

機械学習を用いた3次元モデルからの様式化された線画の生成

内田光洋[†] 齋藤豪^{††}

[†] 東京工業大学 情報工学科 ^{††} 東京工業大学 情報理工学院

1 はじめに

3次元モデルから強弱の付いた線を生成する手法は多く存在するが、実際の手描き線のような複雑な強弱を定式化することは難しい。本研究では畳み込みニューラルネットワーク(CNN)を用いることでデータに基づく強弱の決定を行い、3次元モデルから様式化された強弱付きの線画を生成する手法を提案する。

2 関連研究

前田ら[1]はCNNの入力に3次元モデルのカメラに対する法線、深度を抽出した画像、出力に強弱付き線画を用いる手法を提案している。この手法はシルエット線を太くし、インク溜まりを再現した線画の生成に対して良好な結果を得ている。本研究はより多様な線画を生成可能にするために、この手法を拡張する。

3 提案手法

本手法では、(1)CNNの入力に用いる画像を3次元モデルから生成し、(2)CNNにより太さの強弱の付いた線画像を得て、(3)その線画像から線の強弱を抽出するという3つの段階に分けられる。様式の描画法を抽出した強弱に適用することで、強弱のある様式化された線画を得る。手法の概要を図1に示す。

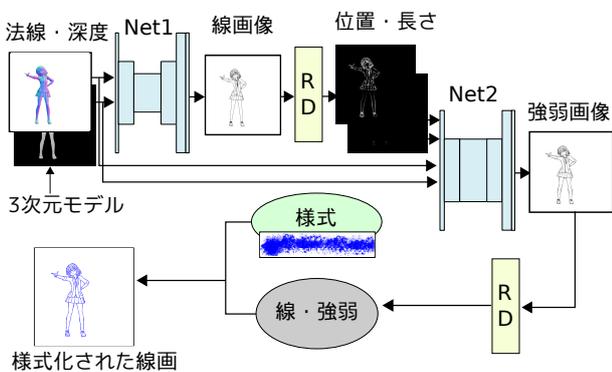


図 1: 手法概要

3.1 一様な線の生成

一様幅の線画像を教師画像として用いて前田ら[1]の提案したCNN(以後 Net1)を学習させる。

学習済みの Net1 で生成した線画像に Ridge Detection[2]を適用して線の稜線をポリラインとして

Stylized line drawings from 3D models using machine learning

[†] Mitsuhiro UCHIDA

^{††} Suguru SAITO

Faculty of Engineering, Tokyo Institute of Technology ([†])
School of Computing, Tokyo Institute of Technology (^{††})

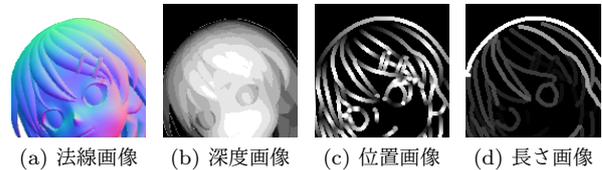


図 2: Net2 の入力画像

検出し、稜線から幅 5px の範囲のピクセルに次の値を格納したグレースケール画像 2 枚を生成する。

- 線の端点を最小、中点を最大として、線の軌跡方向に三角の変化を持つ値 (図 2(c))
- 線の長さに比例する値 (図 2(d))

3.2 3次元モデルからの強弱生成

前節の2チャンネルの画像と Net1 の入力に用いた法線画像(3チャンネル)と深度画像(1チャンネル)を合わせて6チャンネルの入力とし、強弱を線の幅で表す線画像を出力とする新たなCNN(以後 Net2)の学習を行う。Net2 の入力画像の例を図2に示す。ただし、原稿上の見やすさのために深度画像と長さ画像はトーンカーブを補正している。

Net2 の構成を表1に示す。入力層、出力層以外にはバッチ正規化を適用している。

Net2 の学習に用いる損失関数は、教師画像 T と出力画像 O により式1で定義する。

$$loss(T, O) = (\|T - O\|_2)^2 \quad (1)$$

学習の際は、入出力の画像に回転、反転処理を施すことで、データセットの数を増やす。上記の変換を施した画像を 128×128 のサイズで切り抜いた画像 10 枚によるミニバッチ学習を行う。

表 1: ネットワークの構成

カーネル	ストライド	出力画像	活性化関数
-	-	$6 \times W \times H$	-
5×5	2×2	$64 \times W/2 \times H/2$	ReLU
3×3	1×1	$128 \times W/2 \times H/2$	ReLU
3×3	1×1	$128 \times W/2 \times H/2$	ReLU
3×3	1×1	$256 \times W/2 \times H/2$	ReLU
3×3	1×1	$256 \times W/2 \times H/2$	ReLU
3×3	1×1	$512 \times W/2 \times H/2$	ReLU
3×3	1×1	$512 \times W/2 \times H/2$	ReLU
3×3	1×1	$256 \times W/2 \times H/2$	ReLU
3×3	1×1	$256 \times W/2 \times H/2$	ReLU
4×4	$1/2 \times 1/2$	$128 \times W \times H$	ReLU
3×3	1×1	$128 \times W \times H$	ReLU
3×3	1×1	$64 \times W \times H$	ReLU
3×3	1×1	$32 \times W \times H$	ReLU
3×3	1×1	$1 \times W \times H$	Sigmoid

3.3 強弱の抽出

学習済みの Net2 で生成した線画像に Ridge Detection を適用すると、その線のポリライン座標とそれらの点での線の法線方向、及び幅が得られる。特定の幅の範囲を [0:1] の強弱として扱う。

3.4 線の様式化

様式の描画法に対し前節で抽出した線の情報、及び強弱の情報を用いることで、目的の線画を得る。描画は線のテクスチャ画像の変形に基づいている。

4 実験結果と評価

Net1 及び Net2 の学習にはプロ生ちゃん [3] の 3次元モデルを用いる。Net1 の学習の教師画像には 2px の一様幅の線画像 20 枚を用いた。Net2 の学習の教師画像として以下の条件を満たす強弱付き線画像を 20 枚を用いた。

- 法線、深度画像の微分値が一定以上の箇所に線を描く
- 線の幅は最小 2px、最大 4px にする
- 一定以上の深度差がある箇所に曲線を描く場合、曲率が大きい場所で幅を最大にする
- 線の幅は連続的に変化させ、他の線に接していない端点は細くする

ミニバッチ学習により Net1 を 150,000 回学習させた後、Net2 を 200,000 回学習させた。

Ridge Detection で抽出される幅は描いた際の幅からずれることを考慮して、今回は教師画像に 2px から 4px の幅を用いたため、幅 [1px:5px] を強弱 [0:1] と定義した。Net2 の出力画像と対応する教師画像、及びテスト画像から抽出した強弱を比較した結果を表 2 に示す。各値は教師画像（テスト画像）の強弱に対する出力画像の強弱の誤差の割合の平均である。

表 2: 幅の誤差

評価	教師画像	テスト画像
誤差	1.94%	2.38%

曲率に対する線の幅のグラフを図 3 に示す。横軸は曲率の対数、縦軸が強弱である。緑が教師画像、赤が出力画像の平均と標準偏差を表す。教師画像には曲率が一定以上になると強くなる傾向があり、出力画像にも同様の傾向がある。

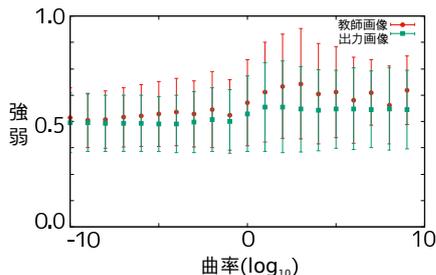


図 3: 曲率に対する線の幅

また、出力画像の強弱は教師画像の強弱に比べて小さくなる傾向があるが、これは教師画像に比べて出力画像はぼやけた線になりやすいという CNN の特徴に起因すると推測される。本手法では一度強弱を抽出するため、弱くなった強弱も調整することができる。その例を図 4 に示す。

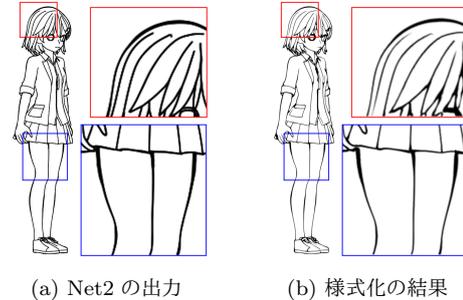


図 4: 強弱の調整

強弱情報に様式を与えた結果を図 5 に示す。図 5(a) は毛筆のような線を表現した線画である。図 5(b) は強弱を濃淡として表現した線画である。

5 まとめ

CNN の出力の線画像を強弱の情報と見なし、その情報から強弱のある様式化された線画を生成する手法を提案した。1 つの強弱情報に対して複数の様式を適用することができるため、多様な表現が可能である。

今後の課題として、Net2 への入力の種類追加が考えられる。線の端点が他の箇所の強弱に影響を及ぼす場合もあり、線の長さや線中の位置だけではその情報を伝達できない。したがって、ある線の端点の種類を線中の各点に与えるなどすれば、必要な情報を伝達できるため、より多様な強弱の決定が可能になることが期待される。

参考文献

- [1] 前田浩輔, 齋藤豪. 畳み込みニューラルネットワークを用いた 3dcg による強弱有り線画像の生成. 研究報告コンピュータグラフィックスとビジュアル情報学 (CG), Vol. 2018-CG-169, No. 7, pp. 1-7, feb 2018.
- [2] Mark Hiner Thorsten Wagner. An unbiased detector of curvilinear structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 2, pp. 113-125, 1998.
- [3] <http://3d.nicovideo.jp/works/td8608>.

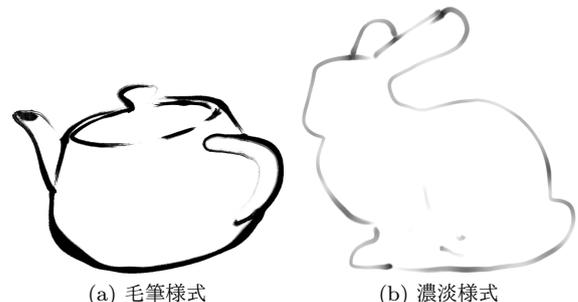


図 5: 様式化の結果