

キャラクターアニメーションにおけるシナリオ駆動型動作合成

江村 恒一*** 青木 輝勝* 安田 浩*

* 東京大学 先端科学技術研究センター 〒153-8904 東京都目黒区駒場 4-6-1

** 松下電器産業 (株) ネットワークシステム開発センター 〒140-8632 東京都品川区東品川 4-5-15

E-mail: * {emura, aoki, yasuda}@mpeg.rcast.u-tokyo.ac.jp, ** emura.koichi@jp.panasonic.com

あらまし これまでに筆者らは、特別なスキルを持たないユーザが容易にかつ短時間で3次元コンピュータグラフィックス(CG)のコンテンツを制作できるDMP(Digital Movie Producer)を提案してきた。本稿では、DMPの表現力向上を目的とし、シナリオ駆動型動作合成の手法を提案する。本手法では、動作を表す動詞句と3次元CG人物キャラクターの複数の基本動作を拡張感情空間上で関連付けることにより、自然文のシナリオから複数の基本動作へ変換し、時間同期を行うことで感情を伴う様々な動作を合成する。更に本手法では、フォントの大きさなどの文字修飾と動作の強調とを関連付けることにより、ユーザが直感的に即時に動作の強調レベルを編集することを可能とする。

キーワード 3次元コンピュータグラフィックス, キャラクターアニメーション, 動作合成, 感情空間

Scenario-driven Motion Synthesis for Character Animation

Koichi EMURA*** Terumasa AOKI* and Hiroshi YASUDA*

* Research Center for Advanced Science and Technology, The University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo, 153-8904 Japan

** Network Systems Development Center, Matsushita Electric Industrial Co., Ltd.
4-5-15, Higashi Shinagawa Shinagawa-ku, Tokyo, 140-8632 Japan

E-mail: * {emura, aoki, yasuda}@mpeg.rcast.u-tokyo.ac.jp, ** emura.koichi@jp.panasonic.com

Abstract We have already proposed DMP (Digital Movie Producer) system that facilitates non-professional users to create 3D CG movies in a short time. In this paper, for the further improvement of DMP, we newly propose the scenario-driven motion synthesis for character animation that converts a scenario into basic motions using the relation between a verb phrase and basic motions in the extended emotional space, synthesizes basic motions with time synchronization, and also enables users to modify the level of emphasis for the synthesized motion intuitively in real time by changing a text character attribute (e.g., font size).

Keyword 3D Computer Graphics, Character Animation, Motion Synthesis, Emotion Space

1. はじめに

インターネットの出現、ブロードバンド化、携帯端末などノン PC 機器が接続するユビキタスネットワーク化によって、少数のプロフェッショナルが制作したコンテンツを多くの視聴者へ配信するブロードキャスティングから、多数のノンプロフェッショナルが作るコンテンツを、興味を一にする複数の個人へ配信するパーソナルキャスティングへとコンテンツ流通が大き

くパラダイムシフトを始めている。

個人が作り発信するコンテンツは、テキストから写真、動画へと進化し、より豊かな自己表現ができるメディアとしてコンピュータグラフィックス(CG)を活用する方向へ進化すると考えられる。

現在最もCGが使われているコンテンツに、映画やアニメーション、ゲームがある。これらの制作者は、Alias, Discreet, Softimage に代表されるデジタルコン

テンツクリエーション(DCC)企業のソフトウェアを用いてコンテンツを制作しているが、このようなソフトウェアはプロフェッショナル向けに設計されており、機能が多いために操作が複雑であり経験やスキルのない一般のユーザには使いこなせない。またコンテンツの映像表現の面で、見る者に劇的に効果を与える演出を施すには、映画制作の演出ノウハウを理解・適応できなければならず、これも一般のユーザには困難である。

これらの課題を解決し、誰でも容易にかつ短時間で3次元CGのコンテンツを制作できることを目的とし、筆者らは、DMP(Digital Movie Producer)を提案し、PC上での実装を通してカメラワークを主とした撮影技法の自動適応を実現、検証してきた。[1][2][3]

本稿では、DMPの表現力向上を目的とし、3次元CG人物キャラクタの動作を複数の基本動作に分割し、動作を表す動詞句と複数の基本動作を拡張感情空間上で関連付けることにより、自然文のシナリオから3次元CG人物キャラクタの動作を合成し、更にフォントの大きさなどの文字修飾と動作の強調とを関連付けることにより、動作の強調表現を制御するシナリオ駆動型動作合成の手法を提案する。

2. 関連研究

3次元CG人物キャラクタのアニメーションを効率よく制作するために、基本動作データを組み合わせる動作合成に関する研究が多くなされている。

動作のキーとなるフレームを複数保存して必要に応じて呼び出すキーフレーム法を拡張し、キーフレーム間を動的に結合する有向グラフで管理し、一連の動作を滑らかに接続するモーショングラフを用いた動作生成[4][5]や、モーショキャプチャされた複数のモーシオン例と物理パラメータを用いてモーシオン間の補間を行うモーシオン補間法[6]は、時間軸上に連続する動作の接続が目的であり、時間軸上に並列する動作の合成には適応できない。

アニメータが直感的にモーシオンを合成できることを目的としたアノテーションによる動作合成[7]は、Walk, Waveなどの動作を時間軸上のマルチトラック上に自由に配置することで、時間軸上に並列する動作の合成を行うが、マウスによる直接操作が必要となり、スキルのない一般のユーザが所望の動作を作るのに適していない。

また自然言語の指示を与えることにより仮想空間内のキャラクタが自律的に動作する研究も多くなされている。

オブジェクト指向の動作データ管理手法[8]は、上位概念モデルの動作データを継承して下位概念モデルの

動作を生成できるが、複数の動作の多重継承には言及していない。

PAR (Parameterized Action Representation)[9][10]は、自然言語をPARとしてオブジェクト、アクション、コンディションなどに変換し、コンディションに合致した場合に、自律エージェントがオブジェクトに作用する動作を行う。PARでは、複合動作も記述可能であるが、自律エージェントに予め与えておくルールであり、ユーザが編集することはできない。

このように従来の動作合成の手法では、プロの制作者がアニメーションを作成するための工数は削減されたが、一般のユーザが自分の思い通りの動作を作るほどに動作のパリエーションを作成できない。特に複数の動作をどのように同期させて新しい動作を作るかについては従来技術がない。また、出来た合成動作を一般のユーザが簡単に編集できない。

3. シナリオ駆動型動作合成

シナリオ駆動型動作合成について図1に示す。本研究では、入力する自然文テキストに含まれる動詞および動詞修飾語からそれぞれ変換する基本動作および感情パラメータ、文字修飾から変換する強調パラメータを用いて、感情や強調の度合いが異なる様々な動作を合成する。そして合成した動作を3次元CG人物キャラクタのモデルデータに適応することにより、キャラクタアニメーションを生成する。

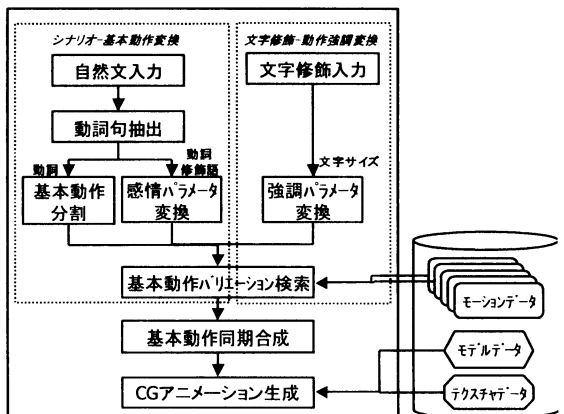


図1 シナリオ駆動型動作合成

3.1. 拡張感情空間

一般的なキャラクタアニメーションでは、感情は表情の変化で表すが、本研究では、拡張感情空間上で管理する基本モーシオンデータを合成することで、Happiness, Surprise, Sadness, Anger, Fear, Disgustの6種類の感情を伴った動作生成を可能とする。拡張感情空間

間は、表 1 に示すように誘意度 (Valence) と覚醒度 (Arousal) の 2 つの感情パラメータを軸として構成する感情空間 [11][12] に新たに強調度 (Emphasis) の軸を拡張しており、同じ感情パラメータにおいて動作の強調度の異なる複数のモーションデータを関連付けることができる。ここで誘意度は感情の positive/negative の度合い、覚醒度は感情の興奮度合い、強調度は感情を動作として強調する度合いを表す。

表 1 感情を表す感情パラメータ

感情	覚醒度 (Arousal)	誘意度 (Valence)
Happiness	0	1.0
Surprise	1.0	0
Sadness	-0.2	-1.0
Anger	0.8	-0.8
Fear	0.5	-0.5
Disgust	-0.5	-0.5

拡張感情空間上への基本モーションデータの関連付けを具体例で示す。MVML 記述は、Motion 要素の name 属性値に基本動作名、Emotion 要素の a, v, e 属性値にそれぞれ拡張感情空間における覚醒度、誘意度、強調度、Resource 要素の src 属性値に基本モーションデータへのリンク情報を記述することで基本モーションデータを拡張感情空間へ関連付ける。例として、走る動作 (run) に、2 種類の驚き (Surprise) の感情を伴う基本モーションデータを関連付ける記述を図 2 に示す。

```
<!-- Motion Variation -->
<MVML>
  <Motion name="run">
    <Instance>
      <Emotion a="1.0" v="0.0" e="0.0"/>
      <Resource src="run01.mtra"/>
    </Instance>
    <Instance>
      <Emotion a="1.0" v="0.0" e="0.5"/>
      <Resource src="run02.mtra"/>
    </Instance>
  </Motion>
  <Motion name="lookback">
    <Instance>
      <Emotion a="1.0" v="0.0" e="0.0"/>
      <Resource src="lookback01.mtra"/>
    </Instance>
    <Instance>
      <Emotion a="1.0" v="0.0" e="0.5"/>
      <Resource src="lookback02.mtra"/>
    </Instance>
    <Instance>
      <Emotion a="-0.2" v="-1.0" e="0.0"/>
      <Resource src="lookback11.mtra"/>
    </Instance>
  </Motion>
</MVML>
```

図 2 MVML 記述

この 2 種類の基本モーションデータは図 3 に示すよ

うに拡張感情空間上に関連付けられている。

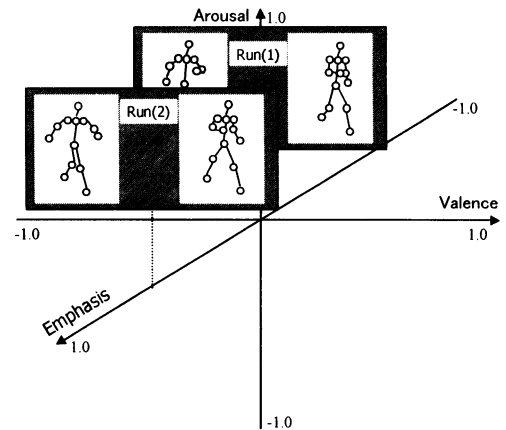


図 3 拡張感情空間

以上のように、基本モーションデータを拡張空間上に関連付けることにより、感情パラメータに最も近い基本動作のバリエーションを検索し、感情を伴う動作を合成することが可能となる。

3.2. シナリオ-基本動作変換

入力する自然文テキストを自然言語解析し、動詞および動詞修飾語を抽出する。次に V2MML 記述を用いて動詞から複数の基本動作に変換する。自然文から基本動作への変換を具体例で示す。V2MML 記述は、Verb 要素の name 属性値に動詞名、Motion 要素の name 属性値に基本動作名を記述する。また par 要素の範囲内の Motion 要素は時間的に並列に起こる動作を示し、Motion 要素の begin 属性値は動作の開始時間の相対的な初期値を記述することで動詞に対応する複数の基本動詞を時間軸上に関連付ける。例として、動詞「逃げる」を、基本動作 run と lookback の並列複合作とし、時間同期の初期値を lookback が run 開始時間の 1 秒後から開始すると定義した記述を図 4 に示す。

```
<!-- Verb to Motion -->
<V2MML>
  <Verb name="逃げる">
    <par>
      <Motion name="run"/>
      <Motion begin="1s" name="lookback"/>
    </par>
  </Verb>
</V2MML>
```

図 4 V2MML 記述

「ネギーマンは驚いて逃げた。」という入力文がある場合、自然言語解析により動詞として「逃げる」、動

詞修飾語として「驚いて」を抽出する。動詞「逃げる」は図 4 の V2MML により基本動作 run と lookback に分割される。また動詞修飾語「驚いて」は、表 1 に示す 6 種類の感情のうち Surprise と一意に対応するので Surprise の感情パラメータ (覚醒度 1.0, 誘意度 0.0) を取得する。次に基本動作 run および lookback の複数のバリエーションから感情パラメータに適応する実体のモーションデータを検索する。図 2 の MVML 記述を参照し、基本動作名がそれぞれ run と lookback で覚醒度 1.0, 誘意度 0.0, 強調度 0.0 [デフォルト値] となるモーションデータへのリンク情報 run01.mtra と lookback01.mtra を取得する。

以上のように、動詞から複数の基本動作、動詞修飾語から感情パラメータを抽出し、抽出した感情パラメータに最も近い基本モーションデータへ変換することが可能となる。

3.3. 基本動作同期合成

一般的な動作合成には、モーションデータを時間軸上に連続して接続して新しい動作を生成する手法と、顔の動作と身体の動作を時間軸上に並列に制御し感情を伴ったジェスチャーを生成する手法があるが、本研究では、感情を伴った複数の身体動作を時間軸上に並列に制御することにより新しい動作を生成する。具体的には、図 5 に示すように、モーションデータ run と lookback を同時に同じモデルデータに適応する際に、lookback が t_{start} の時間遅れて適応するようにする。

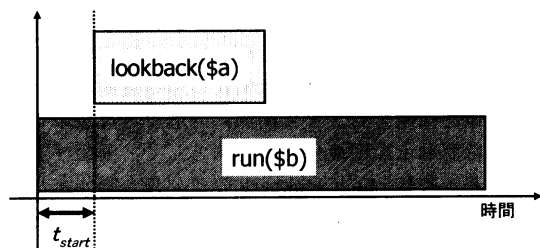


図 5 基本動作同期合成

2 つの動作がどれだけの時間間隔で発現するかは、感情パラメータの覚醒度への依存が大きいため、 t_{start} は、覚醒度の値 $arousal$ と、図 4 の V2MML 記述で指定する動作の開始時間の相対的な初期値 $t_{initial}$ とを引数とする(1)式でモデル化する。

$$t_{start} = f(arousal, t_{initial}) \quad \dots (1)$$

3.4. 文字修飾-動作強調変換

感情の強さをキャラクタアニメーションの自動生

成に反映させるには、前後の文脈や状況判断が必要となり処理や条件判断が複雑になる。本研究では、一般のユーザが電子メールで自分の表現したい感情を絵文字やフェイスマークで書くように手軽に感情の制御を行うため、文字修飾から動作の強調度への変換を行う。

具体的には、入力する自然文テキストの文字修飾としてフォントの大きさがデフォルト値より変化するポイント数を強調度のレンジ (-1.0~1.0) へ正規化して変換することにより、強調度の異なる基本モーションデータへの変換を行う。感情パラメータをからの基本モーションデータ変換と同様に、感情パラメータと強調度が表す拡張感情空間の位置に最も近い基本モーションデータへ変換することが可能となる。

4. 実装と検証

CG をハードウェアで高速に描画するための GPU (Graphics Processing Unit) が CPU (Central Processing Unit) よりも早い速度で進化しており、携帯端末への GPU 搭載および携帯端末用 3 次元 CG API として OpenGL ES 2.0 や Mobile 3D Graphics が規格化されたことにより携帯 3 次元 CG コンテンツの環境が整いつつある。

我々はこれまで PC 上で動作する TVML プレーヤ II [13] を用いて DMP を実装・検証してきたが、今回 iappli Development Kit for DoJa 3.5 [14] を用いてシナリオ駆動型動作合成を実装・検証した。

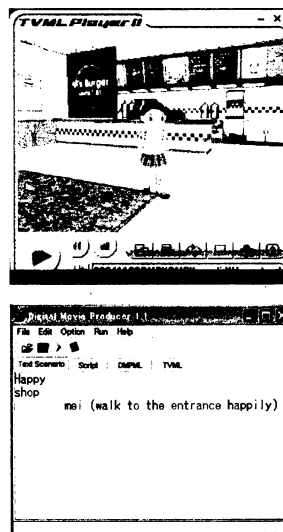


図 6 従来の DMP 動作画面例 (TVML プレーヤ II 利用)

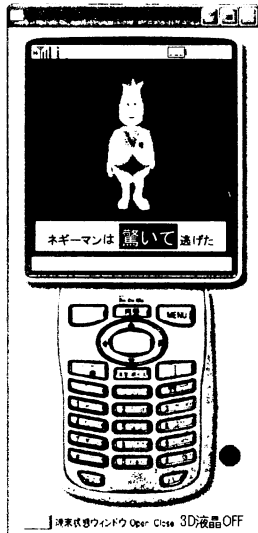


図 7 シナリオ駆動型動作合成の動作画面例 (iappli Development Kit for DoJa 3.5 利用)

従来の DMP の画面を図 6 に示す。右のシナリオ入力部において入力された感情を表す副詞 happily を TVML のコマンド character: expression(...type=happy...)へ変換し、歩く動作と同時に実行するように制御している。

表情での感情の表現は、上半身を写す Bust Shot などの構図では感情が伝わるが、全身を写す Long Shot などの構図では違いが判りにくいという問題があった。これに対し、本研究で実現した感情を伴った動作合成を用いることにより、全身での感情表現が可能となった。

5. 今後の課題

今後の課題として、様々な基本モーションデータを用いた基本動作同期合成の評価と改善、モデルデータの構造を考慮した動作合成の最適化が挙げられる。

また、モデルデータや基本モーションデータなどの部品データがネットワークを介して流通し、一般のユーザがこれらの部品データを再利用して短時間に新しいコンテンツを制作するためには、モデルデータと基本モーションデータのフォーマットおよび再生環境が統一され、動作制御インタフェースが共通化される必要がある。

6. 謝辞

本研究の一部は、文部科学省の科学技術振興調整費重要課題解決型研究「デジタルシネマの標準技術に関する研究」の一環として推進している。

文 献

- [1] 宮崎誠也, 申金紅, 青木輝勝, 安田浩, “自動演出生成方式によるシナリオ入力型 CG 動画制作システム”, 信学技報, IE2003-135, pp.31-34, Dec. 2003.
- [2] Shen Jinhong, Seiya Miyazaki, Terumasa Aoki, Hiroshi Yasuda, “Filmmaking Production System with Rule-based Reasoning”, Proceedings of Image and Vision Computing New Zealand 2003, pp.366-371, New Zealand, Nov. 2003.
- [3] Seiya Miyazaki, Shen Jinhong, Terumasa Aoki, Hiroshi Yasuda, “A System from Script to Motion Picture”, The 9th International Conference on Information Systems Analysis and Synthesis: ISAS '03 (CCCT2003), Vol. IV, Florida, USA, Jul. 2003.
- [4] 斎藤豪, 井元崇之, 中嶋正之, “モーショングラフによる自律型キャラクターの動作アニメーション生成手法”, 信学技報, CG104, pp.47-52, Sept. 2001.
- [5] Lucas Kovar, Michael Gleicher, Frédéric Pighin, “Motion Graphs”, In Proceedings of ACM SIGGRAPH 02, Jul. 2002
- [6] Charles Rose, Michael F. Cohen, Bobby Bodenheimer, “Verbs and Adverbs: Multidimensional Motion Interpolation”, IEEE Computer Graphics and Applications 18, 5, pp.32-40, Sept. 1998.
- [7] Okan Arikan, David A. Forsyth, James F. O'Brien, “Motion Synthesis from Annotations”, Proceedings of SIGGRAPH 2003, CA, Aug. 2003.
- [8] 寺崎卓也, 上川千絵, 尾下真樹, “オブジェクト指向による動作データの管理手法と自然言語からのアニメーション生成システム”, Visual Computing/グラフィックスと CAD 合同シンポジウム, Jun. 2004.
- [9] Norman Badler, Rma Bindiganavale, Juliet Bourne, Martha Palmer, Jianping Shi, William Schuler, “A Parameterized Action Representation for Virtual Human Agents”, In Workshop on Embodied Conversational Characters, Lake Tahoe, CA, Oct. 1998.
- [10] Tonia Bleam, Martha Palmer, K. Vijay-Shanker, “Motion Verbs and Semantics Features in TAG”, TAG+ 4 Workshop, Philadelphia, PA, Aug. 1998.
- [11] 森純一郎, Helmut Prendinger, 土肥浩, 石塚満, “ユーザ感情のモデル化と感性的インタラクション”, 人工知能学会 第 3 回 MYCOM オンライン資料, pp.9-16, 2002.
- [12] Hua Wang, Helmut Prendinger, Takeo Igarashi, “Communicating Emotions in Online Chat Using Physiological Sensors and Animated Text”, ACM CHI '04 April, Vienna, 2004.
- [13] TVML プレーヤ II, <http://www.nhk.or.jp/strl/tvml/japanese/player2/index.html>
- [14] DoJa3.5 プロファイル向け i アプリ作成ツール, http://www.nttdocomo.co.jp/mc-user/i/java/tool_fom_a2.html