

# 自然言語による CG 創作環境の一手法

菅原 衣織 (指導教員: 伊藤 貴之)

## 1. 概要

3次元コンピュータグラフィックス(CG)の制作環境には、商用では Maya や 3D Studio Max など、またウェブ上のツールとしては Google Sketch Up など、既に多数のものが知られている。しかし、いずれも習熟するためには相応の知識や技術が必要であり、CGに関する専門知識の無いユーザや小さな子供にとって親しみやすいツールとは言いがたい。そこで近年ではスケッチインタフェースなどを導入した初心者ユーザ向けの3次元CG制作環境の研究も進んでいる。我々も同様に、専門的な知識のあるユーザではなく初心者ユーザを対象とした手軽なCG制作環境の構築を目標にしている。

近年のスマートフォンの普及に伴って、iPhone4Sから対応になったSiriや、Androidの音声認識APIに代表されるように、音声認識技術の発達が著しい。我々はこの音声認識技術を用いて3次元CGの感覚的な制作環境を構築することを目標にしている。音声入力を導入することで例えば、マウスやタッチパネルなどの2次元操作デバイスがもつ3次元操作の難しさが緩和されることが期待される。また、スマートフォンやポータブルゲーム機などの小型端末ゆえの操作の困難さからも解消されると期待される。さらに、一同に介した多数のユーザによる協調的な制作が容易になると期待される。

本研究はその前段階として、自然言語を入力とした3次元CG制作環境の構想と、その実装について論じる。本手法では入力された自然言語からキーワードマッチングにより、あらかじめデータベースに登録されている3次元形状モデルとテクスチャ画像を抽出し、これらを組み合わせて表示する。

## 2. 関連研究

### 2.1 自然言語による3次元CG制作

自然言語による幾何形状モデリング手法は以前にも発表されている[1]。この研究では英語のSVOの文法に則って球や立方体などのプリミティブな3次元モデルを操作、配置し、組み合わせて物体を制作する「3次元グラフィックシステム Do」を提案している。CGの専門的な知識のあるユーザと知識のないユーザを対象とした実験を行っており、共通の物体を作成する際に掛かった時間や使った単語数を比較している。

## 3. 提案手法

本章では提案手法について述べる。まず本手法における3次元CGのプログラミングモデルを述べ、それに対応する自然言語について論じる。続いて本手法で自然言語を入力した際の処理手順について述べる。

### 3.1 3次元CGのプログラミングモデル

我々の実装ではOpenGLというプログラミング環境をベースにして3次元CGを表示する。このプログラミング環境では以下のような処理手順を要する。

**初期化:** ウィンドウの設定、視点や光源の設定など。

**物体単位の処理:** 物体形状(頂点情報および位相情報)、物体に貼りつけるテクスチャ画像の設定、物体の座標変換(拡大縮小、平行移動、回転など)。

現時点での我々の実装は、この2種類の処理手順のうち物体単位の処理について、以下の2種類の自然言語入力を想

定している。

1) 「AであるB」(Aは修飾語、Bは物体名)という構文により、Bに相当する物体形状と、Aにふさわしいテクスチャ画像を検索して表示する。

2) 座標変換を表す予約語の入力により、物体の拡大縮小、平行移動、回転などを指定する。

本手法における自然言語入力の例と、それに対応して構築されるデータ構造の例を、図1に示す。本手法ではまず物体形状とそのテクスチャ画像を自然言語から検索し、続いて必要に応じてその座標変換を自然言語から設定する。

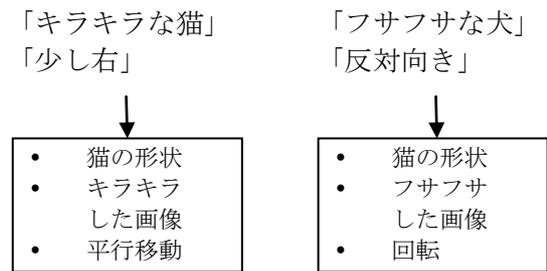


図1. 自然言語入力とデータ構造

### 3.2 3次元形状モデルとテクスチャ画像

本手法は自然言語入力のため既存のCG制作ツールと同様に精度の高い形状モデリングは困難である。既存研究[1]においてもプリミティブな3次元形状の組み合わせだったために創作の範囲が限定されていた。そこで我々は、本ツール自体で3次元形状モデルを設計するのではなく、インターネット上にある3次元形状モデルデータベース(例えば3D Model Search Engine[2])からダウンロードした3次元形状を用いる。これらの3次元形状モデルはWRL形式やSTL形式といった標準的な書式で記述されている。

我々の実装では、後述するタグを付与された3次元形状モデル群とテクスチャ画像群がデータベースに登録されていることを前提とする。そして自然言語を入力すると、その入力文を構成する単語とのマッチングにより、3次元形状モデルとテクスチャ画像を検索する。以上の処理によって定義されたテクスチャ画像付きの3次元形状モデルを、以下「キャラクタ」と称する。

### 3.3 タグとのマッチング

本手法で扱う3次元モデルやテクスチャ画像には、その内容を表す物体名や形容詞をユーザが手動でタグ付けすることを想定する。そしてシステムは、手動でタグ付けされた単語の類似語や上位概念語を割り当て、またテクスチャ画像においては色などの特徴量に対応する単語を自動で割り当てる。以下にタグの付与方法を記す。

**物体名:** ユーザが手動で付与する。1つの3次元形状モデルやテクスチャ画像に対して複数の物体名を付与することも想定する。また、同じ物体名が複数の3次元形状モデルやテクスチャ画像に対して付与されることも想定する。

**形容詞:** ユーザが手動で付与する。ユーザの主観や趣向を反映するために有効である。1つの3次元形状モデルやテクスチャ画像に対して複数の形容詞を付与することも想定

する。また、同じ形容詞が複数の 3 次元形状モデルやテクスチャ画像に対して付与されることも想定する。

**同義語や上位概念語:** 手でタグ付けされた物体名をもとに、概念辞書を参照して自動的に付与する。

**特徴量に対応する単語:** テクスチャ画像に対して色や周波数などの特徴量を算出し、それに対応する単語を自動的に付与する。

なお我々の実装では、概念辞書として日本語 WordNet [3] を利用している。WordNet では各単語の概念が「synset」という単位にまとめられており、さらに関連のある synset とリンクされている。我々の実装では synset に属する同義語 (synonym) と、synset のリンクをたどって得られる上位概念語を用いる。

以上に記したタグ情報をもとにして、本手法は以下の手順で最適な 3 次元モデルとテクスチャ画像のマッチングを図る。

1. 入力文中の物体名 (またはその同義語) をタグを持つ 3 次元形状モデルを検索する。
2. 1. で該当する 3 次元形状モデルが複数存在する場合には、手動でつけた 3 次元形状モデルにつけられた形容詞を比較し、1 個でも多く共通の形容詞を入力文に含む 3 次元形状モデルを選ぶ。
3. 入力文中の修飾語 (またはその同義語) をタグを持つテクスチャ画像を絞り込む。
4. 3. で該当するテクスチャ画像が複数存在する場合には、3 次元形状モデルと同一の上位概念語を共通に有するテクスチャ画像を優先的に選ぶ。
5. 4. で上位概念語を共通に有するテクスチャ画像が複数存在する場合には、手動で 3 次元形状モデルとテクスチャ画像につけられた形容詞を比較し、1 個でも多く共通に有するテクスチャ画像を優先的に選ぶ。

図 2 にタグ情報を格納してあるデータベースの設計の一例を示す。各ポリゴンの物体に対し代表名を設けておき、その代表名のテーブルを親エンティティとする。URL や名前、概念語、形容詞、特徴量のテーブルは親エンティティと 1 対 n の関係になる子エンティティとする。また主キーはアスタリスクで、外部キーは FK で記す。

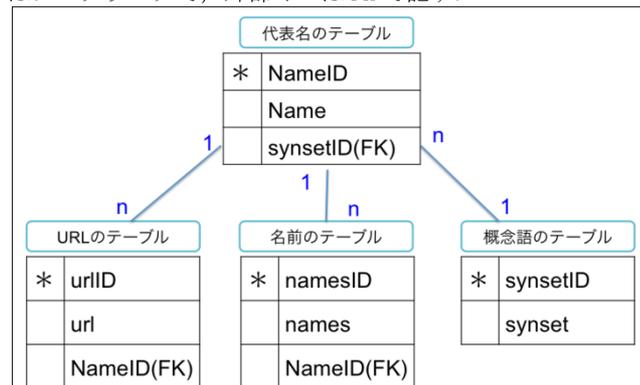


図 2. DB の設計例

### 3.4 座標変換

本手法ではキャラクタに対する平行移動や回転、拡大縮小の指示を数値の表現ではなく「まわれみぎ」や「おおきく」といった口語的な表現を用いて指示する。現在は約 20 種類の指示を用意しているが、今後ユーザアンケートを通して一般的にイメージする口語表現と操作の組み合わせを調査していく予定である。

### 3.5 GUI 機能

我々の実装における GUI を図 3 に示す。画面左側には、キーワードを入力するためのテキスト記入欄や、表示され

たキャラクタの拡大縮小、回転を行うためのチェックボックスを備えている。画面右側には、タグによる優先度が次点以下だったものを表示するためのスペースを備えている。優先度が次点以下だったものはクリックにより切り替えが可能となっている。中央は実行結果を表示する画面である。

## 4. 実行結果

本手法を利用して実際に自然言語を入力し、試験的に実行結果を確認した。図 3 は実行結果の一例である。



図 3. GUI 上での表示例

## 5. まとめと今後の課題

本研究では、自然言語入力によって 3 次元形状モデルとテクスチャ画像を検索することで、感覚的に 3 次元 CG を制作する一手法を示した。

今後の課題として、以下の点に取り組みたい。

**キャラクタ生成のためのマッチング手法の改良.** 現在実装しているマッチング手法は非常に単純なものであり、3 次元形状モデルやテクスチャ画像の数が膨大になっても満足度のいく結果が得られる確証はない。この点について検討を進めたい。

**座標変換の操作方法.**

ユーザアンケートを通して口語表現と座標変換の組み合わせを検討していく。

**小さい子供の興味を引くような GUI の改善.** 本手法は CG の専門知識のないユーザを対象にしている。GUI をより親しみやすいデザインにすることで、いろいろな人にとって使いやすいシステムを開発したい。

**色調や現実性を考慮したモードの設定.** 形容詞や特徴量に関するタグを利用して多様な作風を目指せる機能も追加しようとして検討している。例えば似たような色合いや雰囲気のテクスチャが優先的に選び、統一感のある空間を創ることや、上位概念語を利用してあえて現実には起こりえない組み合わせを選ぶことで、現代アート風の空間を創ることが期待できる。

**ユーザテストの実施.** 本手法のシステムの改良を進めていくにあたってユーザビリティの向上は必須の課題であると考える。一例として、ある同一の風景を制作させ、制作時間とユーザの満足度を測る、というような形でのユーザテストを検討している。

## 参考文献

- [1] 河合善之, 岡田稔, 自然言語による幾何形状モデリングと画像合成の一手法, 情報処理学会論文誌, 42(5), 1161-1169, 2001.
- [2] Princeton Shape retrieval and Analysis Group, 3D Model Search Engine, <http://shape.cs.princeton.edu/search.html>
- [3] 独立行政法人情報通信研究機構 (NICT), 日本語 WordNet, <http://nlpwww.nict.go.jp/wn-ja/>