

人工知能画家 静 第3版

迎山 和司

公立はこだて未来大学

人工知能画家・静(しずか)は、プログラム自体が絵を自律的に描くソフトウェアである。本稿ではその第3版目にあたるソフトウェアを紹介する。第3版はユーザの描いた線描を解釈して、それに人工知能が「考えた」線描を新たに加えるという「連想」という人間の想像力に注目した「お絵かきソフト」である。本稿ではまず概要と特徴を説明し、さらにデモンストレーションでの2つの対照的なユーザ評価を紹介したのち、今後の展望を述べる。

SHIZUKA, the AI painter version 3.0

Kazushi Mukaiyama

Future University-Hakodate

SHIZUKA is one of software which can draw a picture autonomously. It, however, doesn't obtain to invent a convenient machine which can draw better than human beings. This final goal is to understand how we can draw, see and understand our visual representations.

In this version, I developed the interactive drawing software. This software can add new shape, understanding users' drawing interactivity, as if human associate memory.

1. はじめに

人工知能画家・静（しずか）は、プログラム自体が絵を自律的に描くソフトウェアである。しかし、これは人が絵を描かなくてもよい便利な機械を作ることを目的としているわけではない。むしろ、人が絵を描く行為を深く考察することを目的としている。描画行為をプログラムに行わせることは、逆に絵を描くこととはどういうことなのかを深く考えることにつながる。たとえば、四角形を描くにはどのようなプログラムをつくれればいいだろうか？利用するプログラム言語にもよるが始点を指定し高さ幅を指定して四角形を描く命令がほとんどのプログラム言語には用意されている。しかし、同じ直線を4回描き、それぞれの終点で90度回転させて描いてもその結果は四角形である。このように描画のプロセスを考えることが、描画行為をプログラムに行わせる大きな理由である。筆者はこの試みによって、なぜ人間は絵を描くことが出来、また鑑賞し、理解できるのかということでの絵画が本来的に持つ役割を改めて考えてみたいと思っている。創作行為のモデル化自体は非常に大変なことで、今後も制作を続けていく予定でいる。今回はユーザの描いた線描を解釈して、それに人工知能が「考えた」線描を新たに加えるという「連想」という人間の想像力に注目した「お絵かきソフト」の開発を目指した。

2. 概要

静の主な機能は、描かれた線描を構造化して、その線描にユーザに意味の喚起を促す、つまり描かれたものがまったくわけの分からないものではなく、何か意味のあるような形状に解釈できるような、あらたな線描を追記することである。たとえば、顔のような輪郭を描いた場合、静はそこに目に当たる線描を描き加えることをする。そのために静は大きく2つのことを学習している。

1. ユーザによって与えられた線描を、意味を表象しない状態にまで分割された領域に分割し、これをひとつの部品として近隣の他の部品との相対関係を保持して保存する。この相対関係を保持した部品を認知プリミティブと呼ぶ。
2. 保存した認知プリミティブの形状を、自己組織化マップ (Self-Organizing Map-SOM^[2]) を使って分類する。この分類学習により、近い形状をもつ認知プリミティブは、2次元マップ上の近い地点に分類される。

この2つの学習データを元に、まずユーザが描いた線描に近い形状をSOMから検出する。検出された認知プリミティブ候補がもつ相対関係にあたる認知プリミティブを列挙し、線描の上に合成する。例えば、ユーザが顔の輪郭のような線描を描いたとすると、多くの場合、すでに学習した相関関係情報から目にあたる認知プリミティブが選出され、選出された認知プリミティブがユーザの描いた顔の輪郭のような線描の上に合成される。

静は大きく分けて以下の4つの部分から構成されている。

- 認知プリミティブ作成部分
- 自己組織化マップ(SOM)学習、認識部分
- 追加線描作成部分
- ユーザインターフェース、演出部分

以下の節で、これらの各部分についての概要と実装方針を説明する。

2.1. 認知プリミティブ作成

認知プリミティブという単語は、静を説明する上での造語である。形状からなんらかの意味を喚起するものではあるが、それそのものがはっきりとした意味を表象しない視覚形状を指す。例えば図1の場合、分割された各々のブロックが認知プリミティブとなる。

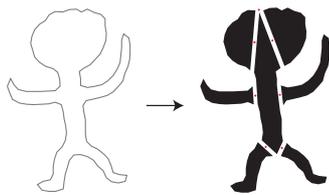


図 1 認知プリミティブによる表現

認知プリミティブは大きく 2 つの情報で構成されている。ひとつは形状を現す画像情報。もうひとつはその認知プリミティブにつながっている他の認知プリミティブの接続点情報と大きさである(図 2)。

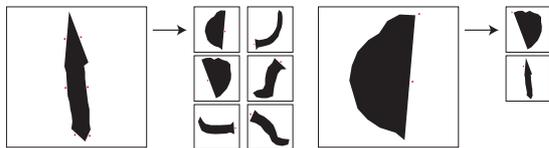


図 2 それぞれの認知プリミティブとリンク先

このように、認知プリミティブは、お互いの相対関係を保持しており、ひとつの認知プリミティブを選ぶとその認知プリミティブを起点として、ユーザの描いた線描を再表現することが出来る。

2.2 . 自己組織化マップ(SOM)学習、認識

作成された認知プリミティブが増えていくと選択肢が増え、ユーザに意味の喚起を促す形状を追加選択することに妥当性がなくなってくる。そのために自己組織化マップ(SOM)を使い、認知プリミティブ同士の分類および近似する認知プリミティブを見つけている。

SOM はコホネンによって考え出された動物や人間の視覚に関係する細胞(とくに大脳皮質)で行われる自己組織化過程のモデルである。SOM は任意のニューロン細胞に当たるノードを 2 次元マップ空間に持ち、任意のベクトルの集合を SOM に入力すると、ノード間で競合学習を行って、入力された各ベクトルを 2 次元マップに投影します。学習によって網の均衡状態が収束すると、結果は入力さ

れたデータの近似する分類マップになる。

静における SOM の学習対象となる認知プリミティブのデータはグレイスケールの 256 × 256 ピクセルのビットマップ画像情報である。図 3 はテストのために 49 個の認知プリミティブだけに限定し、20 × 20 のノードをもつ SOM を使って学習させた結果であるが、「おおよそ」近い形状同士が集まるように分類されていることがわかる。

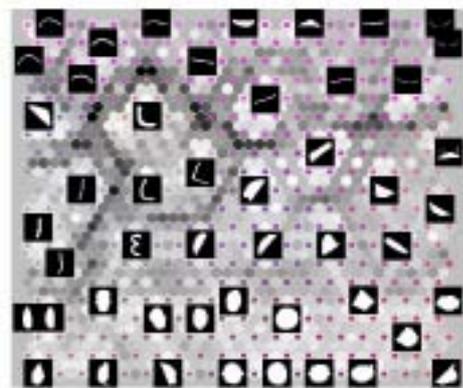


図 3 SOM による認知プリミティブの分類例

認知プリミティブの分類に SOM を使うことで期待できることは、近似するデータをプログラムが自律的に集めることが出来ることである。人間はイメージを連想するとき、似たような形から喚起されることが多い。たとえば長い紐を見たときに、蛇をイメージするといったようなことがある。また、SOM は多くの同じようなデータを学習すればするほど、そのデータにかかる重みが強くなるので、それを選択しやすいマップが出来上がる。これもまた、人間がよく見るものをよく思い出しやすいモデルに近いと言える。

2.3 . 追加線描作成

ユーザの描いた線描から、SOM 認識により選択された認知プリミティブと、それに接続しているすべての認知プリミティブを元にして、追加線描を作成する。作成に当たっては合成先の認知プリミティブの位置と大きさに

合わせて、合成する認知プリミティブを変換する。変換された認知プリミティブのうち、選択された認知プリミティブ以外のものすべてが追加線描作成のための形状画像になる。

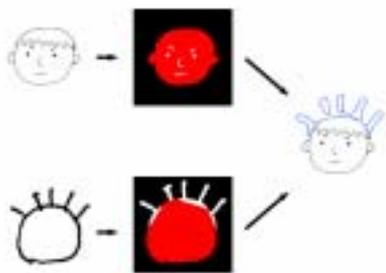


図 4 合成．赤い部分が SOM で近いと判断された部分

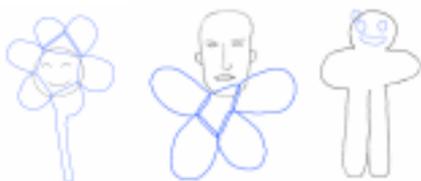


図 5 合成結果の例

線描作成に当たって認知プリミティブがもつグレイスケール画像から境界線を抽出するが、そのまま境界線を抽出すると、分割された部分の直線や、変換したときの誤差から、きれいに画像が一致することはないために、不自然なずれが生じる。このため画像にぼかしフィルタを掛けて境界線をわざとあいまいにしている。この状態の画像の境界線を抽出することにより、手書き風の形状ではずれが不自然にみえなくなる。

得られた境界線は、そのまますべてを一括して表示していない。いちど、境界線の持つ点の座標情報を順に配列に確保する。その配列を順に読み込むことによって線が逐一、描かれていく。これは静が自律的に描いている様子を演出するためである。

2.4 . ユーザインターフェース、演出

どんなに革新的なアイデアを持ったソフトウェアでも、ユーザにとって操作はわかりやす

く親しみやすいものでなければ触ってもらえない。特に美術館やコンベンションなどの展示では、一点だけではなくほかにも多くの作品が展示される。その場合、ユーザの興味を引き、気軽に触ることが出来るソフトでなければ、多くの作品の中からユーザは足を止めてくれない。

したがって、多くのユーザがそのソフトウェアの本質を理解する以前に、親しみやすく楽しく操作してもらうことは非常に重要である。静では、そのために以下の点に注意している。

- 興味を引く画面構成
- 平易な操作体系
- 魅力的なアニメーションとサウンドによる演出

基本的には、メイン画面に描画可能なスクリーンがあるだけである(図 6)。ユーザは主にこの画面に対して、線描を描くことになる。画面は黒板をイメージしたものになっているので、ユーザがこの画面を見たとき、画面に絵が描けることがすぐに理解できる。



図 6 画面構成

また、画面上には常にマスコットキャラクター「静ちゃん」(図 7)がいる。このソフトウェアは「人工知能画家・静」という名称のように、コンピューターを擬人化している。静ちゃんは、ユーザによりその擬人化をイメージしてもらうためのシンボルとして設定した。



図 7 マスコットキャラクター「静ちゃん」

静の操作はいたってシンプルである。「描く」、「消す」、「理解させる」の 3 点しかない。静は、線描を自由に描画できる電子黒板の形態をとっているが、洗練された絵を描くための描画ツールではなく、描画された絵を理解して新たな絵を描くソフトウェアなので、複雑な操作を求めない。むしろ、展示という状況では、ユーザ(ここでは鑑賞者)は展示作品に対して足を止めて触れる時間は、個人用のコンピュータ上で操作する場合に比べて非常に短いものである。限られた時間でそのコンセプトを理解してもらうには、操作が複雑なことは致命的である。

3 . 特徴

人工知能画家・静はユーザが描いた線描から人間がわかる意味のある線描を追加する。

人工知能画家の先行研究として Harold Cohen ^[1]の AARON(アーロン)がある。アーロンは、画家である Cohen の描画知識に基づいて作られている。つまり、人というものは頭があってその下に胴体があり、胴体には上に肩があってそれには腕がつながっているといったものを Cohen が知っており、Cohen はその知識をアーロンに与えている。このような知識をデータとしてもっており、関節の曲がり具合などを乱数により無作為に決定して、一枚のテーマを持った絵を描く。アーロンは Cohen が教えた形を確実に描くがそれ以上のものは描かない。また、一度描き始めると外部からのリアクションを受け付けずに単独で描ききる。

静はアーロンの考え方を基本におきつつも、描かれる対象を知識として持っているのではなく、対象の獲得の仕方を知識として持っている。

つまり人物の知識を持ってはいない。ただ、いろいろな形状をたくさん取り込んでいくだけである。しかし、闇雲に取り込むのではなく人物の形状ならば人物の形状のグループに、花ならば花に近い形状のグループに分類している。この蓄えた中から、形によって近いものを連想してさまざまな組み合わせを提示する。つまり、とりこむ形状を多様化することによって、開発者である筆者にも予想できない形を作り出す可能性がある。また取り込むことを前提とするので、単独ではなく開発者以外のユーザとの共同制作を可能にした。

人工知能画家・静は今回で 3 版目となる。はじめは人間の想像力を simulate(模倣)しようとすることを目標としていた。しかし、結局のところどうやっても人の心の中の仕組みを解き明かすことは今の筆者にはできなかった。そこで見方を変えて絵それ自体に注目した。

我々はある程度同じ意味を持つ図形を共有している。例えば、顔を表す図形はほぼ同じように輪郭の中に目、鼻、口などを描く。このことに注目して、認知プリミティブに変換してみると、その相対関係に一定の規則があることが分かった(図 8)。

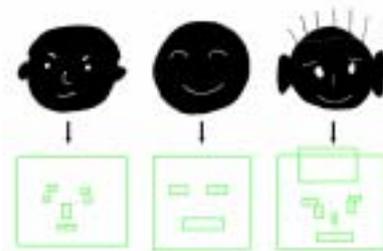


図 8 顔のイメージの相対関係

さらに、この規則を守れば、図形がどんなも

のになろうとも、その規則による意味は保持されることが分かった(図 9)。

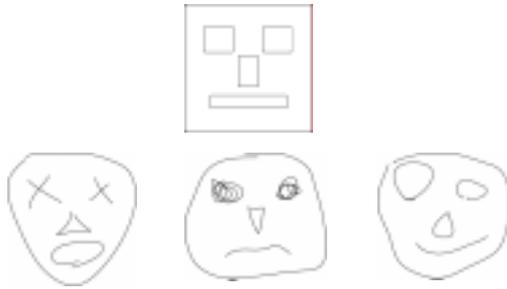


図 9 配置が同じ状態での図形変更例

このことから、図形が意味を表すためには、図形の形状ではなく、図形と図形の相対関係つまり「配置」が重要であることが分かった。配置さえ同じであれば中の形状はなんでもかまわない。とすれば、静は、人間とコンピューターのそれぞれと得意なところを生かした成果といえる。

人間は状況に応じて柔軟に判断をすることに優れているが、同じことを繰り返し大量にすることは苦手である。反対にコンピューターはプログラムされたこと以外に決定を変えることは苦手であるが、決まったことを繰り返し、手当たり次第に組み合わせることは優れている。静はさまざまな絵を断片として記録し、その断片の形から可能性のあるものを組み合わせる。出来上がった絵がどんな意味を持つものなのかは人間の解釈しだいである。静は人間の絵から意味を読み取ろうとする癖を stimulate(刺激)している。

4. デモンストレーション

平成 15 年 12 月 05 日に京都市立芸術大学、平成 16 年 02 月 15 日に東京国際フォーラムにてそれぞれデモンストレーションを行った。

デモンストレーションにはスマートボードと呼ばれる大型のタッチパネルプラズマディスプレイを用いた(図 10)。



図 10 スマートボードによるデモンストレーション

このデモンストレーションによって得られた鑑賞者の反応から、連想追記による反応と、ユーザインターフェースの操作性に関する反応の 2 つを考察することにより、評価とする。

このデモによって得られたユーザ(ここでは鑑賞者)の特徴的な反応は以下のようなものであった。

- 自分の描いた線描の注目してほしいところがあるのに選ばれなかったら残念に思った。
- 自分の描いた線描に意味の読み取りにくい追記がなされると戸惑った。
- 自分の描いた線描にあきらかに意味のわかる追記がなされると喜んだ。
- 静がどういう追記をするかいろいろ試した。

京都で行ったデモでは 20 代の女性 3 人(図 11)から以下の順番で反応が得られた。



図 11 京都でのデモンストレーション

1. 自分の描いた線描に意味の読み取りにくい追記がなされると、それに意味をつけようとした。
2. さらに追記を行い、あきらかに意味のある追記(縦長の楕円に頭と手足をつける)がされた瞬間、感嘆の声を出して喜んだ。
3. さらに、さまざまに静と自分の描画を書き加え、意味をつけようと深く考え始めた。
4. 最終的に、頭に当たる部分に、鑑賞者自身が目や口を書き加え補完をするようになった。

この鑑賞者は、はじめのうちはなにげない線描をおこなっていたが、いったんプログラムが意味のある絵を追記することがわかると、次第に絵について深く考えるようになり、ついにプログラムと共同で絵を完成させるということをした。このような鑑賞者は開発者の「絵について、普段何気なく描いていたことを深く考え、そしてコンピューターと共同制作してほしい」ということを行ってくれたもっとも期待していた鑑賞者であり、「人工知能画家・静」の目標は達成したといえる。

反対に東京で行ったデモでは、6歳の男の子から筆者には意外な反応が得られた(図 12)。



図 12 東京でのデモンストレーション

1. はじめからなにも意図を持って描かない。
2. そのため意味の明らかな追記が生まれに

くく、本人はこのプログラムは何をやっているのか理解できない。

3. 追記にいたるプログラム内部の進行過程もわからないので、待ち時間がストレスになり、静ちゃんにさわったりしてほかのことで遊びはじめる。

筆者はこのプログラムの目標の一つに連想によって想像力をゆたかにするという教育的な効果を考えていた。しかし、子供は大人以上に、より直接的な意味を理解できるものが返ってこない、まったく意図を汲み取ってくれないことがわかった。

インターフェースの操作性に関して言えば、親しみやすさをもってもらうように心がけたおかげで、アニメーションキャラクターに関しては好感を持ってもらうことができた。しかし「静ちゃん」にふれることができるかどうかわからないので、何も説明がないと操作方法がわからないユーザもいた。

5 . 展望

描画プログラムでは以下のような課題がある。

まず、学習させる形状パターンの多様化である。今回は3種類の絵(人の体、人の顔、その他)というテーマで100人の人に描いてもらったものをデータとして使った。パターンをふやせばそれだけ組み合わせがふえるのでより多様性のある絵が生まれると考える。

ただし、意味が解釈できる絵の生成を目標とするならば、闇雲にイメージを覚えこませても思ったより効果がない。効果的な結果を生みやすいパターンの蓄積方法は2つあると考えられます。ひとつは明確な意味をもったイメージを開発者が選定してあたえる方法。もうひとつは、闇雲に獲得したイメージから明確な意味をもつイメージを静が発見する方法です。最後の方法はデータマイニングの分野に活路を見出せるかもしれない。

つぎに着色である。線描だけでなく、色を入れることははずせない目標であると考えている。着色はすでにアーロンで実現されており、その理論は明快であるので実装は難しくないと思う。着色する場合には、明度に注目する。白黒写真でわかるとおり、領域のまとまりぐあいは濃淡で十分わかる。よって明度を守れば色相と彩度は変更しても領域判断はくずれない。

また、構図も考えている。構図というのは絵画における描くものの配置の仕方である。これにもバランスのよく配置する方法というものがある。構図は写真を学んだ経験からトリミングの理論を生かせば可能であると考えている。これは絵画のフレームの矩形の辺に対しての平行や対角線、角への線の流れなどを考慮することである。

さらに、平面だけではなく立体物の生成を試みたいとも考えている。現在の実装は2次元の平面ですが、これを3次元にするだけでこれは実現できる。入出力に3Dスキャナー・プリンターを使用すれば、人工知能彫刻家が可能であると考えている。

謝辞

ソフトウェア開発にあたり、多大な助力をいただいた上野剛志氏、池本昌弘氏、ならびに萩原宏治氏に深く感謝する。

本研究の一部は、情報処理推進機構、平成15年度未踏ソフトウェア創造支援事業の補助による。

参考文献

- [1] Harold Cohen: <http://crca.ucsd.edu/hcohen/>, 人工知能画家 AARON の製作者
- [2] Timo Honkela: “*Self-Organizing Maps in Natural Language Processing*”, Helsinki University of Technology, Espoo, 1997