

深層学習による物体識別を考慮した 手描きスケッチからの三次元CGシーンの自動生成

水野 慎士¹ 新地 洋一¹

概要：筆者らが開発している「不思議なスケッチブック」では、紙にペンで自由にお絵描きをすると、描画オブジェクトを個別に3DCGモデル化して、お絵描き全体から3DCGシーンを自動的に生成する。しかし、各描画物体の種類の識別は行っておらず、必ずしも適切な3DCGシーンが生成されるとは限らなかった。そこで、本研究では深層学習を用いて手描きオブジェクトを識別する手法の開発を行った。開発手法では、家、車、雲など10種類の手描き物体を93%以上の精度で識別することを実現した。そして、識別結果を不思議なスケッチブックに適用することで、家や車はCG空間中の地面に配置される一方で雲や太陽は空中に配置されたり、車は移動して花は揺れるといった動きが付与されるなど、オブジェクトの種類に応じた適切な配置や動きの自動設定が行われることを確認した。その結果、従来の不思議なスケッチブックに比べて、手描きスケッチからより適切な三次元CGシーンを自動的に生成することを実現した。

Automatic Generation of 3DCG Scenes from Hand-drawn Sketches with Object Identification by Deep Learning

SHINJI MIZUNO¹ YOICHI SHINCHI¹

1. はじめに

お絵描きはペンや紙があれば、いつでもどこでも始められ、最も身近な芸術創作の一つである。そのため、CG制作にお絵描きをインターフェースとして利用することは、CG制作を多くの人にとってより扱い易いものにすると考えられ、デジタル技術を用いてスケッチベースで3DCGモデリングを行う様々な手法が開発されてきた。例えば、輪郭からの距離変換ベースで3DCGを簡易的にモデリングする手法[1]、スケッチの陰影を分析して3DCGモデリングを行う手法[2]、建物などを対象にしてプロシージャルな方法で3DCGモデリングを行う手法[3]、などが提案、開発されている。

お絵描きに基づく3DCG制作手法として筆者らが開発した「不思議なスケッチブック」がある[4][5]。不思議なスケッチブックでは、紙にペンで自由にお絵描きをすると、描画オブジェクトを個別に3DCGモデル化して3D空間

に配置することで、お絵描き全体から3DCGシーンを生成する。このとき、描画オブジェクトの種類の識別は輪郭形状や色に基づく簡易的なもので詳細な識別は行っていない。3DCGシーンを生成するにはCGオブジェクトを適切に3D空間に配置する必要があるが、現状ではオブジェクトのスケッチ中の位置をベースにして3DCGシーン中の配置位置を決定している。そのため、手描きスケッチから適切な3DCGシーンが生成されない場合もある。

不思議なスケッチブックで描画オブジェクトの種類を個別に識別できれば、適切な3DCGモデル化と空間配置が可能となり、結果として適切な3DCGシーンの生成が実現する。そこで、本研究では不思議なスケッチブックに各描画オブジェクトの識別機能の追加を行うことで、手描きスケッチからより適切な3DCGシーンを自動的に生成することを目指す。

描画オブジェクトには様々な種類が含まれており、その描画スタイルも様々である。そこで、本研究では描画オブジェクトの種類の識別方法として、近年様々な画像認識能力が示されている深層学習を用いる。著者らはこれまで不

¹ 愛知工業大学大学院経営情報科学研究科
Graduate School of Business Administration and Computer
Science, Aichi Institute of Technology

不思議なスケッチブックを用いたイベントを多数開催しているため、収集したスケッチデータは4万枚以上あり、オブジェクトデータは10万点を超える。そこで、筆者が保持するスケッチデータの中からスケッチでよく描かれるオブジェクト種類を事前に選別して、それらを識別するための識別器をCNNで構築する。

ここで、手描きスケッチの種類は様々であるのに加えて、描画スタイルには色が塗られたものと輪郭だけのものがあり、それらが混在する場合もある。そして、これら2つの描画スタイルは画像的な性質が大きく異なる。本研究では様々な手描きスケッチに対応させるため、色が塗られた物体と輪郭だけのオブジェクトのどちらの特徴も持つ学習データを構築や、各オブジェクトの色塗りの有無を判定する識別機と物体の種類を判定する分類器との組み合わせを行う。

2. 関連研究

写真などで撮影されたオブジェクトの種類を識別する研究は多く存在しており実用化もされている。一方、手描きスケッチはバリエーションが非常に多く、その種類を識別する手法は写真に比べれば少ないが、いくつか報告されている。

GoogleはCNNを使って手描き線画の種類をリアルタイムで識別する「Quick Draw!」を公開している[6]。ただし、色が塗られたスケッチには対応していない。Sasakiらは、写真とイラストを組み合わせた学習データを使ってCNNを学習させることで手描きイラストの種類を識別精度の向上する手法を提案しており、20種類の動物を76%の精度で識別することを実現している[7]。Furusawaらは、オブジェクトの識別とは少し異なるが、CNNを用いて手描きイラストの顔や髪などのパーツを識別して自動的に着色する手法を開発している[8]。

本研究では、線画と色付き画のどちらの描画スタイルの手描きスケッチにも対応しながら、高い精度かつ高速に描画オブジェクトの種類を識別できる手法の開発を行う。そして、開発手法を不思議なスケッチブックに適用して有用性を検証する。

3. 不思議なスケッチブックについて

不思議なスケッチブックは、様々なオブジェクトが描かれた手描きのスケッチをカメラの前に差し出すと、スケッチ中に含まれる各手描きオブジェクトをそれぞれCG化するとともに3D空間に配置することで、手描きスケッチから3DCGシーンをリアルタイムで生成する。図1に不思議なスケッチブックにおける手描きスケッチからの3DCGシーン生成例を示す。

不思議なスケッチブックでスケッチから3DCGシーンを生成するため、まずスケッチ中の手描きオブジェクトを輪



(a) 手描きスケッチ



(b) 生成された3DCGシーン

図1 不思議なスケッチブック



図2 不思議なスケッチブック体験者によるスケッチ例

郭ベースで抽出している。そして、各オブジェクトの色や輪郭形状に基づいて簡単な種類の識別を行って、オブジェクト輪郭の押し出し、オブジェクト領域の距離変換、オブジェクト領域の盛り上げなどによってオブジェクトから2D/3DCGオブジェクトを生成する。そして、生成されたCGオブジェクトを3D空間に配置することで3DCGシーンを生成する。各CGオブジェクトの配置位置はスケッチ中の位置によって自動的に決定しており、スケッチ中の上方に描いてあるオブジェクトは3D空間の空中に配置する。ただし、オブジェクトの下方に直線を引くと地面を示している線と判定して、上方に描いてあるオブジェクトであっても3D空間の地面に配置する。

不思議なスケッチブックはこれまで様々なイベントで展示しており、これまでに4万人以上の体験者がいる。主な体験者は未就学児や小学校低学年生であり、描かれたスケッチは関連研究で扱っていたスケッチと比べると複雑で多彩である。描画オブジェクトには様々な種類が含まれており、描画スタイルも色が塗られたものと線画のものが混在している。図2にイベントで体験者が描いたスケッチ例を示す。

4. CNN を用いた手描きオブジェクト識別

4.1 概要

3章で述べたように、不思議なスケッチブックで描かれるスケッチはバリエーションに富んでいる。本研究ではこのように描き方が多彩な手描きスケッチから抽出されたオブジェクトを識別する手法を開発する。そして識別結果を不思議なスケッチブックに適用することで、手描きスケッチからのより適切な3DCGシーンの自動生成を実現する。

本研究では手描きオブジェクトの種類を識別する手法として、近年画像処理分野で様々な分類問題に関する研究に用いられている深層学習に着目する。多層のニューラルネットワークによる深層学習は、2章で述べたように手描きスケッチの識別においても有効な結果が得られている。しかし、不思議なスケッチブックで描かれたオブジェクトは非常に多彩であり、しかも色塗りオブジェクトと線描オブジェクトが混在するという、関連研究でもあまり取り扱わないイラストを識別対象にしている。

そこで、本研究では識別対象として不思議なスケッチブックでよく描かれる10種類のオブジェクトを選択した。また、色塗りオブジェクトと線描オブジェクトでは画像特徴が大きく異なるため、色塗りオブジェクトデータのみで学習した識別器、線描オブジェクトデータのみで学習した識別器、色塗りオブジェクトと線描オブジェクトの両方の特徴を持つデータで学習した識別器を構築して、精度の比較を行う。識別器は関連研究でも用いられたCNNを用いる。

4.2 データセット

本研究で用いるデータセットは、不思議なスケッチブックでよく描かれる「雲」「太陽」「星」「車」「家」「橋」「ビル」「木」「花」「人」の10クラスオブジェクトをベースとする。不思議なスケッチブックを用いて実施したイベントで収集したスケッチから、オブジェクトを各クラス300点ずつ、合計3,000点用意した。画像サイズは128×128画素である。これを色塗りオブジェクトのデータセットとする(図3)。

線描オブジェクトのデータセットは、色塗りオブジェクトのデータセットに対してモルフォロジー勾配処理を適用してエッジの抽出を行うことで生成した(図4)。データ数は色塗りオブジェクトのデータセットと同じで、各クラス300点ずつ、合計3,000点となっている。画像サイズは色塗りオブジェクトデータと同様に128×128画素である。

色塗りオブジェクトと、それに対応する線描オブジェクトを重畳することで、色塗りオブジェクトと線描オブジェクトの両方の特徴を持つ複合オブジェクトデータセットを生成する。重畳するには線描オブジェクトに対して膨張処理を施す。また、色塗りオブジェクトはオリジナルに加



図3 10クラスの色塗りオブジェクトのデータセット



図4 色塗りオブジェクトデータと線描オブジェクトデータ



図5 色塗りオブジェクトと線描オブジェクトの両方の特徴を持つ複合オブジェクトデータ

えて、彩度を50%および70%低下させて、3種類のデータセットを生成した(図5)。これは色に関する情報の重みを低下させるためである。複合オブジェクトデータセットは各クラス300点ずつ、合計3,000点となっており、色塗りオブジェクトの彩度の違いで3種類を生成している。

4.3 識別器

本研究では事前学習済みの識別器を用いて、再学習を行うことで不思議なスケッチブックで描かれる多彩なオブジェクトに対応した識別器を構築することとする。識別機のベースには、ImageNetによって学習した重みが公開されているVGG16[9]を用いる。

表 1 各データセットで学習した識別器での識別精度

学習データ	識別データ	認識精度 (%)
色塗りオブジェクト	色塗りオブジェクト	96.5
	線描オブジェクト	90.0
線描オブジェクト	色塗りオブジェクト	91.6
	線描オブジェクト	94.3
複合オブジェクト (オリジナル)	色塗りオブジェクト	96.2
	線描オブジェクト	92.1
複合オブジェクト (彩度 50%低下)	色塗りオブジェクト	94.7
	線描オブジェクト	92.6
複合オブジェクト (彩度 70%低下)	色塗りオブジェクト	94.3
	線描オブジェクト	92.9

VGG16 は畳み込みブロックと全結合ブロックの計 6 ブロックで構成される CNN である。本研究では VGG16 の 4 番目の畳み込みブロックまでを ImageNet により事前学習された重みで固定し、それより後層の畳み込みブロックと全結合ブロックの重みを 4.2 節で述べたデータセットを用いて再学習することで、不思議なスケッチブックで描かれたオブジェクトに対応した識別器の構築を行う。なお、事前学習された重みをそのまま用いるブロックについては仮学習データを用いて実験的に決定した。

4.4 実験

色塗りオブジェクトデータセット、線描オブジェクトデータセット、および両方の特徴を持つ 3 種類の複合オブジェクトデータセットを個別に用いて識別器を構築して、手書きオブジェクトの種類の識別を行う実験を行った。使用した PC のスペックは以下の通りである。

- OS: MacOSX Catalina
- CPU: Intel Core i9 2.3GHz
- GPU: AMD Radeon Pro 5500M
- メモリ: DDR4-16GB 2667MHz

実装は Python で行い、Tensorflow と OpenCV を使用している。

実験の評価は交差検証法で行った。10 クラスで各クラス 300 点、合計 3,000 点のデータで構成される各データセットについて、各クラスを均等に 5 つに分割することで、5 つのサブデータセットに分割した。従って、各サブデータセットは、各クラス 60 点、合計 600 点のデータで構成されている。そして、4 つのサブデータセットを用いて識別器の学習を行い、残り 1 つのサブデータセットを識別する実験を行った。これをサブデータセットの組み合わせを変えながら 5 回行うことで、認識精度の平均値で識別器の評価を行った。なお、学習回数は 50 回とした。また、識別実験で用いるサブデータセットは、色塗りオブジェクトデータセットと線描オブジェクトデータセットを用いた。

表 1 に実験結果を示す。実験では、色塗りオブジェクトの種類を最も精度よく識別できる識別器は、色塗りオブ

ジェクトデータセットを用いて学習させた識別器（認識精度 96.5%）であり、線描オブジェクトの種類を最も精度よく識別できる識別器は、線描オブジェクトデータセットを用いて学習させた識別器（認識精度 94.3%）であることを確認した。複合オブジェクトデータセットを用いた識別器では、色塗りオブジェクトと線描オブジェクトのいずれもそれぞれ異なる描画スタイルのデータセットで学習した識別器よりは認識精度は向上したが、同じ描画スタイルのデータセットを用いた場合の精度には至らなかった。

なお、識別に要する処理時間は、1 つの描画オブジェクトに対して 0.08~0.2 秒であった。

5. 描画スタイル分類器との組み合わせ

5.1 概要

4.4 節の実験結果から、色塗りおよび線描という 2 種類の描画スタイルのオブジェクトの種類の識別には、同じ描画スタイルのデータセットで学習した識別器が有効であることを確認した。そのことから、初めにオブジェクトの描画スタイルを精度よく分類することができれば、その描画スタイルのオブジェクトの識別に適した識別器を適用することで、より高い精度で描画オブジェクトの種類の識別を行うことが可能となると考えられる。

そこで、新たに描画スタイル分類器を導入した手法を開発した。初めに、識別対象描画オブジェクトを描画スタイル分類器に適用して、そのオブジェクトの描画スタイルを分類する。そして、その分類結果を用いて、そのオブジェクトと同じ描画スタイルのデータセットで学習した識別器に適用することで、描画オブジェクトの種類の識別を行う。

本研究では SVM を用いた描画スタイル分類器の構築を行った。分類に使用する特徴量には、グレースケール化した画像の平均値、標準偏差、中央値、最頻値の 4 つを用いる。平均値は画像全体の輝度値の平均を求めることで明るさの違いを分析して、標準偏差は輝度値のばらつきを分析する。また、平均値が外れ値の影響を受けやすいことを考慮して、もう一つの画像全体の輝度値の指標として中央値を用いる。最頻値は、線描オブジェクトであれば画像全体に占める背景部分が多くなることを利用している。

5.2 実験

オブジェクトの描画スタイル分類器を構築して実験を行った。使用したデータセットは 4.2 節で述べた色塗りオブジェクトデータセットと線描オブジェクトデータセットである。そして、4.4 節と同じようにデータセットを 5 つのサブデータセットに分割して、4 つのサブデータセットを学習用、残りのサブデータセットを検証用として実験を行い、交差検証法で評価を行った。使用した PC も 4.4 節の実験で使用したのと同じである。

実験では、オブジェクトの描画スタイル分類の精度は平

均で99.4%であり、非常に高い精度で描画スタイルが分類できることを確認した。1つのオブジェクトの描画スタイルの分類に要する処理時間は約0.002秒であり、描画オブジェクトの種類に要する処理時間に比べて1/40以下という非常に小さいものであった。

本実験、および4.4節の実験の結果から、初めにオブジェクトの描画スタイルの分類を行い、次に分類結果のスタイルで学習したオブジェクト識別器を適用することで、個別の描画スタイルで学習したオブジェクト識別器や複合オブジェクトで学習したオブジェクト識別器に適用するよりも高い精度で描画オブジェクト種類の識別を行うことができることを確認した。

6. 不思議なスケッチブックへの適用実験

6.1 概要

本研究で開発して手描きオブジェクトの識別手法を不思議なスケッチブックに適用して実験を行った。オブジェクト識別手法は5章で述べた描画スタイル分類機とオブジェクト識別器を組み合わせたものである。使用したPCは4.4節と5.2節で使用したものと同一である。不思議なスケッチブックの実装は筆者らの従来研究と同じであり、C++でOpenGLとOpenCVを用いている。スケッチはA4用紙に手描きされ、Webカメラを用いて1280×720画素で取り込む。

オブジェクト識別器と不思議なスケッチブックは一つのPCで別プロセスとして起動しており、オブジェクト識別器は識別結果を不思議なスケッチブックに送る。そして、不思議なスケッチブックは各オブジェクトをCGモデル化して3DCG空間に配置する際、各オブジェクトの種類に識別結果に応じて、以下のように配置位置の変更や動きの付与を行った。

- 「雲」：空中に配置、垂直振幅運動を付与
- 「太陽」「星」：空中に配置、回転運動を付与
- 「車」：水平振幅運動を付与
- 「人」：垂直振幅運動を付与
- 「木」「花」：振り子運動を付与

6.2 結果と考察

絵が得意でない成人被験者2人にそれぞれ絵を描いてもらい、手描きオブジェクトを識別して不思議なスケッチブックに適用して3DCGシーンを自動的に生成する実験を行った。図6、図7に実験結果を示す。

図6の実験では描画オブジェクト数は11点、図7の実験では描画オブジェクト数は14点であったが、すべてのオブジェクトの種類は正しく識別された。そして、それぞれの識別結果を適用した3DCGシーンが自動で生成された。雲、太陽、星は空中に配置されて、それぞれに応じた動きも付与された。また、車、人、木や花もそれぞれに応じた

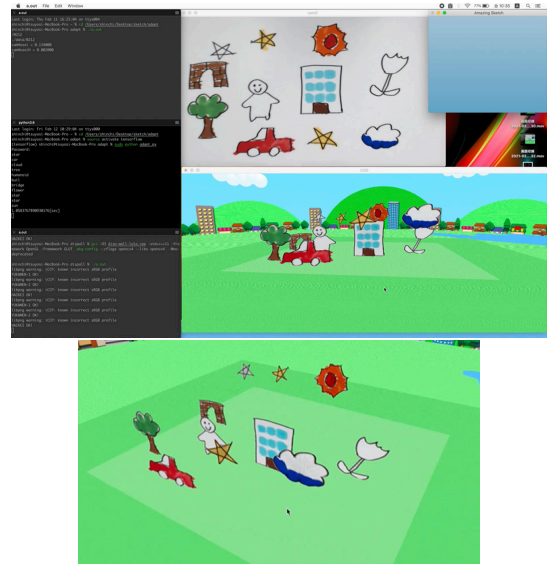


図6 描画オブジェクト識別結果の不思議なスケッチブックへの適用実験(1)



図7 描画オブジェクト識別結果の不思議なスケッチブックへの適用実験(2)

動きが付与された。これらは従来の不思議なスケッチブックでは実現できなかったことであり、提案手法によって不思議なスケッチブックによってより適切な3DCGシーンが自動的に生成されるとともに、動きによる表現の幅も広がった。

スケッチを取り込んでから3DCGシーンが生成されるまでの処理時間は約1.0秒であった。従来の不思議なスケッチブックでは30fpsのリアルタイム処理を行っているため、処理時間の大部分は描画オブジェクトの識別に要するものである。図2のように通常は1つのスケッチに10~30点の描画オブジェクトが含まれており、4.4節で述べたように各オブジェクトの識別に約0.08秒の処理時間を要するため、現状では1つのスケッチから3DCGを生成するの

に1~3秒程度が必要となる。そのため、従来の不思議なスケッチブックと同様にスケッチを介した3DCG空間とのインタラクションを行うには処理時間の低減や必要に応じた識別処理の適用が必要である。

7. まとめ

本研究ではCNNを用いて手描きオブジェクトの種類を識別する手法の開発を行った。そして、その識別結果を不思議なスケッチブックへ適用させる実装を行った。提案手法によって、不思議なスケッチブックでよく描かれる10種類のオブジェクトを高い精度で識別して、識別結果を用いて従来より適切な3DCGシーンを自動的に生成できることを確認した。

今後の課題としては、識別できる描画オブジェクトの種類、描画オブジェクトの3DCGモデル生成へのオブジェクト識別結果の適用、各描画オブジェクトの識別結果に基づくシーン全体の識別と3DCGシーン生成への適用などが挙げられる。

なお、本研究の一部は公益財団法人大幸財団研究助成金によって行った。

参考文献

- [1] T. Igarashi, S. Matsuoka and H. Tanaka: Teddy: A sketching interface for 3d freeform design, Proc. ACM SIGGRAPH '99, pp. 409-416 (1999).
- [2] 松田浩一, 鈴木俊博, 静春樹, 近藤邦雄: スケッチインタプリタシステム: 手描き陰影による3次元形状制御法, 情報処理学会論文誌, Vol. 44, No. 11, pp. 2547-2555 (2003).
- [3] G. Nishida, I. Garcia-Dorado, D. G. Aliaga, B. Benes, A. Bousseau: Interactive sketching of urban procedural models, ACM Trans. Graph., Vol. 35, Issue 4, No. 130 (2016).
- [4] 近藤菜々子, 水野慎士: スケッチブックでのお絵描きを三次元CGで拡張する映像ツールの提案とその実現方法, 情報処理学会論文誌・デジタルコンテンツ, Vol. 1, No. 1, pp. 1-9 (2013).
- [5] 水野慎士: 輪郭検出を用いた「不思議なスケッチブック」のCG表現拡張とイベント展示のためのビューア開発, 芸術科学会論文誌, Vol. 17, No. 4, pp. 72-82 (2018).
- [6] Google: Quick Draw!, 入手先 (<https://quickdraw.withgoogle.com/>) (参照 2021-05-01).
- [7] K. Sasaki, M. Yamakawa, K. Sekiguchi, T. Ogata: Classification of photo and sketch images using convolutional neural networks, Proc. of ICANN 2016, pp. 283-290 (2016).
- [8] C. Furusawa, K. Hiroshiba, K. Ogaki, Y. Odagiri: Comicolorization: Semi-Automatic Manga Colorization, Proc. of SIGGRAPH Asia 2017, Technical Brief (2017).
- [9] K. Simonyan, A. Zisserman: Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv:1409.1556 (2014).