

N グラムモデルを用いた輪郭線補完

平野 雅人[†] 山村 毅[†]

本研究では、輪郭線抽出処理の結果、途切れた輪郭線や、主観的輪郭のような、物理的には存在しないが知覚される輪郭を生成する問題を扱う。主観的輪郭については様々な研究がなされており、それによれば、主観的輪郭の生成には、輪郭成分の方位検出のような低次のプロセスと、形状記憶にかかわる高次のプロセスからなるとされている。

この2つのプロセスを確率モデルを用いることで実現する。低次のプロセスは有効線分の抽出という形で表現し、高次のプロセスはその有効線分の組み合わせをNグラムモデルとトリガーモデルを線形補間したもので表現する。このような手法により、自然な線分によって補完できることを示す。

Contour completion by an N-gram model

MASATO HIRANO[†] and TSUYOSHI YAMAMURA[†]

In this paper, we deal with contour completion problem, that is, interpolation of contour parts as well as complement of contours which are perceived without physical stimuli, like subject contours. Various studies on subject contours have been done, and according them, a subject contour is generated through low and high level processes; the low level process includes detection of contour orientation and the high level one is related to shape reconstruction.

We model these two processes based on probability. The low and high level processes are realized by extraction of valid segment and combination of those valid segments, respectively. Experimental results show that our method can generate natural contours.

1. はじめに

二次元画像から物体抽出や物体認識を行う場合、途切れた輪郭線が問題となることがある。ここでいう途切れた輪郭とは次の2つの種類に大別できる。

1つ目はフィルタリングやテンプレートによる輪郭線の抽出を行った場合に生じるものである。一般に輪郭線抽出処理は、濃淡値などの物理的量のみに依存するため、光源の位置や物体表面の反射率などの外的要因により物理量の差分が小さくなると、輪郭線が途切れてしまう。このように輪郭線が途切れてしまうと物体の境界がわからないため、物体の抽出を行うことができなくなる。途切れた領域が小さな場合は収縮・伸張による補間を行う方法が、大きな場合はスプラインなどによる線形近似により補間を行う方法がある。

2つ目は、人間の視覚が自然に行う視覚的補間である¹⁾。視覚補間は非感覚的補間 (amodal interpolation) と感覚的補間 (modal interpolation) に分けられる。人

間は、物体が重なっている状態を観察しているとき、遮蔽された領域が実際に見えなくとも、遮蔽する領域の背後まで広がって存在すると認識する。このような働きを非感覚的補間という。一方、境界線や明暗の差などの物理的刺激がないにもかかわらず知覚する輪郭を主観的輪郭 (subjective contour) と呼ぶ。補間された輪郭が実際に目に見えることから、感性的補間とも呼ばれる。

安田ら²⁾は、主観的輪郭生成の工学的モデルを提案している。それによれば、ある面がほかの面や線を覆い隠すとき、その重なり部分にはT型頂点が生じるが、覆い隠す面が主観的輪郭によって囲まれる図形の場合は、それがL型頂点やI型頂点になる。従ってこのような頂点を画像中から抽出して滑らかにつなぐことで、主観的輪郭を生成することができる。この手法は、石寺ら³⁾が指摘しているように、脳内のモデルを正しくモデル化しているわけではない。しかし、少なくともこの手法を用いれば、主観的輪郭が生成できるのは確かであり、また、非感覚的補間にも応用可能である。実際、Teranishiら⁴⁾及び山村ら⁵⁾はこの考え方を、重なり合った物体の内部輪郭抽出に応用している。

[†] 愛知県立大学
Aichi Prefectural University

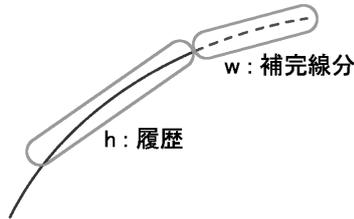


図 1 補完線分
Fig. 1 Complement lines

しかし、これらに代表される主観的輪郭に関する研究では、適切な頂点を見つけ出し、それらをスプラインによって補間を行っているため、滑らかにつながる曲線のような近似はできても、長方形の角が欠けているような、急激な変化を含む図形に対しては対応できない。

本研究では、確率的モデルを導入し、学習データから特徴的な線分とその線分の共起関係を学習することで、自然な輪郭の生成を試みる。主観的輪郭は、輪郭成分の方位検出のような低次プロセスと形状記憶にかかわる高次のプロセスの2つに分かれるとされている。本研究では、低次のプロセスは有効線分の抽出で表現し、高次のプロセスはその有効線分の組み合わせを、Nグラムモデルとトリガーモデルを線形補間したもので表現する。実験は主に感性的補間を対象とし、隠蔽された輪郭を補完する線分が、いかに人間の視覚心理に近いかを考慮し評価する。

以下、2では特徴的な線分の学習方法と、その線分を基本単位として補完する線分の評価法について述べる。そして3では、実際に作成したシステムの流れを説明し、4で実験を行う。

2. 輪郭生成の確率モデル

今、ある端点から補完する線分を推測することを考える。

補完すべき線分は、1. 緩やかに変化する、2. 一般的によく見る比較的単純な線分である、と仮定できる。このような特徴から、補完すべき線分は、対象となる入力画像の実際の輪郭形状に大きく依存していると考えられる。したがって、補完線分を求めることは、図1のよう

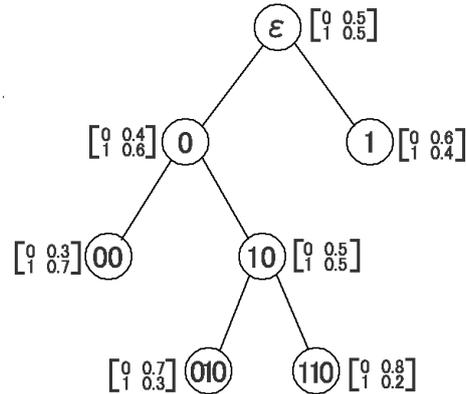


図 2 予測接尾木
Fig. 2 Prediction suffix tree

に入力画像の一部を履歴 h 、補完線分を w としたとき、

$$\operatorname{argmax}_w P(w|h) \quad (1)$$

となるような補完線分 w を推定する問題と考えることができる。この w を推定するために、基本単位として有効線分という概念を導入する。

2.1 有効線分の抽出

出現確率の高い線分を有効線分と呼ぶことにする。

今、コード化された輪郭線のコードを $x_1^n = x_1 \dots x_n$ と表す。与えられた x_1^n のコード列の生成確率 $P(x_1^n)$ を、確率論の乗法定理を用いると、式(2)のようになる。

$$\begin{aligned} P(x_1^n) &= P(x_1)P(x_2|x_1) \dots P(x_n|x_1^{n-1}) \\ &= \prod_{i=1}^n P(x_i|x_1^{i-1}) \end{aligned} \quad (2)$$

式(2)より、条件付き確率 $P(x_i|x_1^{i-1})$ を求めればコード列の生成確率を求めることができるが、組み合わせ数が膨大となるため、このまま用いることは現実的ではない。そこで、Nグラムモデルを用いて、コード列の生起が(N-1)前までのコードの並びにのみ依存すると仮定して近似を行う。これを用いると、式(2)は以下のようにになる。

$$P(x_i|x_1^{i-1}) \approx P(x_i|x_{i-N+1}^{i-1}) \quad (3)$$

$$P(x_1^n) \approx \prod_{i=1}^n P(x_i|x_{i-N+1}^{i-1}) \quad (4)$$

このようにNグラムモデルを使用することで、履歴を使用した線分の推定を行うことができるが、履歴として使用するデータの長さが問題となる。

Nが大きくなればなるほど x_i を特定しやすくなるため、精度が向上すると考えられる。しかしNが大きくなると、そのようなデータが学習データ内に複数回

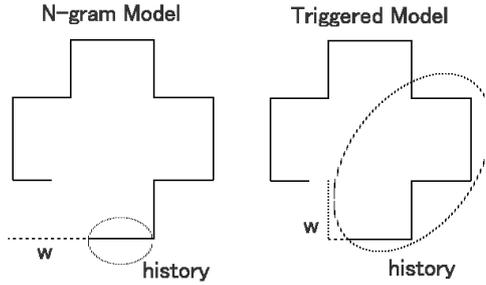


図3 N グラムモデルとトリガーモデルの違い
Fig. 3 Difference between N-gram Model and Triggered Model

でくる可能性が低くなり、モデル自体の信頼性が低くなる。実際、言語処理の分野では、 $N > 3$ にしても精度が向上しないことが報告されており⁶⁾、パラメータ推定と精度の観点から $N=2$ (バイグラム) あるいは $N=3$ (トライグラム) を用いることが多い。

しかし、本研究では、有効線分を生成するために使用するデータに相対的な方向を表すチェーンコードを用いているので、データの単位としては画像全体に比べ非常に局所的な変化に注目している。そのため、バイグラムやトライグラムでは情報量が少なく、有益な生起確率を計算することができない。そこで、より精度の高い線分を得るために N の値を必要時に変化させて扱う可変長 N グラムモデルを使用する。

可変長 N グラムを実現するには様々な研究がなされているが、ここでは Singer⁷⁾ による予測接尾木による表現を使用する。

予測接尾木の例を図2に示す。各節点は、ラベルとして条件付き確率の履歴を持ち、生起確率の分布を $P_n(x)$ として持つ。このように表現された予測接尾木を用いたとき、コード列 x_1^n の生成確率は式(5)で表される。

$$P(x_1^n) = \prod_{i=1}^n P_{n(i-1)}(x_i) \quad (5)$$

ここで $n(i)$ は接尾木の根から $x_i \dots x_1$ とたどったときに到達できる最も深い節点を表す。このように構築したモデルから、生成確率を求め、式(6)を満たすコード列を、有効線分とする。

$$P(x_1^t) > \max(0.125^t, 0.125^D) \quad (6)$$

ここで D は探索を打ち切るための定数である。

2.2 補完線分の評価

2.1 で生起確率の高い要素(有効線分)について述べた。この有効線分を単位として、その組み合わせで輪郭線の補間を行う。これを有効線分のバイグラムモデルとトリガーモデルでモデル化する。

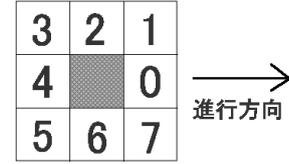


図4 コード変換テンプレート
Fig. 4 Code transform template

2.2.1 バイグラムモデルによる輪郭線の生成

今、与えられた輪郭線(履歴)が有効線分の列として $h_1^n = h_1 h_2 \dots h_n$ と表すことができたとする。補完すべき輪郭線を表す有効線分列を $w_1^m = w_1 w_2 \dots w_m$ とすると、その生起確率は、

$$\begin{aligned} P(w_1^m | h_1^n) &= P(w_1 | h_1^n) P(w_2 | h_1^n w_1) \dots P(w_m | h_1^n w_1^{m-1}) \\ &= \prod_{i=1}^m P(w_i | h_1^n w_1^{i-1}) \end{aligned} \quad (7)$$

となる。各 w_i の生成にバイグラムを用いれば、

$$P(w_1^m | h_1^n) \approx P(w_1 | h_n) \prod_{i=2}^m P(w_i | w_{i-1}) \quad (8)$$

となる。この確率を最大にするように有効線分列 $w_1 w_2 \dots w_m$ を選ぶ。

2.2.2 トリガーモデルによる輪郭線生成

バイグラムモデルを用いることで局所的にもっともらしい線分を生成することができるが、このままでは全体的なバランスを考えた線分の生成は難しい(図3)。そこで、より長距離の履歴を考慮するためにトリガーモデルを導入する。トリガーモデルは式(9)で計算される。

$$P_T(w_n | w_{n-M}^{n-1}) = \frac{1}{M} \sum_{m=1}^M \alpha(w_n | w_{n-m}) \quad (9)$$

ここで $\alpha(w_n | w_{n-m})$ は、線分 w_{n-m} と線分 w_n の間の共起の強さを表しており、これは学習データから抽出されたトリガー対から計算される。トリガー対の抽出方法は様々あるが、本研究ではユニグラムレベル・トリガーによる抽出を行った。

まず、基礎となるモデルとしてユニグラム $P(w)$ を考える。ユニグラムは N グラムモデルの一種であり、 $N=1$ の場合のモデルである。これは線分が前の線分に影響されずに独立に生起することを表している。次に、このモデルにトリガー対 $w_a \rightarrow w_b$ を追加した

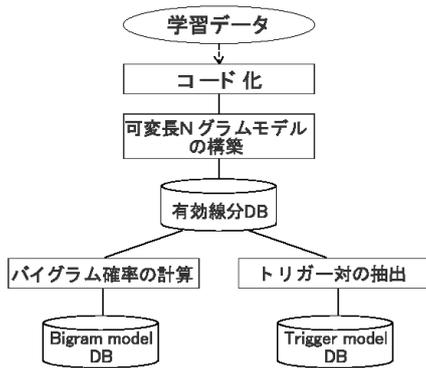


図 5 学習の流れ
Fig. 5 The flow of study

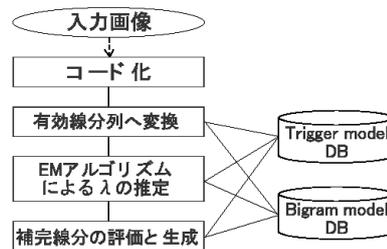


図 6 補完の流れ
Fig. 6 The flow of compliment

ときにどれくらいモデルが改善されるかを計算することで、そのトリガー対が重要なかを判断する。この考えに基づき計算を行うと共起の強さを表す未知パラメータ $p(w_b|w_a)$ は、

$$p(w_b|w_a) = \frac{C(w_a, w_b)}{C(w_a, w_b) + C(w_a, \bar{w}_b)} \quad (10)$$

で計算できる。共起の強さを表す $\alpha(\cdot)$ は、 $p(w_b|w_a)$ を正規化することで得ることができる。

$$\alpha(w_b|w_a) = \frac{p(w_b|w_a)}{\sum_{w'_b} p(w'_b|w_a)} \quad (11)$$

これにより、長距離間の線分の共起の強さを得られる。

2.3 モデルの合成

バイグラムモデルとトリガーモデルを合成して、補完線分の生成確率を求める。すなわち、それぞれのモデルによる生成確率を線形で補間する。

$$P(x) = \lambda P_N(x) + (1 - \lambda) P_T(x) \quad (12)$$

ここで λ は、学習データの尤度を最大化するような補完係数 ($0 \leq \lambda \leq 1$) で、EM アルゴリズムを用いて求めることができる。

3. 構築システム

ここでは構築したシステムについて説明する。システムは大きく分けて、学習と補完の2つのフェーズに分けることができる。共通の処理として、コード化という処理がある。まずはその処理について説明し、そのあとに学習と補完のアルゴリズムに関して説明する。

3.1 コード化

学習データや入力画像の輪郭情報を扱うために、輪郭追跡を行って、チェーンコードに変換し、そのあと図4のような方向を表すコードに変換する。今、輪郭追跡によって得られたチェーンコード $C = c_1 \dots c_n$ を相対的なチェーンコード $X = x_1 \dots x_{n-1}$ に変換するこ

とを考える。このとき、変換式は、

$$x_i = \begin{cases} c_{i+1} - c_i, & c_{i+1} - c_i \geq 0 \\ (c_{i+1} - c_i) + 8, & c_{i+1} - c_i < 0 \end{cases} \quad (13)$$

のようになる..

このように画像に対して相対的なコードを用いることで、直線は直線、曲線は曲線のように線分情報を方向に関係なく記述でき、また状態数を減らすことが可能である。

3.2 学習

学習は、有効線分の抽出・バイグラム確率の計算・トリガー対の抽出を行う。学習の流れを図5に示す。

まず、学習データから得られた輪郭線を3.1の手法によりコード化する。このコード列を使用し2.1の手法で有効線分を抽出する。有効線分の抽出にあたり、可変長 N グラムモデルを使用するが、システムに実装する上で確率有限オートマトンに変換をする。予測接尾木を用いて単語の条件付き生成確率を求めるためには、単語の現れる履歴を同定するために接尾木の根から子節点への経路を次々にたどらなければならない。一方、確率有限オートマトンでは、状態が履歴に対応しているため、わざわざ履歴を同定する必要がない。この点から確率有限オートマトンの方が実装の上でより効率的な表現方法であるといえる。このように抽出された有効線分を基本単位とし、学習データから得られた輪郭をコード化したものを使用して、バイグラム確率の計算とトリガー対の抽出を行う。

3.3 補完

ここでは、未定係数 λ の推定と補完を行う。補完の流れを図7に示す。推定には入力画像によって与えられた途切れた輪郭情報を有効線分の列に変換したものを使用する、有効線分の列への変換には、学習で構築した

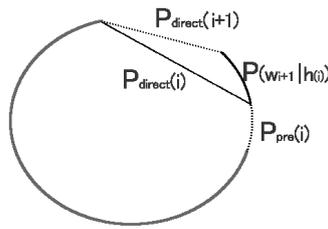


図7 輪郭補完のアルゴリズム
Fig. 7 Contour completion algorithm

バイグラムモデルによって行い、最も尤度の高くなる並びを使用する。その後、未定係数 λ を推定し、補完を行う。今、 i 回目の補完候補線分の生起確率を $P_{pre}(i)$ 、始点から終点を直接結ぶ線分の生起確率を $P_{direct}(i)$ と定義する (図7)。補完のアルゴリズムは以下のように行う。

- (1) 初期化の処理として $P_{pre}(i) = 1.0$ とする。
- (2) 現在、補完を行いたい始点と終点を最短で結び、そのときの長さを $L_d(i)$ としたとき、その線分の生起確率を

$$P_{direct}(i) = P_{pre}(i) \times 0.125^{L_d(i)} \quad (14)$$

として計算する。

- (3) 前回の推定から得られる輪郭の履歴 $h(i)$ から推定される、補完線分の候補 $w_1, \dots, w_j, \dots, w_n$ 各々について、

$$P_{comp}(i+1) = P_{pre}(i) P(w_{i+1}|h(i)) \times 0.125^{L_d(i+1)} \quad (15)$$

を計算する。 $L_d(i+1)$ は現在の候補を補完線分として採用したときの始点と終点を直接結んだ最短直線の長さである。

- (4) $P_{direct} < P_{comp}$ を満たすならば、補完候補 w_{i+1} を集合 W に追加する。
- (5) $W \in \phi$ ならば、現存の推定した線分の中で、もっとも生起確率の高い候補を補完線分とし終了。そうでなければ補完線分候補の集合 W のなかで、生起確率の高い5候補を次回の探索候補として (2) 以降を繰り返す。これにより得られた結果を、本研究の最終結果とする。

4. 実験

4.1 実験方法

実験の画像として、次のようなものを使った。

- 8bit 256 階調濃淡画像
- サイズ:450x270
- 学習データ

長方形, 四角形 : 大きさ, 回転したものを各 5 点ずつ

円, 楕円 : 大きさを変えたものを各 5 点ずつ

- 入力画像
 - 学習データでは用いなかった図形 2 種類
 - 他物体に覆われた物体を想定

4.2 実験結果

実験の対象として図 8,11 のような画像を想定し、補完すべき対象図形をそれぞれ図 9,12 とする。それぞれの結果を図 10,13 に示す。対象図形は、前もって手動により画像から抽出した。

図からわかるように、人間の視覚心理の結果にそっている。特に図 13 の結果は、履歴として使っている輪郭線の特徴を補完をする上でうまく使っているといえる。このことから、本手法が途切れた輪郭線の補完に有効であるということがわかるであろう。

5. まとめ

本研究では、確率モデルを使用することで途切れた輪郭線を補完する方法を示した。基本単位となる線分として、有効線分という単位を導入し、可変長 N グラムモデルにより抽出を行った。補完候補を N グラムモデルとトリガーマデルを線形補完したモデルにより評価を行った。本手法が、感性的輪郭を代表とする途切れた輪郭線を補完する手法として有効であることがわかった。

本研究では、実験の対象として感性的輪郭の補完を行ったが、主観的輪郭にも対応できると考える。本手法の特徴として、その輪郭線分がいかに有益であるかということを確認率という側面から評価することができる。主観的輪郭は、図形の完成度やバランスに大きく依存していることがわかっている。本手法を使用することで、それらを生起確率という方法で評価できる。

しかし一方で、結果が学習データの種類や数、パラメータに大きく依存するため、人間の予測結果と大き

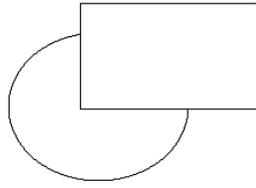


図 8 原画像 (1)
Fig. 8 Original image (1)

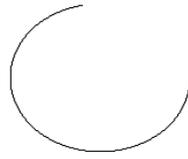


図 9 対象図形 (1)
Fig. 9 Target image (1)

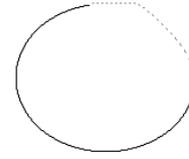


図 10 結果 (1)
Fig. 10 Result (1)

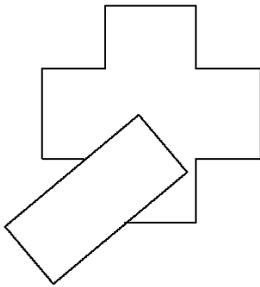


図 11 原画像 (2)
Fig. 11 Original image (2)

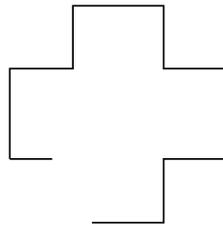


図 12 対象図形 (2)
Fig. 12 Target image (2)

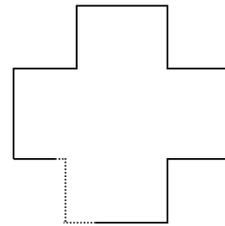


図 13 結果 (2)
Fig. 13 Result (2)

くかけ離れる心配がある。たとえば、学習データと補完対象となる入力データとの間で解像度が大きく異なる場合、良好な結果が得られるとは限らない。

これらの問題に対処する一つの方法として、有効線分の特徴をうまく記述できるクラスによってクラスタリングをおこない、パイグラムモデルではなく、パイグラムクラスモデルを使用することで対処する方法が考えられる。今後これらの方法についても検討し問題に対処したい。

参 考 文 献

- 1) 川人光男, 行場次郎, 藤田一郎, 乾 敏郎, 力丸 裕 : "岩波講座 認知科学 3 視覚と聴覚", 岩波書店, (1994)
- 2) 安田浩之, 安藤和久, 大西 昇 : "物理的に存在しない輪郭線の抽出", 電子情報通信学会論文誌 (D-II), J73-D-II, 6, pp906-913(1990-06).
- 3) 石寺永記, 荒井祐之, 土屋雅彦, 宮内裕子, 高橋信一, 栗田正一 : "主観的輪郭の形成に関する視覚情報処理モデル", 電子情報通信学会論文誌 (D-II), J76-D-II, 4, pp.873-880(1993-04)
- 4) Teranishi M., Ohinishi N. and Sugie N. : "A Computational Model for Generating Subjective Contours and Its Application to Detection of Multiple Object Overlapping", Int. Symp. on Neural Informaiton Processing, pp.134-137(1992)
- 5) 山村 毅, 滝沢 和明, 大西 昇, 杉江 昇 : "主観的輪郭に示唆を得た, 重なり合った物体の内部輪郭の生成", 電子情報通信学会論文誌 (D-II), J78-D-II, 5, pp880-883(1995-05)
- 6) Weintraub, M., Aksu, Y., Dharanipragada, S., Khudanpur, S., Ney, H., Prange, J., Strolcke, A., Jelinek, F. & Shriberg, L. : "Fast Training and Protability", Language Modeling Summer Research Workshop Technical Reports, CLSP Research Note No.1, Jhons Hopkins University, (1996)
- 7) Singer, Y : "What has been will be again", A Machine Learning Approach to the Analysis of Natural language. Ph.D Thesis, Hebrew University, (1995)