

オブジェクトの状態変化に基づくストーリーを有した絵集合の生成

The Sequential Pictures Generation System based on the Distribution of Objects Transition

上野 未貴*¹ 森 直樹*¹ 松本 啓之亮*¹
 Miki Ueno Naoki Mori Keinosuke Matsumoto

*¹大阪府立大学, 工学研究科
 Graduate school of Engineering, Osaka Prefecture University

Picture understanding by computer has become important in computer science. However, to define the information of picture semantics is still difficult and there are no model to represent picture completely. In this paper, we propose the novel method which generates sequential pictures using image objects databases. We also investigate the transition types of scenes in existing picture books to show the effectiveness of proposed method.

1. はじめに

計算機が小説, 漫画, 音楽など, 人のように知的な産物を生成するという事は人工知能研究の大きな目的の一つである. その中で, 絵を計算機上で扱う研究は, 画素分布に基づく画像認識の発展の一方で, 絵に込められた意図や感情などの意味的情報を扱う技術は十分に研究されていない. しかしながら, 絵は情報量が多く, 計算機による絵理解が重要視されている.

以上の点を背景として, 筆者らは絵を導入した対話システムである絵情報共有型会話エージェント (Picture Information Shared Conversation Agent: Pictgent) [1] を提案してきた. Pictgent では絵と絵の内容を表す情報を絵モデルとして定義し, 対話に利用しているが, 絵モデルの完備な定義はなされていない.

これらの問題点を解決するため, 本研究では絵集合として, 4コマ漫画や, 絵本など, ストーリーに沿って, 連続的に状態変化する2枚以上の絵を対象にすることにより, 絵の中で共通する要素を抜き出して, 少ないコストで多くの絵を生成する手法を提案する. その後, 既存の絵本のページを対象とし, その内容を提案手法で再現するために, 必要な情報および状態遷移オペレータについて考察する.

2. 提案手法

本章では, 提案する絵の自動生成手法について述べる.

2.1 絵モデル

絵モデルと呼ばれる絵に関する情報データを用いて, 計算機による絵の扱いを実現する. 絵モデルでは, 絵内の各構成要素すべてをオブジェクトと定義する. すべての絵に対して合理的な絵モデルの定義は困難であるため, 本研究では対象とする絵を限定して絵モデルを用いる.

まず, 最初の絵 p_0 をシステムが生成する. p_0 は白紙もしくは背景画であり, 定められたいくつかの領域を持つ. ユーザは絵に対して操作を加えることができ, これをユーザ入力とする. 時刻 t におけるユーザ操作は, \mathbf{u}_t として表し, \mathbf{u}_t は p_t に対するあらゆる種類の操作を含む. p_{t+1} は \mathbf{u}_t を p_t に適用することで生成される. 以下にこの関係を示す.

$$p_{t+1} = \prod_{u_i \in \mathbf{u}_t} u_i p_t \quad (1)$$

ここで, u_i は時刻 t における, 各ユーザ操作を表す.

以下に, $p_t, o_t, \mathbf{u}_t, \{O_t\}, \{U_t\}$ の関係を示す.

$$\{O_t\} = \{o_t \in p_t\} \quad (2)$$

$$\{U_t\} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_t\} \quad (3)$$

$$p_{t+1} = \prod_{u_i \in \mathbf{u}_t} u_i p_t = \prod_{u_i \in \mathbf{u}_t} u_i \prod_{u_j \in \mathbf{u}_{t-1}} u_j p_{t-1} \quad (4)$$

$$\dots \quad (5)$$

$$= \prod_{j=0}^t \prod_{u_i \in \mathbf{u}_j} u_i p_0 \quad (6)$$

$u_i \in \mathbf{u}_t$ は, 描画オペレータ ϕ と p_t 中の被演算オブジェクトを表す. 各 ϕ は, 色, 角度などの特定の引数を持つ.

オブジェクト o_t はオブジェクトの状態変化を表すことができる. n コマ目中のオブジェクト o_n は以下の式で表されるようなマルコフ性を持つ.

$$o_n = \prod_{i=0}^{N_n^{\max}} \phi_i o_0 \quad (7)$$

ここで, N_n^{\max} は o に適用した ϕ の合計数である. 各 u_i は少なくとも1つの ϕ を持つため, $N_n^{\max} \geq n-1$ である.

2.2 基底描画オペレータ $\hat{\phi}$

画像に変形などの操作を加えることのできる機能を描画オペレータ ϕ と定義する. 複数のオペレータの積もまたオペレータとなるが, それ以上分解できない描画オペレータを基底描画オペレータ $\hat{\phi}$ として定義する. この基底描画オペレータを増やして, あらかじめ用意した画像を変形し組合せることによって生成可能な画像の種類を増やす. これにより, 事前に用意するオブジェクトの種類を削減可能である.

2.3 描画オペレータ ϕ の実装

複数の基底描画オペレータ $\hat{\phi}$ の積で表現されたオペレータを描画オペレータと定義する. 本研究ではシステムを Java により実装しており, デザインパターンにおける Command パターンでオペレータを表現している. 基底描画オペレータは Command, 描画オペレータは MacroCommand の関係にある. すなわち, 基底描画オペレータのキューが, 描画オペレータである. 例えば, 割れるという状態変化を表す ϕ_{crack} という描画オペレータは, 基底描画オペレータ $\hat{\phi}_{\text{divide}} \hat{\phi}_{\text{move}} \hat{\phi}_{\text{rotate}}$ からなる. 右から被演算子であるオブジェクト o を作用させると (8) 式のように3つの基底オペレータが順に適用される.

$$\phi_{\text{crack}} o = \hat{\phi}_{\text{rotate}} \hat{\phi}_{\text{move}} \hat{\phi}_{\text{divide}} o \quad (8)$$

連絡先: 上野 未貴, 大阪府立大学工学研究科, 〒599-8531 堺市中区学園町1番1, E-mail: ueno@ss.cs.osakafu-u.ac.jp

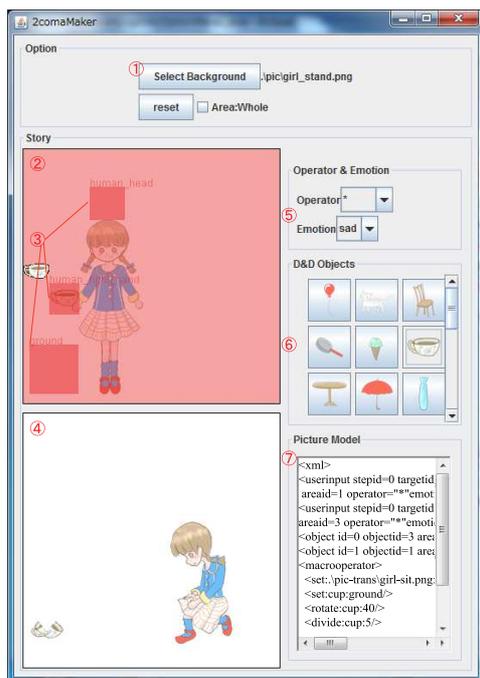


図 1: 2 コマ漫画生成システムの外観

2.4 システムの動作

以下では、(1)式における p_t に対して、ユーザがオブジェクトを設置するなどの操作をした直後の絵を p'_t で表す。ユーザの操作により p'_t に適用される描画オペレータは一意に定まり、それらを p'_t に適用した結果が p_{t+1} となる。

- 1 コマ目の背景画像を指定する。この画像を p_0 とする。
- システムはユーザの入力を受け付ける。ユーザは、 p_0 に対していくつかの操作をする。本研究では、ユーザの操作はオブジェクトを配置するのみである。ユーザは気に入ったオブジェクトを p_0 のいずれかの領域に、操作と感情の情報を付加して配置する。各領域に 1 つまでのオブジェクトが存在するという条件下で、いくつかのオブジェクトを配置する。ユーザがオブジェクトを置いたとき、新たな領域が絵に加わることがある。
- 全てのユーザ入力 u を p_0 に適用して、 p'_0 を得る。この p'_0 が 1 コマ目となる。
- ユーザがあるオブジェクトに対して指定した描画オペレータや感情の情報を元に、システムは p'_0 中のオブジェクトを変化させる。変化するオブジェクトの新たな画像は、状態変化のデータベースを参照して決定する。
- 全ての変化を p'_0 に適用して、 p_1 を得る。 p_1 は 2 コマ目を意味する。2 コマ漫画のときは p_1 が最終状態となるが、 n コマの場合は、 p'_1 が p_2 となり、再度入力を受け付け、 p'_{n-1} が p_n となるまで上記を繰り返す。

2.5 ユーザ入力 u

状態変化を起こす要因を以下に示す。なお、3. 章で示すアプリケーション中では、これらはユーザ入力に相当する。

オブジェクト o ユーザが配置したオブジェクトの情報。

領域 a o を配置した領域。

操作 m 配置したオブジェクトに適用する操作の種類。

感情 e 配置したオブジェクトにユーザが割り当てる感情。

3. アプリケーション

提案手法を使って、実際に 2 コマ漫画生成システムを作成した。図 1 にシステムの外観を示す。

図 1 中に振った番号と対応づけて、各部を説明する。

- 1. 画像指定** 1 コマ目のベースとなる絵モデルを指定する。白紙を表す空の絵モデル、もしくは既にオブジェクトが配置された絵モデルを指定可能である。
- 2. 1 コマ目画像** 1 コマ目の画像を表示する。ユーザ入力後、 p_0 は p'_0 となる。
- 3. オブジェクト配置可能領域** 1 コマ目の画像に付加できる領域を表示する。
- 4. 2 コマ目画像** 1 コマ目に対して加えられた情報を元に、生成された 2 コマ目の画像を表示する。
- 5. 描画オペレータ** 配置するオブジェクトに適用可能な操作と感情の種類。
- 6. 配置可能オブジェクト** 3. で示した領域にドラッグ & ドロップで加えることのできるオブジェクトを示す。
- 7. 絵モデル** 2 コマ目の絵モデル XML を表示する。

4. 考察

2.2 で示した基底描画オペレータを用いて、既存の絵本上で見られた状態変化の表現方法の一部を示す。

転がる 回転, 移動

窓に人の顔が映る トリミング, 拡大・縮小

崖から落ちる 移動, 回転

仲間が増える 複製, 移動

絵本ではストーリーが理解可能か否かという観点から、対象年齢という明確な区分けがあるため、使用される状態変化の共通割合が比較的高い。描画オペレータの組合せで表せる状態変化の種別については発表時に示す。

また、以下の情報の項目を検討し、絵の定量化を進める。

絵本特有の状態変化 寝る, 泣く, 歩く, 抱く, 話す, 手を繋ぐ, 手を組む, ドアを開ける, 入口に立つ→対象年齢が理解しやすい情報が中心的なため偏りがある。

絵に特有の状態変化 木陰などから体の一部が出ている, 窓に人の顔が映る, 群れの中で 1 個体だけが他と異なる。→空間的広がりを持つ情報や、言語で表すと無粋な情報

状態変化しやすい部位 手, 口→絵本では絵中の人物の感情を理解しやすい部分に変化することが多く見られる。

5. まとめと今後の課題

本研究では、計算機の特性を生かして描画オペレータを使用した絵の生成手法を示した上で、実際の絵本に適用するために、どのようなオペレータが必要かを考察した。今後の課題として、絵の意味的情報の中で定量化できる箇所を増やすため、既存の作品を対象として、必要な情報を検討する。

本研究は JSPS 特別研究員奨励費 (課題番号 25・10978) の助成を受けたものです。

参考文献

- [1] Miki Ueno, Naoki Mori, Keinosuke Matsumoto : *Picture Information Shared Conversation Agent: Pictgent*, Distributed Computing and Artificial Intelligence Advances in Intelligent and Soft Computing, Springer, Volume 151, pp.91-pp.94, (2012)