

## 深層学習を用いた数学グラフ画像のベクタ画像変換方法

大嶋健悟（富山県立大学）・高木昇（富山県立大学）・澤井圭（富山県立大学）

・増田寛之（富山県立大学） ・本吉達郎（富山県立大学）

## 1. はじめに

グラフは、数学や科学の関数や方程式を表すために使用される。これらのグラフは視覚的であり、視覚障害者は図にアクセスすることが困難である。視覚障害者向けの電子書籍の国際標準を開発、維持、促進する DAISY (Digital Accessible Information System) では、SVG を活用した数式や図表等の科学技術文献への対応が推奨されている[1]。画像をベクタ形式で保存することで様々なメリットが存在するが、ラスタ形式の画像をベクタ形式に変換する研究は少ない。本稿では数学グラフに焦点を置き、実践成分のベクタ画像変換手法について提案する。

## 2. 提案手法

ラスタ画像をベクタ形式に変換する際、成分ごとに分割する必要がある。提案する手法は、数学グラフにおける実線成分を各成分にセグメンテーションすることを目的とする。本手法は Kim ら[2]の先行研究である Vectornet を用いて大きく分けて 3 ステップの処理を行う(図 1 参照)。

**Step1** はじめに、交点領域の検出を行う。入力画像  $I$  に細線化処理(Zhang-Suen)を適用し、細線化画像すべてのピクセル  $p$  について、 $3 \times 3$  のフィルタを使用して、ピクセル  $p$  が細線化画像の交点であるかどうかを判別する。ここで、入力画像  $I$  の各成分は複数のピクセルの線幅を持つ。ゆえに細線化画像には、1 つの交点領域に複数の交点が存在する。したがって入力画像の 1 つの交点領域に存在する複数の交点に階層型クラスタリングを適用する。適用した階層型クラスタリングはウォード法である。クラスタ間のクラスタ重心のユークリッド距離が閾値以上になるとクラスタリングを終了する。

**Step2** Step1 で得られた交点領域を入力画像  $I$  から  $64 \times 64$  の領域でパッチを切り出す。切り出したパッチの集合  $S_p$  を  $S_p = \{P_1, P_2, \dots, P_j\}$  とする。  $j$  は交点領域の数で、  $P_j$  は切り出したパッチ画像である。パッチの集合  $S_p$  は、すべて Vectornet へ入力し、パッチ画像のセグメンテーション結果  $S_p = \{P_1, P_2, \dots, P_j\}$  を得る。

**Step3** 入力画像  $I$  から Step1 で得られた交点領域  $32 \times 32$  の領域を除去した画像  $I_r$  とする。ここでピクセルの連結した成分  $C$  を  $C = \{p_1, p_2, \dots, p_n\}$  とし、画像  $I_r$  内の成分の集合  $S_{I_r}$  を  $S_{I_r} = \{C_1, C_2, \dots, C_i\}$  とする。ただし、  $p_n$  は画像  $I_r$  内の成分  $C$  上のピクセルであり、  $i$  は画像  $I_r$  内の成分数である。次に Step2 より、  $P_j \in S_p$  の同じラベルが割り当てられた成分を取り出し、1 パッチから得られる成分の集合を  $S_{P_j} = \{C_1, C_2, \dots, C_i\}$  とする。ここで、  $l$  は  $P_j$  のラベル数である。そして、  $(C_i \in S_{I_r}) \cap (C_l \in S_{P_j}) \neq \emptyset$  であるとき、成分  $C_i$  と成分  $C_l$  は連結した成分であるため、成分  $C_i = C_i \cup C_l$  とする。これらの処理を、すべての  $P_j \in S_p$  に関して行う。

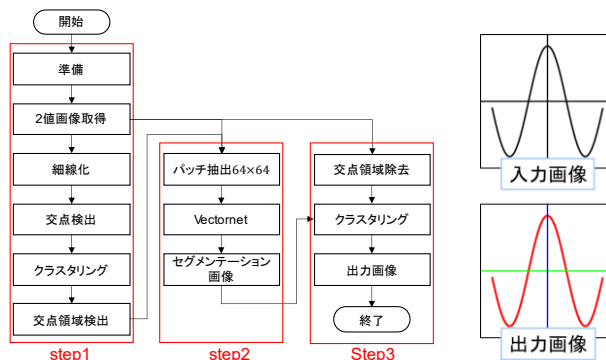


図 1 提案手法の処理の流れ

## 4. 実験

提案した手法の有効性を調査するために、検証を行った。Vectornet の学習に用いるデータセットは直線及びベジェ曲線をランダムに生成することで 45000 枚作成した。

検証に使用するデータセットは、LaTeX のライブラリである Tikz を用いて実線のみの中初等数学で用いられる数学グラフを 20 枚作成した。

評価方法は、本手法から得られる各成分に分けられた出力が、Tikz の成分と同様の出力になっていれば正解、そうでなければ不正解として、著者が目視で判断した。

## 5. 結果及び考察

実験を行った結果 20 枚中 5 枚の画像を正確に成分ごとに分割することができた。分割に失敗したほとんどの画像が、パッチ画像のセグメンテーション結果が良好でなかった。直線及び、ベジェ曲線で生成したデータセットが、数学グラフのパッチ画像に現れる線のかさなりとは大きく異なることが問題であると考えられる。

## 6. まとめ

本稿では、視覚障害者支援のための数学グラフのベクタ画像変換手法について提案し、検証した。実験は、Tikz で 20 枚の実線のみの中初等数学の数学グラフを作成し、提案した手法を評価したが、実験を行った結果、20 枚中 5 枚が正解という望ましい結果とはならなかった。原因として、学習するデータセットに問題があることが想定される。したがって今後は、数学グラフから抽出されるパッチ画像に近いデータセットを作成することで、さらに精度が向上すると考えられる。

## 参考文献

- [1] 河村宏著: デジタル・インクルージョンを支える DAISY と EPUB, 東京科学技術振興機構, p. 311, 2011.
- [2] Byungsoo Kim, Oliver Wang, A. Cengiz Öztireli and Markus Gross, "Semantic Segmentation for Line Drawing Vectorization Using Neural Networks," *COMPUTER GRAPHICS FORUM*, Vol.37, pp. 329-338, 2018.