

# マルチメディア理解システムIMAGES-Mにおけるアニメーションと自然言語文の相互変換

古川 忠昭 笠 晃一 横田 将生  
(福岡工業大学)

人間とコンピュータ間のマルチメディアコミュニケーションを可能とするシステムIMAGES-Mの一部である動画(アニメーション)の相互変換を行うサブシステムについて報告する。IMAGES-Mは自然言語、知識表現、心像(イメージ)、世界を包括的に取り扱う心像意味論(MIDST:Mental Image Directed Semantic Theory)に基づいている。このサブシステムは、入力された自然言語文を心像のモデルである抽象的な属性空間の軌跡として意味解釈し、環境条件等の制約を満足するように具体化した後、アニメーションに変換し、逆に心像のアニメーションを論理式化された軌跡を介して自然言語文に変換する。本システムにより人体モデルを用いた動作の自然言語表現とアニメーションとの間の相互変換の実験を行った結果、かなり一般性のある処理が実現できる見通しを得た。

## Mutual conversion between animation and natural language in multimedia understanding system IMAGES-M

Tadaaki FURUKAWA Koichi RYU  
Masao YOKOTA  
Fukuoka Institute of Technology

This paper reports the mutual converter between natural language and animation which is one of the subsystems of multimedia understanding system IMAGES-M. IMAGES-M is based on MIDST(Mental Image Directed Semantic Theory), which is such a total theory that deals with natural language, knowledge representation, mental image and worlds. This subsystem interprets input sentences as loci in the abstract spaces, so called Attribute Spaces, it instantiates the interpretations in accordance with the environmental or contextual restrictions and, translates them into animations, and animations translates into sentences via logically formulated loci. Several mutual converting experiments using human body models on this subsystem have given a fairly good proof of generality of the processing.

### 1 はじめに

現在までに、自然言語の意味に関する研究が数多くなされてきている。しかしながら、自然言語

理解処理における人間の理解過程を十分に説明でき、しかもコンピュータ処理にも適するようなも

のではなかった。そこで自然言語の意味を人間の心像現象と対応づけ、コンピュータにも適した理論(心像意味論と呼ぶ)の構築を行った。これは人間の自然言語解釈は心像(イメージ)に基づくものとし、心像のモデルを与えることにより自然言語意味記述および理解処理方法を提案しているものである。本研究は、コンピュータと人間との自然言語と動画(アニメーション)によるコミュニケーションの実現を目的としている。人間によって入力される自然言語文から、理解結果を視覚的なアニメーションとして表示することによるコミュニケーションをはかる。心像意味論により入力される自然言語文を心像のモデルである抽象的な属性空間の軌跡として意味解釈する。次にこの解釈結果の表現(軌跡式)を環境条件等の制約を満足するように具体化し、アニメーションに変換する。また心像のアニメーションを入力することで、自然言語文を出力することができる。

本研究は人体モデルを用いた言語(軌跡式)とアニメーションの相互変換を行っている。そのため、まず人体モデルに対しての動作に関する言語表現の一般化を行った。そこで以下では2章でここで行った動作に関する言語表現の一般化について説明し、3章でシステム構成、4章でシミュレーション、5章においてまとめについて述べていく。

## 2 動作に関する言語表現の一般化

人間の動作は後述する制約条件により図1のように区別することができる。ここでまず動作命令を以下のように一般化した。

(人間)が(人体の部位)を([C])動かす。

[ ]は任意要素であることを示している。動作属性(C)を表1に示す。但しこの表1において“V”はベクトル(多次元)、“S”はスカラー(1次元)を表す。

表 1: 動作に関する属性

番号	属性	性質	属性表現例
A12	位置	[V]	p から q へ
A13	運動方向	[V]	上, 下
A14	向き	[V]	北, 南
A15	運動軌跡の形	[V]	ジグザク
A16	速さ	[S]	速い, 遅い
A17	移動距離	[S]	50cm
A38	回数	[S]	1回, 2回

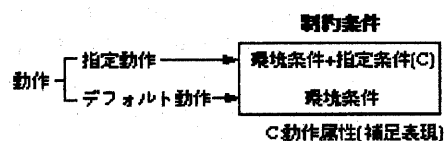


図 1: 動作パターン

## 3 システム構成

システムは、意味表現変換モジュール・アニメーション入出力モジュールにより構成している(図2)。自然言語からアニメーションへの変換は、人間により入力される自然言語文をIMAGES-Mの文字言語処理部において自然言語意味表現(軌跡式)に変換させ、次にその軌跡式を制約条件を考慮しコンピュータで処理できるコマンド列に意味表現変換モジュールにおいて具体化し、アニメーション化する。アニメーションから自然言語への変換は、アニメーション入出力モジュールにおいて入力されたアニメーションをコマンド列に変換し、意味表現変換モジュールにおいて軌跡式に変換する。軌跡式は、文字言語処理部にて自然言語文に変換される。

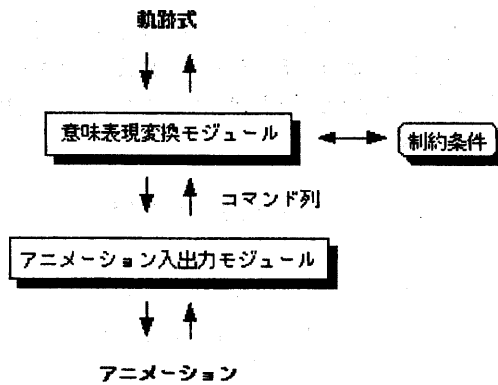


図 2: システム構成

### 3.1 意味表現変換モジュール

軌跡式とコンピュータで扱いやすいコマンド列との相互変換を行う。その際、環境条件等の制約を考慮する。

#### 3.1.1 制約条件

一般に、人間の心(精神)の世界を内界、それが感知する物理的な世界を外界と呼んでいる。外界物は、物質や物体である物質体、および、それらが存在する物理的領域よりなる。外界事物は、そのような外界物が関与する事象で、物の移動や変形などを指示する。

外界事物である人間が、その存在する物理的領域で動作をする場合には図で示したように、大きく分けて環境条件と指定条件の2つの条件があると考えられる。

#### 環境条件

外界事物における物理的領域と、人体の部位の制約条件がある(図3)。部位の制約条件とは部位同士のつながりや、部位それぞれの可動性(動かせる方向、速さなど)である。

#### 環境条件

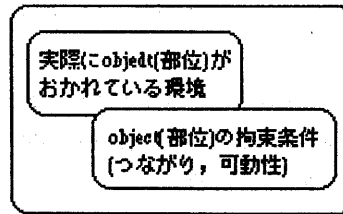


図 3: 環境条件

(例)  $connect(x, y) \wedge connect(y, z)$

$\wedge$  手(x)  $\wedge$  腕(y)  $\wedge$  胴(z)

#### 指定条件

表1に示したような動作に関する属性(補足表現)を指す。

(例) 左へゆっくり歩く。

#### 3.1.2 コマンド列

コマンド列を以下に示す。

```
{ continue, return, number, model,
  {{ part, frame { xbegin, xdou, xend },
    { ybegin, ydou, yend }, { zbegin, zdou, zend } } } }
```

それぞれのパラメータは以下を表す。

- continue : 継続動作回数
- return : 往復動作回数
- number : 動作する部位数
- model : 動作するモデル
- part : 動作する部位番号
- frame : 1動作のフレーム数

- xbegin,xdou,xend : 縦方向の開始角度, 縦方向1フレームの動作角度, 縦方向の終了角度
- ybegin,ydou,yend : 縦軸方向回転の開始角度, 縦軸方向回転1フレームの動作角度, 縦軸方向回転の終了角度
- zbegin,zdou,zend : 横方向の開始角度, 横方向1フレームの動作角度, 横方向の終了角度

### 3.1.3 軌跡式およびコマンド列間の相互変換

本研究では人体モデルを障害のない空間に配置していると仮定した。そのため、環境条件としてはobject(部位)の拘束条件のみとなる。後、コマンド列への相互変換で考慮することは指定条件となる。これらの相互変換に際して具体的な例を挙げて説明する。

#### (a) 軌跡式からコマンド列への変換

(例) locus(ICHIRO migite ue ue a13)

この軌跡式に対するコマンド列は以下のようになる。

```
{0,0,1,1,{{3,18,{0,10.000000,180},
{0,0.000000,0},{0,0.000000,0}}}}
```

軌跡式は継時的連言でないため、continue、return は0となる。また同時的連言でもなく、動作の対象となる事物は‘migite’のみであるため、number は1となる。さらに、速さに変化を持たせるような軌跡式もないため、frame はデフォルトで用意しているパラメータを代入している。動作は属性番号‘a13’の運動方向(‘ue’)上向きとなる。そこで現在の位置よりも縦方向において上への移動となるような動きをすればいいことになる。動作において上向きの動作という他に、指定動作がないため、単純に縦方向のみの動きとしている。例えばこの例の軌跡式の他に、運動方向

が‘migi’や‘hidari’などの語彙が含まれるようになる。部位の制約条件より縦軸方向回転の角度や横方向の角度などが加わりコマンド列が複雑になる。

#### (b) コマンド列から軌跡式への変換

(例) {0,0,1,2,{{3,36,{0,10.000000,0},
{0,0.000000,0},{0,0.000000,0}}}}

このコマンド列に対する軌跡式は以下のようになる。

locus(TARO migite en en a15)

continue、return が0のため、軌跡式に継時的連言がない。model が2のため、対象のモデルは‘TARO’、またnumber は1のため同時的連言もなく、部品番号3‘migite’のみである。動作方向は縦方向へ10度ずつ36フレームで動き、開始角度、終了角度が同じであるため、部品‘migite’は開始時と同じ所へ戻ってきている。これにより動作は属性番号‘a15’の運動軌跡が‘円’となる。

以下に属性とそれに対応するコマンド列のパラメータを表2に示す。

表2: 属性とパラメータの対応表

	位置	運動方向	向き	運動の軌跡	連言	移動距離	図表
continue	0	0	0	0	0	0	0
return	0	0	0	0	0	0	0
number	0	0	0	0	0	0	0
frame	0	0	0	0	0	0	0
part	0	0	0	0	0	0	0
xbegin	0	0	0	0	0	0	0
xdou	0	0	0	0	0	0	0
xend	0	0	0	0	0	0	0
ybegin	0	0	0	0	0	0	0
ydou	0	0	0	0	0	0	0
yend	0	0	0	0	0	0	0
zbegin	0	0	0	0	0	0	0
zdou	0	0	0	0	0	0	0
zend	0	0	0	0	0	0	0

### 3.3 アニメーション入出力モジュール

各部位に部位番号の与えられた人体モデルを、障害のない空間にあらかじめ配置しておく。意味表現変換モジュールよりコマンド列を受け取ったアニメーション入出力モジュールは、コマンド列に指定されている部位をアニメーションさせる。

ユーザーがアニメーションを入力した場合、その入力したアニメーションをコマンド列に変換し、意味表現モジュールに出力する。ユーザーは、マウス・キーボードを用いてアニメーションの入力を行う。

#### 4 シミュレーション

「両手をゆっくりあげる」という文を入力した場合について説明する。

意味言語処理部より次のような軌跡式を受け取る。

```
locus(ICHIRO migite ue ue a13)
□locus(ICHIRO hidarite ue ue a13)
□locus(ICHIRO migite yukkuri yukkuri a16)
□locus(ICHIRO hidarite yukkuri yukkuri a16)
```

これをコマンド列に変換する。

```
{0,0,2,1
{{3,36,{0,5.000000,180},{0,0.000000,0},
{0,0.000000,0}},
{4,36,{0,5.000000,180},{0,0.000000,0},
{0,0.000000,0}}}}
```

'a16' の 'yukkuri' という副詞が入っているため、デフォルトで用意されている一動作の角度を小さくし、フレーム数を増やして動作をゆっくりさせている。このシミュレーションのアニメーション例を図4に示す。

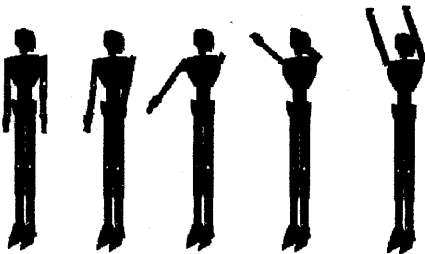


図4: アニメーションの出力例1

また、「タローは手を回す」という文を入力した場合、軌跡式は以下ようになる。

```
locus(TARO migite en en a15)
コマンド列は以下ようになる。
{0,0,1,2,{{3,36,{0,10.000000,0},
{0,0.000000,0},{0,0.000000,0}}}}
```

対象のモデルがタローなので、modelは '2' となる。また 'a15' の運動軌跡が '円' となるので、動作する部位の終了位置は開始位置と同じになる。

このアニメーションを図5に示す。

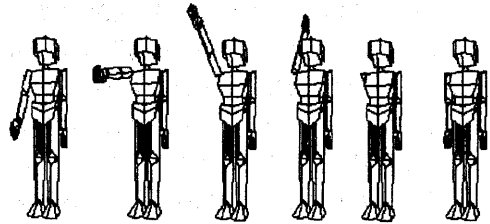


図5: アニメーションの出力例2

次に、図6のアニメーションに対応するコマンド列を以下に示す。

```
{2,0,1,1,
{{1,6,{0,0.000000,0},{0,-10.000000,-60},
{0,0.000000,0}},
{{1,12,{0,0.000000,0},{-60,10.000000,60},
{0,0.000000,0}},
{{1,6,{0,0.000000,0},{60,-10.000000,0},
{0,0.000000,0}}}}
```

変換された軌跡式は以下ようになる。

```
locus(ICHIRO kubi migi migi a14)
• locus(ICHIRO kubi hidari hidari a14)
• locus(ICHIRO kubi migi migi a14)
```



図 6: アニメーションの入力例

## 5 まとめ

人間の自然言語の解釈による心像(イメージ)は様々であるが、このシステムにより誰もが理解できる解釈結果のアニメーションによる入出力ができた。本システムはまだ試作の段階で完璧な相互変換が行えるような単語登録や意味辞書を構築していない。今後、より複雑な動作の単語登録や意味辞書の構築をし、モーションキャプチャー等を用いて、アニメーションの入力を用意にすることが必要であると考えられる。本システムにより人体モデルを用いた自然言語とアニメーションの相互変換の実験を行った結果、かなり一般性のある処理が実現できる見通しを得ることができた。

## 参考文献

- [1] 横田将生:視覚情報と言語, 画像応用技術専門委員会サマーセミナー'95 Vol.4,1995.9.1~2
- [2] 横田将生, 心像意味論に基づくマルチメディア・コミュニケーション理論, 自然言語処理 106-10,1995
- [3] 横田将生, 西村靖司, 白石正人, 笠見一:心像意味論に基づく日本語連接名詞の構文および意味分析, 電子情報通信学会(D),J77D,1,pp.131-142,1994
- [4] 倉田志津, 横田将生:自然言語動作概念のアニメーション表現について, 電気関係学会九州支部連合大会,1328,1996
- [5] 倉田志津, 横田将生:自然言語動作概念のアニメーション表現について, 情報処理学会研究報告,97-NL-118,p67-72,1997

[6] 古川忠昭, 笠見一, 横田将生:マルチメディア理解システム IMAGES-M -アニメーション処理部, 第55回情報処理学会全国大会,6W-4,4-82-83,1997