GRAPH CUTS BY USING LOCAL TEXTURE FEATURES OF WAVELET COEFFICIENT FOR IMAGE SEGMENTATION

Keita FUKUDA

Graduate School of Engineering,
Kobe University
fukuda@me.cs.scitec.kobe-u.ac.jp

Tetsuya TAKIGUCHI, Yasuo ARIKI

Organization of Advanced Science and Technology,
Kobe University
takigu@kobe-u.ac.jp   ariki@kobe-u.ac.jp

ABSTRACT

This paper proposes an approach to image segmentation using Iterated Graph Cuts based on local texture features of wavelet coefficient. Using Multiresolution Analysis based on Haar Wavelet, low-frequency range (smoothed image) is used for n-link and high-frequency range (local texture features) is used for t-link along with color histogram. The proposed method can segment the object region with noisy edges and colors similar to the background, but heavy texture change. Experimental results illustrate the validity of our method.

Index Terms— Image Segmentation, Graph Cuts, Multiresolution Analysis, Local Texture Feature

1. INTRODUCTION

Extracting the foreground objects in static images is one of the most fundamental tasks in image content analysis, object detection and image editing. The task can be formulated as an image segmentation problem.

In recent years, the image segmentation problem is formalized as an optimal solution problem. The graph cuts technique proposed by Boykov [1] [2] provides a globally optimal solution for segmentation. In Snakes [3] and Level Set Method [4], the local minimum solution for boundary property cost function is computed. On the other hand, in graph cuts, it is possible to compute the global minimum solution, and the cost function is general enough to include both region and boundary properties of the segments.

However, it is difficult to segment an image with complex noisy edges because local noisy edges influence the cost of n-link calculated between neighboring pixels. To solve this problem, Nagahashi [6] proposed a coarse-to-fine approach to detect true boundaries using graph cuts.

In addition, it is difficult to segment an image with an object whose colors is similar to the background. To solve this problem, in this paper, we propose to employ image segmentation using Iterated Graph Cuts based on the smoothed image obtained from the low frequency range in the wavelet coefficient and the local texture features of wavelet coefficient for

t-link as well as color likelihood. the Nagahashi's method [6] is also included in the proposed method.

2. GRAPH CUTS

2.1. Theory of Graph Cuts

This section describes the Graph Cuts based segmentation proposed by Boykov and Jolly [1][2]. An image segmentation problem can be viewed as a binary labeling problem. From a given image, we can construct a weighted graph $G = (V, E, W)$ that consists of nodes V, edges E and nonnegative weights (costs) W. The nodes are pixels p on the image P and the edges have adjacency relationships with four or eight connections between neighboring pixels $q \in N$. The labeling problem is to assign a unique label A to each node. $A = (A_1, A_2, ..., A_p, ..., A_{|P|})$ can be obtained by minimizing the energy $E(A)$ in Eq.(1). A is a binary vector i.e. $A_p \in \{"obj", "bkg"\}$.

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

The coefficient $\lambda(\geq 0)$ in Eq.(1) specifies the relative importance of the region properties term $R(A)$ expressed as Eq.(2), to the boundary properties term $B(A)$ expressed as Eq.(3).

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

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

The term $R(\cdot)$ may reflect on how the intensity of pixel p fits into a known intensity model of object and background. The term $B(A)$ comprises the" boundary" properties of segmentation. $B(A)$ should be interpreted as a penalty for discontinuity between pixels p and q. $B_{p,q}$ is normally large when pixels p and q are similar. $\delta(A_p, A_q)$ is 1 if $A_p \neq A_q$, otherwise $\delta(A_p, A_q)$ is 0.

2.2. The Image Viewed as a Graph

The general approach to construct a graph from an image is shown in Fig. 1. Each pixel in the image is viewed as a

node in a graph. Edges are formed between nodes with the edge costs corresponding to how similar two pixels are (i.e. neighbor-link). The two terminal nodes, the source (S) and the sink (T) do not correspond to any pixels in the image but instead are viewed as representing the object and background respectively. The terminal edge costs are computed using models for the object and background (i.e. terminal-link).
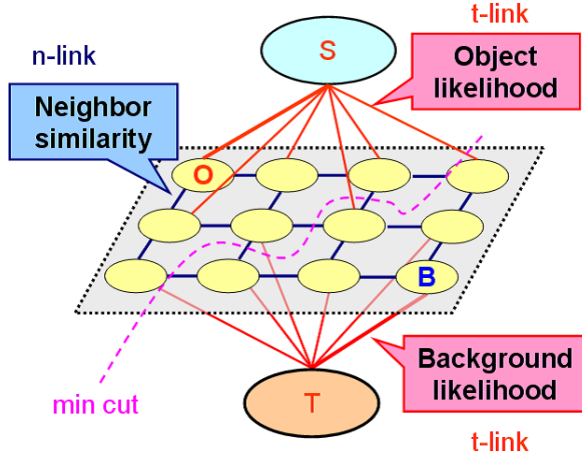

Fig. 1. Graph representing a 3-by-4 image.

Table 1. Edge cost

edge		cost	for
n-link	$\{p,q\}$	$B_{\{p,q\}}$	$\{p,q\} \in N$
t-link	$\{p,S\}$	$\lambda \cdot R_p(\text{"}bkg\text{"})$	$p \in P, p \notin O \cup B$
		K	$p \in O$
		0	$p \in B$
	$\{p,T\}$	$\lambda \cdot R_p(\text{"}obj\text{"})$	$p \in P, p \notin O \cup B$
		0	$p \in O$
		K	$p \in B$

Table 1 lists the edge costs of the graph. The region term and boundary term in Table 1 are calculated by:

$$\begin{cases} R_p(\text{"}obj\text{"}) = -lnPr(C_p \mid O) \\ R_p(\text{"}bkg\text{"}) = -lnPr(C_p \mid B) \end{cases} \quad (4)$$

$$B_{\{p,q\}} \propto exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \cdot \frac{1}{dist(p,q)} \quad (5)$$

$$K = 1 + \max_{p \in P} \sum_{q:\{p,q\} \in N} B_{\{p,q\}} \quad (6)$$

Let O and B define the "object" and "background" seeds respectively. The seeds are given by the user. I_p and C_p are brightness value and RGB color features at pixel p. The boundary between the object and the background is found by searching the minimum cost cut [5] on the graph G.

2.3. Problem with Graph Cuts

It has been difficult to segment images that include complex noisy edges in interactive graph cuts (Problem 1 in Fig.2). This is because the cost of the n-link becomes larger due to noisy edges than that of t-link. The edge has a strong influence when there is a large n-link. This problem was solved by using iterated graph cuts based on multi-scale smoothing to avoid the noisy edges[6]. But new problem was found that it was difficult to segment images with an object whose color is similar to the background (Problem 2 in Fig.2). This is because t-link is calculated based only on the color likelihood of the object and background. This problem can be solved by employing texture likelihood as well as color likelihood.

In this paper, we present an approach to image segmentation using iterated Graph Cuts based on local texture features as well as low frequency features of wavelet coefficient.
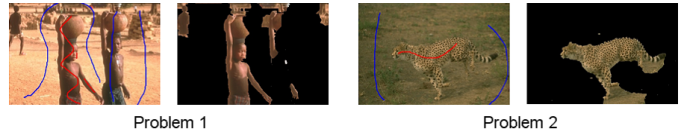

Fig. 2. Problem with Graph Cuts

3. PROPOSED METHOD

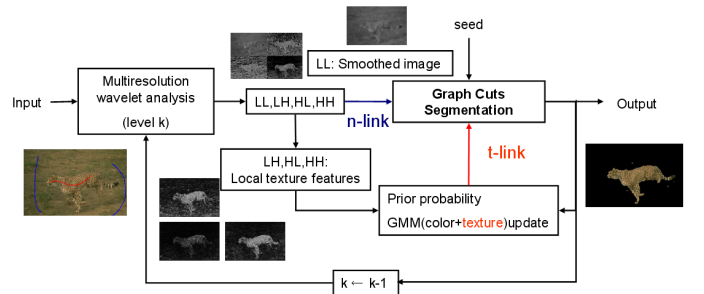### 3.1. Overview of Proposed Method

Fig. 3. Proposed method

The proposed method is shown in Fig. 3. "Object" and "background" seeds are given by the user on an input image. After initializing the level k, the input image is decomposed into subbands (LL, LH, HL, HH) by using Multiresolution Wavelet Analysis at the level k. Smoothed image defined in low-frequency range (LL) is used for n-link, and local texture features defined in high-frequency range (LH, HL, HH) are used for t-link. The likelihood are derived from local texture features as well as color features. The prior probabilities are defined by a distance transform using the previous segmentation result and the posterior probability obtained by multiplying one prior probability with the feature likelihood is set to t-link edge as the edge cost. Graph Cuts segmentation is carried out, and these processes are repeated until $k = 0$.

3.2. Multiresolution Analysis based on Wavelet Transform

Multiresolution Analysis using Wavelet Transform starts from the resolution level +1. Low frequency and high frequency in resolution level k are shown as follows:

$$c_n^{(k)} = \frac{1}{2}(c_{2n}^{(k-1)} + c_{2n+1}^{(k-1)}) \tag{7}$$

$$d_n^{(k)} = \frac{1}{2}(c_{2n}^{(k-1)} - c_{2n+1}^{(k-1)}) \tag{8}$$

The signal is down-sampled after Haar Wavelet transform, so that each subband contains one quarter of the pixels of the input image. Three subbands contain high-frequency information in different orientations: vertical (LH), horizontal (HL), and diagonal (HH). The remaining subband (LL) contains low-pass information. In Fig. 4, for example, signal is first decomposed into subbands HL1, LH1, HH1, and LL1. Image LL1 is further decomposed into HL2, LH2, HH2 and LL2.
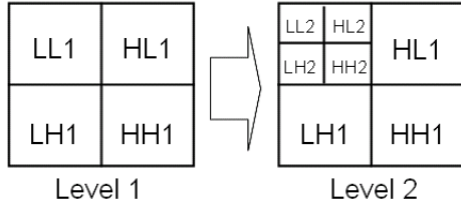

Fig. 4. Multiresolution analysis

3.3. Multi-scale Smoothing (n-link)

Low-frequency image (LL k) obtained by multiresolution analysis at the level k in the smoothed image and is used for n-link. Global to local segmentation is performed by iterating Graph Cuts process with multiresolution analysis using coarse to fine level in the similar way proposed by Nagahashi [6]. The difference is that Nagahashi employed Gaussian Smoothing and we employed the LL image obtained by multiresolution analysis.

3.4. Local Texture Features (t-link)

High-frequency (LH k, HL k and HH k), wavelet images are obtained by using multiresolution analysis at the level k. Local texture features are defined by wavelet coefficient of LH k, HL k and HH k and are used for t-link along with color features as described in 3.5. Local texture features are defined by averaging the absolute wavelet coefficient in the window (3×3) surrounding pixel p as follows:

$$T_p = \frac{1}{9} \sum_{p,q \in N} | d^{(k)} | \; for \; LHk, HLk, HHk \tag{9}$$

Local Texture features T_p are larger in complex region, and smaller in flat region.

3.5. The Posterior Probability (t-link)

Six dimensional features $Y_p = \{C_p, T_p\}$ are derived from RGB color features C_p and local texture features T_p. In Eq.(4), the edge costs in t-link are transformed to the posterior probability for further accuracy as follow:

$$\left\{ \begin{array}{l} R_p("obj") = -lnPr(O \mid Y_p) \\ R_p("bkg") = -lnPr(B \mid Y_p) \end{array} \right. \tag{10}$$

The posterior probability is proportional to the product of the prior probability and the features likelihood according to Bayes' theorem.

The feature likelihoods $Pr(Y_p \mid O)$, $Pr(Y_p \mid B)$ are derived by using Gaussian Mixture Model, and the prior probabilities $Pr(O)$ and $Pr(B)$ are derived by the distance transform of the segmentation result image at one previous multiresolution level. The brief shape information can be used as the prior probability. Distance d from the border between the object and background is normalized from 0.5 to 1.0. The prior probability is defined by using d_{obj} and d_{bkg} as follow:

$$Pr(O) = \left\{ \begin{array}{ll} d_{obj} & if \; d_{obj} \leq d_{bkg} \\ 1 - d_{obj} & otherwise \end{array} \right. \tag{11}$$

$$Pr(B) = 1 - Pr(O) \tag{12}$$

The flow of estimating the posterior probability of $\{p, T\}$ is shown in Fig.5. The posterior probability $\{p, S\}$ side is estimated in the same way using the background.
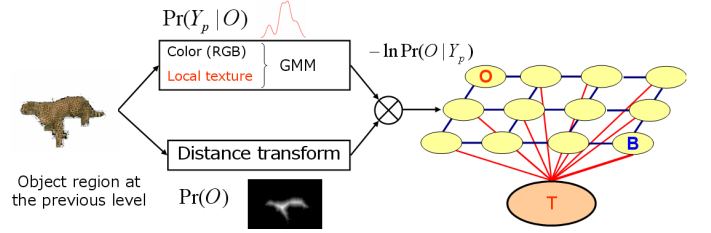

Fig. 5. The Posterior Probability for {p,T}

In Eq.(10), edge costs are calculated in {p,S} and {p,T} t-link. When one level of Graph Cuts segmentation finished, multiresolution level is set to $k - 1$. This process is repeated until $k = 0$.

4. EXPERIMENT

4.1. Experimental Conditions

Segmentation experiments were carried out using 50 images provided by Grab Cuts Database [7]. We compared Interactive Graph Cuts (conventional method (1)), iterated Graph Cuts using Smoothing (conventional method (2)) and the proposed method using the same seeds. The segmentation error rate is defined as:

$$Err[\%] = \left(\frac{miss \; detected \; pixels \; in \; the \; object}{image \; size} \right.$$
$$\left. + \frac{miss \; detected \; pixels \; in \; the \; background}{image \; size} \right) \times 100 \tag{13}$$
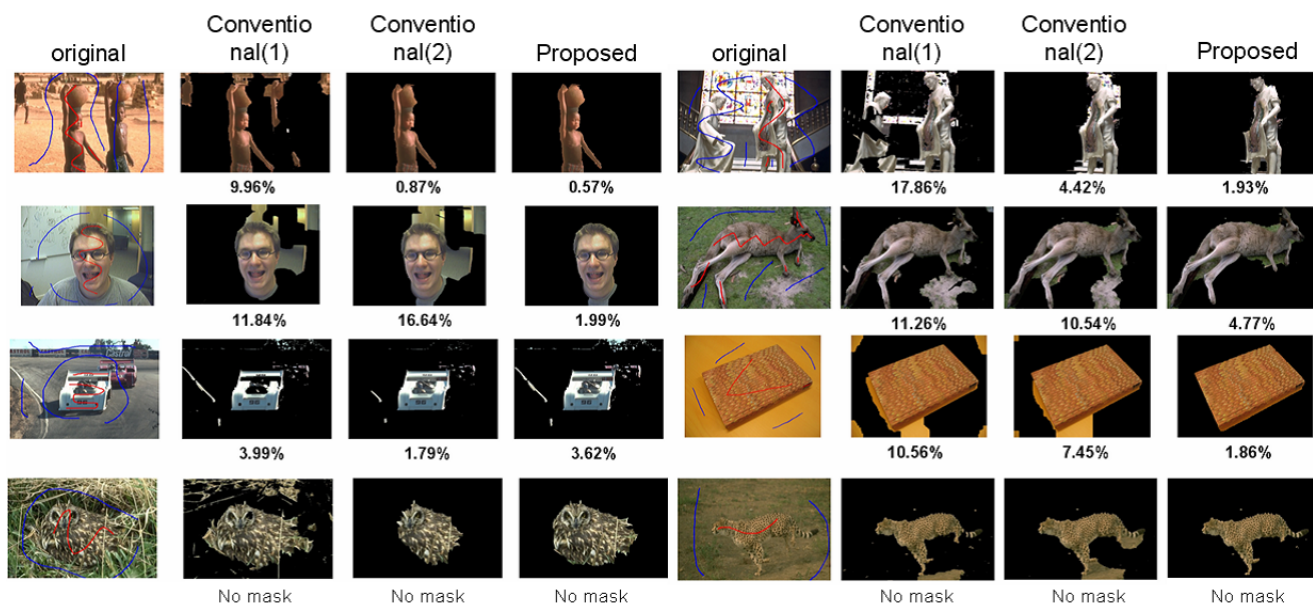
Fig. 6. Segmentation Result and Error rate

4.2. Experimental Result

Table 2. Error rates [%]

Conventional(1)	Convectional(2)	Proposed
4.05 %	2.42 %	1.57 %

The error rate in segmentation results is shown in table 2. The proposed method improved the error rate from 4.05 % to 1.57 % and 2.42 % to 1.57 % compared to the conventional method (1) and (2) respectively. Reduction rate is 61.2% from the conventional method (1) and 35.1% from the conventional method (2). The reason why the absolute difference of the error rate between the conventional and the proposed method is small in that the database includes many images where the convectional method (1) and (2) could achieve the low error rate. Fig. 6 shows the examples of the segmentation result with high error rate.

The conventional method (2) and the proposed method can achieve the better image segmentation for the images with complex edges than conventional method (1) (examples of the first line in Fig.6). They can remove noises because the prior probability has object shape information and smoothing process can reduce influence of strong edges.

The proposed method can achieve the better image segmentation for the images with object colors similar to background than conventional method (1) and (2) (examples of the other lines in Fig.6). Although the proposed method is effective for images with complex edges such as natural object images, the proposed method is ineffective for a flat image such as artifact images due to a few edges (i.e. automobile in Fig.6).

5. SUMMARY

This paper proposes an approach to image segmentation using Iterated Graph Cuts based on local texture features of wavelet coefficient. Using Multiresolution Analysis based on Haar Wavelet, low-frequency range (smoothed image) is used for n-link and high-frequency range (local texture features) is used for t-link along with color histogram. The proposed method improved the segmentation error rate compared to the conventional method. Future works include optimization of the weight to texture and color for segmentation.

6. REFERENCES

[1] 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.

[2] Y.Boykov, G.Funka-Lea, "Graph Cuts and Efficient N-D Image Segmentation, " International Journal of Computer Vision, 70(2):109-131, 2006.

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

[4] 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.

[5] 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.

[6] T.Nagahashi, H.Fujiyoshi, T.Kanade, "Image Segmentaion Using Iterated Graph Cuts Based on Multi-scale Smoothing, "

[7] http://research.microsoft.com/vision/cambridge/ i3l/segmentation/GrabCut.htm