

## 自然言語からの仮想人間の複合動作アニメーションの生成

森脇 晃† 國分 剛† 尾下 真樹 †

九州工業大学大学院 情報工学研究科 情報科学専攻†  
九州工業大学 情報工学部 システム創成情報工学科†

〒820-8502 福岡県飯塚市川津 680-4

E-mail: moriwaki@cg.ces.kyutech.ac.jp, kokusan@cg.ces.kyutech.ac.jp

我々は、入力された自然言語による文章からアニメーションを生成するシステムを開発している。本論文では、複数の動作を並列・連続して行うような入力文に対応した手法を提案する。本手法では、入力された自然言語文中の動作間の時間関係を表す要素に注目した意味解析を行い、各動作をどのような組み合わせ・順序で実行するという情報を持つ動作時間テーブルを生成する。その動作時間テーブルに基づき、データベースから検索された動作データを接続または合成することで自然なアニメーションを生成する。

## Generating Animation of Composite Motions from Natural Language Text

Akira MORIWAKI Tsuyoshi KOKUBU Masaki OSHITA

Department of Systems Innovation and Informatics Kyushu Institute of Technology  
680-4 Kawazu, Iizuka-Shi, Fukuoka, 820-8502, Japan

E-mail: moriwaki@cg.ces.kyutech.ac.jp, kokusan@cg.ces.kyutech.ac.jp

We have been developing an animation system that converts a natural-language input text to an animation. In this paper, we present new methods to a handle input text that includes a series of serial and parallel motions. Our system retrieves the temporal relationship between verbs in an input text and generates a time table which represents the execution-schedule of corresponding motions. A synthesized animation is generated by connecting and compositing motion data that are searched from a motion database.

### 1. はじめに

近年、映画やTV等における、3DCGアニメーションの利用が普及化している。3DCGアニメーションでは、キャラクターの動きの元となる動作データが使われる。しかし、映画等で必要となる動作のデータは膨大な数となり、動作の作成作業には多くの手間隙がかかる。そこで、一度使用した動作データを再利用したいという要求がある。し

かし、多種多様な動作データを、従来の階層型ファイルシステム、単純なテーブル方式のリレーショナルデータベース等で管理することは困難である。

また、最近では、アニメーションの作成ツールの高度化、多機能化にともない、ツールを使うために必要な専門知識も多く、一般人には扱いにくいものとなっている。この問題に対し、誰でも手軽に扱える作成ツールの要求がある。

本研究では、大量の動作データを管理し、入力文より動作データを検索することで、アニメーションの生成を行うシステムの開発を行っている[4]。本システムでは、自然言語で記述された文章を与えると、文章にあった動作データ、キャラクターをデータベースより検索する。検索されたデータを用いて自動的にアニメーションの生成を行うことを可能としている。しかし、これまでは複数の動作を並列に行う複合動作に対応していなかった。

複数の動作の連続・並列実行の実現に取り組んでいる[5]。本研究では動作同士の時間・依存関係を表す「動作時間テーブル」を提案する。さらに動作時間テーブルに基づき動作データを自動で接続または合成することで複合動作アニメーションを生成する手法を提案した。今回は、動作を並列して行う動作合成についての手法を新たに提案する。動作合成とは、ある動作の一部を別の動作の部位に切り替える場合を言う。また切り替え部位による全身が受ける影響を抽出し、適用する。

第2章では本システムの概要について述べる。また、第3章で自然言語解析について、第4章で動作接続・合成手法について述べる。第5章で実験と考察を行い、第6章で本論文のまとめを行う。

## 2. システムの概要

### 2.1 システムの構成と処理の流れ

図1に全体のシステム構成と全体の処理の流れを示す。本システムでは、入力文を与えると、文に応じたアニメーションを生成する。まず、入力文章の自然言語処理を行い、動作データ検索に必要なキー情報を抽出する。データベース内の動作データの候補に対して、キー情報と、動作データに付随した動作フレームを比較、評価することにより、最適な動作データを検索・出力する[4]。今回は、動作データ検索・出力に並行して時間関係に注目した意味解析を行い、抽出された時間制約条件を出力する。次に、この時間制約条件と検索された動作データから、動作時間テーブルを作成する。その後、動作時間テーブルに基づいて複合動作のアニメーションを生成する。

### 2.2 複合動作

本研究における複合動作とは、複数の動作が時間的に重なりを持つものを示す。つまり連続動作や並列動作を示す。連続動作は「Aをした後しばらくしてBをする」「Aをし終わると同時にBを始める」などであり、並列動作は「AをしながらBをする」などである。

Input: 「走って、ボールを蹴る。そして喜ぶ。」

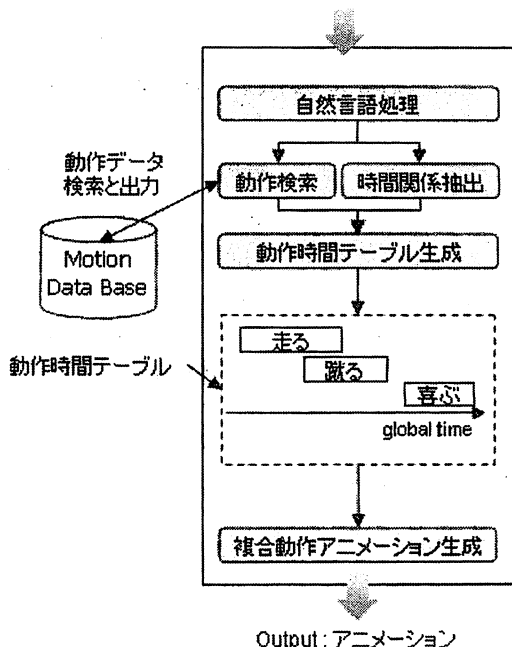


図1 システムの構成図

### 2.3 動作時間テーブル

本研究では、複合動作の生成を実現するために、動作データと動作の時間関係を持つ動作時間テーブルを定義した。動作時間テーブルの各要素は、1つの動作を表し、以下の表1のデータを持つ。これを入力文中の動作の数だけ生成し、動作時間テーブルとした。本来、並列動作の場合、動作データとともに動作の主要部位を文章から取得するが、現在の手法では取得できていないため、手で動作データにメタ情報として付加している。

表1 動作時間テーブルの1要素のデータ構造

データ構造	説明
動作主	動作を行うキャラクター(人物)
動作データ	各時刻の全関節角度 切り替え部位 (合成の場合)
接続・合成方法	前要素との繋がり方 Connection, Transition, Composition
前要素への参照	接続・合成に必要
開始時間	動作の開始タイミング

## 3. 自然言語からの複合動作スケジューリング

本章では、自然言語入力から動作時間テーブル

を生成するための方法として、自然言語入力からの時間制約条件抽出と、時間制約条件からの動作時間テーブル生成について説明する。

自然言語からのアニメーション生成として馬場らの研究[3]がある。このシステムではキャラクターの移動に注目し、動作の時間関係とオブジェクトの配置関係についての条件を文章から抽出しているが、本研究では動作の時間関係のみに注目して研究を行っている。

### 3.1 時間制約条件

入力文から動作の時間関係を設定する際に、本システムでは、2動作間の適切な関係を設定するために、それらの要素ごとに時間制約条件として抽出する。表 2 に時間制約条件の構造を示す。接続方法の制約は、動作が連続して行われたり、並行に行われたりする場合について表す。タイミングの制約は、接続方法の制約が連続の場合に、前後の動作にオーバーラップが有るか無いかを表す。時間制約は、動作 2 の開始タイミングの細かい指定が出来る場合、数値で表す。

表 2 時間制約条件の構造

データ構造	内容
関連動作	動作 1, 動作 2
接続方法の制約	連続 or 並列
タイミングの制約	オーバーラップ有・無
時間制約	動作 2 の開始タイミング

### 3.2 時間制約条件抽出ルール

本研究では、入力文中の接続助詞や句読点に注目し、時間制約条件を抽出するルールを設定した。以下にそのルールを示す。

動作 1 の動詞を持つ文節にこれらの要素が含まれている場合、条件を以下のように抽出する。

- 接続助詞「て」 or 読点 (,)  
連続に接続, オーバーラップ有  
動作 2: 文節の係り先の動作
- 句点 (.)  
連続に接続, オーバーラップ無  
動作 2: 文中の次の動作
- 接続助詞「ながら」 or 「つつ」  
並列 (合成)  
動作 2: 文節の係り先の動作

### 3.3 動作時間テーブルの生成

自然言語入力から抽出した時間制約条件(3.1 節)と、動作検索[4]によって検索された動作データを用いて動作時間テーブルを生成する。(図 1 参照)

接続制約が連続で、オーバーラップが無い場合には接続・合成方法を Connection, 時間制約が与

えられていればその値を開始時間へ反映させる。また、接続制約が連続でオーバーラップする場合は、接続・合成方法を Transition とする。この場合、自然言語からは時間の指定を行うことができない。接続制約が並行の場合は、接続・合成方法を Composition とする。

## 4. 複合動作アニメーションの生成手法

### 4.1 全体の流れ

本手法では、まず、自然言語解析結果から作成された動作時間テーブルと動作データを入力として受け取る。次に、動作時間テーブルに付随している動作同士の時間関係と接続・合成手法の情報に基づき、接続や合成の対象となる動作データから動作切り替えタイミングを決定する。そして、動作間に推移動作を生成し、滑らかな複合動作アニメーションを生成する。

### 4.2 動作接続・合成手法の分類

本研究では、動作を連続して行う場合は動作接続、並列に行う場合は動作合成手法を用いる。動作接続・合成手法は数多く研究[1]されているが、手法の名称が統一されていない。そのため本研究では、複合動作を生成する際に必要となる手法を 3 つに分類した。

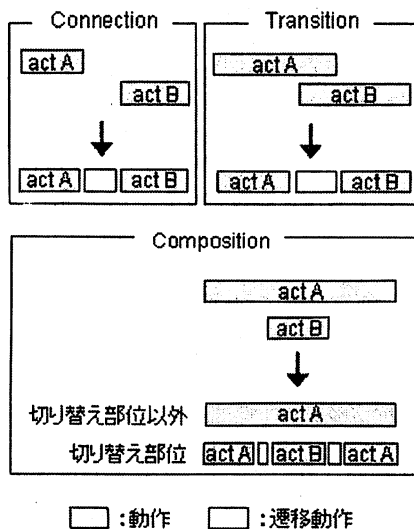


図 2 動作接続・合成手法の分類

[動作接続]

- Connection  
動作間を一定時間空けて接続する場合の手法

- Transition

actionAの一部とactionBの一部を重ねて接続する場合の手法。「歩く」と「走る」の様な、同様のサイクルを持った動作を接続する。

[動作合成]

- Composition

actionAの一部(切り替え部位:動作データにメタ情報として付随している)をactionBに切り替える手法

### 4.3 動作接続手法

この節では、動作接続の手法について述べる。本手法では、足の拘束条件に注目し、動作を切り替えるタイミングを検出する。その後、姿勢ブレンドまたは動作ブレンドを用いて遷移動作を生成する。また、遷移動作中、足が地面を滑る動作が生成されることを防ぐために、地面に接している足の位置を拘束する。

#### 4.3.1 動作接続における、動作切り替えタイミング決定手法

本手法では、キャラクターの足と地面の接触条件から、動作切り替えタイミングを決定する。前動作の終了付近と後動作の開始付近において、足と地面の接触条件を検索する。同じ接触条件を持ったフレームを切り替えタイミングとする。完全に一致した接触条件は必要せず、少なくとも片足の接触条件が同じであれば良い。

キャラクターの姿勢を切り替えタイミングの決定に用いることをせず、この様な手法を用いる理由は、遷移動作中に足が地面を滑る動作が生成されることを防ぐことが容易になるためである。ここで決定した切り替えタイミングを利用して、動作間に遷移動作を生成することで、対象とした動作データを接続する。遷移動作生成手法は次節で述べる。

#### 4.3.2 動作接続における遷移動作生成手法

本研究では、決定した切り替えタイミングにおけるキャラクターの姿勢を用いて、動作間に遷移動作を生成する。

Connection の場合は、決定した切り替えタイミングにおけるキャラクターの関節角度をブレンドする。時間経過に応じてブレンド比率を変えることで、動作間に遷移動作を生成する。

Transition は「歩く」と「走る」の様な、同様のサイクルを持った場合、「走って、蹴る」といった連続動作の場合の動作接続手法である。前者の場合は、「足を上げる→足を着く」の様なサイクルが両動作に存在する。本研究では、「足を上げる」と「足を着く」を接続対象となる両動作

データから検出し、キーフレームとする。その後、Time Warping により、両動作のサイクルの時間を合わせ、両動作のキーフレーム間のキャラクターの関節角度をブレンドすることで遷移動作を生成する。後者の場合は、Connection と同様に扱う。

また、遷移動作中の足の滑りを解決する必要がある。本手法では、遷移動作中、足の空間位置を固定し、固定中の足の関節角度を、Inverse Kinematics 計算により求める。遷移動作中に固定した足については、遷移動作後に後動作へ遷移させる必要がある。本手法では、遷移動作中に固定していた足が、後動作において、宙に浮いている間に関節角度のブレンドによって遷移させる(式(3))。

### 4.4 動作合成手法

この節では、動作合成の手法について述べる。例えば、「歩く」動作の途中の右腕の動きのみを、「右手を振る」動作の右腕の動きに切り替えることで「右手を振りながら歩く」という新しい動作を生成する。しかし、単純に2つの動作データを組み合わせるだけでは、最終的な動作が必ずしも自然にはならない。例えば、「右手を振る」動作では、右手の動きが主な動きの要素であるが、右手を振ることによって、背中や腰などの部位にも微妙な揺れが生じている。右手の動きだけを別の動作と組み合わせると、合成後の動作にはこのような揺れが生じないため、最終的に生成される動作が不自然となってしまふ。

本手法では、2つの動作における指定部位の位置・速度や合成開始・終了時刻の間隔を条件として評価することで、適切な合成開始・終了時刻を決定する。また、切り替え部位の動きから、それ以外の部位が受ける影響を抽出して、それを合成動作に適用することで、揺れの影響を考慮した、より自然な動作を生成する。

#### 4.4.1 動作合成における、動作切り替えタイミング決定手法

本手法では、キャラクターの腰を基準とした指定部位の位置・速度と、開始・終了時刻の間の時間と元の動作の時間の伸縮度合いの2つの条件を考慮して、適切な合成開始・終了時刻を決定する。

Gardon[2]らは、2つの動作の前後をなめらかに合成することで新たな動作を生成する場合に、合成を足と地面の間の拘束条件を考慮して開始・終了時刻を計算する手法を提案している。しかし、彼らの手法では、全身の動作の遷移を想定している。2つの動作を前後に接続するだけであるため、動作の伸縮による時間変化などは考慮していない。本手法では、動作合

成を対象とするため、従来手法とは異なる評価関数を使用する。

$$F(t_a, t_b) = \alpha |P_1(t_a) - P_2(t_b)| + \beta |V_1(t_a) - V_2(t_b)| + \gamma \cdot T_f \quad (1)$$

Fは評価値、 $\alpha, \beta, \gamma$ は各要素重み、 $P_n(t_x)$ 、 $V_n(t_x)$ は動作nのxフレーム目のエンドエフェクタの位置と速度を示す。T<sub>f</sub>は、切り替え動作の伸縮時間を表し、第2動作の長さ、開始・終了時間の差が近いほど小さな値になるように、それぞれの長さの比から計算する。

本手法では、まず、第1動作と第2動作の初期区間において、評価値Fが最小となる時刻の組を、切り替え開始タイミングとして決定する。このとき、式1の第3項目は用いない。次に、第2動作の終了区間と絶対時間上でそれに近い第1動作の区間において、評価値Fが最小となる時刻の組を切り替え終了タイミングとして決定する。このとき、第2動作の伸縮度合いを考慮するために、式1の第3項目が必要になる。

#### 4.4.2 動作合成における遷移動作生成手法

本手法では、指定部位を切り替える場合、遷移動作を生成することで滑らかに切り替える。まず、遷移動作の時間的長さ *blend\_duration* を求める。

$$\text{blend\_duration} = \frac{|P_{a,i} - P_{b,j}|}{V_{a,i}} \quad (2)$$

次に、遷移動作を指定部位の関節角度の線形補間によって生成する(式(3))。

$$\text{angle} = (1 - \text{ratio}) \times m1\_angle_i + \text{ratio} \times m2\_angle_j \quad (3)$$

$$\text{ratio} = \frac{\text{time}}{\text{blend\_duration}} \quad (4)$$

$mX\_angle$  は動作 X の関節 i の角度、*ratio* はブレンド比率、*time* は遷移動作開始からの経過時間を表す。

#### 4.4.3 揺れの混ぜ合わせ

本手法では、指定された切り替え部位以外の部位の動きから揺れの情報を抽出し、合成後の動作に適用することで、自然な動作を生成する。このとき、合成する2つの動作では、姿勢や動きが異なるため、単純に関節角度をそのまま適用することはできない。本手法では、第1動作の関節角度軌道を滑らかにすることで、揺れが生じなかったときの関節角度軌道とする(図3の左図、以下基本軌道)。具体的には、基本軌道の開始時点の角度・角速度、終了時点の角度・角速度から、エルミート関数によって、この関節角度軌道と、実際の関節角度軌道との差分を、揺れとして抽出する(図3の左図点線)。そして、揺れの成分のみを、

合成動作の切り替え部位以外の関節角度軌道に加算する。

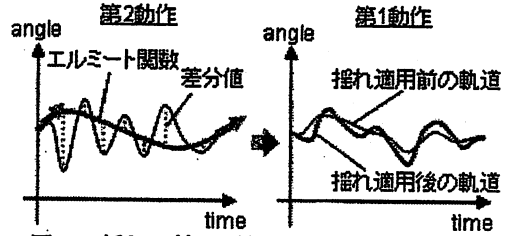


図3 揺れの抽出(左図)と適用(右図)

## 5. 実験

### 5.1 実験内容

今回の実験は、まず、自然言語の入力文を解析する。その後解析結果から動作時間テーブルを作成する。その動作時間テーブルを元に複合動作アニメーションを生成した。

### 5.2 自然言語解析実験結果・考察

自然言語からの複合動作スケジューリングの提案手法についての実験結果を以下に示す。

- 入力 (テキスト形式)  
「太郎は走って、ボールを蹴る。そして喜ぶ。」
- 出力

#### [要素 1]

動作主格	太郎
動作データ	A33
動作名	走る
接続形式	初期動作

#### [要素 2]

動作主格	太郎
動作データ	F20
動作名	蹴る
接続形式	Transition
前要素	要素 1

#### [要素 3]

動作主格	太郎
動作データ	I16
動作名	喜ぶ
接続形式	Connection
前要素	要素 2

以上のように、動作が連続(Transition, Connection)する場合は正しい結果が出力された。

並列(Composition)に関しては、別の実験において正しい結果の出力を確認したが、切り替え動作(動作2)の開始タイミングの設定方法が決定していないため、現在は同時に開始する。また、「ガッツポーズする」などといった、「〇〇する」という形の動詞の場合、正しく動作として抽出す

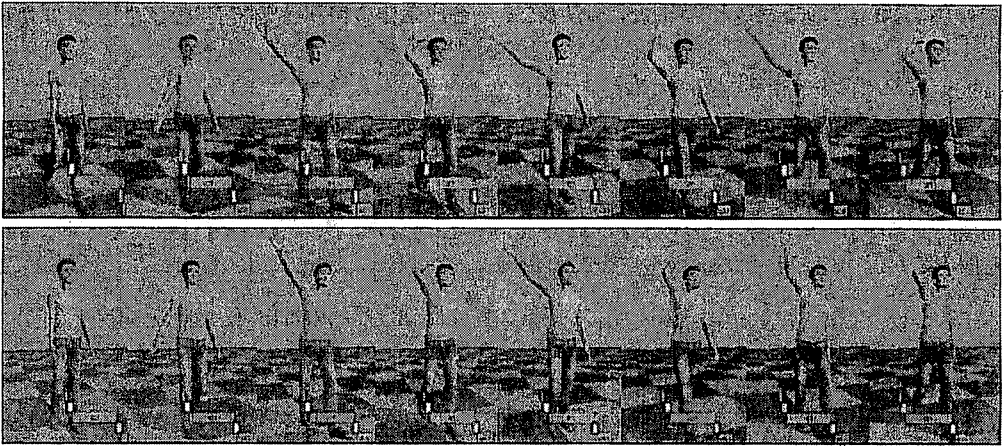


図 4: Composition 「歩く」 + 「右手を振る」。上段: 揺れ適用前, 下段: 揺れ適用後

ることが出来なかった。「する」を自立語として判定できていなかったためだと考えられる。

### 5.3 アニメーション生成実験結果・考察

今回の実験では、自然言語解析によって生成された動作時間テーブルを入力とし、Transition と Connection 「走って、ボールを蹴る。喜ぶ。」を実現したアニメーションを生成した。

Transition と Connection の例に関しては、足の拘束条件に基づいて切り替えタイミングを決定し、動作間に遷移動作を生成することで、滑らかに接続することができた。しかし、Connection では、一定速度で遷移をする様な現実の人間が行わない動作が生成される場合があった。これは遷移動作が関節角度の補間のみで生成されていることが原因である。つまり全く力学を考慮していない動作が生成されている。

Composition 「歩きながら右手を振る」のアニメーション生成の結果を図 4 に示す。図 4 上段では、右手を切り替えただけで、そのほかの部位は「歩く」動作であり、不自然な動作となる。図 4 下段では、右手の切り替えに加え、全身に揺れを混ぜ合わせた。そのことで、右手の振りによって腰や首が自然に揺れている。結果より、右手を振ることによって全身が受ける影響を抽出し、それを「歩く」動作に適用することで、より人間らしいアニメーションを生成に成功したと言える。

## 6. まとめ

本論文では、オブジェクト指向データベースを用いた自然言語からのアニメーション生成システムとして、時間関係に注目した意味解析による

動作時間テーブルを用いた動作スケジューリング手法と、複合動作アニメーション生成のための動作接続・合成手法を提案した。

今後の課題として、自然言語解析については、並列動作を含む場合について、入力文から適切な切り替え時間を抽出する手法についての検討、対応できていない動詞についての検討が必要である。アニメーション生成については、Connection の例において人間的でない遷移動作が生成される問題について、力学を考慮して動作を生成する手法について検討が必要である。

## 参考文献

- [1] Charles Rose, Brian Guenter, Bobby Bodenheimer, Michael F. Cohen, "Efficient Generation of Motion Transitions using Spacetime Constraints", SIGGRAPH 1996
- [2] P. Ghardon, R. Boulic, D. Thalmann, On-line Adapted Transition between Locomotion and Jump, In Proceeding of Computer Graphics International (CGI), Stony Brook, NY, USA, June 2005.
- [3] Hiromi BABA, Tsukasa NOMA, Naoyuki OKADA, "Visualization of Temporal and Spatial Information in Natural Language Descriptions", IEICE Transaction on Information and Systems, Vol. E79-D, No. 5, pp. 591-599, May 1996.
- [4] 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 2004
- [5] 國分剛, 森脇晃, 寺崎卓也, 尾下真樹, "自然言語からの仮想人間の複合動作アニメーションの生成", 火の国シンポジウム 2005