

人物を対象とする表現技術の研究動向

問瀬 健二 渡部 保日児 末永 康仁

NTT ヒューマンインタフェース研究所

人物像をコンピュータグラフィックスで生成・表示する手法に関する研究について解説し、その動向を紹介する。髪や皮膚のレンダリング技術を中心に、顔のアニメーションを生成する方法や人物の動きを表示するために必要な技術に関してサーベイする。

A Survey on Rendering of Human Image

Kenji Mase Yasuhiko Watanabe Yasuhito Suenaga

NTT Human Interface Laboratories

This paper presents an overview of trends and issues on generation and rendering methods of human image. We focus on realistic rendering of human hair and skin, kinematics, and related issue on facial animation.

1 まえがき

本文では、人物像をコンピュータグラフィックス（以下CG）で生成・表示する手法に関する研究について解説し、動向を紹介する。人物像は我々が生活する上でもっとも親しみを感じる画像であり、人とのコミュニケーションにおいて大事な役割を果たしている。この対象をCGで表現しようという試みは、極めて自然である。それによって、近い将来、CG人間の相手と会話をしたり、CGでできたアナウンサーが番組を進行したり、教育の場で人体を探究するためのモデルとして使われたり、手術の計画立案や事故を再現するシミュレーションに使ったり、科学、産業、娯楽のあらゆる場面での利用が想定される。

CGによる人物の表示は、すでに1970年代初頭いくつかの試みがある[1, page 3]。E. Catmull[2]が1972年に見せた5本の指をもつ人間の手のアニメーションは我々に大きな衝撃を与えた。F. Parke(現NYIT)は71年頃からI. SutherlandのいるUtah大学で顔のアニメーションに取り掛かっている。

出版された論文等から、歴史を振り返ると、Platt & Badler[34]の筋肉モデルを使った表情のアニメーション、Parke[20]のパラメータモデル、Waters[35]の新しい筋肉モデルによる表情がある。人物全体では、Thalman夫婦のアニメーション“Rendez-vous à Montréal”[7]に現れるマリリン・モンローとハンフリー・ボガート、PIXARのアカデミー賞受賞作品“Tin Toy”の赤ちゃんが注目を浴びた。NHKは最近「驚異の小宇宙・人体」[8]というCGを駆使した番組でCG人間を示した。

グラフィックスハードの進歩により、ポリゴン数が少なければ、表情のアニメーションもリアルタイムで可能となり、SIGGRAPH'88のビデオショーでは、B. deGrafらがデザインしたCG顔“Mike, the Talking Head”が実演された。SIGGRAPH'89にはCGシンガーも現われた。

本文では、顔のアニメーションをつくるためのレンダリング技術を中心に、関連するモデリング技術もふくめ、解説する。表情を含む顔のアニメーションのためのモデリングについては、SIGGRAPH'89のコースノート(#22)[1]に解説と歴史、さらに重要な論文が掲載されている。ま

た、Magenat-Thalman[3]はCG人物像をつくるための技術の動向を総合的にサーベイしている。本報告でも一部参考にした。なお、人物を含むアニメーションには、ディズニーの作品に代表されるセル画によるものがあり、キーフレームの補間などのCG技術によってこのセル画の作成を簡略化する手法もあるが、ここでは扱わないことにする。また、衣服の表示[4]についても、扱わない。モデルとなる形状データと動きデータの入力は重要な問題であるが、ここでは簡単にあつかうことにする。とくに動きはロボットの行動計画の問題としても捉えられ、その分野の論文も参考となろう。

2 人物像表示に関わる技術の概観

人物像の表示技術は次のように分類できる。

1. 人体（全体）の幾何モデリングおよび質感表現
2. 人体の動作の入力および記述、幾何モデルの変形
3. 頭部（顔も含む）の幾何モデリングおよび質感表現
4. 表情動作の入力および記述、幾何モデルの変形
5. 動作による外界（衣服を含む）との（物理的）インタラクション
6. 音声情報との合成
7. 衣服のモデリングおよび質感表現（本文では扱わない）

これらは、幾何モデルの作成・記述、質感の表示、幾何モデル変形のための動作モデルの作成・記述と分けることもできる。

幾何モデルの記述

幾何モデルの記述には一般にサーフェスを三角パッチで表現したり、パラメータ曲面を使う方法や、2次曲面体などの幾何関数プリミティブの集合演算で表現したりする。

衣服のモデリングに見られるように、素材の物理的振舞いをシミュレーションするアプローチもある。

質感の表現

質感表現は光の反射特性をモデル化することが必要であるが、膚に適した反射モデルはいまのところない。しかしながら頭部にかぎらず、膚の質感を出す技術が必要で、これまでは主に実画像をテクスチャとするテクスチャマッピングの手法がとられてきたが、このテクスチャを発生させることが今後の課題であろう。頭部には髪、目、口、歯などの特徴的な部位があり、これらのリアルな生成も大きな課題である。自然物体のレンダリング [5] の問題と同じく、これらのリアル感を

- (1) 幾何モデルによって実際に物体を作って再現するか
- (2) マクロな反射モデルを使って近似的に表現するか
- (3) 実写した画像を使うか

などの選択肢があり、使う側の立場でどれかを選ぶことになる。

動きの表現

動きの表現はモデルの記述と大きな関わりがあるが、人物全体では部位に関する階層構造をつくり、適当な物理モデルで運動を表現するのが一般的である。顔の動きには表情の発出と発声によるものがある。テクスチャの書き換え、モデルのパラメータによる単純な変形、筋肉モデルや皮膚の表層モデルを使った表皮の変形などのレベルにわたることができる。

3 人物全体の表示

人物モデルを記述するには、全体が滑らかで、骨格と関節の動作が表現でき、その際の形状の変化が自然である必要がある。小松 [6] は双 4 次 Bézier パッチで人体全体を表現し、関節部を Gregory パッチで表現した。これは滑らかな形状を表現できるが、個人性を出すには向いていない。一方、3 角パッチによる表現はデータ量が膨大となるため、モデルの作成、変形が困難である。Thalman 夫妻は局所変形に適した Joint-dependent Local Deformation (JLD) 演算の考え方を示した [7]。Thalman 夫妻は Synthetic Actors (合成俳優) の作成に熱意を示し、前述のマリリン・モンローに続いて、エルビス・プレスリー、マドンナなどを手掛けている。メタボールによる方法 [8] は滑らかな曲面を簡単に表現することができるが不自然な

融合が起きてしまうという欠点がある。変形可能な表面の記述法には Barr [9]、Sederberg ら [10] がある。

3.1 筋肉と骨格の表現

人体をその解剖学からの知識によって筋肉と骨格で表現することには次のような意義がある。

- (1) 動きの表現可能な構造を記述できる [11]。
- (2) 動作時の変形を自然に (エレガントに) 表現できる [12]。
- (3) 運動力学を適用できる。
- (4) 軌跡からの動作パラメータ逆計算の問題 (inverse kinematics) を表現できる。
- (5) 障害物の検出と行動計画ができる。

これらは次に述べる運動の表現と密接な関係がある。

3.2 体全体の動きの表現

体の動きを表現するには、まず、各パーツの動きをキーフレームで表して適当に補間する方法がある。初期のアニメーションでは人間の動きを実写して、各部の動きを手作業で拾うことでキーフレームのパラメータを入力した。Lee & Kunii [13] は画像処理によってこのデータ入力の省力化をはかった。次の段階は、inverse kinematics [14] である。これは例えば、コップをつかもうと腕を伸ばすときに各関節はどの角度で回転すべきかという問題を解くことであり、ロボットの制御と同じ問題になる。

現在は筋肉を一種のパネと仮定して、各部品の質量も考慮した運動方程式を解く物理モデリング (physically-based modeling) による表現が盛んである。物理モデリングには次の利点がある。

- (1) アニメータが入力する必要があるのは物理パラメータと行動のプランでよい [15]。
- (2) 動作がより自然に表現できる [16]。
- (3) 障害物などの物体とのインタラクションを表現できる。

Bruderlin らは、SIGGRAPH'89 で人間の歩行のリアルなアニメーションを作る手法とシステムを発表した [16]。KLAW (Key-frame Less Animation for human

Walking) と名付けられたこのシステムは、人間の脚の骨格と筋肉とを解剖学的に忠実に表現するのではなく、上肢と下肢をそれぞれ自動車に用いられるサスペンションで近似し、全身を片方の脚で支えるモードと両脚で支えるモードの系列で記述することにより、歩幅、ピッチ、歩行速度を制御し、かなりリアリスティックな歩行アニメーションを作り出す。このように、物理モデリングといっても、必ずしもすべてを忠実にモデリングしているわけではない。

Thalmann 夫妻の最近の作品で釘を打つ両腕のシーンがでてくるが、左手で支えた釘を打った反動でハンマーが跳ね返り、それに伴って右手も動くなど、かなりリアルな動きになっている。ただし、物理モデリングは一般には計算時間の問題があったり、かならずしも解が閉形式で求まらないという問題がある。

4 頭部・顔の表示

人体の中で頭部はもっとも印象が強い。そのためリアル感を出すことは重要な要素である。

4.1 顔形状の表現

人体全体をパラメトリック曲面で表すのが主流であるのに対して、頭部は3角パッチが良く使われる [17]。これは、頭部の形状が複雑で、パラメトリック曲面では細かい変化をうまく表現しにくい [18] ことと、人体ほどデータ量が大きくなることとおおきい。頭部については、CT像をもとにボクセルデータで表現して使う応用もある [19]。

データの収集はいろいろな方法があり、機械的に測る方法、2面図を使う方法 [20] などが用いられてきた。最近では各種のレンジセンサがあり、精度の高い形状データを簡単にとることができる [21]。しかし、いまのところこれらの装置は大がかりで、一般化モデルを変形して、自動処理により個人の形状に合わせる技術 [22] も必要である。

4.2 皮膚の表現

皮膚を表現する良いモデルはこれまでなかった。色を慎重に選んで、従来のシェーディングモデルで表現するか、

TVカメラでとった顔のテクスチャをはりつける [23,24] かどちらかであった。皮膚のキメを作るために、石井ら [25] のボロノイ分割によるアプローチは、皮膚の構成原理からみて興味深い。皮膚の反射特性は従来のシェーディングモデルとは一致しないので、反射スペクトルの特性を調べる [26] などの必要がある。

4.3 髪の毛の表現

髪の毛は、動物の毛皮 [27,28] を表現するという一般化された問題の特殊なケースとしてとらえられる。毛は非常に本数が多く、細く小さいオブジェクトであり、かつ柔軟物体であるため、表現が難しい。すなわち、マクロにとらえてテクスチャとして表現すると変形を表現しにくくなるし、ミクロにとらえるとモデリングが複雑・膨大になるという問題がある。

山名らは髪全体を異方性反射物体としてとらえリアルなテクスチャを表現し、変形を表現するために異方性のパラメータを制御する手法をとった [29]。モデリングを単純化する方法として、Pearce [30] はパーティクルシステムを採用し、渡部らは3角柱モデルを採用した。渡部らの方法はグラフィックスWSに標準的に装備されているZ-Bufferを使って、数万本もの髪の毛を数分で表示できる [31]。これは幾何モデルで表現するため変形操作が可能である。また、モデリングを単純化するためにWispモデルも採用している。Kajiya & Kay [28] の3次元テクスチャを用いる手法は、極めてリアルな熊のぬいぐるみの表示に成功したが、計算時間が非常に長い(数時間)という問題を解決しなくてはならない。また、 α -ブレンディング法でエイリアスのない髪の毛を表示する報告もある [3]。リアル感を求めるためには、self-shadowingの問題と、髪の毛同士あるいは髪の毛と膚、衣服の間の物理的関係を記述して自然な流れを作るという課題がある。中嶋らは髪の毛の動きの表現を試みている [32]。

4.4 表情の表現

皮膚が内部の筋肉の動作によって変形を受けると表情が現れる。単に3角パッチを3次的に移動する [20,33]

だけでも効果があるが、筋肉の動きを適当にシミュレートする [34,35] と、指示が簡単になる。この時に、心理学者が提唱する筋肉構造も考慮した Facial Action Coding System (FACS) [36,37] がしばしば使われる [34,35,38]。これはアクションユニットと呼ぶ動作単位を組み合わせて表情を記述しており、どの部位をどう動かせばどのような表情を作るかということがわかる点で都合がよい。

表皮を数層のパネモデルで表わし、再下層の筋肉の動きで表皮を動かすと、自然な変形を表現できる [39]。Pieper [39] は、これを手術後の皮膚の変形のシミュレーションに使った。

表情は発声によっても作られる。アニメーションを表示する際に音と同期させて口を動かすことが必要になる [40]。各音素に対する変形ルールを作って話す表情を作る試みが行われている [41,42]。一方、Lewis & Parke [43]、森末ら [44] は音素を認識し、それに対応する口形状を生成した。

5 おわりに

我々が、話相手として CG 人物像を、違和感なく受け入れるようになるまでには、まだ解決すべき課題が山積みされている。人物が身近であればあるだけ、CG でリアルな画像を出せば出すほどすこしの違いが増幅されて違和感として映ることになるかもしれない。実はその違和感を払拭するものは、実世界の現象を忠実に再現するモデルではなく、まったく別のものかもしれない。このおそらく、感性に訴えるものを見つけるのは容易ではないだろう。しかし、この人類の新しい友人達に我々と同じ姿を与える仕事は非常に楽しい仕事でもある。

参考文献

- [1] "State of art in facial animation", *SIGGRAPH'89 course notes, no. 22*, Aug. 1989.
- [2] Edwin Catmull, "A system for computer-generated movies", *Proc. ACM Annual Conference*, pp. 422-431, 1972.
- [3] Nadia Magnenat-Thalmann, "New trends in the direction of synthetic actors", In T.S.Chua and T.L.Kunii, editors, *CG International '90*, pp. 17-35, Springer-Verlag, 1990.
- [4] 今岡 春樹, "服地・布の生成に関する研究動向", 情処研資, *CG40-14*, Aug. 1989.
- [5] 中島 正之, "CG による自然物体の生成技法", 情処研資, *CG40-10*, 1989.
[人体の表示]
- [6] 小松 功児, "キャラクタ・アニメーションのための人体の曲面モデル", 情報処理学会論文誌, **29**, 1, pp. 7-19, 1988.
- [7] Nadia Magnenat-Thalmann and D Thalmann, "The direction of synthetic actors in the film rendez-vous à montreal", *IEEE CG&A*, **7**, 12, pp. 9-19, 1987.
- [8] 伊藤 博文, "NHK 「驚異の小宇宙・人体」におけるサイエンティフィック、ビジュアライゼーション", *NIC GRAPH'89*, 1989.
- [9] A.H. Barr, "Global and local deformations of solid primitives", *Computer Graphics*, **18**, 3, pp. 21-30, 1984.
- [10] T.W. Sederberg and S.R. Parry, "Free-form deformation of solid geometric models", *Computer Graphics*, **20**, 4, pp. 151-160, 1986.
[動きの表示]
- [11] David Zeltzer, "Motor control techniques for figure animation", *IEEE CG&A*, **2**, 9, pp. 53-59, 1982.
- [12] J.E. Chadwick, D.R. Haumann, and R.E. Parent, "Layered construction for deformable animated characters", *Computer Graphics*, **23**, 3, pp. 243-252, 1989.
- [13] M.W. Lee and T.L. Kunii, "Animation design: A database-oriented animation design method with a video image analysis capability", In N. Magnenat-Thalmann and D. Thalmann, editors, *State-of-the-art in Computer Animation*, Tokyo, 1989, Springer
- [14] M. Girard and A.A. Maciejewski, "Computational modeling for computer generation of legged figures", *Computer Graphics*, **19**, 3, pp. 263-270, 1985.
- [15] A. Witkin and M. Kass, "Spacetime constraints", *Computer Graphics*, **22**, 4, pp. 159-168, 1988.
- [16] Armin Bruderlin and Thomas W. Calvert, "Goal-directed, dynamic animation of human walking", *Computer Graphics*, **23**, 3, pp. 233-242, 1989.
[顔の表示]
- [17] 秋本高明, 間瀬健二, 末永康仁, "形状の自動変形による表情を持つ顔画像の生成方法の検討", 情処研資, *CG 28-14*, Aug. 1987.
- [18] 小松 功児, "アニメーション制作のための顔の曲面モデル", 情報処理学会論文誌, **30**, 5, pp. 633-641, 1989.

- [19] 横井 茂樹, “手術シュミレーター”, 計測と制御, 28, 7, pp. 629-632, 1989.
- [20] I. Frederic Parke, “Parameterized models for facial animation”, *IEEE CG&A*, pp. 61-68, 1982.
- [21] 末永 康仁, 渡部 保日児, “A synchronized cylindrical range and color data scanner and its application to 3d face data acquisition”, 情処研資, CV67-5, Jul. 1990.
- [22] 秋本 高明, リチャード ワレス, 末永 康仁, “正面・側面像を用いた頭部形状の自動推定”, 90 春季信学全大, 1990.
[皮膚の質感]
- [23] 秋本高明, 末永康仁, “頭部の正面・側面像と基本 3 次元形状を用いた顔画像生成手法”, 信学技報, PRU88-47, Jul. 1988.
- [24] 相沢清晴, 原島 博, 斉藤隆弘, “構造モデルを用いた画像の分析合成符号化方式”, 信学論, J72-B-I, 3, pp. 200-207, Mar. 1989.
- [25] 石井智海, 倉地忠彦, 安田孝美, 横井茂樹, 鳥脇純一郎, “ボロノイ分割を利用した生体皮膚の質感表現に関する研究”, 情処研資, CG46-9, 1990.
- [26] 石野喜信, “膚の反射スペクトルに及ぼす照明および測定角度の影響”, 1990 春季信学全大, D-542, 1990.
[頭髮]
- [27] G. Miller, “From wire-frame to furry animals”, *Graphics Interface '88*, pp. 138-145, 1988.
- [28] J.T. Kajiya and T.L. Kay, “Rendering fur with three dimensional textures”, *Computer Graphics*, 23, 3, pp. 271-280, 1989.
- [29] 山名 岳志, 末永 康仁, “関数制御型異方性反射モデルを用いた頭髮の質感表現”, 信学論, J73-D-II, 6, pp. 880-886, Jun. 1990.
- [30] A. Pearce, B. Wyvill, G. Wyvill, and D. Hill, “Speech and expression: a computer solution to face animation”, *Graphics Interface '86*, pp. 136-140, 1986.
- [31] 渡部保日児, 末永康仁, “三角柱と房のモデルによる頭髮像の生成”, 信学論, J73-D-II, 3, pp. 367-373, Mar. 1990.
- [32] 中嶋 正之, 三輪 喜良, 安居院 猛, “確率モデルによる頭髮の動き表現”, 信学技報, IE89-60, 1989.
- [33] 秋本高明, “あごの 3 次元モデルと表皮の自動変形による表情と動きをもつ顔画像の生成”, *NICOGRAPH'86* 論文集, pp. 207-213, Nov. 1986.
[表情]
- [34] S.M. Platt and N.I. Badler, “Animating facial expressions”, *Computer Graphics*, 15, 3, pp. 245-252, Aug. 1981.
- [35] K. Waters, “A muscle model for animating three-dimensional facial expression”, *Computer Graphics*, 21, 4, pp. 17-24, 1987.
- [36] P. Ekman and W. V. Friesen, “The Facial Action Coding System”, Consulting Psychologists Press, Inc., San Francisco, CA, 1978.
- [37] 工藤 力 訳, “表情分析入門”, 誠信書房, P.Ekman W.Friesen 著.
- [38] 崔昌石, 原島 博, 武部 幹, “分析合成符号化における表情の記述と合成について”, 信学技報 IE88-84, 1988.
- [39] Steve Pieper, “Physically-based animation of facial tissue”, In *SIGGRAPH'89 course note #22[1]*, (see also his Ph.D thesis, MIT).
- [40] Frederic I. Parke, “A model for human faces that allows speech synchronized animation”, *Computer & Graphics*, 1, pp. 3-4, 1975.
- [41] David R. Hill, Andrew Pearce, and Brian Wyvill, “Animating speech: an automated approach using speech synthesised by rules”, *The Visual Computer*, 3, pp. 277-289, 1988.
- [42] 森島 繁生, 岡田 信一, 原島 博, “知的インタフェースのための顔の表情合成法の一検討”, 信学論, J73-D-II, 3, pp. 351-359, Mar. 1990.
- [43] J.P.Lewis and F.I.Parke, “Automated lip-synch and speech synthesis for character animation”, *CHI+GI'87* pp. 143-147, 1987.
- [44] 森末 映史, 津田 俊隆, 松田 喜一, “音声情報に基づいた口の動画像生成に関する一検討”, 信学技報, CS87-130, 1987.