

複雑なテクスチャを持つ物体の輪郭線の抽出

姚 鳳会⁺ 邵 桂鳳⁺⁺ 玉木 明和⁺⁺⁺ 加藤 清史⁺ 山田 博章⁺

⁺東亜大学工学部 ⁺⁺西南学院大学商学部 ⁺⁺⁺九州工業大学工学部

あらまし 物体の輪郭線は、物体の抽出、認識、分類などの後処理においてとても重要である。室内風景のような複雑な背景において、単純のテクスチャを持つ対象物体の輪郭線の抽出についてはさまざまな方法が提案されている。しかし、この問題の逆問題、すなわち、均一な背景において複雑なテクスチャを持つ物体の抽出についての議論はまだ少ない。本稿では、我々はこの問題に焦点をあて、一つの抽出方法を提案する。本手法では、まず、入力画像からマスク画像と評価画像を生成する。次に、マスク画像を収縮させながら、輪郭線の検出と評価を行う。評価画像を使って、マスク画像の収縮処理の終了条件を与える。このように、複雑なテクスチャを有する物体の輪郭線を検出する。また、実画像を使って、本手法の有効性を確認した。本手法では、連続性、平滑性、正確性と高速性において、良好な結果が得られた。

Detecting Contours of Objects with Complicated Textures

Fenghui Yao⁺ Guifeng Shao⁺⁺ Akikazu Tamaki⁺⁺⁺ Kiyoshi Kato⁺ Hiroaki Yamada⁺

⁺Faculty of Engineering, University of East Asia

⁺⁺Faculty of Commercial Science, Seinan Gakuin University,

⁺⁺⁺Faculty of Engineering, Kyusyu Institute of Technology

ABSTRACT The object contours are very important in object extraction, object recognition, object classification and so on. There are a lot of literatures about extracting the contours of objects with simple texture from the complicated background such as indoor scene. However, there is few discussion about its opposite problem, that is, extracting the contours of objects with complicated textures from the uniform background. We focus on this problem and relate a method to solve this problem. The processing procedures in this method are as follows. Firstly, the mask image and evaluation image are generated from the input image. Then, the mask image is shrunk, and the contours are detected and evaluated after each shrinking operation. The evaluation image is used to determine termination of the shrinking operation. In this way, the contours of objects with complicated textures are detected, which are employed to extract the target objects. The effectiveness of this method is confirmed by using the real-world images. We obtained good results in continuousness, smoothness, preciseness and processing time.

1. はじめに

物体の輪郭線は、物体の抽出、認識、分類などの後処理においてとても重要である。室内風景のような複雑な背景において、簡単なテクスチャを持つ対象物体の輪郭線の抽出についてはさまざまな方法が提案されている¹⁾。しかし、この問題の逆問題、すなわち、均一な背景において複雑なテクスチャを持つ物体の抽出についての議論はまだ少ない。人間の顔の認識などのような後処理を考えると、背景はできるだけ簡単にすることは好ましく、背景を複雑にする必要はないと思われる。ここで、我々はこの逆問題に焦点をあて、一つの抽出方法を提案する。背景は均一であるので、単純な2値化処理で対象物を抽出できると思われがちである。しかし、対象物のテクスチャが複雑になると、この問題は困難になる。例えば、図1(a)に示したフルカラーの入力画像に対して、自動閾値選定法で2値化したものは同図(b)に、ガウシアン関数でぼかした後、自動閾値選定法で2値化したものは同図(c)に示す。この2つの図から分かるように、単純な2値化処理の後、正確な輪郭線を抽出するのは無理である。また、同図(d)にLoGオペレータを使って得られたエッジ画像である。同様に、このエッジ画像から輪郭線を検出するのは困難である。

近年、動的輪郭モデル Snakes が提案され²⁾、さまざまに応用されている。図1(a)の入力画像に対して、このモデルを使って得られた輪郭線は図1(e)に示す。この図から分かるように、Snakes は次のような欠点を持っている。(1) 得られた輪郭線は連続でない(4近傍³⁾ あるいは8近傍の意味での連続)。(2) 得られた輪郭線は変形された。(3) 得られた輪郭線は平滑でない。これ以外に、(4) この方法は非常に時間がかかる。物体の形状認識などの後処理にとっては、2番目の欠点は致命的である。

そこで、我々は物体の形状認識などの後処理を考慮して、複雑なテクスチャを有する物体の輪郭

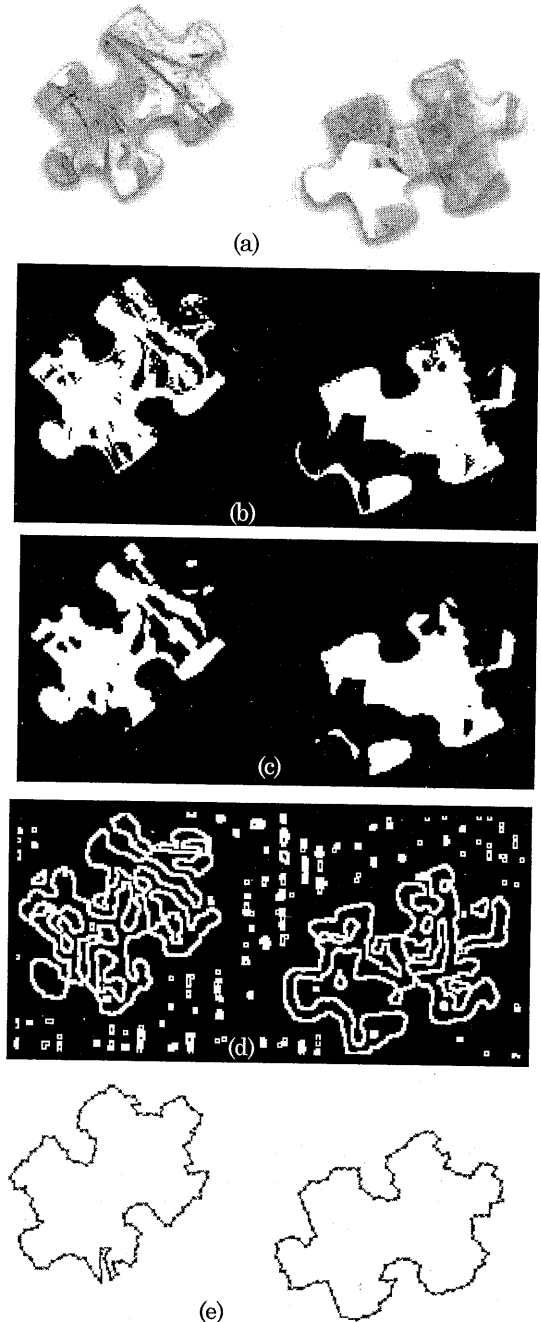


図1 (a)入力画像、(b)2値化画像、(c)ガウス関数でぼかした後の2値化画像、(d)LoGオペレータを使って得られたエッジ画像、(e)Snakes を使って得られた輪郭線

線の抽出法を提案する。本手法では、まず、視覚刺激を定義する。次に、入力画像からマスク画像と評価画像を生成する。マスク画像を収縮させながら、輪郭線を検出する。評価画像を使って、マスク画像の収縮処理の終了条件を与える。本手法の目標は

- ・ 連続（4近傍或いは8近傍の意味での連続）である、
- ・ 平滑である、
- ・ 変形されていない

輪郭線を求めることである。本研究の最終目的は複雑なテクスチャを有する対象物体の切り出しとその形状認識である。

2.1 マスク画像の生成

ここで、新たに視覚刺激を定義し、視覚刺激を検出することによって、マスク画像を生成する。

人間の視覚システムにおいては、目は対象物の何らかの特別なファクターをキャッチすることによって、対象物を検知している。言い換えれば、対象物の持っている何らかの特別なファクターが人間の視覚神経を刺激し、目の注意を引いている。この何らかの特別なファクターは視覚刺激と呼ばれている。原画像をガウス関数でぼかしてからラプラシアン演算を適用する方法は、人間の視覚に関する生理学的なものに対応がつかないといわれている。しかし、ガウス関数でぼかすと、対象物体の

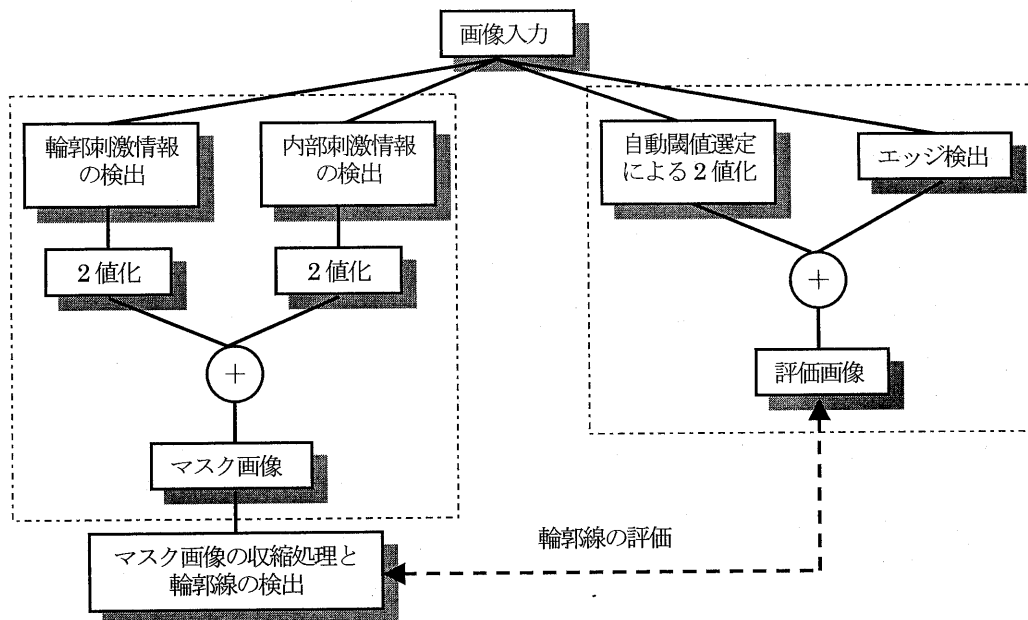


図2 本手法の処理の流れ

2. 本手法の処理の流れ

本手法の処理の流れを図2に示す。全体の処理はマスク画像生成部、評価画像生成部と輪郭線抽出評価部を含む。入力画像は、図1(a)に示すような背景は均一で、複数の複雑なテクスチャを有する物体を含むフルカラー画像とする。

エッジは鈍くなる。つまり、エッジの持っている視覚刺激は失われる。そこで、我々は新たに視覚刺激を定義する。入力画像を $I(x, y)$ ($x=1, 2, \dots, X; y=1, 2, \dots, Y$) とすると、点 (i, j) における視覚刺激は、この点だけの特徴でなく、その周りのある小さな範囲内の点の特徴にも左右される。この小さな範囲を点 (i, j) におけるローカルウィンドウと呼ぶことにする。物体領域は何らかの特徴を持って、背景領域と違うように見える。また、物体と

背景のエッジにおいて、輝度の変化が存在する(物体のテクスチャは複雑なので、エッジにおいては、必ず急な輝度変化があるとは言えない)。ここで、点 (i, j) においてのローカルウィンドウ内部の統計特徴量を点 (i, j) の視覚刺激と定義する。具体的に、点 (i, j) においてのローカルウィンドウ内部の平均値と分散は点 (i, j) の視覚刺激と定義し、次のように表す。

$$S_{int} = \frac{1}{M \times N} \sum_{j=0}^N \sum_{i=0}^M f(i, j) \quad (1)$$

$$S_{con} = \sum_{j=0}^N \sum_{i=0}^M (f(i, j) - S_{int})^2 \quad (2)$$

ここで、 M と N はローカルウィンドウのサイズで、 $f(i, j)$ は点 (i, j) の R、G、B 値、或いは Y、I、Q 値、または D、S、H 値である。式(1)は、物体領域の視覚刺激をよく検出できるが、これはガウス関数と同じく、エッジにおいて視覚刺激は失われる。 S_{in} は内部視覚刺激と呼ぶことにする。式(2)は、輪郭とエッジにおいて大きな値が取る。 S_{con} は輪郭視覚刺激と呼ぶことにする。式(1)と式(2)を全体の入力画像に適用すると、内部視覚刺激マップと輪郭視覚刺激マップが得られる。

輪郭視覚刺激は次のように2値化する。

$$S_{con}^B = \begin{cases} 0 & \text{if } S_{con} > k_1 \times T_{con} \\ 1 & \text{others} \end{cases} \quad (3)$$

ここで、 T_{con} は輪郭視覚刺激の平均値で、 k_1 は重み係数で、 S_{con}^B は2値化した輪郭視覚刺激である。

内部視覚刺激は次のように2値化する。

$$S_{int}^B = \begin{cases} 1 & \text{if } S_{int} > k_2 \times T_{int} \\ 0 & \text{others} \end{cases} \quad (4)$$

ここで、 T_{in} は、 S_{con}^B の値が0である点に対応している点の内部視覚刺激の平均値で、 k_2 は重み係数で、 S_{in}^B は2値化した内部視覚刺激である。

S_{con}^B と S_{in}^B の論理和はマスク画像 F_{mask} と定義する。すなわち、

$$F_{mask} = S_{con}^B \oplus S_{in}^B \quad (5)$$

2.2 評価画像の生成

評価画像を生成するために、できるだけ輪郭情報を失わないようにする必要がある。具体的には、評価画像 F_{eval} は2値化された濃淡画像と2値化されたエッジ画像の論理和と定義する。すなわち、

$$F_{eval} = F^B \oplus F_{edge}^B \quad (6)$$

ここで、 F^B は自動閾値選定法で入力画像 $I(x, y)$ を2値化した画像で、 F_{edge}^B は同様な手法で2値化したエッジ画像である。

エッジ画像を生成するには、Sobel オペレータ、Prewitt オペレータ、Roberts オペレータ、LoG オペレータなどの手法がある。得られたエッジの質、ノイズ、処理時間を考慮して、ここではRoberts オペレータを使用する。

2.3 輪郭線の抽出と評価

上に述べた方法で生成したマスク画像と評価画像を使って、輪郭線の抽出と評価を行う。

マスク画像に対して、全ての輪郭線を抽出する。得られた輪郭線に対して次の処理を行う。

(1) 穴埋め処理

内側の輪郭線はマスク画像上の穴に対応しているので、その領域を白画素で塗りつぶす。この輪郭線を削除する。

(2) ノイズ除去

輪郭線の長さが予め決めた閾値より短いとき、この輪郭線で囲まれた領域は孤立ノイズと見なし、削除する。

このような処理を行った後、残った輪郭線はそれぞれ一つの対象物体に対応している。

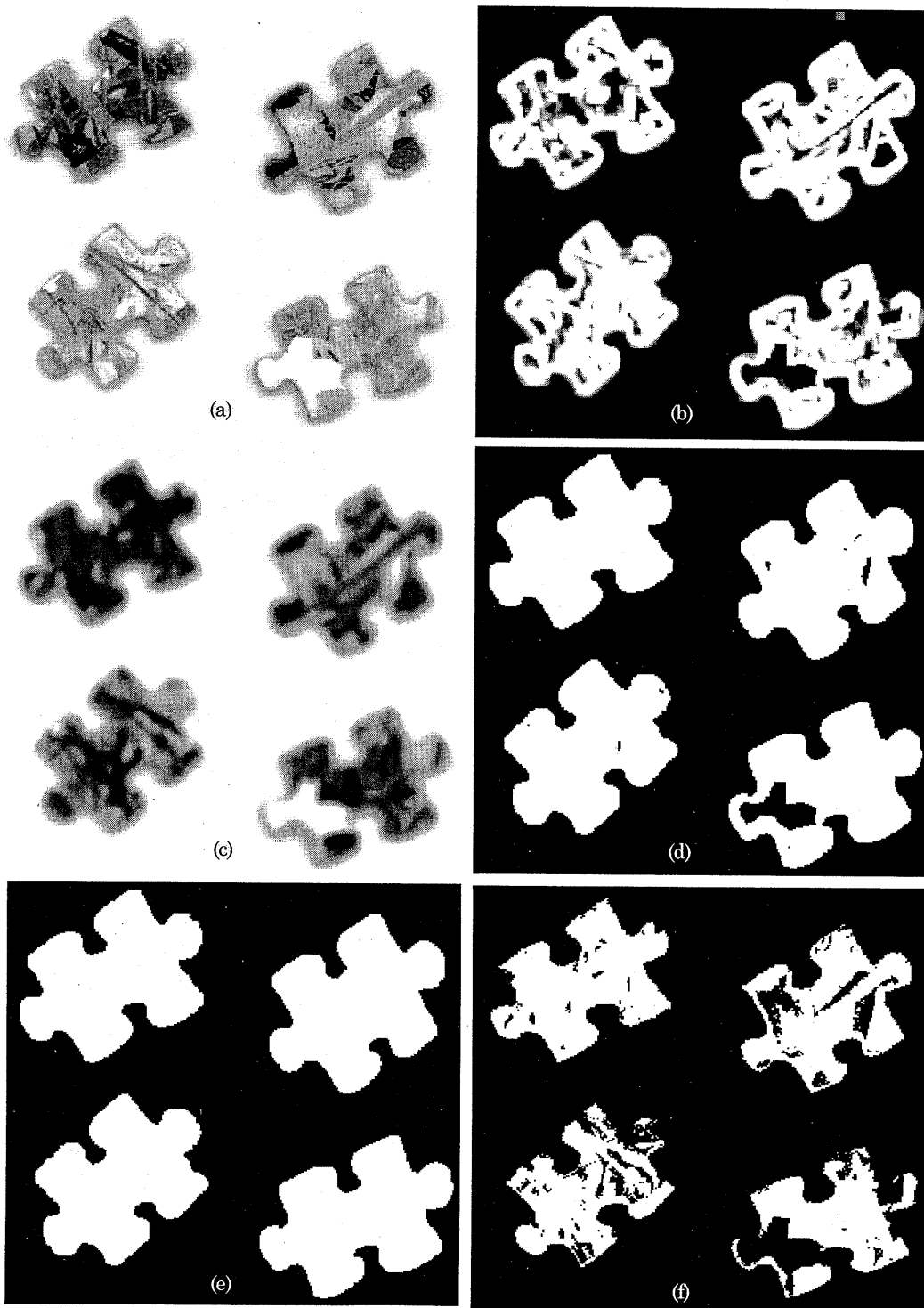


図3 (a)入力画像、(b)輪郭視覚刺激、(c)内部視覚刺激、(d)マスク画像、
 (e)穴埋め処理後のマスク画像、(f)評価画像

次に、各輪郭線と評価画像の一致度 η を計算する。 η は次のように定義する。

$$\eta = \frac{N_{eval}}{N_{total}} \quad (7)$$

ここで、 N_{total} は輪郭線上の全ての点の数で、 N_{eval} は N_{total} の中の評価画像に含まれた点の数である。

η は閾値 η_{thres} より小さいとき、マスク画像を収縮させてから、再度 η を計算する。この処理は、 η が η_{thres} を超えるまで繰り返される。

3. 実験結果

実験画像はイメージスキャナで作成したフルカラー画像である。視覚刺激を求めるには、画像の R、G、B 値を使用した。実験は 133MHz の Pentium の DOS/V 互換機上で行われた。OS は Windows NT Workstation 4.0 で、使用言語は Microsoft Visual C++ V.4.1 である。

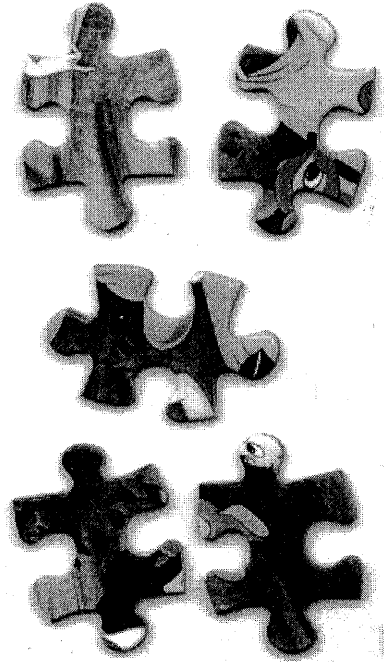


図4 入力画像

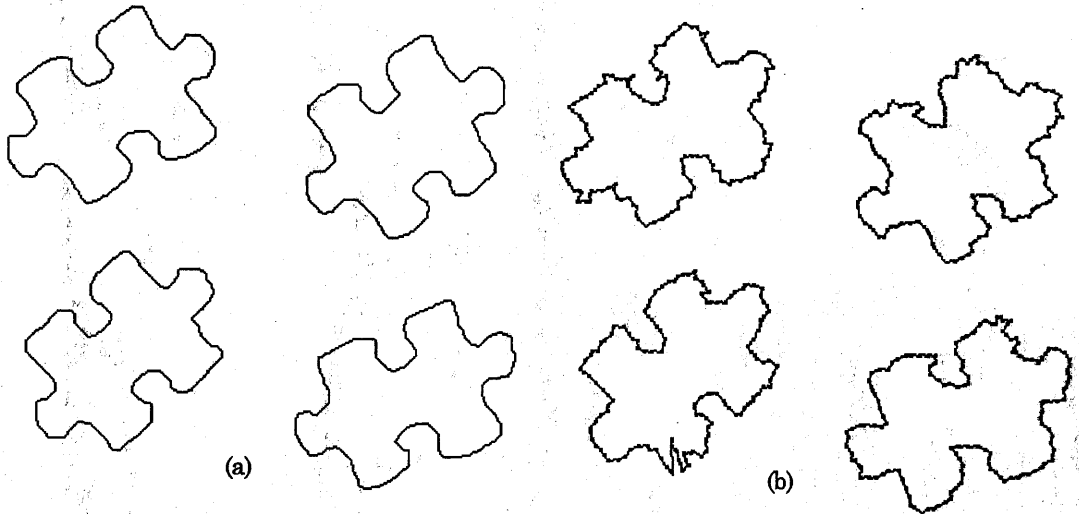


図5 (a)本手法で得られた輪郭線、(b)Snakes で得られた輪郭線

最初の実験の入力画像は図3(a)に示す。この画像の中には、4つの複雑なテクスチャを有する物体が存在する。検出された輪郭視覚刺激と内部視覚刺激は、それぞれ同図(b)と(c)に示す。この二つの視覚刺激を2値化されてから、論理和を行い、同図(d)のようなマスク画像が得られる。穴埋め処理を行った後のマスク画像が同図(e)に示す。

Roberts オペレータで検出したエッジ画像と濃淡画像をそれぞれ2値化し、その論理和画像(評価画像)は同図(f)に示す。得られた輪郭線は図5(a)に示す。

次の実験の入力画像は図4に示す。この画像の中には、5つの複雑なテクスチャを有する物体が存在する。得られた輪郭線は図6(a)に示す。

4. 本手法の評価

本手法の評価として、その処理結果と Snakes の処理結果との比較を行った。Snakes は動物体の追跡などに幅広く応用できるが、本手法はできない。この点を除いて、以下の項目を比較する。

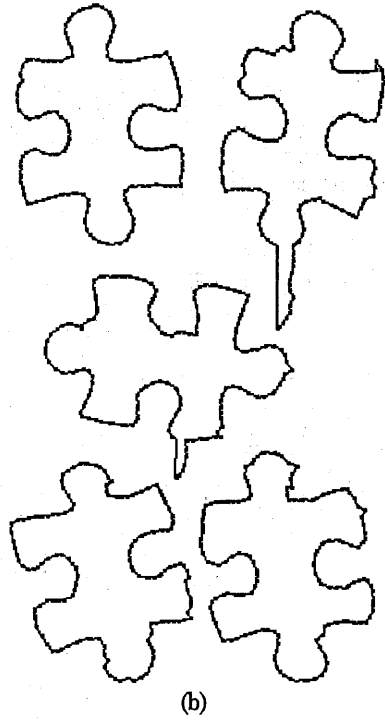
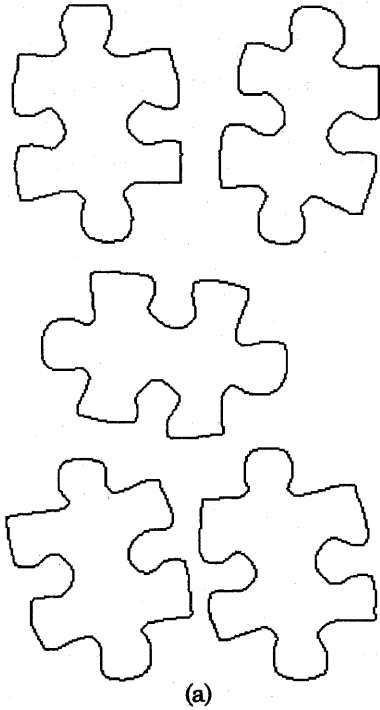


図6 (a)本手法で得られた輪郭線、(b)Snakes で得られた輪郭線

また、比較しやすくするために、図3(a)と図4(a)に示す入力画像に対して、Snakes を使った処理結果はそれぞれ図5(b)と図6(b)に示す。黒点は Snakes の制御点を示す。

(1) 連続性

図5と図6から分かるように、本手法で得られた輪郭線は連続である(8近傍の意味での連続)。しかし、Snakes を使って得られた輪郭線は連続でない(図5(b)と図6(b)の黒点で示す)。Snakes で連続な輪郭を求めようとする場合は、初期状態では、Snakes の制御点間の距離を1にする必要がある。この場合、隣り合う制御点同士のエネルギーは互いに大きく影響し合っ、エネルギー関

数は発散する場合は多い。また、たとえエネルギー関数が収束できるとしても、莫大な時間がかかる。

(2) 正確性

本手法では、Snakes の結果と比較すると、正

確の輪郭線が得られた。図5(b)と図6(b)の Snakes の結果を見るとわかるように、Snakes で得られた輪郭線は、正確に言えば、変形されている。また、図6(b)の右上と真ん中の物体の輪郭線は大きく変形されている。これは、収束とは言えず、発散している。すなわち、入力画像に複数の対象物が含まれるとき、同じ条件の下で、Snakes はそれぞれの対象物に収束すると限らない。言い換えれば、複数の対象物に対して、Snakes はそれぞれの対象物に収束することは保証できない。

(3) 平滑性

本手法で得られた輪郭線は滑らかである。これに対して、Snakes で得られた輪郭線はぎざぎざ

である。これは対象物体のテクスチャに起因している。対象物体のテクスチャが複雑なので、Snakesの制御点は強いエッジの近くに吸引される。

以上の(1)~(3)より、Snakesで得られた輪郭線は、そのままでは物体の形状認識などの後処理に使えない。

表1 本手法とSnakesの処理時間 単位: mS

		処理内容	実験 I	実験 II
本手法	マスク画像の生成	視覚刺激の検出	51544	49611
		視覚刺激マップの生成	511	170
		視覚刺激の2値化	220	430
		マスク画像生成	60	70
		マスク画像の輪郭線抽出	60	70
		マスク画像の穴埋め処理	3225	3295
	評価画像の生成	濃淡画像の生成	160	160
		エッジ画像の生成	351	391
		エッジ画像の2値化	711	761
		濃淡画像の2値化	741	771
		評価画像の生成	110	150
	輪郭線の抽出と評価		431	370
合計		58124	56249	
Snakes		186068	286332	

(4) 処理時間

実験 I と II において、本手法と Snakes の処理時間を表 1 にまとめた。Snakes の場合、初期状態の制御点間の距離を 6 と、エネルギー最小化の繰り返し処理の回数の上限を 60 と設定する。エネルギーを最小化するには、Greedy アルゴリズム⁹を使用した。このアルゴリズムの処理時間は $O(n \times m)$ である。ここで、 n は制御点の数、 m は制御点の近傍点の数である。この場合、Snakes の処理時間は本手法の 3~5 倍である。表 1 から分かるように、視覚刺激検出は本手法の処理時間の大部分を占めている。他の処理の時間は省略し、視覚刺激検出の処理だけを考えると、本手法の処理時間は $O(K \times L \times M \times N \times T_A)$ である。ここで、 K と L は入力画像のサイズで、 M と N はローカルウィンドウのサイズで、 T_A は一回の足し算、引き算と掛け算の合計時間である。実験 I と II においては、 M と N の値は 3 とした。

5. むすび

本稿では、複雑なテクスチャを有する物体の輪郭線の抽出について、一方法を提案した。本手法では、連続性、正確性、平滑性、高速性において良好な結果が得られた。

物体は複雑なテクスチャを持っている場合、エッジと輪郭においては、輝度は弱変化から急な変化までさまざまである。弱変化でも視覚刺激として大きく検出できるように、式(1)を使用した。また、物体内部の変化も視覚刺激として大きく検出できるように、式(2)を使用した。実験を通して、式(1)と式(2)は有効に視覚刺激を検出できることが確認された。また、視覚刺激を 2 値化するには、式(3)と式(4)を使用した。これは有効であることが確認できた。

重み係数 k_1 , k_2 と一致度の閾値 η_{thres} はそれぞれ実験的に 0.45, 0.9 と 0.3 と決めた。実験結果から、これらの値は妥当であることは分かる。

今後、本手法の結果を複雑なテクスチャを有する物体の形状認識に応用する予定である。

謝辞 本研究の一部は平成 10 年度 (財) 柏森情報科学振興財団の助成を受けた。記して謝意を表す。

参考文献

- 1) 福井: "領域間の分離度に基づく物体輪郭抽出," 信学論(D-II), Vol. J80, No.6, pp.1406-1414, Jun. 1997.
- 2) 荒木: "複数物体の抽出を目的とした交差判定により分裂する動的輪郭モデル," 信学論(D-II), Vol. J79, No.10, pp.1704-1711, Oct. 1997.
- 3) Kass, M. K., Witkin, A. and Terzopoulos, D: Snakes: Active Contour Models, Proc. 1st ICCV, pp. 259-268, 1987.
- 4) Williams, D. J. and Shah, M.: A Fast Algorithm for Active Contours, Proc. 3rd ICCV, pp.592-595, 1990.