画像セグメンテーションにおけるウェーブレット 係数の局所テクスチャ特徴量を用いた GraphCuts

福田 恵太[†] 滝口 哲也^{††} 有木 康雄^{††}

†神戸大学大学院工学研究科 〒 657-8501 兵庫県神戸市灘区六甲台町 1-1
 †† 神戸大学工学部 〒 657-8501 兵庫県神戸市灘区六甲台町 1-1
 E-mail: †fukuda@me.cs.scitec.kobe-u.ac.jp, ††{takigu,ariki}@kobe-u.ac.jp

あらまし 本稿では,ウェーブレット係数から局所テクスチャ特徴量を定義した Graph Cuts による画像セグメンテー ション手法を提案する.提案手法では,Haar ウェーブレット変換による多重解像度解析を行い,低周波数領域から平 滑化処理画像を得て n-link に用い,さらに高周波数領域からテクスチャ特徴量を定義し,色分布と共に t-link に用い て繰り返し Graph Cuts セグメンテーションを行う.これにより,複雑なエッジへの対応と共に,物体領域と背景領 域に類似色を含む画像やテクスチャ変化の強い画像に対してもセグメンテーションが可能となる.評価実験より,提 案手法による画像セグメンテーションの誤検出率は1.57% となり,従来手法の Graph Cuts 2.42% に比べ,誤検出率 を 35.1% 削減することができた.

キーワード 画像セグメンテーション, Graph Cuts, 多重解像度解析, 局所テクスチャ特徴量

Graph Cuts by using Local Texture Features of Wavelet Coefficient for Image Segmentation

Keita FUKUDA[†], Tetsuya TAKIGUCHI^{††}, and Yasuo ARIKI^{††}

† Graduate School of Engineering, Kobe University, Rokkodaicho 1-1, Nada-ku, Kobe, Hyogo 657-8501 Japan
 †† Faculty of Engineering, Kobe University, Rokkodaicho 1-1, Nada-ku, Kobe, Hyogo 657-8501 Japan
 E-mail: †fukuda@me.cs.scitec.kobe-u.ac.jp, ††{takigu,ariki}@kobe-u.ac.jp

Abstract This paper proposes a approach to image segmentation using Graph Cuts based on local texture features of wavelet coefficient. Using Multiresolution Analysis based on Haar Wavelet, low-frequency range (scaling image) is used for n-link and high-frequency range (local texture features) is used for t-link with color histogram. Iterated Graph Cuts, the propose method can segment the region of object for images with a lot of edge, with colors similar to foreground and background, and with heavy texture change. Experimental results showed our method (Err:1.57%) reduced segmentation error rates by approximately 35.1% compared to previous methods (Err:2.42%). **Key words** Image Segmentation, Graph Cuts, Multiresolution Analysis, Local Texture Feature

1. はじめに

画像処理に関する重要な問題の一つとして画像のセグメン テーションがある.セグメンテーションとは,一枚の画像から 対象となる領域を抽出することであり,一般物体認識や画像合 成など画像処理技術の前処理として利用されている.

近年,セグメンテーション問題をエネルギー最小化問題と捉 えて解を求める手法が多く提案されている。その一つの手法 として Graph Cuts によるセグメンテーションが挙げられる. 同じくエネルギー最小化問題から解く Snakes [1] や Level Set Method [2] では境界線に対してのエネルギー関数から局所最小 解を求めるのに対し Graph Cuts では,各領域からエネルギー 関数を定義し,大域最小解を求めることが可能で,領域と境 界の両方の特性を用いたセグメンテーションが実現する.また Graph Cuts はセグメンテーションだけでなく,画像復元,動 画解析など幅広くのコンピュータビジョン分野に応用されてい る手法 [8] である.

Graph Cuts を用いたセグメンテーション手法として, Boykov らにより Interactive Graph Cuts [3] [4] が提案されている. Interactive Graph Cuts では,ユーザーが画像に対して物体と背 景にそれぞれラベルを与え,画像からノードとエッジを持つグ ラフを作成し,minimum cut/max flow algorithm [6] で定義 したエネルギー関数の最小化を行う.また Rother らは,矩形 領域により物体領域と背景領域を指示することでラベルを与え, セグメンテーション結果を再学習し Graph Cuts を繰り返し行 う Grab Cuts [5] を提案している.

この Graph Cuts には、ピクセル間の輝度値から計算され る n-link の影響により、局所的なエッジを乗り越えることが 困難である問題があった.この問題を解決する手法として、永 橋らが平滑化処理の繰り返しによる画像セグメンテーションの Graph Cuts [9] を提案している.さらに学習された色分布を基 とした物体と背景の尤度から計算される t-link では、物体領域 と背景領域に類似色を持つ画像では、色分布からのセグメン テーションが困難であるという問題がある.

そこで本研究では,多重解像度解析を行い,低周波数領域から平滑化処理画像を得て n-link に,さらに高周波数領域から局所テクスチャ特徴量を定義し,色分布と共に t-link に用いて Graph Cuts セグメンテーションを繰り返し行う手法を提案する.

提案手法により,n-link に対しては多重解像度レベルが高い 低解像度画像(低周波数成分)から解像度を上げていくことで 平滑化の繰り返し処理[9]と同じく,大域的なセグメンテーショ ンから局所的なセグメンテーションを可能とし,また t-link に 対してはそのレベルでの高周波成分から局所テクスチャ特徴を 定義することで,色分布のみに依存しないセグメンテーション が可能になる.

2. Interactive Graph Cuts

本章では従来手法である Interactive Graph Cuts [3] につい て述べる.

2.1 Graph Cuts のエネルギー関数

画像 P に対する各ピクセルを $p \in P$ としたとき, ラベル を $A = (A_1, ..., A_p, ..., A_{|p|})$ と二値ベクトルとし, 各 A_p に は物体 ("obj") か背景 ("bkg") のラベルが与えられる.また p の近傍のピクセルを $q \in N$ と定義する. Graph Cuts では MAP-MRF [7] と同様に,エネルギー関数を式 (1) のように定 義する.

$$E(A) = \lambda \cdot R(A) + B(A) \tag{1}$$

 $\lambda(\geq 0)$ は,領域に対するペナルティ関数 R(A)と,境界に対す るペナルティ関数 B(A)の比率を表す.またそれぞれのペナル ティ関数は式 (2) (3)のように定義する.

$$R(A) = \sum_{p \in P} R_p(A_p) \tag{2}$$

$$B(A) = \sum_{\{p,q\} \in N} B_{\{p,q\}} \cdot \delta(A_p, A_q)$$
(3)

and

$$\delta(A_p, A_q) = \begin{cases} 1 & (ifA_p \neq A_q) \\ 0 & (otherwise) \end{cases}$$
(4)

 $R(\cdot)$ はピクセルpが物体もしくは背景のモデルにどれだけ適合

しているかを示す関数である . $B_{\{p,q\}}$ は近傍のピクセルとの関係を示し , $p \ge q$ の輝度値が似ていれば大きな値を出力する関数である . 以上によりエネルギー関数 E(A)を最小とするラベル Aを求めることで , セグメンテーションを行う .

2.2 グラフ作成とカット

画像からグラフを図 1 のように作成する . グラフは有向グ ラフであり,画像の各ピクセルに対応したノードと, sink と source と呼ばれるターミナルを用意する . 各ピクセルに対応 したノード間を結ぶエッジを n-link と呼ぶ . また各ノードか らターミナル (sink(T) と souce(S)) を結ぶエッジを t-link と 呼ぶ .



図 1 グラフの作成

次に,全てのエッジに対して重みを与える.n-link とt-link のエッジに与える重みは表1に従う.Oは物体,Bは背景を意味し, $p \in O, p \in B$ はユーザーが予め物体もしくは背景とラベルを与えたピクセルのことである.以下ラベルを与えられた ピクセルのことを,seed と呼ぶ.

エッジの重み 表 1 edge $\cos t$ for n-link $\{p,q\}\in N$ $\{p,q\}$ $B_{\{p,q\}}$ $\lambda \cdot R_p("bkg")$ $p \in P, p \notin O \cup B$ $\{p, S\}$ K $p\in O$ t-link 0 $p \in B$ $\lambda \cdot R_p("obj")$ $p \in P, p \notin O \cup B$ $\{p, T\}$ 0 $p\in O$ K $p\in B$

表 1 の変数を以下の式 (5), (6), (7) で定義する. C_p はピク セル p での RGB 値, I_p は輝度値である.また dist(p,q) は近 傍との距離を表す.

$$B_{\{p,q\}} \propto exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \cdot \frac{1}{dist(p,q)}$$
(6)

$$K = 1 + \max_{p \in P} \sum_{q:\{p,q\} \in N} B_{\{p,q\}}$$
(7)

-2 -



図 2 提案手法の流れ

作成したグラフに対して,最小カットを求める.式(8)は, カット C が通る S 側から T 側へ向かうエッジの重み w_e の和 を表し,この最小化問題を解く.なお式(1)におけるラベル Aの配置エネルギーとそれに対応するカットのコストは等価[3] である.

$$\mid C \mid = \sum_{e \in C} w_e \tag{8}$$

min cut/max flow algorithm [6] を用いることでこの問題を解 き,物体と背景のグラフを分割する.カットの後に,Sと繋がっ ているピクセルは物体,同様にTと繋がっているピクセルは背 景とすることで,画像のセグメンテーションが実現する.

2.3 問題点と提案手法

t-linkの影響が強くなると色分布に依存が強くなる.そのため突発的な誤検出が生じる.また n-linkの影響が強くなると エッジ情報に依存が強くなる.そのため複雑なエッジが存在する場合,局所的なエッジを乗り越えることが困難である.よって比率λの調整が問題となっていた.さらに物体領域と背景領域に類似色を持つ画像では,色分布に依存しているt-linkが上手く機能しないという問題がある.

一つ目の解決策として永橋らにより提案された平滑化処理の 繰り返しによる Graph Cuts [9] で,この問題が解決した.二つ 目の解決策として色分布のみではなく局所テクスチャ特徴量を 定義することで,色分布のみに依存しないセグメンテーション を実現する.

これら二つの問題点を解決するため,本研究では多重解像度 解析を用いることで,平滑化と局所テクスチャ特徴を並列に処 理する画像セグメンテーション手法を提案する.

3. 提案手法

本章では多重解像度解析を用いることで,平滑化と局所テク スチャ特徴を並列に処理する画像セグメンテーション手法につ いて述べる.

3.1 提案手法の流れ

図 2 に,提案手法の流れを示す.はじめに,入力画像に対し てユーザは物体(赤線)と背景(青線)の seedの入力を行う.次 に画像をレベル n まで多重解像度解析を行い,低周波成分であ る平滑化画像は n-link に,また高周波成分から局所テクスチャ 特徴量を定義し,色特徴量と共にGMM に適用し,物体及び背 景の尤度関数を求める.1つ上の多重解像度レベルにおける処 理結果画像に距離変換を行い事前確率を推定し,事前確率と尤 度関数から求めた事後確率を t-link に用いる.Graph Cuts に よるセグメンテーションの後に,多重解像度レベルを1つ下の レベルにする.すなわち初期値n = 3のとき,レベル3のとき の処理結果をレベル2でのセグメンテーションに用い,次のレ ベル1のセグメンテーションでは,レベル2の処理結果を用い る.この処理をn = 0になるまで繰り返し行う.

3.2 Haar ウェーブレット変換による多重解像度解析 本研究においては, Haar 関数のウェーブレット変換を用い て多重解像度解析を行った.そのスケーリング関数とウェーブ レット関数を式 (9)(10) に示す.

$$\Phi_H(x) = \begin{cases} 1 & if \ 0 \le x < 1 \\ 0 & otherwise \end{cases}$$
(9)

$$\Psi_{H}(x) = \begin{cases} 1 & if \ 0 \le x < \frac{1}{2} \\ -1 & if \ \frac{1}{2} \le x < 1 \\ 0 & otherwise \end{cases}$$
(10)

また分解アルゴリズムは,一回の変換をレベル1とし,その階 層レベルを n とすると,

$$c_k^{(n)} = \frac{1}{2} (c_{2k}^{(n-1)} + c_{2k+1}^{(n-1)})$$
(11)

$$d_k^{(n)} = \frac{1}{2} (c_{2k}^{(n-1)} - c_{2k+1}^{(n-1)})$$
(12)

-3-

のようになり,式(11)から低周波数成分,式(12)から高周波 数成分が得られる.



図 3 多重解像度解析

画像データに対して, Haar ウェーブレットを適用させる.式 (11) と式(12) により横方向と縦方向にダウンサンプリングを 行うと,図3の左のように低周波数成分(LL)と縦方向,横方 向,斜め方向の高周波数成分(HL,LH,HH)が得られる.この 処理を再帰的にLLに対して行うことで,多重解像度解析を行 う.図3はレベル1からレベル2での分解を行った例であり, 図4は実際に画像に適応させた結果である.



図 4 実画像に対する多重解像度解析

3.3 平滑化処理 (n-link)

多重解像度解析で得られたレベル n のときの低周波数成分 (LLn),すなわち平滑化画像を n-link に用いる.セグメンテー ション終了後に1つ下の多重解像度レベルにする (n-1) こと で平滑化の度合いが小さくなっていくことから,大域的なセグ メンテーションから局所的なセグメンテーションを実現する.

3.4 局所テクスチャ特徴量 (t-link)

前節と同じく多重解像度解析で得られたレベル n のときの各 高周波数成分 (LHn,HLn,HHn) からピクセル p に対して局所 テクスチャ特徴量 T_p を定義し,色特徴量 C_p と共に t-link に 用いる.局所テクスチャ特徴量は,式(13)のように各成分で注 目ピクセル p 周辺(3×3)のウェーブレット係数 $d^{(n)}$ から絶 対値の平均値を定義する[10].

$$T_p = \frac{1}{9} \sum_{p,q \in N} |d^{(n)}| \qquad for \ LHn, HLn, HHn \qquad (13)$$

図 5 ではそれぞれの成分で得られた Level 1 (n = 1)の局所 テクスチャ特徴量を画像として表示した.テクスチャが複雑な 領域では特徴量の値が大きく,また平坦な領域では特徴量の値 は小さくなるという傾向がある.



図 5 局所テクスチャ特徴量

3.5 事後確率の推定 (t-link)

3 次元の RGB 色特徴量を C_p とし,前節で定義した LH,HL,HH の局所テクスチャ特徴量 T_p と合わせることで 6 次元の $Y_p = \{C_p, T_p\}$ を定義する. Y_p を用いて式 (5) で与え られたエッジ重みをここでは事後確率として以下のように定義 する.

$$\begin{cases} R_p("obj") = -lnPr(O \mid Y_p) \\ R_p("bkg") = -lnPr(B \mid Y_p) \end{cases}$$
(14)

事後確率はベイズの定理より,式 (15)(16)のように尤度と事前 確率から求めることができる.ただし $Pr(Y_p)$ はどちらの式に も共通の項なので省略し,式 (14)を求めた.

$$Pr(O \mid Y_p) = \frac{Pr(Y_p \mid O)Pr(O)}{Pr(Y_p)}$$
(15)

$$Pr(B \mid Y_p) = \frac{Pr(Y_p \mid B)Pr(B)}{Pr(Y_p)}$$
(16)

尤度 $Pr(Y_p \mid O)$, $Pr(Y_p \mid B)$ は RGB 色特徴量 C_p と前節で 定義した局所テクスチャ特徴量 T_p から求め, 事前確率 Pr(O), Pr(B) は一つ上の多重解像度レベルのセグメンテーション結果 から距離変換を用いて求める.物体領域の情報から事後確率を 推定する流れを図 6 に示す.なお背景側も同様に背景領域情報 から事後確率を推定する.



図 6 事後確率の推定

3.5.1 尤 度

尤度 *Pr*(*Y_p* | *O*), *Pr*(*Y_p* | *B*) はユーザが予め seed としたピ クセル集合を基に,GMM (Gaussian Mixture Model)を用い て求める.RGB の色特徴量 3 次元と高周波数成分 LH,HL,HH の局所テクスチャ特徴量 3 次元で合計 6 次元となる.よって



図7 セグメンテーション結果と誤検出率

式 (18) は, d = 6 となる.またモデルパラメーターの推定 ($\alpha_i, \mu_i, \Sigma_i$)には, EM algorithm [12] を用いた.式 (17) により Y_p の尤度を求める.また GMM の混合分布数 K は [5] と同様 に K = 5 とした.

$$Pr(Y_p \mid \bullet) = \sum_{i=1}^{K} \alpha_i p_i(Y_p \mid \mu_i, \Sigma_i)$$
(17)

$$p(Y_p \mid \mu, \Sigma) = \frac{1}{(2\pi)^{d/2} \mid \Sigma \mid^{1/2}}$$

$$\cdot exp\left(-\frac{1}{2}(Y_p - \mu)^T \Sigma^{-1}(Y_p - \mu)\right)$$
(18)

3.5.2 事前確率

1つ上の多重解像度レベルの GraphCuts によるセグメンテー ション結果から,事前確率 Pr(O), Pr(B) を求める.得られた 物体と背景の境界付近では,物体と背景が入れ替わる可能性が 高く,反対に境界から離れている位置では物体と背景が入れ替 わる可能性は低いと考えられる.そこで物体領域と背景領域に 対して位置変換し,境界からの距離 $d \ge 0.5$ から 1 に正規化し, 物体と背景の距離 d_{obj} , d_{bkg} を用いて事前確率を以下のように 定義する.

$$Pr(O) = \begin{cases} d_{obj} & \text{if } d_{obj} \leq d_{bkg} \\ 1 - d_{obj} & \text{otherwise} \end{cases}$$
(19)

$$Pr(B) = 1 - Pr(O) \tag{20}$$

GMM により得られた尤度と距離変換により得られた事前確 率で物体と背景それぞれに対して事後確率を求める.式 (14) により、 $\{p,S\}$ と $\{p,T\}$ 間の t-link のエッジコストを計算し, Graph Cuts によるセグメンテーションを行う. セグメンテー ションが終了すると、多重解像度レベル n を一つ下 (n-1) に 設定し,再び 3.2 節以降の処理を繰り返す.

4. 実 験

4.1 実験概要

一枚の静止画像に対して,同じseedを与えて従来手法と提案 手法の比較実験を行った.従来手法である Interactive Graph Cuts (以下,従来手法(1)),平滑化処理の繰り返しを利用し た Graph Cuts (以下,従来手法(2))と提案手法である多重 解像度解析により平滑化処理と局所テクスチャ特徴量を利用 した Graph Cuts を比較した.評価用の画像は Grab Cuts Database [13] により提供されている 50 枚の画像(人,動物, 物など)を使用する.予め用意されているマスク画像(Human Segmentation Image)を正解とし,各手法のセグメンテーショ ン結果の誤検出率を求めた.誤検出率は以下の式から求める.

誤検出率 [%] =
$$\left(\frac{物体領域の誤検出ピクセル数}{ \pounds ピクセル数} + \frac{背景領域の誤検出ピクセル数}{ \pounds ピクセル数} \right) \times 100$$
 (21)

なおパラメータは , いずれの手法も表 1 の $\lambda=0.05$, 式 (6) の $\sigma=12.0$ とした . また多重解像度のレベルは初期値を n=3 とした .

4.2 実験結果

50 枚の画像に対するセグメンテーション結果を表 2 に示す.

従来手法 (1)	従来手法 (2)	提案手法
$4.05 \ \%$	2.42~%	1.57~%
表 2 誤 検 出 率		

表2より提案手法と従来手法で誤検出率に大きな差は表れて いないが,これは従来手法のGraph Cutsでも十分なセグメン テーション結果が得られた画像を多く含んでいたためである. また,図7より誤検出率が2,3%ほどの差でもセグメンテー ション結果には大きく表れる.従来手法と提案手法において, セグメンテーション結果に差が表れたものを図7に提示する.

4.3 考 察

物体領域と背景領域に類似色を持つ画像やテクスチャ変化の 強い画像に対して,従来のGraph Cutsより高精度な結果が得 られ,また従来のGraph Cutsで十分なセグメンテーション結 果が得られた画像に対しても同精度のセグメンテーション結果 が得られた.

従来手法(1)に比べ,従来手法(2),提案手法では突発的な ノイズが少ない.これは一つ上の多重解像度レベルの処理結果 から事前確率が与えられているため,大まかな物体領域と背景 領域をもつ t-link が得られていることから,突発的なノイズが 除外されていると考えられる.事前確率の影響で,早い段階で 生じた突発的なノイズが後の段階で拡大することも考えられる が,初めの段階では平滑化処理により突発的なノイズが生じに くくなっているため,そのような問題が解決されている.

次に,平滑化と同時に局所テクスチャ特徴を定義している提 案手法では従来手法(1),(2)と比べ,局所テクスチャ特徴量を t-link で定義しているので,色分布のみで物体領域と背景領域 の区別が困難な領域に対してセグメンテーションが可能である. 図7では,本や人の例では物体領域と背景領域ではテクスチャ 変化が大きいため,物体領域と背景領域でしテクスチャ 変化が大きいため,物体領域と背景領域との色変化が小さい境 界も分かれている.しかし,車の例ではテクスチャ変化が少な く色変化が大きいため,効果が表れなかった点が問題である. 画像によって有効な特徴量を考慮するため,色分布もしくは局 所テクスチャ分布のどちらがはっきり分かれているか判別する ことが必要であると考えている.今後,複雑でセグメンテー ションが困難な画像データを対象とすることで有効性をさらに 検討する必要がある.

5. ま と め

本稿では,多重解像度解析を行い,低周波数領域から平滑化 処理画像を得て n-link に,さらに高周波数領域から局所テク スチャ特徴量を定義し,色分布と共に t-link に用いて繰り返し Graph Cuts セグメンテーションを行う手法を提案した.物体 領域と背景領域に類似色を持つ画像やテクスチャ変化の強い画 像に対して,従来の Graph Cuts より高精度な結果が得られた. さらに複雑なデータを対象として有効性を検討する必要がある. また今後は t-link に有効な特徴量を考慮し,また t-link に特定 の物体モデル特徴を含めることで自動セグメンテーションなど を検討する予定である.

献

 M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models, "International Journal of Computer Vision, Vol. 1, No. 4, pp. 321-331, 1988.

文

- [2] James A. Sethian. "Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry Fluid Mechanics", Computer Vision, and Materials Science. Cambridge University Press, 1999.
- [3] Y. Boykov, M.P.Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images", In IEEE International Conference on Computer Vision and

Pattern Recognition, volume 2, pages 731-738, 2004.

- [4] Y.Boykov, G.Funka-Lea, "Graph Cuts and Efficient N-D Image Segmentation.", International Journal of Computer Vision, 70(2):109-131, 2006.
- [5] C. Rother, V. Kolmogorov, A. Blake, "GrabCut: Interactive Foreground Extraction using Iterated Graph Cuts", ACM Transactions on Graphics (SIGGRAPH'04), 2004
- [6] Y. Boykov, V. Kolmogorov, "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision", PAMI, vol. 26, no. 9, pp. 1124-1137, Sept. 2004.
- [7] G.R. Cross and A.K. Jain, "Markov Random Field Texture Models", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-5, No. 1, pp. 25-39, Jan. 1983.
- [8] 石川博, "グラフカット", 情報処理学会研究報告, 2007-CVIM-158-26, pp. 193-204, 2007.
- [9] 永橋知行,藤吉弘亘,金出武雄,"平滑化処理の繰り返しによる 画像セグメンテーションのためのグラフカット", MIRU2007, pp. 241-248, Jul, 2007.
- [10] 福田盛介,廣澤春任,"ウェーブレットによるテクスチャ情報の 抽出と多周波ポーラリメトリックな SAR 画像の分類手法につい て",電子情報通信学会技術研究報告.SANE,宇宙・航行エレ クトロニクス, Vol.98, No.7(19980420) pp. 11-16.
- [11] H.Lombaert, Y.Sun, L.Grady, C.Xu, "A Multilevel Banded Graph Cuts Method for Fast Image Segmentation" ICCV05,pp. 259-265, 2005.
- [12] Arthur Dempster, Nan Laird, and Donald Rubin, "Maximum likelihood from incomplete data via the EM algorithm ", Journal of the Royal Statistical Society, Series B, 39(1):1-38, 1977
- [13] http://research.microsoft.com/vision/cambridge/ i3l/segmentation/GrabCut.htm