Vol.2013-MUS-99 No.48
2013/5/12

ピアノ演奏CGアニメーションの自動生成 ～肌質感リアルタイムレンダリングと演奏モーションのヒューマナイズ～

村松 昇[1,a)]　飛谷 謙介[1]　平田 純也[1]　巳波 弘佳[1]　古屋 晋一[2]　長田 典子[1]

概要：CG（Computer Graphics） 技術を用いたピアノ演奏アニメーションは，商業アニメ制作や音楽教育など，様々な分野で注目されている．本研究の目的は，楽譜から演奏動作を自動生成し，リアルな肌質感を持ったピアノ CG アニメーションを制作することである．本研究では，事前計算 GI テクスチャ技術をピアノ CG アニメーションシステムに実装し，肌質感を向上させた．また演奏スキルの知見に基づき演奏動作のヒューマナイズ処理を行い，より自然で人間らしい演奏動作を実現した．

キーワード：リアルタイム，アニメーション，Global illumination

1. 緒論

ピアノ演奏を CG アニメーションで表現する技術は，映像コンテンツ制作や音楽教育など多様な分野で期待されている．しかし，これまでのピアノ演奏に関する研究の多くは音楽的観点からのものや，手の構造，動きという観点からのもの [1][2][3] であり，ピアノ演奏の動作そのものに着目した研究は多くない．とりわけ演奏動作の生成や，CG 表現技術など，機械的な指の動きを扱ったものが多く，人間らしい手指の動きを実現した研究は少ない．

我々はリアルで自然なピアノ演奏動作の解析や CG 表現の研究 [4] を進めている．本研究のゴールの 1 つは 3DCG のキャラクタが自動的にピアノ演奏をしてくれる「ヴァーチャルピアニスト」であり，その技術課題は 2 つに大別される．1 つがレンダリングのリアリティ，他の 1 つが演奏モーションのリアリティの問題である．

レンダリングには，半透明な肌質感を表現するため Sub-Surface Scattering（以下 SSS） 技術 [5] や GI （Global Illumination） 技術が開発されているが，リアルタイム処理のためには GPU （Graphics Processing Unit） を用いて高速かつ高品質のレンダリングを実現する必要がある．演奏モーションの生成手法については，従来はモーションキャプチャ技術によって人間の動きを計測したが，価格・時間・データ容量ともに高コストであるという欠点がある．我々は，楽譜のみからピアノの運指および指軌道を計算する方法を提案したが，ヒューマナイズ処理（人間らしさの向上）など多くの課題が残っている．

本研究によって，ピアノ演奏の CG アニメーションを自動的・半自動的に生成することが可能となれば，アニメーション制作における時間とコストを大幅に削減でき，また，デジタルコンテンツ制作や電子ピアノ利用者のための教育アプリの提供が可能となる．

2. ピアノ演奏 CG の概要

本研究のコンセプトである「ヴァーチャルピアニスト」は，楽譜を入力として与えると，3DCG のキャラクタが曲を演奏するアニメーションを自動的に出力するシステムである．

入力として与える楽譜は，MIDI（Musical Instrument Digital Interface）に変換して読み込む．また，従来通りのモーションキャプチャシステムで取得した演奏モーションデータも読み込み可能である．楽譜を用いた演奏動作の自動生成を行う利点は，高価な計測機材や演奏を行う人間が必要ではないことである．また，モーションキャプチャシステムを用いる利点は，人間の動きをそのまま再現可能なことである．本研究では，楽譜を用いた演奏動作の自動生成を採用し，与えられた MIDI データにより，運指を決定し，最適化手法により演奏動作を自動生成する．次に，予

[1] 関西学院大学
Kwansei Gakuin University
[2] ハノーファー音楽演劇大学
Hannover University of Music and Drama
[a)] sho.m@kwansei.ac.jp

ⓒ 2013 Information Processing Society of Japan

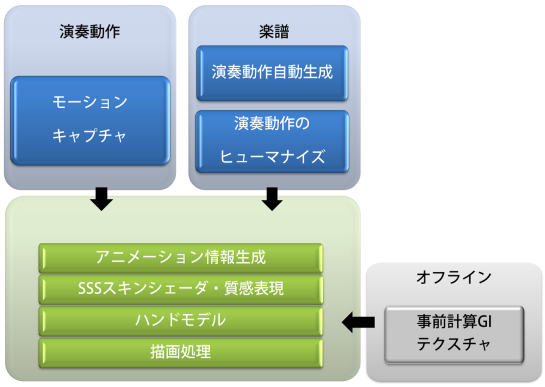


図 1　ピアノ演奏 CG システムの概要

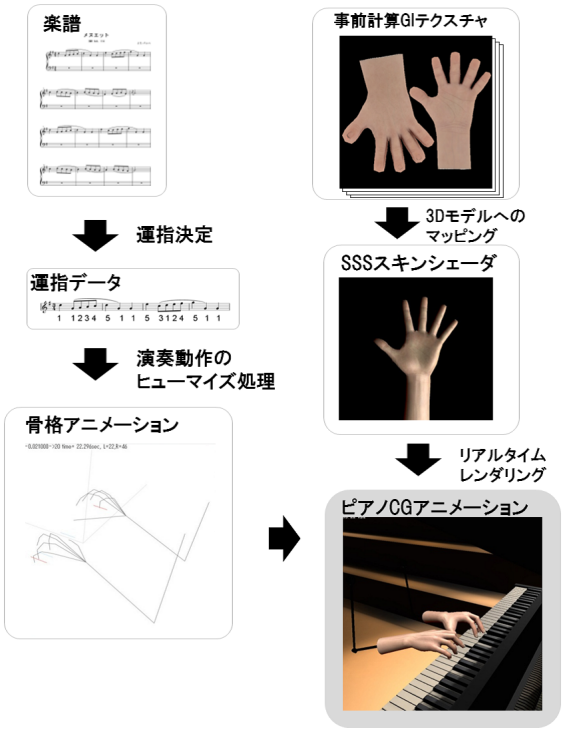


図 2　楽譜を用いた演奏動作の自動生成

め用意した 3DCG のハンドモデルに最適化された演奏動作を与え，演奏アニメーションが出力される．図 **1** にピアノ演奏 CG システムの概要，図 **2** に楽譜を用いた演奏動作の自動生成の概要を示す．

3.　肌質感リアルタイムレンダリング

リアルタイムレンダリングにおいて，高品質なモデルを実現しようとすると CPU や GPU に対して高い負荷がかかってしまう．また，先行研究 [6] に報告されているように，肌質感のリアルタイムレンダリングは，リアルタイム技術が優先されていて，質感に関しては十分に研究されていない．そこで，本章では，人間の肌の質感 [7] を保ちつつ，高い負荷をかけずに安定したフレームレートを維持したレンダリングが可能となる，CG レンダリングにおける事前計算 GI テクスチャの適用について述べる．

3.1　事前計算 GI テクスチャの生成

本手法における、人間の肌のテクスチャ生成には，市販されているテクスチャマッピング付きの高品質で汎用的な 3DCG のハンドモデルを用いる．このモデルに対してライティングの事前計算や SSS シミュレーションを行う。

3.1.1　ライティングの事前計算

本研究では，全方位ライティングと指向性ライティングを用意し，ラジオシティによって GI の計算を行う．全方位からの完全拡散光を用いたライティング環境での，GPU によるシェーディングを行うと，光が届きにくい複雑な場所のリアルな陰影は再現が難しい．

例としては，事前計算で得たテクスチャは指の股などにあまり光が入射していないので，リアルタイムでレンダリングする際に，その部分に局所的なライティングを行なっても影を完全に消すことができない．そこで本研究では，全方位からの完全拡散光ライティングだけでなく各角度からの指向性ライティング環境における事前計算も行い，そのテクスチャを差し替えることによって，高い再現性を実現する．今回は水平方向に 1 度間隔でのサンプリングを行う．

3.1.2　補間テクスチャによるデータ量削減

指向性を持ったライティングに対応したテクスチャ画像を全て事前計算で得るためには，仮に水平方向のみを 1 度間隔でサンプリングしても３６０枚の画像ファイルが必要となる．計算機の性能向上により，近年ではメモリの容量も大容量になりつつあるが，CG のインタラクティブ性を保つためにも，データ量の削減は必須である．今回は，２枚のテクスチャから中間のテクスチャを作りだす，補間テクスチャでのデータ量削減を行う．

補間テクスチャは２つのテクスチャ画像からその中間の角度のテクスチャをアルファ合成することによって得る．例えば，１０度間隔でサンプリングを行った場合，０度，及び１０度の角度からのライティング環境での事前計算テクスチャを合成することによって，擬似的に５度の角度でライティングされたテクスチャを得る．この場合，テクスチャファイルの量は，１度間隔の時に比べて１／１０となる．図 **3** に得られたテクスチャ，図 **4** に補間テクスチャの概略図を示す．

3.1.3　光源色への対応

従来の事前計算によるリアルタイム GI では，光源色の変更が反映されないという欠点があったが，事前計算によって得られたテクスチャ画像を R,G,B の３つのチャンネルに分離し，比率を変えて再合成する事によってテクスチャの色の変更を可能とする．光源色に変更があった場合，その RGB の比率を求め，テクスチャ画像の RGB 比率を光源のものと同じにする事によって，テクスチャの色の光源色によるリアルタイムでの変化にも対応する（図 **5**）．

ⓒ 2013 Information Processing Society of Japan

3.1.4 Subsurface Scattering

可能な限りリアルな肌の質感表現を行うために，SSS をと呼ばれる技術を採用した．SSS は半透明な物体の内部で光が複数回に渡って反射し，入射点とは異なる場所から出射する現象である．人間の肌をはじめとした透明感のある物体の表現に利用される技術である．

本研究では、3 層の SSS シミュレーション（ 図 **6**）をリアルタイムに行うことで，ラジオシティによる複雑な入射を考慮する．これにより正確な陰影が得ることが可能である．

以上 3.1 節で述べた処理により，より人肌としてリアルな質感と，指の形状による複雑な陰影を持ったテクスチャ画像を得ることができる．また，レンダリングに用いるハンドモデルを変更しても，以上の事前計算を行えば，モデルにあったテクスチャの生成も可能である．

図 **7** に事前計算 GI テクスチャを用いてリアルタイムレンダリングを行った結果を示す．

4. 楽譜を用いた演奏動作自動生成

本システムでは，楽譜情報の MIDI 番号や音の鳴るタイミングを入力することにより，ピアノの演奏動作を自動生成することが可能である．しかしながら，モーションキャプチャで計測されるような，人間らしい演奏モーションを生成することは難しい．本章では，機械的な動作をより人間らしく行わせる，演奏動作のヒューマナイズ処理について述べる．

4.1 ヒューマナイズ処理

ヒューマナイズ処理とは，機械的で単調な動作などを，人間らしく行わせることであり，様々な分野において研究されている．ロボット工学の分野では，人間らしい動作を行うために人間の筋骨格を模したロボットなども開発されている．さらに CG の分野では，顔表情 CG などの分野で重要性が認識されている [8][9] ．そこで，我々はスキル向上を目的とした研究をベースに，手指の動作に対してヒューマナイズ処理を行うことにより，より人間らしい CG アニメーションの生成を行う．

4.2 演奏動作のヒューマナイズ処理

本研究では，楽譜から自動生成された演奏動作に対して，各関節の初期位置や，関節の硬さをパラメータとして設定した．このパラメータを操作することにより，自動生成された機械的な演奏動作を，より人間らしい動作を生成すると同時に，演奏の熟練度の違いを表現することが可能となる．

4.3 演奏動作の熟練度の違い

古屋ら [10] はプロピアニストと，アマチュアの筋活動を

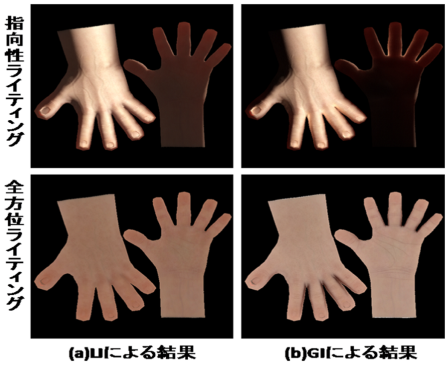


図 **3** 得られたテクスチャ画像



図 **4** 補間テクスチャ生成概略図

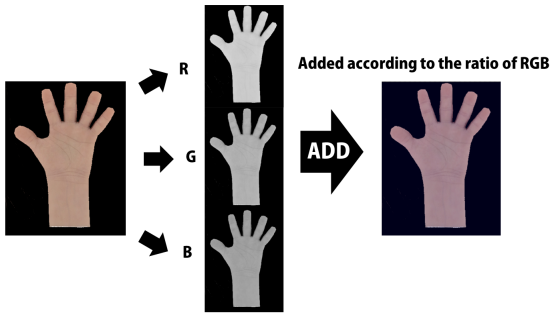


図 **5** テクスチャ画像分離合成手順

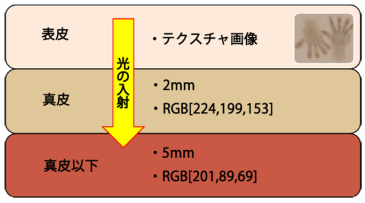


図 **6** スキンシェーダの概念図

比較することにより，ピアノ演奏を行う際に使う筋肉，関節を特定した．プロピアニストは，特に筋負荷量が増大する高速度の打鍵動作において，近位関節の動きを大きく利用することによって遠位関節の運動量を減らし，疲労しやすい末梢の筋肉の負荷量を低減していることが示唆され

ⓒ 2013 Information Processing Society of Japan

た．古屋らの知見をベースに各関節の使用頻度を設定することにより，演奏動作に熟練度の違いを CG アニメーションで再現する．

5. 結果と考察

5.1 レンダリング環境

本研究では表 1 のレンダリング環境で行った．

表 1 動作環境

CPU	Intel(R) Core(TM) i7 CPU 2.80GHz
GPU	NVIDIA GeForce GTX 480
OS	Microsoft Windows 7 Professional
メモリ	8GB RAM
DirextX バージョン	DirectX 11.0
解像度	1440 × 900

5.2 レンダリング結果

レンダリング結果を 図 7，図 8 に示す．



図 7 レンダリング結果

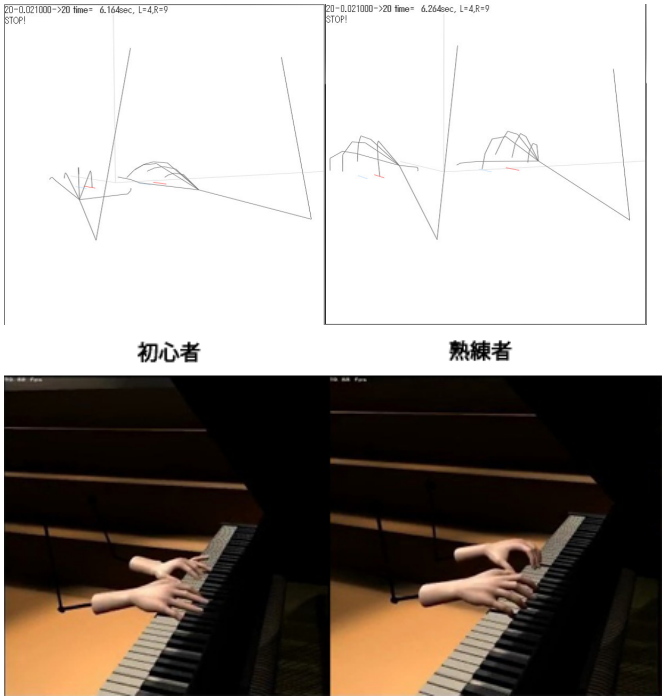


図 8 上：骨格アニメーションでの素人と熟練者の比較
下：3DCG アニメーションでの素人と熟練者の比較

5.3 考察

結果より，事前計算 GI テクスチャを用いることで，従来のピアノ CG アニメーションよりさらにリアルな肌質感を表現可能となった．事前計算 GI テクスチャは，アニメーションで用いるハンドモデルを別のハンドモデルに変更しても，変更したハンドモデルに対して，ライティングなどの再計算することにより，テクスチャの再生成することが可能である．従来のテクスチャは特定のハンドモデルでしか用いることができなかったが，この手法を用いることによって，汎用性が増したといえる．しかしながら，再生成するためにはさまざまな手順をふまないといけないため，汎用性があっても手間がかかってしまうという問題点がある．また，本研究では事前計算 GI テクスチャをハンドモデルに適用したが，本手法を様々な CG オブジェクトへ，適用可能である．

また，演奏モーションのヒューマナイズ処理に関しては，図 8 に示すように，楽譜から自動生成された演奏動作がより人間らしい動作を行うようになった．これによって，熟練者が演奏しているような動作や，素人の演奏を表現することが可能である．熟練者の演奏が楽譜から生成することが可能であれば，生成された演奏動作をお手本としたピアノの学習が可能になる．また，カメラの位置，角度を変更することにより，さまざまなハンドモデルの姿勢を参考にすることができ，さらに，素人の演奏動作と，上手い演奏動作の比較を行うことができれば，上達に向けて手指のどこの部位に注目すればいいかを確認できる．

6. 結論

本研究では，事前計算 GI テクスチャをピアノ CG アニメーションに実装することにより，肌質感が大幅に向上し，よりリアルなアニメーションが制作可能となった．また，ヒューマナイズ処理を行うことにより，機械的であった演奏動作がより，人間らしい動作を行うようになり，また，ピアノの熟練度合に応じた演奏動作の表現も可能となった．

今後の課題としては，モーションキャプチャで取得した演奏動作と楽譜から自動生成された演奏動作の比較を行い，モーションキャプチャで取得した高品質な人間らしい演奏動作に近づけていくことや，本手法の手指以外の 3DCG モデルへの展開があげられる．

ⓒ 2013 Information Processing Society of Japan

参考文献

[1] John L, Ying W and Thomas S. H. (2000). Modeling the constraints of human hand motion. Human Motion,. Proceedings. Workshop on,vol. 2, no. 1, pp. 52-77, 2000.
[2] 舟橋健司, 安田孝美, 横井茂樹, 鳥脇純一郎. (1998). 仮想空間における両手による協調操作モデル. 情報処理学会論文誌, vol. 36, no. 5, pp. 1334-1342.
[3] George E and Karan S. (2009). Handrix: Animating the Human Hand. ACM SIGGRAPH2003. 110-119.
[4] N. Kugimoto, K. Takai, R. Miyazono, K. Omori, T. Fujimura, S. Furuya, H. Katayose, H. Miwa, and N.Nagata., (2009). ”CGAnimation for Piano Performance”, ACM SIGGRAPH2009, Aug. 3-7.
[5] M. Hasan, M. Fuchs, W. Matusik, H. Pfister and S. Rusinkiewicz. (2012). : ”Physical reproduction ofmaterials with specified subsurface scattering”, ACM SIGGRAPH 2010 papers, SIGGRAPH’10, New York, NY, USA, ACM, pp. 61:1-61:10.
[6] Masaki O and Kenji S. (2009). Artist-Oriented Real-Time Skin Deformation Using Dynamic Patterns. Arts and Technology, pp 88-96.
[7] J. Jimenez, D. Whelan, V. Sundstedt., and D. Gutierrez. (2010). Real-Time Realistic Skin Translucency. IEEE Computer Graphics and Applications, Vol. 30(4), pp. 32-41.
[8] Villagrasa, S., and Susin Sanchez, A. (2009). FACe! 3D Facial Animation System based on FACS. IV Iberoamerican Symposium in Computer Graphics - SIACG 2009, pages 203-209.
[9] Richard B, Pakorn K, and Martin L., (2002). Jeremiah : the face of computer vision. SMARTGRAPH’02, pages 124-128.
[10] Furuya, S., Goda, T., Katayose, H., Miwa, H., and Nagata, N. (2011). Distinct interjoint coordination during fast alternate keystrokes in pianists with superior skill. Frontiers in Human Neuroscience, 5:50.

ⓒ 2013 Information Processing Society of Japan