-idle fingers, or $\Theta_{\mathbf{C}}$ can be determined from $\mathbf{X_C}$ by *inverse kinematics* [Schilling90] as follows. As for thumb (resp. other fingers), the number of DOFs is 5 (resp. 4), but there are 2 (resp. 1) constraints between the joint angles [Rijpkema91,Lee95], which make the effective number of DOFs be 3. It can be uniquely determined given the position/orientation of the wrist and the tip position. The Newton-Rhapson method [Press 88] is employed for the calculation. The possibility of the occurrence of 2 or more solutions [Asada86] can be eliminated by the joint limit constraints and taking initial value as the previous value in the Newton-Rhapson method. Thus we have

$$\Theta_{\mathbf{C}} = f^{-1}(\mathbf{X}_C, \mathbf{S}_W)$$

, where f^{-1} indicates the inverse kinematics function. Then $\Theta_R, \mathbf{S}_W$ can be determined by

$$(\Theta_R, \mathbf{S}_W) = \arg\min_{\Theta_R, \mathbf{S}_W} G(\Theta_{\mathbf{C}}, \Theta_R, \mathbf{S}_W) = \arg\min_{\Theta_R, \mathbf{S}_W} G\left(f^{-1}(\mathbf{X}_C, \mathbf{S}_W), \Theta_R, \mathbf{S}_W\right)$$

In other words, $\Theta_R, \mathbf{S}_W$ are determined such that $G\left(f^{-1}(\mathbf{X}_C, \mathbf{S}_W), \Theta_R, \mathbf{S}_W\right)$ is minimized for given $\mathbf{X_C}$. If G and f^{-1} are continuous as above, $\Theta_R, \mathbf{S}_W$ is also a continuous function for given continuous $\mathbf{X_C}$, because the domain of ($\Theta_R, \mathbf{S}_W$) is *compact*.

We use *direction set method* [Press 88] to find $\Theta_R, \mathbf{S}_W$ with the previous value as the initial value. But the dimension of ($\Theta_R, \mathbf{S}_W$), or the number of DOFs of ($\Theta_R, \mathbf{S}_W$), is $3 \times |\text{idle finger}| + 6$, which can be 21 maximally. Thus we employ a heuristic that each idle fingers except thumb assumes the same corresponding joint angles if feasible, which makes the dimension of search space reduce to 12 in the worst case. We also exploit the commonly used term to expedite the calculation.

5. Results

As far as we know, there is no publication on the results of an objective comparison between the preferred fingerings of different pianists, for a polyphonic piece. Thus we choose two pieces from [Cortot], where the recommended fingerings are presented. Table 2 illustrates the experimental results, where the written fingers are those recommended in [Cortot]. We calculate the evaluation function for each written fingering. The fingerings found by our method coincide with the first written fingerings. Figure 2 shows some snap shots of the animation playing the piece A.

	Piece A[1]		Piece B[2]	
	fingering		fingering	
result	(14)(25)(14)(25)	1636	(13)(24)(15)(23)	1312
written finger-ing	*(14)(25)(14)(25)*[3]	1636	*(13)(24)(15)(23)*	1312
	(14)(25)(13)(24)	2161	(13)(24)(15)(24)	1509
	(25)(14)(25)(14)	3018	(13)(24)(13)(24)	2001
	(13)(14)(13)(14)	∞[4]	(24)(35)(14)(23)	2572
			(13)(24)(13)(24)	2036

Table 2: Comparison between the written fingerings and the fingerings found by our method

[1] F. Chopin Etude Op. 25 No. 8 in Db major): 32th measure, first 4 double notes ((F3,Db), (F#3,D), (G3,Eb), (G#3,E))

[2] F. Chopin Etude Op. 25 No. 6 in g# minor): 5th measure, first 4 double notes ((B3,D#), (C#,E), (CX=D,E#), (DX=E,F#))

[3] The written fingerings in italic are those which are to the fingerings of search result

[4] The current implementation is not adapt to execute successive key touch of one finger

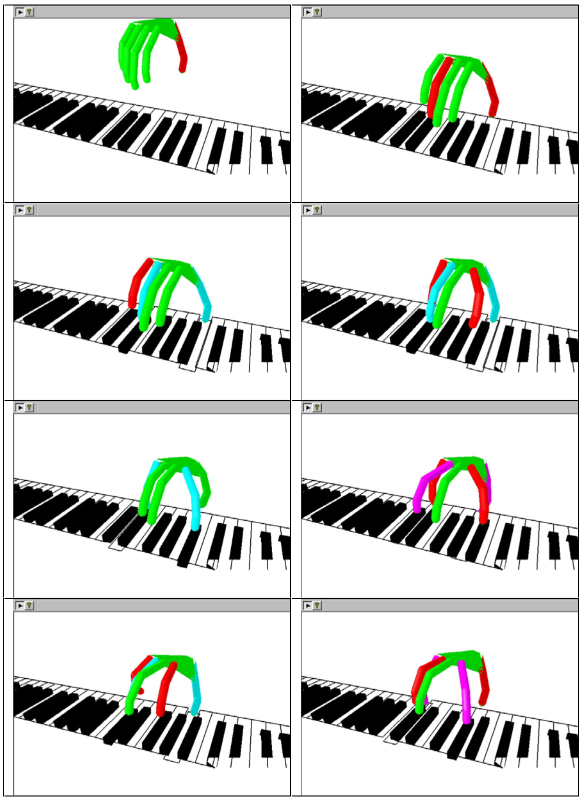

Figure 2: Snap shots in animation of playing the piece A[5]

6. Conclusion

An algorithm for the animation of the pianist's hand and the determination of piano fingering based on the animation is described and tested. In general, the result of the fingering determination accords with the written fingerings. The animation shows feasible movement of hands and fingers.

At present, we adopted a purely kinematic model, which lacks physical reality. We are studying the dynamic model suitable for the pianist's hand. In our algorithm, there are a bunch of parameters which can affect the animation results. They can be adjusted to the given set of score/fingering pairs by a learning method as in [Sayegh89].

7. Reference

[Asada86] H. Asada, J. E. Slotine, *Robot analysis and control*, John Wiley & Sons, Inc. 1986.

[Cortot] Alfred Cortot, *Chopin 12 etudes op. 25 pour piano*, Edition de travail avec commentaires d'Alfred Cortot, Salabert editions.

[Douville96] B. Douville, L. Levison, N. Badler. "Task level object grasping for simulated agents," *Presence* 5(4), pp. 416-430, 1996.

[Huang95] Huang Z., Boulic R., Magnenat Thalmann N., Thalmann D., "A Multi-sensor Approach for Grasping and 3D Interaction," *Proceedings of Computer Graphics International* '95, Leeds, Academic Press, pp.235-254.

[Huang96] P. S. Huang and M. van de Panne. "A Search Algorithm for Planning Dynamic Motions," *Computer Animation and Simulation '96 Proceedings of the 7th Eurographics Workshop on Simulation and Animation*, Springer Verlag. 1996, pp. 169-182.

[Kentner76] Louis Kentner, *Piano*, Collection Yehudi Menuhin, Hatier, 1976.

[Kijima96] Ryugo Kijima and Michitaka Hirose, "Representative Spherical Plane Method and Composition of Object Manipulation Methods", *IEEE VRAIS '96*,

[Latombe91] Jean-Claude Latombe, *Robot Motion Planning*, The kluwer international series in engineering and computer science; SECS 0124. 1991.

[Mas94] R. Mas, D. Thalmann, "A Hand Control and Automatic Grasping System for Synthetic Actors," *Proceedings of Eurographics* '94, Oslo

[Neuhaus71] Heinrich Neuhaus, *L'art du Piano*, Editions vande Velde, 1971.

[Parncutt97] Richard Parncutt, John A. Sloboda, Eric. F. Clarke, ,Matti Raekallio, and Peter Desain, "An Ergonomic Model of Keyboard Fingering for Melodic Fragments," *Music Perception* 97, 4, pp. 341-381.

[Press 88] Wiliam H. Press, Brian P. Flannery, Saul A. Teukolsky, Wiliam T. Vetterling, *Numerical Recipes in C*, The art of scientific computing, Cambridge University Press, 1988.

[Rijpkema91] Hans Rijpkema, "Computer animation of knowledge-based human grasping," *Proceedings of SIGGRAPH* '91, pp. 339-348.

[Lee95] J. Lee, and T. L. Kunii, "Model-Based Analysis of Hand Posture," *IEEE Computer Graphics and Applications*, pp. 77-86, 1995.

[Sayegh89] Sayegh, S. I., "Fingering for string instruments with the optimum path paradigm," *Computer Music Journal*, 13 (3), pp. 76-83.

[Schilling90] Robert J. Schilling, *Fundamentals of robotics : analysis and control*, Prentice-Hall, Inc. 1990.

[Viana98] Alexandre B. Viana, Jose H. F. Cavalcanti, Pablo Javier Alsina, "Intelligent system for piano's fingering learning aid," *ICARCV '98 -Fifth International Conference on Control, Automation, Robotics and Vision*.

[Winston84] Patrick H. Winston, *Artificial Intelligence*, 2nd Edition, Addison-Wesley, 1984.

[5] Green: idle, Red: prefley/postfly, Cyan: push/pull, Blue: sound-generate.