

國立交通大學

電子工程學系 電子研究所碩士班

碩士論文



應用於彩色影像切割之區域性可靠資訊匯集

Local Belief Aggregation for Color Image Segmentation

研究生：詹景竹

指導教授：張添烜

中華民國 九十七年 九月

應用於彩色影像切割之區域性可靠資訊匯集

Local Belief Aggregation for Color Image Segmentation

研究生：詹景竹
指導教授：張添烜 博士

Student: Jing-Chu Chan
Advisor: Tian-Sheuan Chang



A Thesis
Submitted to Department of Electronics Engineering & Institute of Electronics
College of Electrical and Computer Engineering
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of Master
In
Electronics Engineering
September 2008
Hsinchu, Taiwan, Republic of China

中華民國 九十七年 九月

應用於彩色影像切割之區域性可靠資訊匯集

研究生：詹景竹

指導教授：張添烜 博士

國立交通大學電子工程學系 電子研究所碩士班

摘要

在彩色影像切割裡，馬可夫隨機場理論被用來解決如何給予畫面像素適當標籤的問題。在此論文裡，我們以區域的內部特性以及區域和區域之間的相關性來建立起馬可夫模型。然而，龐大的切割標籤數量，對於使用信任傳遞 (Belief Propagation, BP) 演算法來近似以馬可夫隨機場理論為基礎的彩色影像切割法遇到一些困難。這些困難包含了以下兩點：計算複雜度過高以及記憶體儲存空間過大而不敷使用的問題。在此論文裡，我們另外提出了一個利用地域性可靠資訊匯集的演算法來解決這些問題。這個方法主要是以限制鄰近點傳送進來的訊息數量為概念來達成目的。我們將此演算法套用到我們提出來的馬可夫模型上，利用近似的方式找出最大事後機率 (maximum a posteriori, MAP) 的結果。跟原本的信任傳遞演算法比較起來，我們提出的演算法可以減少相當多的記憶體儲存空間。在評量影像切割的結果方面，我們選擇與眾所皆知的平均位移 (Mean Shift) 演算法來做比較。在此，我們使用非監督方式的評比方法。這個方法主要是利用色彩視覺差異的特性設計而成的。實驗數據顯示，所提出的彩色影像切割演算法無論在主觀或是客觀的評比上，皆可以得到與平均位移演算法有相似的效果。除此之外，所提出的演算法在運算方面也比平均位移演算法還來的更具平行性。

Local Belief Aggregation for Color Image Segmentation

Student: Jing-Chu Chan

Advisor: Tian-Sheuan Chang

Department of Electronics Engineering & Institute of Electronics

National Chiao Tung University

Abstract

Markov Random Field (MRF) is used to solve the problem of labeling pixels in image segmentation. In this thesis, we formulate the MRF model based on the intra and inter region criteria. However, the enormous number of segment label in color image segmentation causes MRF-based color segmentation algorithm using belief propagation (BP) to suffer from complexity and storage explosion. To cope with this problem, this thesis also proposed a local belief aggregation (LBA) algorithm which restricts the number of messages to be aggregated from a neighboring node, to find the segmentation image that approximate the MAP solution of our MRF model. With the proposed LBA, memory storage is much reduced compared with the original BP algorithm. To evaluate the segmentation results, we compare the segmentation image with the well-known mean shift algorithm. Here, the unsupervised evaluation scheme using visible color difference is used as our objective evaluation metric. Experimental results show that the proposed color image segmentation algorithm can achieve a comparable result to mean shift algorithm both objectively and subjectively. Besides, the computation of LBA possesses more parallelism than the mean shift algorithm.

誌 謝

首先，要感謝我的指導教授—張添烜博士，這兩年來給我的支持和鼓勵，引領我以正確的態度來面對與解決問題。在研究上，總是讓我能自由的發揮，並在遇到瓶頸的時候給予建議與協助。此外感謝老師提供豐富的實驗室資源，使我不但能充分的利用軟硬體設備來進行研究，也能有機會出國參與研討會拓展視野。因此，我對與老師的感激之情溢於言表。

謝謝我的口試委員們，交大電子王聖智教授和清大電機陳永昌教授，感謝你們百忙中抽空來指導我，因為你們寶貴的意見讓我的論文更加完備。

感謝 VSP 實驗室的好伙伴們，特別要謝謝引我入門的張彥中學長，帶領我從零開始，一點一滴的從研究中學習與成長，尤其是當我對困難感到迷惘時，總是不厭其煩的和我討論，並且提供我許多寶貴的意見與想法。感謝林佑昆學長、李國龍學長，你們對於經驗與知識的分享，讓我受用無窮。謝謝李得璋學長、郭子筠學長和廖英澤學長耐心的教導我許多硬體設計的觀念與技巧，也感謝林嘉俊學長與吳私璟學長，你們在研究上認真負責的態度是我學習的榜樣。感謝曾宇晟和蔡宗憲同學不時的給予研究上的幫忙與協助，能跟你們一起討論是一段很難得的過程。感謝張瑋城與戴瑋呈同學，和你們一起待在實驗室熬夜趕作業是一段很有趣的回憶。感謝黃筱珊學妹、陳之悠、沈孟維、許博淵、蔡政君與廖元歆學弟們，有你們的陪伴，我的碩士班生涯充滿了歡笑。謝謝實驗室的所有成員們，和你們一起努力的日子，都是我在交大寶貴的回憶。

謝謝李韋磬同學，和你討論課業外的事情讓我的生活增添一些有趣的變化。感謝陳斯甯同學，謝謝妳的熱心相助，尤其是在最後的階段還抽空協助我檢查英文文法以及多次的聽我練習口頭報告並給我許多的建議。有妳這樣的朋友真是我的榮幸。

最後要感謝默默支持我的家人們，我的爸媽與姊姊，你們的溫暖是我努力最大的支柱。

在此，把本論文獻給所有愛我與所有我愛的人。

Contents

Chapter 1. Introduction.....	1
1.1. Background.....	1
1.2. Motivation and Contribution.....	2
1.3. Thesis Organization	3
Chapter 2. Previous Work.....	5
2.1. Mean Shift Algorithm	5
2.2. Watershed Algorithm	9
2.2.1. Immersion-based Method	10
2.2.2. Toboggan-based Method.....	12
2.3. Markov Random Field based Algorithm.....	14
2.3.1. Markov Random Field Theory.....	14
2.3.1.1. Random Fields	15
2.3.1.2. Markov Random Fields.....	16
2.3.1.3. Gibbs Random Fields.....	18
2.3.1.4. Relation between MRF and GRF.....	19
2.3.1.5. MAP-MRF Framework.....	20
2.3.2. Application in Segmentation.....	21
2.3.2.1. Deng's Work	22
2.3.3. Inference Algorithm Using Loopy Belief Propagation	23
Chapter 3. Color Image Segmentation Algorithm Using MRF Model.....	27
3.1. Algorithm Overview	27
3.2. CIE L*a*b* Color Space	29
3.2.1. Introduction.....	29
3.2.2. Color Transform from RGB to CIE L*a*b*	30
3.2.3. Color Difference	31
3.3. MRF Model Formulation.....	31
3.3.1. Likelihood and Prior	31
3.3.2. Model Approximation	32
3.4. Local Belief Aggregation.....	37
3.4.1. Reliable Message Aggregation	37
3.4.1.1. Four-window Based Local Search Window.....	38
3.4.1.2. One-window Based Local Search Window	40
3.4.2. Maximum Belief Segment State Selection	41
3.4.3. Region Merge Process	43
3.4.4. Convergence of Local Belief Aggregation.....	45

3.5.	MRF Model Comparison	46
3.6.	Complexity Analysis and Comparison.....	48
Chapter 4.	Experimental Results and Analysis	51
4.1.	Introduction.....	51
4.2.	Segmentation Results.....	52
4.2.1.	Chen’s Quantitative Evaluation Metrics	52
4.2.2.	Comparison between One-window Based LBA and Four-window Based LBA.....	54
4.2.3.	Comparison between Different Segmentation Algorithms	58
Chapter 5.	Conclusion and Future Works.....	63
5.1.	Conclusion	63
5.2.	Future work.....	63
Reference	65
Appendix	69



List of Figures

Fig. 2.1. Flow of the mean shift filtering.....	6
Fig. 2.2. Example of mode seeking procedure and density estimation.....	7
Fig. 2.3. Flow of the mean shift segmentation.	8
Fig. 2.4. Visualization of mean shift filtering and segmentation results for gray level data [6]. (a) Input image. (b) Mean shift mode seeking paths. (c) Mean shift filtering result. (d) Mean shift segmentation result.	9
Fig. 2.5. Example of one dimensional signal using immersion-based watershed approach.	10
Fig. 2.6. Example of one dimensional signal using toboggan-based watershed approach.	12
Fig. 2.7. Flow chart of the Fairfield’s toboggan-based watershed segmentation.....	14
Fig. 2.8. Example of neighborhood system. (a) First-order neighborhood system. (b) Second-order neighborhood system.	17
Fig. 2.9. Clique types for first order and second order neighborhood system. (a) First order neighborhood system. (b) Second order neighborhood system.	18
Fig. 2.10. An undirected graphical model with hidden variables and observed variables.	24
Fig. 2.11. Local message passing in a Markov Network [11].....	25
Fig. 3.1. Flow of image segmentation using local belief aggregation.	27
Fig. 3.2. CIE L*a*b* color space.	30
Fig. 3.3. Relation between the value of clique potential function and the value of difference of average segment label color for two different segment labels.	36
Fig. 3.4. Reliable message aggregation using four local search windows.....	38
Fig. 3.5. Reliable message aggregation using one local search window.....	40
Fig. 3.6. Four-window based local reliable message aggregation in a Markov network.	42
Fig. 3.7. One-window based local reliable message aggregation in a Markov network.....	42
Fig. 3.8. Region merge process in four-window based method.....	44
Fig. 3.9. Region merge process in one-window based method.....	45
Fig. 3.10. E_{cw} convergence of four-window based LBA	46
Fig. 3.11. Segmentation results with different MRF model using LBA algorithm. (a) MRF model using (3.12) and (3.13). (b) MRF model using (3.24) and (3.25).....	47
Fig. 3.12. Comparison of each computational unit between different methods used in the proposed LBA.	49
Fig. 3.13. Segmentation result of different LBA algorithms. (a) One-window based LBA. (b) Four-window based LBA.	50
Fig. 4.1. The plot of intra-region visual error v.s. inter-region visual error [23].	54
Fig. 4.2. E_{cw} curve on test image 253055.	55

Fig. 4.3. E_{cw} curve on test image 241004.	55
Fig. 4.4. E_{cw} curve on test image 3096.	56
Fig. 4.5. Segmentation results with the proposed two methods. (a) Test image 253055. (d) Test image 241004. (g) Test image 3096. (b)(e)(h) Segmentation results using one-window based LBA. (c)(f)(i) Segmentation results using four-window based LBA.	58
Fig. 4.6. E_{cw} curve of different segmentation algorithms on test image 253055.	59
Fig. 4.7. E_{cw} curve of different segmentation algorithms on test image 241004.	59
Fig. 4.8. E_{cw} curve of different segmentation algorithms on test image 3096.	60
Fig. 4.9. Segmentation results with different algorithms. (a)(d)(g) Segmentation results using proposed four-window based LBA. (b)(e)(h) Segmentation results using mean shift algorithm. (c)(f)(i) Segmentation results using watershed algorithm.....	61
Fig. A.1. Subjectively determine the relation between reliable message number and convergence iteration in four-window based LBA with test image 253055.	70
Fig. A.2. Objectively determine the relation between reliable message number and convergence iteration in four-window based LBA with test image 253055.	70
Fig. A.3. Intermediate results on 253055 at different iterations using 4 reliable messages. (a) Iteration 1. (b) Iteration 5. (c) Iteration 10. (d) Iteration 25. (e) Iteration 40. (f) Iteration 80.	72
Fig. A.4. Intermediate results on 253055 at different iterations using 9 reliable message number. (a) Iteration 1. (b) Iteration 5. (c) Iteration 10. (d) Iteration 25. (e) Iteration 40. (f) Iteration 44.	73
Fig. A.5. Segmentation results with different reliable message number using proposed 4 window based LBA. (a) 1 reliable messages. (b) 2 reliable messages. (c) 3 reliable messages. (d) 4 reliable messages. (e) 5 reliable messages. (f) 6 reliable messages. (g) 7 reliable messages. (h) 8 reliable messages. (i) 9 reliable messages.	74

List of Tables

Table 3.1. Parameter settings of image 253055 for LBA.	48
Table 3.2. Average execution counts of image 253055 for LBA.....	49
Table 4.1. Effect of color difference in CIE L*a*b* color space for human's visual perception [28]....	53



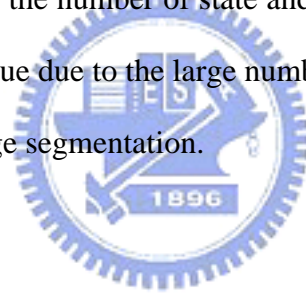
Chapter 1. Introduction

1.1. Background

Image segmentation is an important low-level pre-processing step for image analysis applications such as stereo vision, medical image analysis, and video object segmentation [1]-[4]. It classifies pixels of interest in an image into several non-overlapping regions with unique segment labels. Research on image segmentation has continued for many years and many methods have been proposed. One of the most well known methods is the watershed [5]. It is based on the concept of extracting regions as catchment basins topographically. The simple concept and low computational complexity of the watershed have enabled it to be adopted by many applications. However, the watershed is sensitive to noise. Another well known method is the mean shift method [6]. It is a nonparametric and iterative mode seeking algorithm that works in the joint spatial-range space of a color image. In contrast to watershed's sensitivity to noise, mean shift's mode seeking approach is more robust to noise. Mean shift algorithm has been considered to have the best performance among most low-level segmentation methods.

Despite solving image segmentation problem solely based on the topography and density in the multi-dimensional feature space, segmentation methods based on the Markov Random Field (MRF) model that was originally introduced by Geman and Geman [7] have also attracted attention. MRF-based methods model the segmentation as a labeling problem with a MRF having a maximum a posteriori (MAP) solution

corresponding to the ideal segmentation result. Various MRF models have been formulated in different image segmentation methods [8] [9]. One way to efficiently solve the MAP problem in a MRF is Pearl's belief propagation (BP) [10]. Belief propagation approximates the inference much faster than Gibbs sampler and simulated annealing [7]. BP has been successfully applied to stereo vision [11] with a MRF model whose number of discrete state (number of disparity range) is less than a hundred. However, BP suffers from computational complexity and storage requirement explosion when trying to apply to color image segmentation whose number of state (number of segment label) can reach up to thousands. This is because BP's complexity is quadratically proportional to the number of state, and the storage size is linearly proportional to the number of state and MRF connectivity. As a result, the complexity and storage issue due to the large number of segment label has obscure BP's application in color image segmentation.



1.2. Motivation and Contribution

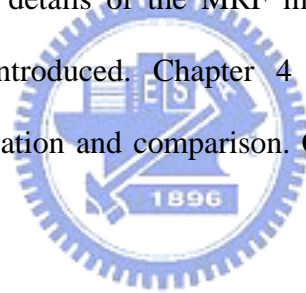
The issues mentioned above motivate us to propose a color image segmentation method with a new MRF modeling of color image segmentation and a local belief aggregation (LBA) algorithm to estimate the MAP inference. The proposed MRF models the likelihood and prior probability based on the concept of the intra and inter region constraint respectively. Such MRF formulation is a more direct formulation in contrast to MRF formulations that only models edge label. Consequently, our MRF possesses large number of labels. To cope with this, the proposed LBA method, which was inspired by the original BP and dynamic quantization (DQ) [12], aggregates only limited number of reliable messages from neighboring nodes iteratively.

The contribution of the thesis includes

1. We formulate the MRF model in a simple manner based on the concept of intra and inter region criteria.
2. We proposed a local belief aggregation (LBA) algorithm to estimate the MAP of MRF model for color image segmentation.

1.3. Thesis Organization

In Chapter 2, we briefly introduce existing important methods in image segmentation. In Chapter 3, we briefly introduce the transformation and color distance of CIE L*a*b* color space. In addition, the details of the MRF model formulation and the LBA estimation algorithms are introduced. Chapter 4 presents the quantitative and qualitative performance evaluation and comparison. Conclusion and future work are given in Chapter 5.



Chapter 2. Previous Work

2.1. Mean Shift Algorithm

Mean shift [6] image segmentation is an unsupervised clustering algorithm using information of feature space to cluster image into several segment regions. Mean shift segmentation algorithm includes two steps. First, the mean shift filtering procedure detects each cluster or the basin of attraction with iterative mode seeking procedure to estimate the density gradient in the probability density function. A mode is defined as the local maxima of the probability density function. The basin of attraction of a mode is defined as a region in which all the data points would converge to this mode through the iterative mean shift procedure. In other words, the points in the same basin of attraction are associated with the same cluster. Later, the cluster delineation step groups together all the clusters of a mode within a Euclidean distance in the feature space into a single cluster.

The mean shift vector is originally deduced from the concept of finding the gradient of probability density function. It is designed to move the point x in the feature space toward the corresponding mode as

$$M_h(x) = \frac{1}{n_x} \sum_{x_i \in S_h(x)} (x_i - x), \quad (2.1)$$

where h is the radius of the hyper sphere $S_h(x)$ in the d -dimensional Euclidean space centered on x with n_x pixels in it. In color image segmentation, a 5-dimensional feature space is used. It contains 2-dimensional spatial information and 3-dimensional range (or color) information. Each pixel in the image represents a vector with its

corresponding 5-dimensional component in the feature space. This information is applied to mean shift vector to find the corresponding cluster mode. By calculating the mean shift vector, the location of the center of the hyper sphere is shifted iteratively according to

$$x^{i+1} = x^i + M_h(x^i), \quad (2.2)$$

and the procedure will continued until the convergence is met at the corresponding mode for a given pixel. The convergence condition is $x^{i+1} \approx x^i$. Fig. 2.1 gives the flow chart of the mean shift filtering procedure. The output result is the smoothed image of the origin image.

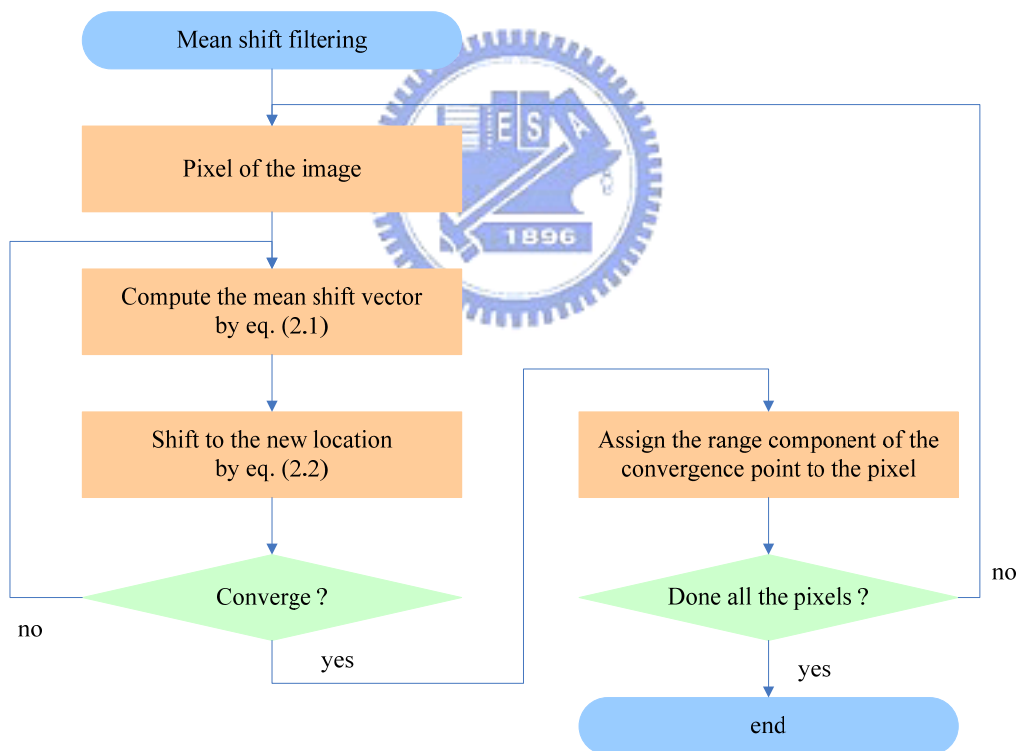


Fig. 2.1. Flow of the mean shift filtering.

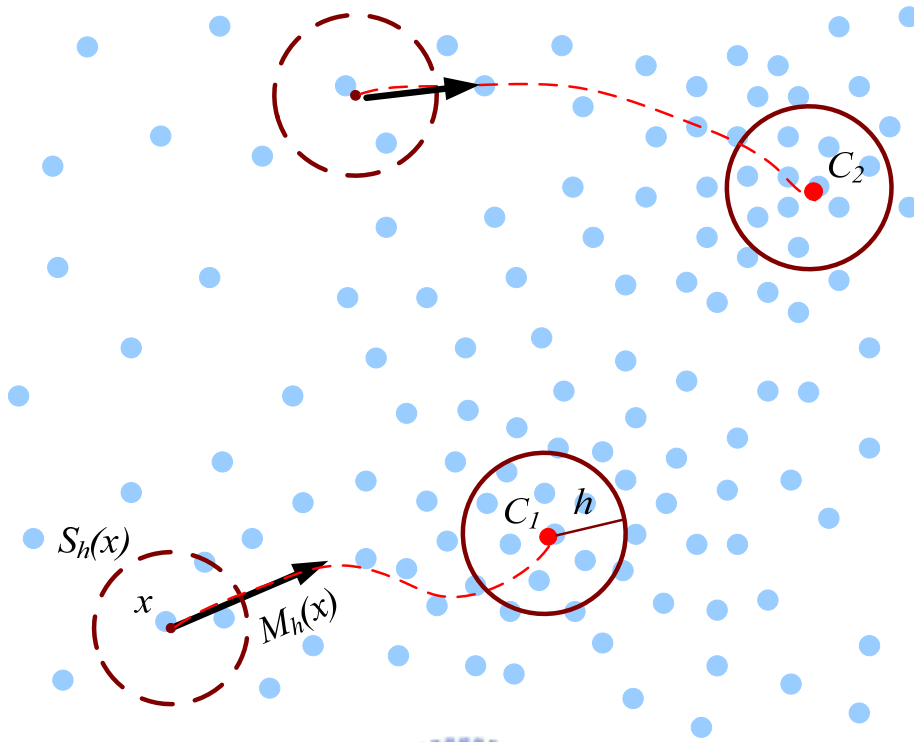


Fig. 2.2. Example of mode seeking procedure and density estimation.

An example of the mean shift vector of a given point moving toward the corresponding local density maximum is given in the Fig. 2.2. In the figure, points inside the sphere $S_h(x)$ with radius h around x is used to estimate the probability density function of x . The direction of the mean shift vector $M_h(x)$ is computed and the new location is shifted iteratively until the point of convergence is reached. The convergence point always has the highest density in the feature space and is colored in red in Fig. 2.2. C_1 and C_2 are the center of the cluster 1 and cluster 2 in the example respectively.

For mean shift segmentation, an additional procedure that groups the clusters with mode distance smaller than h_s in the spatial domain and h_r in the range domain is performed after mean shift filtering. The parameters h_s and h_r are the radius of the window in the spatial and range domain respectively. Finally, an optional procedure

that eliminates regions with area smaller than M pixels is also performed to further improve the quality of the segmentation results. The flow chart of the mean shift segmentation is illustrated in Fig. 2.3.

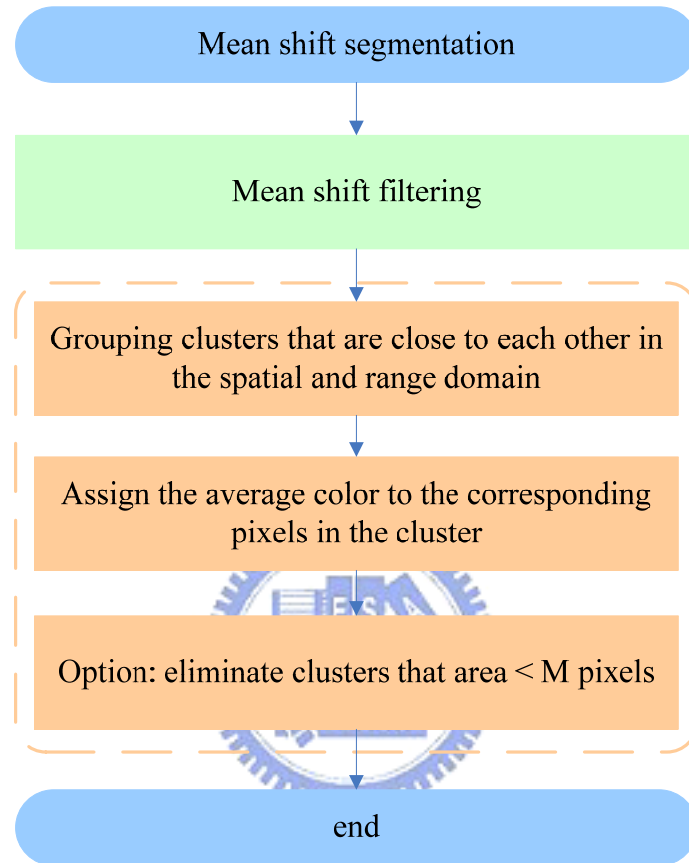


Fig. 2.3. Flow of the mean shift segmentation.

An example of the mean shift segmentation is illustrated in Fig. 2.4 [6]. Fig. 2.4 (a) is a part of image data from *Cameraman* test image. Fig. 2.4(b) demonstrates the intermediate results during the mean shift filtering procedure. In the figure, each pixel is iteratively calculated using (2.1) and (2.2) to find the mean shift path represented as the block line. The black dots are the points of convergence for the corresponding pixels. After the mean shift filtering procedure, the smoothed image is demonstrated in Fig. 2.4 (c). Finally for the mean shift segmentation, clusters that are close to each other

in a predefined threshold are grouped together. Fig. 2.4(d) shows the final segmentation results using mean shift segmentation algorithm.

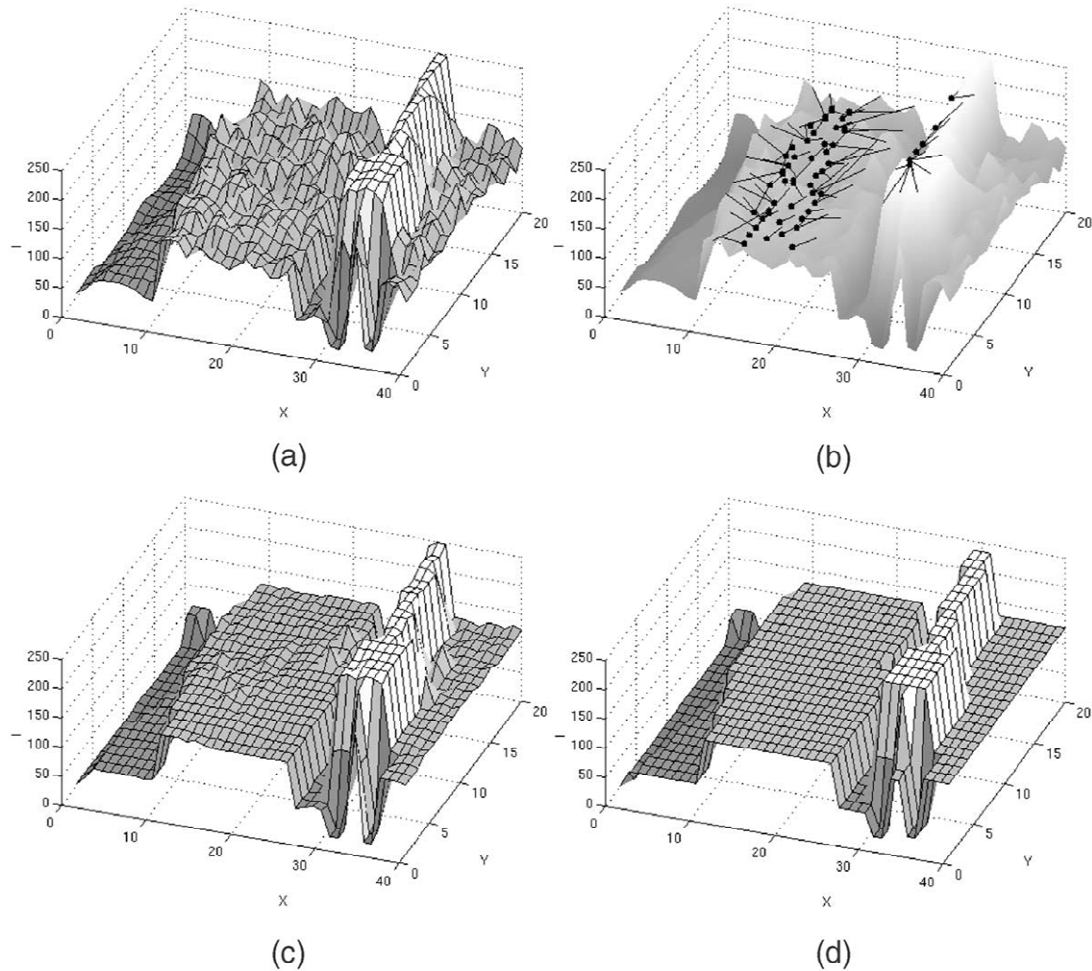


Fig. 2.4. Visualization of mean shift filtering and segmentation results for gray level data [6]. (a) Input image. (b) Mean shift mode seeking paths. (c) Mean shift filtering result. (d) Mean shift segmentation result.

2.2. Watershed Algorithm

Watershed segmentation is a popular and well known algorithm that extracts regions as catchment basins based on the concept of topography. The gradient image of the input image is used as the topographic surface in which the gradient value represents the altitude. The segmentation of an image finds the watershed line on the gradient image

and thus separates each region. There exist two approaches to implement watershed segmentation, one is the immersion-based method and the other is the toboggan-based method. The immersion-based watershed segmentation uses a bottom-up approach while the toboggan-based method uses a top-down approach to find the watershed line on the geography.

2.2.1. Immersion-based Method

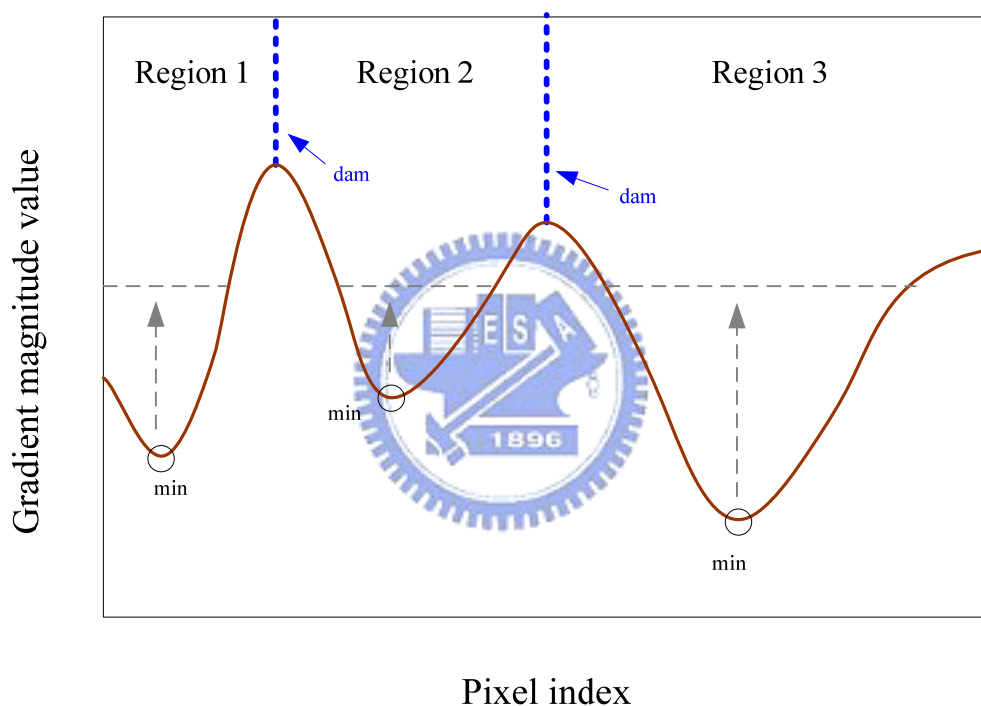


Fig. 2.5. Example of one dimensional signal using immersion-based watershed approach.

Immersion-based method can be explained as an iterative flooding approach. It can be thought as first pierce holes in each regional minimum of the topography surface. Then we slowly immerse this surface into the water. Starting from the regional minimum of the surface, the water will progressively flood up the catchment basins. While the waters from different catchment basins are about to merge, we build the dam to prevent

them from merging. In the end, each catchment basin is separated by the dam. The dams represent the watershed lines and catchment basins represent the segment regions. Take Fig. 2.5 for example. Three regional minimums are found and each of them corresponds to a catchment basin. Two dams which represent the watershed lines in the image are built to delimit the catchment basins. As a result, three segment regions are found.

There are two steps to implement the watershed algorithm proposed by Vincent and Soille [5], sorting and flooding procedure. Sorting procedure first sorts the pixels of an image in the increasing order of the gradient value for the purpose to access the pixels directly in a certain gray level. Then a flooding procedure is performed level by level starting from the minimum level to determine the watershed and catchment basins. At each gray level, pixels belong to the corresponding level h is first marked in label *MASK*. Then the neighboring status of those marked pixels is checked. If at least one of a neighbor of a pixel is labeled from the previous iteration, then the corresponding pixel is inserted in a first-in-first-out (FIFO) queue. Later, a recursive label propagation of each marked pixels in the FIFO is performed. If a pixel is adjacent to at least two different catchment basins, then the pixel is labeled as a watershed. If a pixel is only adjacent to one catchment basin, then it is labeled as the same label with the corresponding catchment basin. In the end, the remaining pixels still marked as *MASK* in the level h belongs to the regional minimum. The pixel and its connected components are given a new label as a new catchment basin. A pseudo code with more details of Vincent's watershed segmentation algorithm is referred to [5].

2.2.2. Toboggan-based Method

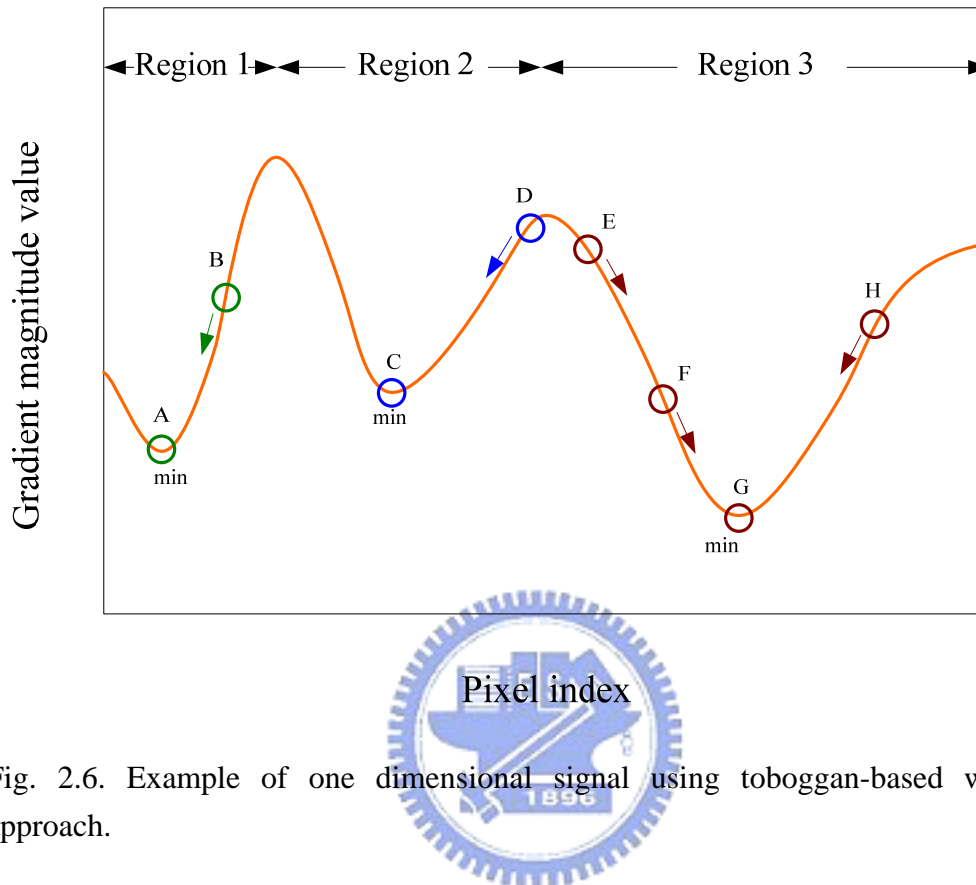


Fig. 2.6. Example of one dimensional signal using toboggan-based watershed approach.

Toboggan-based method can be thought as a rain drop sliding down from the hill by analogy. It tries to find out the downstream path where each rain drop slides down to a regional minimum of the topography surface. Each pixel represents a rain drop in the corresponding altitude of the topography surface. Pixels that slide down to the same regional minimum belong to the same catchment basin and a unique label is given. Fig. 2.6 gives a one dimensional example to describe the concept of the toboggan based method. The gradient value of pixel E 's right hand side is lower than the value of left-hand side, thus pixel E slides down in the right direction toward the regional minimum G . Pixel E , F and H slide down to the same regional minimum, thus belong to the same catchment basin. Finally, three regions are produced. Note that the

toboggan-based algorithm is processed in a raster scan order. Thus in contrast with immersion-based algorithm, there is no need to perform an expensive sorting process which results in an irregular computing order. However, toboggan requires an additional backtracking procedure to solve the labeling problem in the algorithm.

The toboggan algorithm proposed by Fairfield [13] includes two steps, toboggan enhancement and contrast segmentation. In toboggan enhancement step, pixels slide down in the steepest descent according to the gradient value. In this step, pixels belonging to respective catchment basin are determined. This step usually produces an over segmentation result. To achieve better segmentation result, contrast segmentation is used as a post process to the segmentation image produced in the toboggan enhancement step. The contrast segmentation checks the color different between neighbors. If the color difference of two neighbor pixels is less than a pre-defined threshold, then two pixels is connected. This concept is similar to the region merge process. Fig. 2.7 shows the flow chart of the Fairfield's toboggan based watershed segmentation.

Several toboggan-based approaches have been proposed to further improve the quality of the segmentation image based on Fairfield's work. Readers can referred to [14] [15] for further detail information.

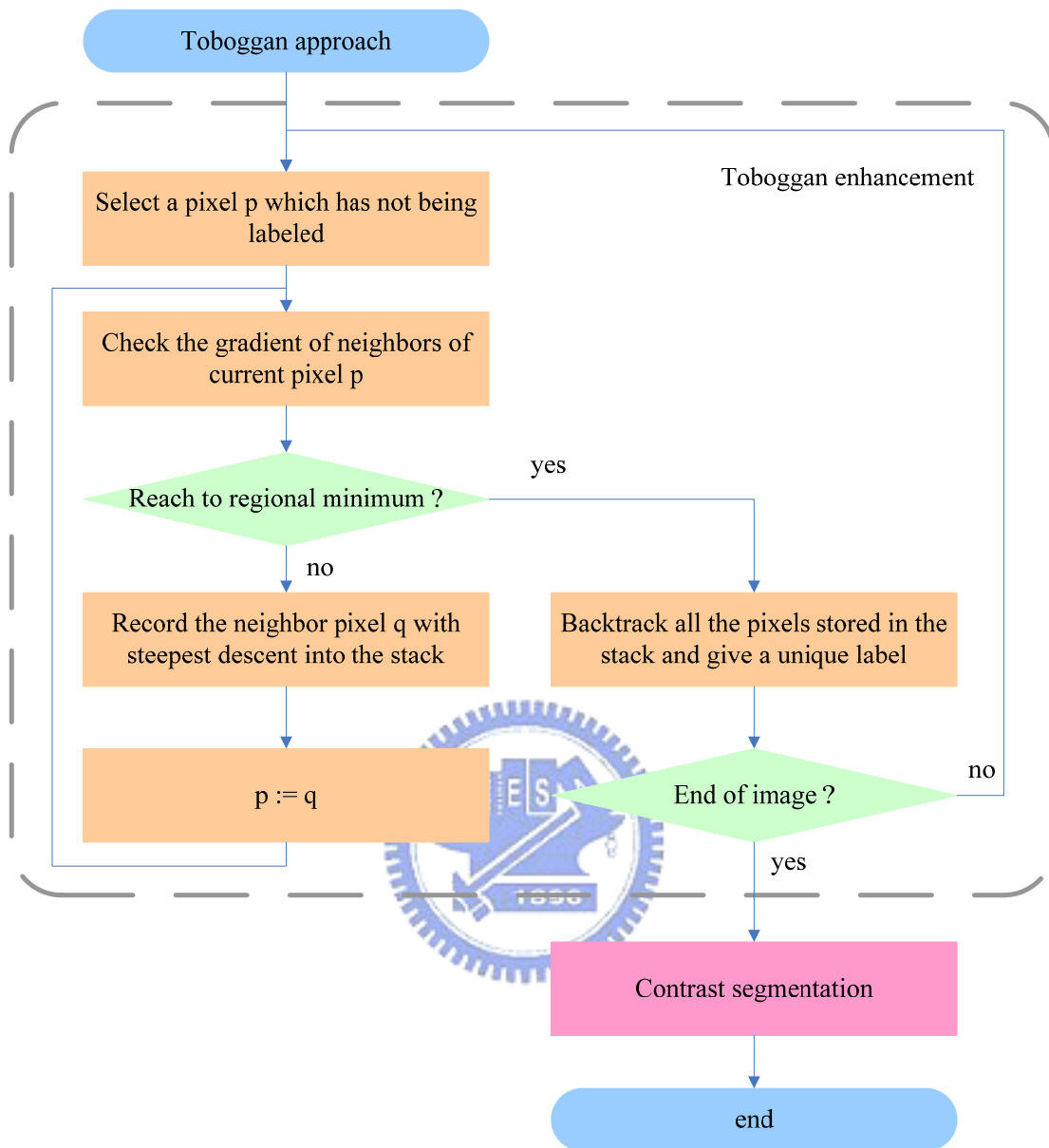


Fig. 2.7. Flow chart of the Fairfield's toboggan-based watershed segmentation.

2.3. Markov Random Field based Algorithm

2.3.1. Markov Random Field Theory

The following content of introduction on Markov random field theory in section 2.3.1 is

referenced from [16] [17]. Readers can refer to it for more details.

2.3.1.1. Random Fields

Many problems can be seen as labeling problems in terms of labels and states (or sites).

A label is an event that may happen to a state and a set of discrete state is defined as

$$DS = \{0, 1, \dots, m-1\}. \quad (2.3)$$

A labeling problem chooses a label from the label set

$$L = \{0, 1, \dots, l-1\}, \quad (2.4)$$

and assigns it to each of the states in DS . The value f_i is regarded as a particular mapping of state i from DS to L and the set

$$f = \{f_0, f_1, \dots, f_{m-1}\}, \quad (2.5)$$

is called a labeling or a configuration of states in DS . The set of all possible configurations is called the configuration space and is defined as

$$\Omega = |L|^m, \quad (2.6)$$

where m is the size of set DS .

Let $F = \{F_0, F_1, \dots, F_{m-1}\}$ be a family of random variables defined on the set of states DS where each random variable F_i take a value from the set of label L . This family F is called a random field. The event that F_i takes the value f_i is denoted as $F_i = f_i$. The probability that a random variable F_i takes a value f_i is denoted as $P(F_i = f_i)$ and abbreviated as $P(f_i)$. The notation $(F_0 = f_0, F_1 = f_1, \dots, F_{m-1} = f_{m-1})$ or simply $F = f$ denotes the joint event where f is a configuration of F and the joint probability is denoted as $P(F = f)$, abbreviated $P(f)$.

2.3.1.2. Markov Random Fields

In a regular lattice case, we consider DS as an image lattice X so that $X = \{x = (p, q) \mid \forall x \in DS\}$. Let N_x denotes the set of states neighboring x . $N = \{N_x \mid \forall x \in X\}$ is said to be a neighboring system if it has the following two properties:

- (i) A state is not neighboring to itself: $x \notin N_x, \forall x \in X$.
- (ii) Neighboring is symmetry: $x \in N_y \Leftrightarrow y \in N_x, \forall x, y \in X$.

The definition of n^{th} order neighborhood set and its neighborhood system is given as

$$v_{(p,q)}^n = \{(k - p)^2 + (l - q)^2 \leq n \mid \forall (k, l) \in X, \forall (p, q) \in X\}, \quad (2.7)$$

$$v^n = \{v_{(p,q)}^n \mid \forall (p, q) \in X\}. \quad (2.8)$$

A first-order and a second-order neighborhood system are given as an example shown in Fig. 2.8. A first-order neighborhood system, also called a four neighborhood system, has four neighbors for each state in the regular lattice. A second-order neighborhood system, also called an eight neighborhood system, has eight neighbors for each state. Order that is higher than two is rarely used because it is complicated and requires a lot of computations in most applications.

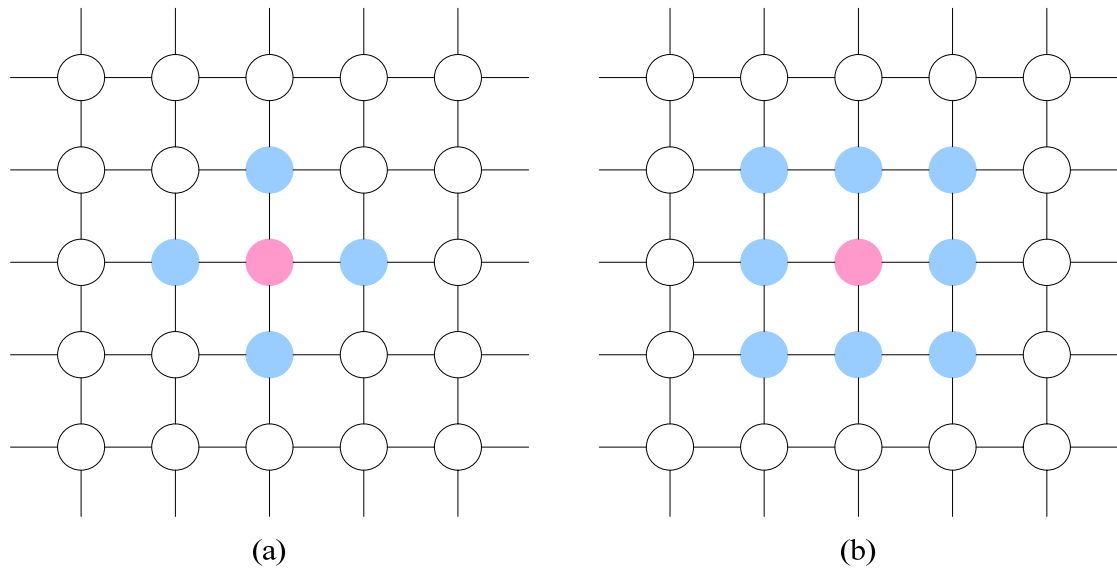


Fig. 2.8. Example of neighborhood system. (a) First-order neighborhood system. (b) Second-order neighborhood system.

A random field F is said to be a Markov Random Field (MRF) on X with respect to a neighborhood system N if and only if the following conditions are satisfied:

- (i) Positivity : $P(F = f) > 0$ for all possible configurations.
- (ii) Markovianity : $P(F_k = f_k | F_i = f_i, i \in X \setminus \{k\}) = P(F_k = f_k | F_i = f_i, i \in N_k)$.

The notation \setminus denotes the exclusive operation, thus the notation $i \in X \setminus \{k\}$ denoted above means that i represents all possible states in set X but the state k . Thus $\{f_i, i \in X \setminus \{k\}\}$ denotes the set of labels of all states but k and $\{f_i, i \in N_k\}$ denotes the set of labels at the states neighboring k . Hence, the Markovianity condition describes the local characteristics of the random field that the probability of a state given a label in X is only affected by its neighborhood system. The positivity condition describes that all configurations are possible.

2.3.1.3. Gibbs Random Fields

A clique c is defined as any state in set X with all its possible pair neighbors in a neighborhood system. The set of all cliques is denoted as C . Fig. 2.9 gives examples of cliques for both the first order and second order neighborhood system on a regular image lattice. The first order neighborhood system contains two types of cliques, single-state clique and pair-state clique, as shown in Fig. 2.9 (a). A single-state clique contains only one state; a pair-state clique contains a pair of neighboring states; a triple-state clique contains a triple of neighboring states; and so on. The line connected the lattices in Fig. 2.9 indicates the neighboring connection between the states.

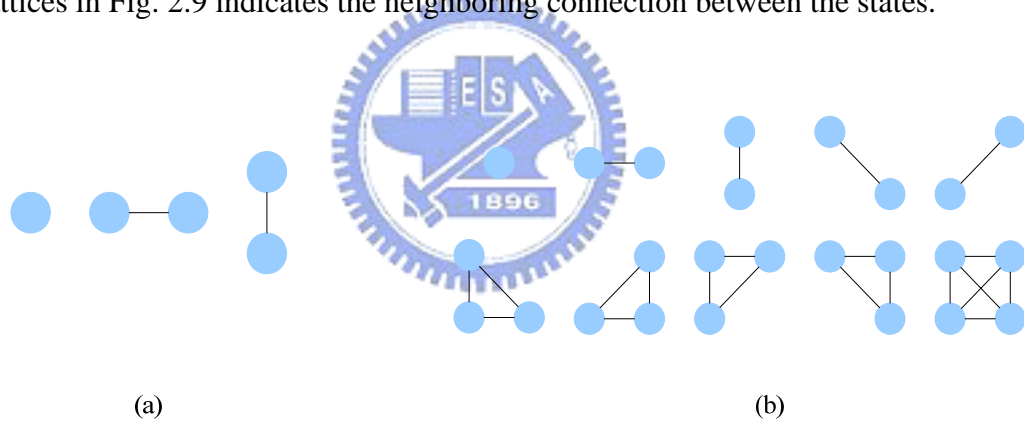


Fig. 2.9. Clique types for first order and second order neighborhood system. (a) First order neighborhood system. (b) Second order neighborhood system.

A random field F is said to be a Gibbs Random Field (GRF) on X with respect to a neighborhood system N if and only if its configurations follows a Gibbs distribution. A Gibbs distribution has the following form

$$P(F = f) = \frac{1}{Z} \exp\left(-\frac{1}{T} U(f)\right), \quad (2.9)$$

where T is a constant named as temperature, Z is a normalized constant called partition

function defined as

$$Z = \sum_{f \in \Omega} \exp\left(-\frac{1}{T} U(f)\right). \quad (2.10)$$

Ω denotes the configuration space of all possible configurations defined in (2.6) and $U(f)$ is the energy function that sums up the clique potential functions $U_c(f)$ of all possible cliques c

$$U(f) = \sum_{c \in C} U_c(f). \quad (2.11)$$

The value of $U_c(f)$ depends on the local configuration on the clique c .

The joint probability $P(F = f)$ in (2.8) measures the probability of the occurrence of a particular configuration. From the definition above, it is clear that the lower the energy of a configuration has, the higher the probable a configuration is.

2.3.1.4. Relation between MRF and GRF

Markov random field follows the Markovianity property, thus it is characterized by the local property. Gibbs random field obeys a Gibbs distribution, thus it is characterized by the global property. The Hammersley-Clifford theorem [18] established the equivalence relationship of these two types of properties. The theorem states that a random field F is a MRF on X with respect to the neighborhood system N if and only if the random field F is a GRF on X with respect to the neighborhood system N . This equivalent provides a simple way to specify the local characteristic property of MRF by specifying the clique potential function which encodes a prior knowledge of interactions between neighbor nodes in image lattice. Thus the problem of finding the joint probability $P(F = f)$ of MRF becomes equivalent to first specifying the clique potential function $U_c(f)$ and then calculating the energy function $U(f)$ as shown in

(2.10).

2.3.1.5. MAP-MRF Framework

The Bayesian approach can be used to solve the problem of image segmentation which can be thought as a labeling problem that gives a label in the segment label set L (2.4) to a state in set X . The result of the labeling problem is the segmentation image which is of interest. Let S be the set for a segment results based on the feature vector extracted from original image I . According to the Bayes' rule, the posteriori probability can be presented as

$$P(S = s | I = i) = \frac{P(I = i | S = s)P(S = s)}{P(I = i)}, \quad (2.12)$$

where $P(I = i | S = s)$ represents the probability distribution of varying the segmentation result S for fixed image color information I and thus is called the likelihood of I . $P(S = s)$ is the priori probability (prior) that defines the joint probability distribution of neighboring segment labels. $P(I = i)$ is the probability of the given image color information and it remains unchanged during the process, thus it is considered as a constant and the posteriori probability (2.12) is equivalent to

$$P(S = s | I = i) \propto P(I = i | S = s)P(S = s). \quad (2.13)$$

To obtain the most probable estimate of interest, a maximum a posteriori (MAP) approach is used. Taking the MAP of the posteriori probability (2.13)

$$s^{MAP} = \arg \max_{s \in \Omega} P(S = s | I = i) = \arg \max_{s \in \Omega} P(I = i | S = s)P(S = s). \quad (2.14)$$

The prior $P(S = s)$ can be expressed as a MRF model. Thus it is served as a sum of the clique potential functions and can be expressed in the form similar to (2.9) as

$$P(S = s) \propto \exp(-U(s)). \quad (2.15)$$

The Likelihood can also be expressed in terms of likelihood energy in the similar form of (2.10) as

$$P(I = i | S = s) \propto \exp(-U(i | s)). \quad (2.16)$$

Thus the posteriori probability can be expressed in terms of energy function as

$$P(S = s | I = i) \propto \exp(-U(s | i)), \quad (2.17)$$

where

$$U(s | i) = U(i | s) + U(s). \quad (2.18)$$

Thus from (2.17) and (2.18), the MAP of the posteriori probability in (2.14) is equal to find the minimize of the posterior energy function

$$\hat{s} = \arg \min_{s \in \Omega} U(s | i). \quad (2.19)$$

With the use of MAP-MRF approach, a segmentation problem which is also referred to the labeling problem can be solved for the following steps. First, define the neighborhood system and the set of cliques. Then define a clique potential function and the likelihood energy function for the estimation of (2.15) and (2.16) respectively. Finally, choose an optimization algorithm to find the optimized MAP solution of the posteriori probability (2.14) or (2.17).

2.3.2. Application in Segmentation

The applications of MRF model have been widely use in a variety of image processing tasks such as image restoration, edge detection, motion analysis and image segmentation. In this section, we will focus on the MRF model established in the field of image segmentation.

2.3.2.1. Deng's Work

Deng's work [8] proposed a simple MRF model for unsupervised image segmentation. The segmentation problem can be expressed in the Bayesian framework (2.12) where the posteriori probability $P(S = s | I = i)$ consists of two components, a region labeling component and a feature modeling component. These two components can be formulated in the MRF model. The prior $P(S = s)$ is referred to the region labeling component and the energy function of the prior using the pairwise multi-level logistic (MLL) model is given as

$$U(s) = \sum_{q \in X} \left[\beta \sum_{q \in N(p)} \delta(s_p, s_q) \right], \quad (2.20)$$

where β is a weighting constant which can be specified a priori, s_p is a labeling condition of state p in the set of image lattice X and a clique potential function is defined as

$$\delta(s_p, s_q) = \begin{cases} 1, & s_p = s_q \\ -1, & s_p \neq s_q \end{cases}. \quad (2.21)$$

Deng assumes that the probability distribution of all feature data for one segment region is a Gaussian distribution. Based on this assumption, the likelihood energy which is referred to the feature modeling component that describes the features of an image is defined as

$$U(i | s) = \sum_s \left\{ \sum_{p \in X} \left[\frac{(i_p - \mu_l)^2}{2(\sigma_l)^2} + \log(\sqrt{2\pi}\sigma_l) \right] \right\}. \quad (2.22)$$

where i_p is the feature information of state p extract from the original image I , μ_l and σ_l are the mean and standard deviation for the segment region labeled l . Note that the number of segment regions is assumed to be known in prior. After defining the components, the energy of the posterior probability $P(S = s | I = i)$ is then defined as

$$U(s|i) = \alpha(t) \times U(i|s) + U(s) \quad (2.23)$$

where $\alpha(t)$ is the variable weighting parameter. Deng claims that with this function-based parameter gives the individually contribution of the two components to the entire energy $U(s|i)$, the proposed simple MRF model is able to automatically estimate model parameters and produce accurate unsupervised segmentation results using expectation-maximization (EM) algorithm and fast simulated annealing (SA).

2.3.3. Inference Algorithm Using Loopy Belief Propagation

There are several methods to solve the MAP-MRF problem such as simulated annealing (SA) [7], iterated conditional modes algorithm (ICM) [19], belief propagation (BP) [11], and graph cut method (GC) [20]. Among all, we are interested in the loopy belief propagation that applies Pearl's algorithm [10] to the graph with loops or undirected graphs. A Markov network is an undirected graph in the literature of probabilistic graph models [21], where the nodes represent variables and arcs which connect the neighboring nodes represent compatibility relations between neighboring nodes. Fig. 2.10 shows an example of an undirected graphical model. Yellow nodes represent the hidden variables and green nodes represent the observed variables. In this section, we will focus on the loopy belief propagation algorithm. We will refer loopy belief propagation as belief propagation for brevity.

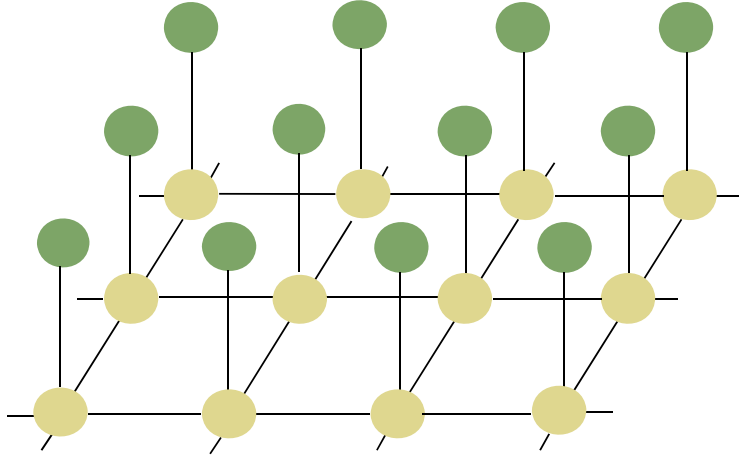


Fig. 2.10. An undirected graphical model with hidden variables and observed variables.

Pearl's algorithm is an exact inference algorithm for graphs without loops or directed graphs. For the graphs with loops or undirected graphs such as the image lattice structure, belief propagation is not guaranteed to find the global optimal solution. Despite of this, several applications have been successfully used in applications such as the one with stereo vision [11]. Belief propagation iteratively propagates messages in the Markov network. Let $m_{pq}^{t+1}(x_p, x_q)$ be the message that propagates from node x_p to x_q in iteration $t+1$, and is defined as

$$m_{pq}^{t+1}(x_p, x_q) \leftarrow \kappa \max_{x_p} \psi_{pq}(x_p, x_q) m_p^t(x_p, i_p) \prod_{x_k \in N(x_p) \setminus x_q} m_{kp}^t(x_k, x_p) \quad (2.24)$$

where $m_p^t(x_p, i_p)$ is the message from observed node i_p to node x_p in iteration t , $\psi_{pq}(x_p, x_q)$ is the compatibility matrix with size $L \times L$ between node x_p and its neighbor nodes x_q . L is the size of all possible labels. Note that both message $m_{pq}^{t+1}(x_p, x_q)$ and $m_p^t(x_p, i_p)$ are vectors with L elements. The belief of node x_p is computed as follows:

$$b_p(x_p) \leftarrow \kappa m_p(x_p, i_p) \prod_{x_k \in N(x_p)} m_{kp}(x_k, x_p), \quad (2.25)$$

$$x_p^{MAP} = \underset{x_p}{\operatorname{argmax}} b_p(x_p), \quad (2.26)$$

where κ is the normalized constant. Note that belief $b_p(x_p)$ is also a vector with L elements. Fig. 2.11 gives an example of local message passing in the Markov network.

The message propagates from node x_1 to node x_2 is $m_{1,2} \leftarrow \kappa \max_{x_1} \psi_{12}(x_1, x_2) m_1 m_{3,1} m_{4,1} m_{5,1}$.

The belief at node x_2 is $b_2 \leftarrow \kappa m_2 m_{1,2} m_{6,2} m_{7,2} m_{8,2}$.

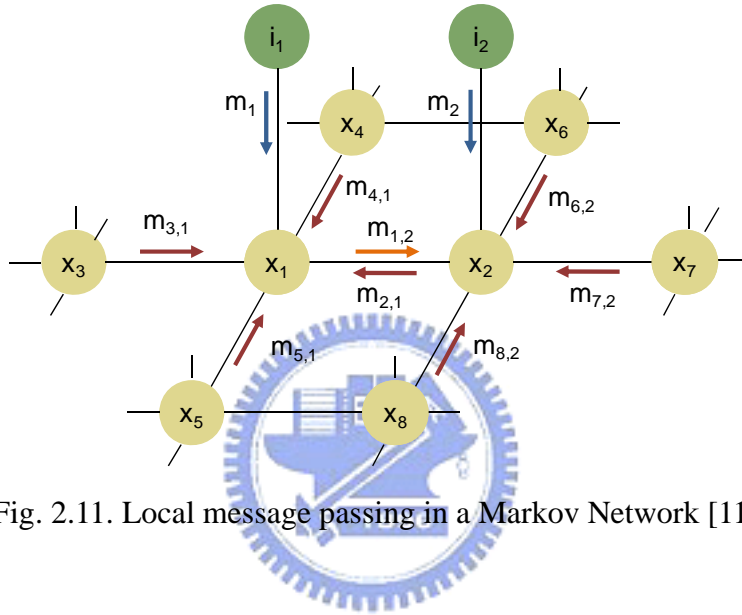


Fig. 2.11. Local message passing in a Markov Network [11]

The belief propagation with max-product update rule maximizes the joint posterior probability $P(X = x | I = i)$ with the MRF model in the following steps:

1. Initialize messages $m_{pq}(x_p, x_q)$ and $m_p(x_p, i_q)$ of each node to a constant and the observed values calculated from the likelihood function respectively.
2. Update messages of each node iteratively using equation (2.24).
3. Computes and determines the belief of the corresponding nodes using equation (2.25) and (2.26) at the end of the BP algorithm.

Chapter 3. Color Image Segmentation

Algorithm Using MRF Model

In this chapter, we propose a color image segmentation method with a new MRF modeling of color image segmentation and a local belief aggregation (LBA) algorithm to estimate the MAP inference. The proposed MRF models the likelihood and prior probability based on the concept of the intra and inter region constraint respectively. The LBA algorithm, which is inspired by the original BP and DQ, is proposed to cope with the problems of the large number of segment labels existing in color image segmentation.

3.1. Algorithm Overview

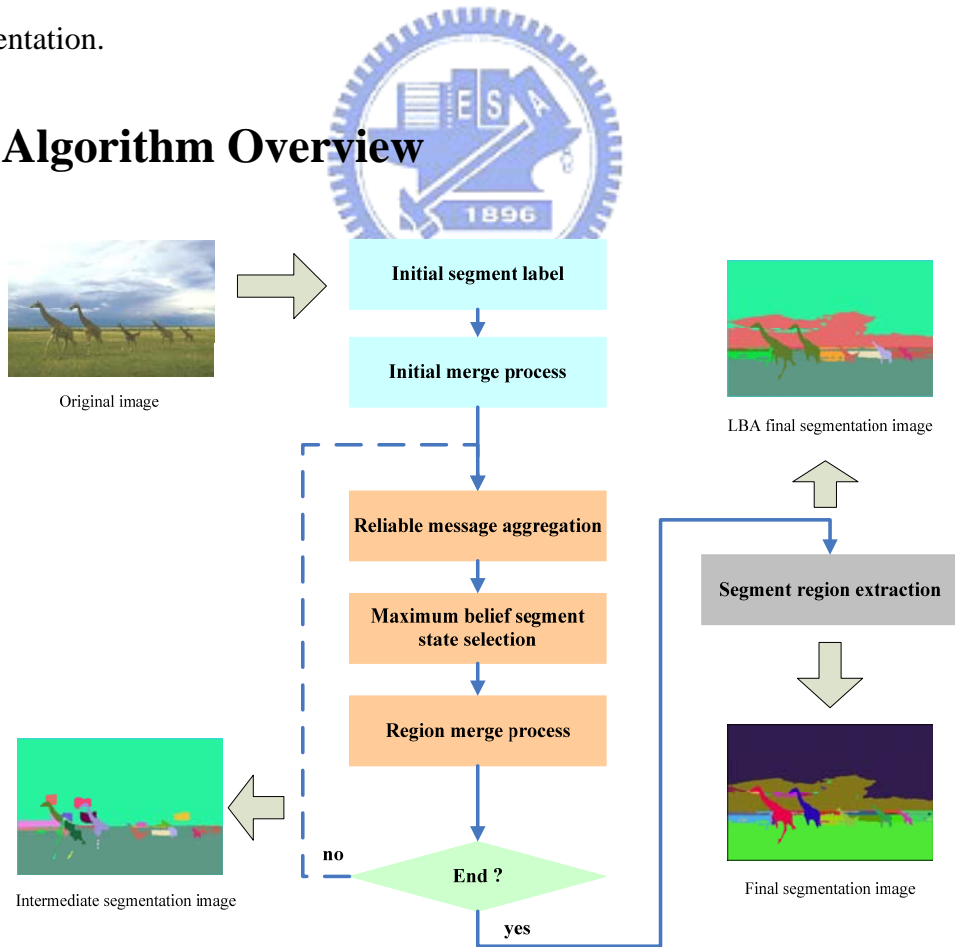


Fig. 3.1. Flow of image segmentation using local belief aggregation.

Fig. 3.1 illustrates the flow of the image segmentation using local belief aggregation. Before using local belief aggregation algorithm to optimize the total energy, an initial segment label is assigned. In this step, a unique segment label is first assigned to each pixel in the test image. If there are $|I|$ pixels in the image, there would be $|I|$ segment labels in the initial segmentation image. Note that there is no restriction on the quality of the initial segment map. However, a more accurate initial guess will lead to faster convergence of the following local belief aggregation algorithm. Besides, segment labels can also be used as the seeds for segment regions. Therefore, an initial merge process is applied to the initial segmentation image. After the initial merge process, the proposed local belief aggregation (LBA) is performed iteratively to find the segmentation image. We model the image segmentation as a labeling problem using a four-connect MRF, in which each node corresponds to a pixel and each state corresponds to a segment label. The LBA consists of reliable message aggregation, maximum belief segment state selection, and region merge process. First, reliable message aggregation aggregates reliable message information from the neighboring nodes of each node in the MRF. At each node, the belief value of each segment state is computed using the reliable message aggregated in the previous step. The segment state with the maximum belief value is chosen. For each node, the corresponding label of the chosen segment state is selected as the segment label. After the maximum belief segment state selection, a region merge is applied to further improve the segmentation image quality. The LBA steps are iteratively performed until convergence or a preset iteration limit is reached. After the LBA, segment region extraction is performed to output the final segmentation image. The segment region extraction assigns a new unique segment label to each region using connected component. This is an option step in the LBA algorithm. Note that we perform two different kinds of initial merge process,

a local partial merge and a global merge, as an initial segment map for the following LBA. A local partial merge that only considers spatial and color information in a 3×3 sliding window is used for the four-window based LBA. A global merge is used for one-window based LBA.

3.2. CIE $L^*a^*b^*$ Color Space

In the proposed color image segmentation algorithm, a precise estimation of color distance is important. Thus a proper choice of color space is important in our case. We adopt the CIE $L^*a^*b^*$ color space from all the existing color space for two reasons: (1) approximately uniform color scale, (2) similar to human visual perception.



3.2.1. Introduction

The $L^*a^*b^*$ color space is developed by the CIE to be approximately perceptually uniform. Color difference between points in the color space corresponds to the visual difference between the colors of the points. The L^* axis represents the lightness (luminance) in this color space with white at $L^* = 100$ and black at $L^* = 0$. The a^* and b^* axes represent color component while a^* encodes the red-green sensation and b^* encodes the yellow-blue sensation. Positive a^* axis indicates amounts of red color and negative a^* axis indicates amounts of green color while positive b^* axis indicates amounts of yellow color and negative b^* axis indicates amounts of blue color. Note that there is no specific numerical limit for these two color components. Fig. 3.2 illustrates a brief plot of the CIE $L^*a^*b^*$ color space.

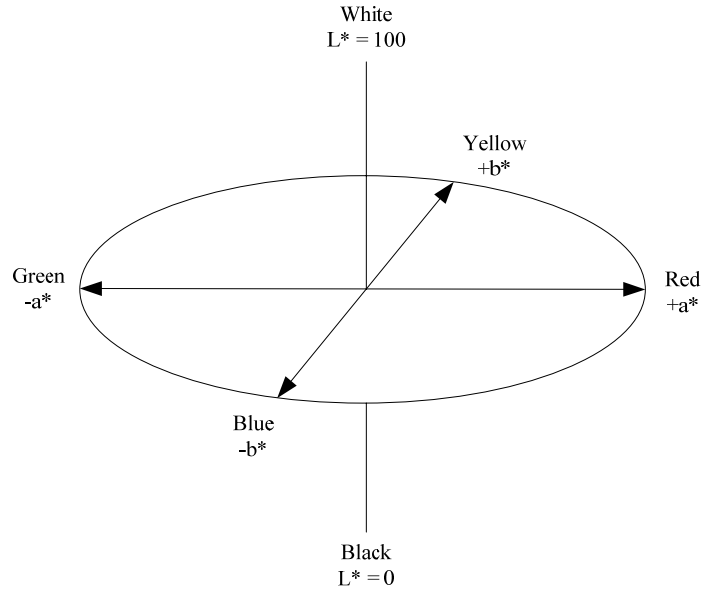


Fig. 3.2. CIE L*a*b* color space.

3.2.2. Color Transform from RGB to CIE L*a*b*

Color transformation from RGB to CIE L*a*b* color space is done by the following two steps. First we transform RGB to CIE XYZ space. This transformation is made by

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (3.1)$$

Then we transform the resulting CIE XYZ space to CIE L*a*b* space. The transformation is defined by

$$L^* = 116 \times \left[f\left(\frac{Y}{Y_n}\right) \right] - 16, \quad (3.2)$$

$$a^* = 500 \times \left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \right], \quad (3.3)$$

$$b^* = 200 \times \left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right], \quad (3.4)$$

where

$$f(t) = \begin{cases} t^{\frac{1}{3}} & \text{if } t > 0.008856, \\ 7.787 \times t + \frac{16}{116} & \text{if } t \leq 0.008856. \end{cases} \quad (3.5)$$

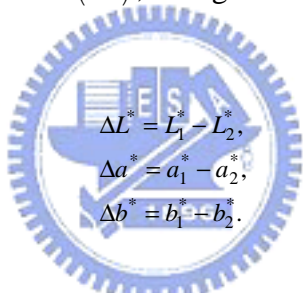
(X_n, Y_n, Z_n) : refernece white

3.2.3. Color Difference

Color difference between two points in the CIE L*a*b* space is given by the Euclidean distance formula

$$\Delta E_{ab}^* = \left[(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2 \right]^{\frac{1}{2}}, \quad (3.6)$$

where the differences in lightness (ΔL^*), red-green sensation (Δa^*) and yellow-blue sensation (Δb^*) is defined as



$$\begin{aligned} \Delta L^* &= L_1^* - L_2^*, \\ \Delta a^* &= a_1^* - a_2^*, \\ \Delta b^* &= b_1^* - b_2^*. \end{aligned} \quad (3.7)$$

3.3.MRF Model Formulation

According to the Bayes' theorem, the posterior probability of a segmentation image S , given the image information I , can be represented as (2.12) and simplified to (2.13) since $P(I = i)$ in (2.12) can be considered as a constant. In this section, we formulate the MRF model of the likelihood $P(I = i | S = s)$ and the prior $P(S = s)$ in order to further estimate the posterior probability to obtain the segmentation result.

3.3.1. Likelihood and Prior

We define the likelihood $P(I = i | S = s)$ as

$$P(I = i | S = s) \propto \prod_{x \in I} \exp(-F(x, s_x, i)), \quad (3.8)$$

where $F(x, s_x, i)$ is the cost function of node x with segment label s_x given the observation I . The prior $P(S = s)$ is defined as

$$P(S = s) \propto \prod_{x \in I} \prod_{y \in N(x)} \exp(-\eta(s_x, s_y, \gamma(s_x, s_y))), \quad (3.9)$$

where $\eta(s_x, s_y, \gamma(s_x, s_y))$ is the clique potential function of segment label s_x and s_y in which node y is the neighbor of node x . $\gamma(s_x, s_y)$ is the line process which penalizes the clique potential according to the relationship between segment label s_x and s_y . By combining (2.13), (3.8) and (3.9), the basic model (2.12) of the image segmentation becomes

$$P(S = s | I = i) \propto \prod_{x \in I} \exp(-F(x, s_x, i)) \prod_{x \in I} \prod_{y \in N(x)} \exp(-\eta(s_x, s_y, \gamma(s_x, s_y))). \quad (3.10)$$

3.3.2. Model Approximation

To estimate the optimal solution of a MRF model, the maximum a posteriori (MAP) approach is used. Taking the MAP of (3.10) gives

$$\begin{aligned} & \max\{P(S = s | I = i)\} \\ &= \max\left\{ \prod_{x \in I} \exp(-F(x, s_x, i)) \prod_{x \in I} \prod_{y \in N(x)} \exp(-\eta(s_x, s_y, \gamma(s_x, s_y))) \right\} \\ &= \max\left\{ \prod_{x \in I} \exp\left[-\left(F(x, s_x, i) + \sum_{y \in N(x)} \eta(s_x, s_y, \gamma(s_x, s_y))\right)\right] \right\} \\ &= \exp\left\{-\min \sum_{x \in I} \left[F(x, s_x, i) + \sum_{y \in N(x)} \eta(s_x, s_y, \gamma(s_x, s_y))\right]\right\}. \end{aligned} \quad (3.11)$$

The question here is to determine the definition of the cost function $F(x, s_x, i)$ and the clique potential function $\eta(s_x, s_y, \gamma(s_x, s_y))$ in the likelihood and the prior respectively.

Haralick and Shapiro [22] pointed out two criteria on the characteristics of segments in a good segmentation image. One measures the intra region uniformity and the other

measures the inter region disparity. These two criteria show consistency with the likelihood and prior function. Thus we model the likelihood and prior function based on the intra and inter region criteria respectively. Specifically, the definition of intra and inter region energy defined in [23] is adopted and modified. We define the cost function and the clique potential function as

$$F(x, s_x, i) = u\left(\|i_x - \bar{s}_x\|_{L^*a^*b^*} - th_v\right), \quad (3.12)$$

$$\eta(s_x, s_y, \gamma(s_x, s_y)) = u\left(th_v - \|\bar{s}_x - \bar{s}_y\|_{L^*a^*b^*}\right) \times \gamma(s_x, s_y), \quad (3.13)$$

where \bar{s}_x and \bar{s}_y represents the average color of segment label s_x and s_y respectively, th_v denotes the threshold for visible color difference, $\|\cdot\|_{L^*a^*b^*}$ denotes the color difference in the CIE L*a*b* space and $\gamma(s_x, s_y)$ denotes the relation between a pair of segment labels and is defined as

$$\gamma(s_x, s_y) = \begin{cases} 1, & s_x \neq s_y \\ 0, & s_x = s_y \end{cases}. \quad (3.14)$$

Thus, with the above definition of cost function and clique potential function, (3.11) can be rewritten as

$$\begin{aligned} & \max\{P(S = s | I = i)\} \\ & = \exp\left\{-\min \sum_{x \in I} \left[F(x, s_x, i) + \sum_{y \in N(x)} \eta(s_x, s_y, \gamma(s_x, s_y)) \right]\right\} \\ & = \exp\left\{-\min \left[\sum_{x \in I} F(x, s_x, i) + \sum_{x \in I} \sum_{y \in N(x)} \eta(s_x, s_y, \gamma(s_x, s_y)) \right]\right\} \\ & = \exp\left\{-\min[E_{intra} \times |I| + E_{inter} \times C \times |I|]\right\} \\ & = \exp\left\{-\min[E_{cw}]\right\} \end{aligned} \quad (3.15)$$

where E_{cw} is the weighted sum of intra-region visual error (E_{intra}) and inter-region visual error (E_{inter}) whose definition from [23] are

$$E_{intra} = \frac{\sum_{x \in I} u\left(\|i_x - \bar{s}_x\|_{L^*a^*b^*} - th_v\right)}{|I|}, \quad (3.16)$$

$$E_{inter} = \frac{\sum_{x \in I} \sum_{y \in N(x)} \alpha(s_x, s_y) \times u\left(th_v - \|\overline{s_x} - \overline{s_y}\|_{L^*a*b^*}\right)}{C \times |I|}, \quad (3.17)$$

$$\alpha(s_x, s_y) = \begin{cases} 1, & s_x \neq s_y \\ 0, & s_x = s_y \end{cases}, \quad (3.18)$$

$$u(t) = \begin{cases} 1, & t > 0 \\ 0, & t \leq 0 \end{cases}. \quad (3.19)$$

$|I|$ represents the total number of pixels in an image, C is a weighting constant. From the inference of (3.15), it is obvious that estimating a MAP of $P(S=s|I=i)$ with the definition of cost function and clique potential function in (3.12) and (3.13) is equivalent to obtaining the minimum energy of E_{cw} . This equivalence is what we desired since [23] claimed that lower E_{cw} implies better image segmentation results.

In addition to the definition of intra-region visual error and inter-region visual error defined in (3.16) - (3.19), Chen also claims that the modified version of these two terms that further include the color distance values can also give a quantitative evaluation of the segmentation images [24]. The definition of the modified intra-region visual error (ME_{intra}) and modified inter-region visual error (ME_{inter}) is given as follows

$$ME_{intra} = \frac{\sum_{x \in I} \left\{ u\left(\|i_x - \overline{s_x}\|_{L^*a*b^*} - th_v\right) \times \left(\|i_x - \overline{s_x}\|_{L^*a*b^*} - th_v\right) \right\}}{|I|}, \quad (3.20)$$

$$ME_{inter} = \frac{\sum_{x \in I} \sum_{y \in N(x)} \left\{ \alpha(s_x, s_y) \times u\left(th_v - \|\overline{s_x} - \overline{s_y}\|_{L^*a*b^*}\right) \times \left(th_v - \|\overline{s_x} - \overline{s_y}\|_{L^*a*b^*}\right) \right\}}{C \times |I|}, \quad (3.21)$$

$$\alpha(s_x, s_y) = \begin{cases} 1, & s_x \neq s_y \\ 0, & s_x = s_y \end{cases}, \quad (3.22)$$

$$u(t) = \begin{cases} 1, & t > 0 \\ 0, & t \leq 0 \end{cases}. \quad (3.23)$$

Again, if we define the cost function and the clique potential function as

$$F(x, s_x, i) = u\left(\|i_x - \bar{s}_x\|_{L^*a^*b^*} - th_v\right) \times \left(\|i_x - \bar{s}_x\|_{L^*a^*b^*} - th_v\right), \quad (3.24)$$

$$\eta(s_x, s_y, \gamma(s_x, s_y)) = u\left(th_v - \|\bar{s}_x - \bar{s}_y\|_{L^*a^*b^*}\right) \times \left(th_v - \|\bar{s}_x - \bar{s}_y\|_{L^*a^*b^*}\right) \times \gamma(s_x, s_y), \quad (3.25)$$

$$\gamma(s_x, s_y) = \begin{cases} 1, & s_x \neq s_y \\ 0, & s_x = s_y \end{cases}, \quad (3.26)$$

and take the MAP of (3.10), the same conclusion will make as in (3.15).

Although maximizing the posterior defined by (3.12), (3.13) and (3.24), (3.25) can minimize E_{cw} as shown in (3.15), an energy distribution that is proportional to color difference is considered to be a more suitable measure. Thus, based on the property of intra region and inter region, we re-formulate the cost function and clique potential function as

$$F(x, s_x, i) = \|i_x - \bar{s}_x\|_{L^*a^*b^*}, \quad (3.27)$$

$$\eta(s_x, s_y, \gamma(s_x, s_y)) = \begin{cases} \|\bar{s}_x - \bar{s}_y\|_{L^*a^*b^*} & s_x = s_y \\ th_s - \|\bar{s}_x - \bar{s}_y\|_{L^*a^*b^*} & s_x \neq s_y \end{cases}, \quad (3.28)$$

where th_s denotes the threshold for maximum difference of average segment label color between two segment labels. From the empirical experiment results, we select threshold th_s to be 150 to simplify the problem of finding different threshold for different test images. Thus, in the case of different segment labels detected in (3.28), we truncate the value of clique potential function to zero if $\|\bar{s}_x - \bar{s}_y\|_{L^*a^*b^*}$ is larger than the 150, which is the predefined threshold th_s as shown in Fig. 3.3.

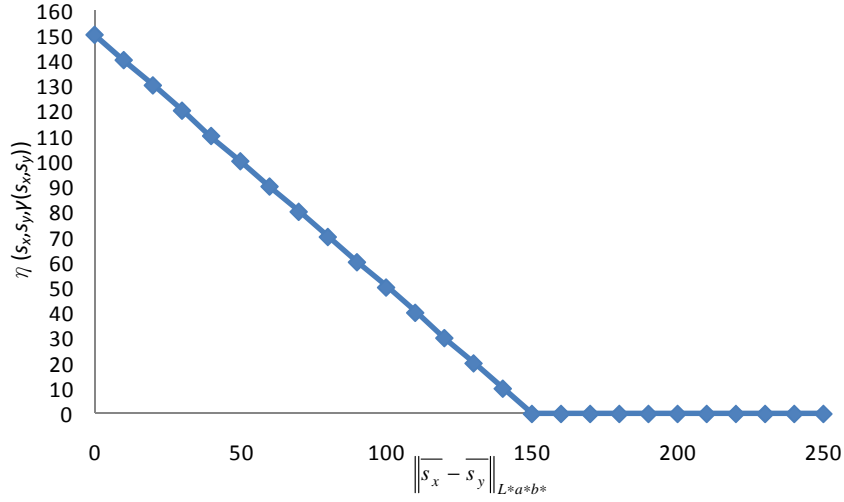


Fig. 3.3. Relation between the value of clique potential function and the value of difference of average segment label color for two different segment labels.

Our formulation adopts Gaussian-like color distribution model; however, [25] has pointed out that such model may not always be true. To accommodate the distribution deviations, two discontinuity preserving robust functions derived from the Total Variance (TV) model [26] are applied to the cost and clique potential function.

$$P(S = s | I = i) \propto \prod_{x \in I} \exp(-\rho_l(s_x)) \prod_{x \in I} \prod_{y \in N(x)} \exp(-\rho_p(s_x, s_y)), \quad (3.29)$$

where the robust functions are defined as

$$\rho_l(s_x) = -\ln \left[(1 - e_l) \exp\left(-\frac{|F(x, s_x, i)|}{\sigma_l}\right) + e_l \right], \quad (3.30)$$

$$\rho_p(s_x, s_y) = -\ln \left[(1 - e_p) \exp\left(-\frac{|\eta(s_x, s_y, \gamma(s_x, s_y))|}{\sigma_p}\right) + e_p \right]. \quad (3.31)$$

Note that parameter σ and e control the sharpness and upper-bound of the function respectively. In our experiment, the parameter σ and e is set to be 2.0 and 0.01 respectively for both cost function and clique potential function.

Finally, the posterior $P(S = s | I = i)$ can be factorized into the following form

$$P(S = s | I = i) \propto \prod_{x \in I} \psi_x(s_x, i_x) \prod_{x \in I} \prod_{y \in N(x)} \psi_{xy}(s_x, s_y), \quad (3.32)$$

where $\psi_x(s_x, i_x) \propto \exp(-\rho_l(s_x))$ is the local evidence for node x , and $\psi_{xy}(s_x, s_y) = \exp(-\rho_p(s_x, s_y))$ is the compatibility matrix between node x and its neighbor nodes y .

3.4. Local Belief Aggregation

To efficiently estimate the MAP of the posterior probability (3.32), a local belief aggregation (LBA) is proposed. Although belief propagation (BP), which is a linear time algorithm proportional to the number of hidden nodes [11], can also be used to solve the MAP problem, there are some difficulties to directly apply BP algorithm in the proposed MRF-based segmentation model. The enormous size of segment labels not only results in heavy computational burden but also leads to a huge memory storage requirement. Both of these constraints restrict the use of BP in color image segmentation. Thus, a local belief aggregation which is modified from the original BP algorithm is proposed to find the MAP segmentation image.

3.4.1. Reliable Message Aggregation

Reliable message aggregation is the first step in the LBA. The reliable message aggregation restricts the number of segment state's message to be aggregated from the neighbor nodes. That is, only a limited number of the most probable segment states, which we considered to be reliable, can send out messages. To decide the most probable segment states, a local search window approach is used. Here we introduce

two methods of using local search window to send out the reliable messages. In section 3.4.1.1, we will introduce the four-window based local search window method. In section 3.4.1.2, one window-based local search window method is proposed to further reduce the computational complexity than the previous method.

3.4.1.1. Four-window Based Local Search Window

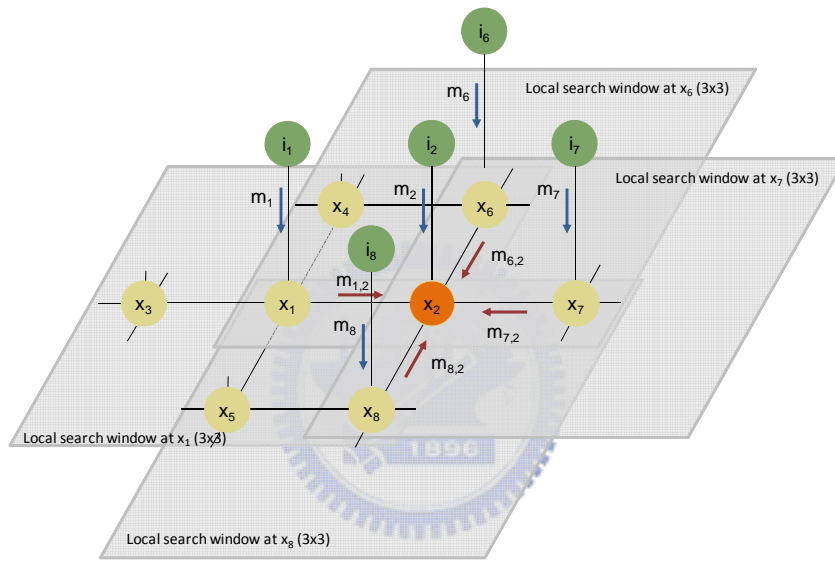


Fig. 3.4. Reliable message aggregation using four local search windows.

Fig. 3.4 shows the concept of four-window based local search window. In this method, each neighboring node has its corresponding local search window in a preset size. For example, nodes x_1 , x_6 , x_7 and x_8 have their corresponding 3×3 local search window shown in Fig. 3.4. During the procedure, the message of node x_1 will be calculated and decided using the local search window center at node x_1 . Once the reliable message is ready, message of node x_1 will propagate to the current node x_2 . The same action is performed for node x_6 , x_7 and x_8 at the same time to propagate the reliable messages to node x_2 .

Let m be the number of all the segment states appeared in the local search window. All messages from these m segment states will be calculated. Let $m_{xy}(s_x, s_y)$ be the message from node x to node y , and is defined as

$$m_{xy}(s_x, s_y) \leftarrow \kappa \max_{s_x} \psi_{xy}(s_x, s_y) v_x(s_x, s_z) m_x(s_x, i_x) \prod_{s_k \in N(s_x) \setminus s_k} r_{kx}(s_k, i_k), \quad (3.33)$$

where $v_x(s_x, s_z)$ is the spatial function considering the spatial relationship of the m segment states in node x and the segment state of node z in the corresponding local search window. Here our spatial function is simply defined as the reciprocal of the spatial distance

$$v_x(s_x, s_z) = \frac{1}{Eu_d(s_x, s_z)}, \quad (3.34)$$

where $Eu_d(s_x, s_z)$ is the Euclidean distance between the segment state of node x and the segment state of node z . $m_x(s_x, i_x) = \psi_x(s_x, i_x)$ is the message from observed node i_x to node x and $r_{kx}(s_k, i_k) = \psi_k(s_k, i_k)$ is the message from node k to node x . If the number of possible segment state m in the local search window is larger than a preset number of reliable segment states n , then only the message of the most probable n segment states out of the m segment states can be transferred according to the message value calculated by (3.33). Higher message value represents higher probability. The other $(m-n)$ number of less probable segment states will be discarded for current node calculation. From (3.33), it is obvious that the message can be calculated on-the-fly, hence no memory storage for previous iteration's messages is required any more.

3.4.1.2. One-window Based Local Search Window

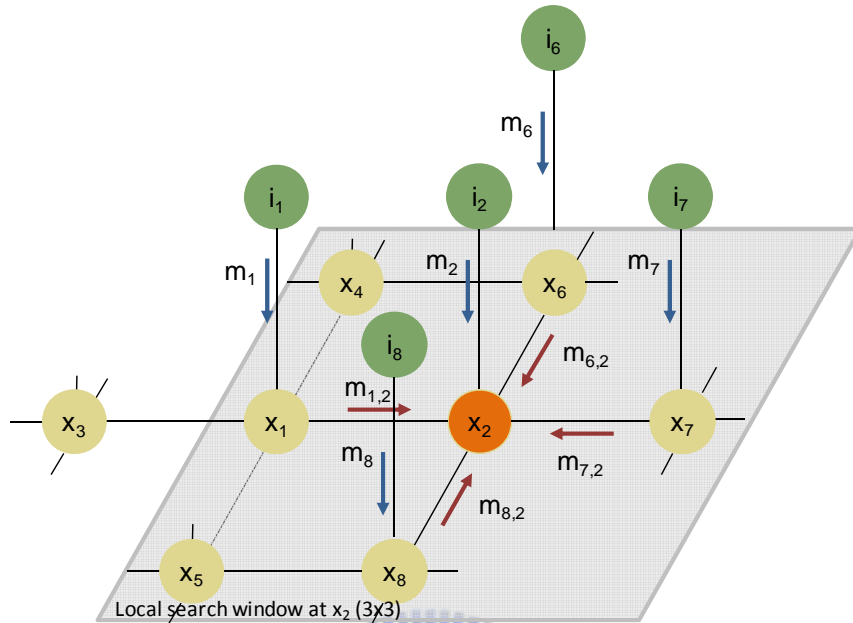


Fig. 3.5. Reliable message aggregation using one local search window.

Fig. 3.5 shows the concept of reliable message aggregation using one-window based local search window. In this method, only one local search window with preset size is required to determine the reliable messages. Let m be the number of all the segment states appeared in the local search window. All messages from these m segment states will be calculated using the same equation (3.33). All the m number of messages in four directions will be considered as reliable messages and will aggregate into the node without being discarded to provide reliable information for further belief value calculation.

3.4.2. Maximum Belief Segment State Selection

The maximum belief segment state selection is performed after the reliable message aggregation. During this procedure, the belief of the node x will be estimated and the segment state with the maximum belief will be selected. The belief is computed as follows:

$$b_x(s_x) \leftarrow \kappa n_x(s_x, i_x) \prod_{s_k \in N(s_x)} m_{kx}(s_k, s_x), \quad (3.35)$$

$$s_x^{MAP} = \arg \max_{s_x} b_x(s_x). \quad (3.36)$$

Note that in the four-window based local search window method, there is a chance that some of the segment state's message is missing while computing the belief value. This is due to the restriction on the number of the reliable message to be sent from the neighboring nodes with their corresponding local search window. To remedy this, the missing message belonging to a specific segment state of the corresponding neighboring node is re-computed. For one-window based local search window method, there is no need to re-compute since no restriction on the number of the reliable message is performed.

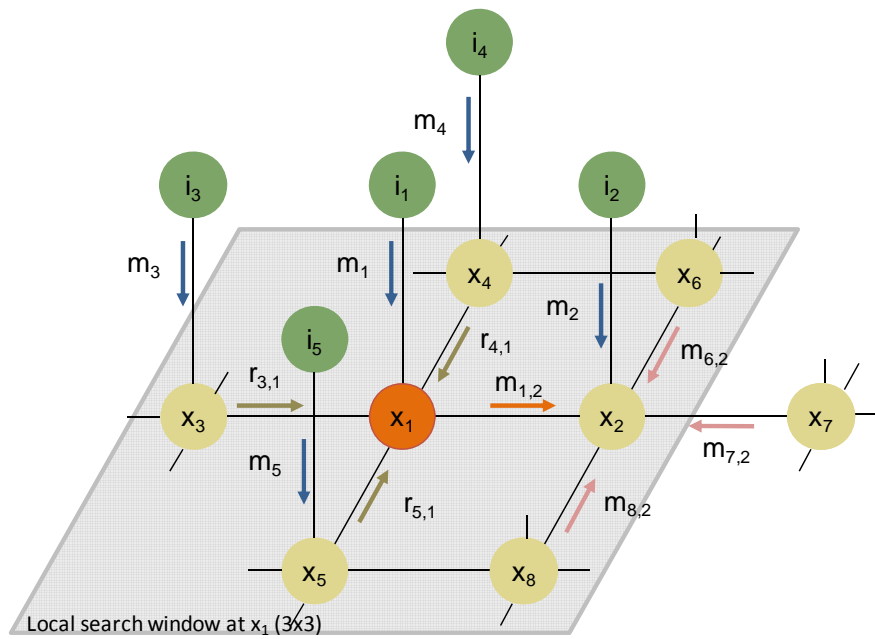


Fig. 3.6. Four-window based local reliable message aggregation in a Markov network.

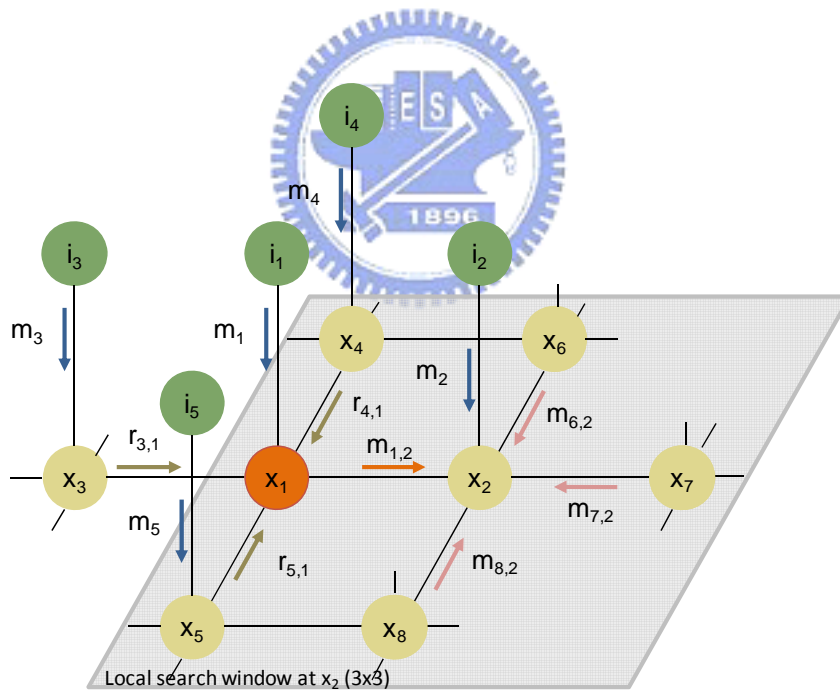


Fig. 3.7. One-window based local reliable message aggregation in a Markov network.

Fig. 3.6 and Fig. 3.7 demonstrate the example of the message aggregation in four-window based local search window and one-window based local search window with belief calculation in a Markov network respectively. In both figures, green nodes

represent the observed variables. Yellow and orange nodes represent the hidden variables. Gray region represents a preset size of local search window which is 3×3 in the example centered at nodes x_1 and x_2 in the four-window based method and one-window based method respectively. Reliable message aggregate from node x_1 to x_2 is $m_{1,2} \leftarrow \kappa \max_{s_{x_1}} \psi_{12}(s_{x_1}, s_{x_2}) \nu_1(s_{x_1}, s_{x_2}) m_1 m_3 m_4 m_5$. The belief at node x_2 is

$$b_{x_2} \leftarrow \kappa m_2, m_{1,2}, m_{6,2}, m_{7,2} m_{8,2}.$$

3.4.3. Region Merge Process

The initial guess of the segmentation image may consist of a large number of unnecessary segment labels. As a consequence, we would be using more than one segment labels to represent a region. This would prevent the overall energy from converging. For the above reason, redundant segment labels should be pruned. Hence, a region merging process is inserted in the end of each iteration. At the end of each iteration, the average color difference of two neighboring segment regions is checked. If the color difference is smaller than a pre-defined threshold, the two segment regions should be the same. In other words, the two different segment labels represent the same region. In this case, one of the two segment labels is replaced by the other.

Additional region merging based on the area of regions is also performed to further improve the quality of the segmentation image. In the four-window based local search window, the additional region merging is performed at the first and last iteration of the LBA as shown in Fig. 3.8. In the one-window based local search window, the additional region merging is performed twice, after the region merge process at the first and last iteration of the LBA as shown in Fig. 3.9. In this additional region merge,

segment regions with an area smaller than a predefined area are merged into their neighboring segment regions that have the smallest color distance with an area larger than a predefined number. The predefined areas are both 20 pixels in our case.

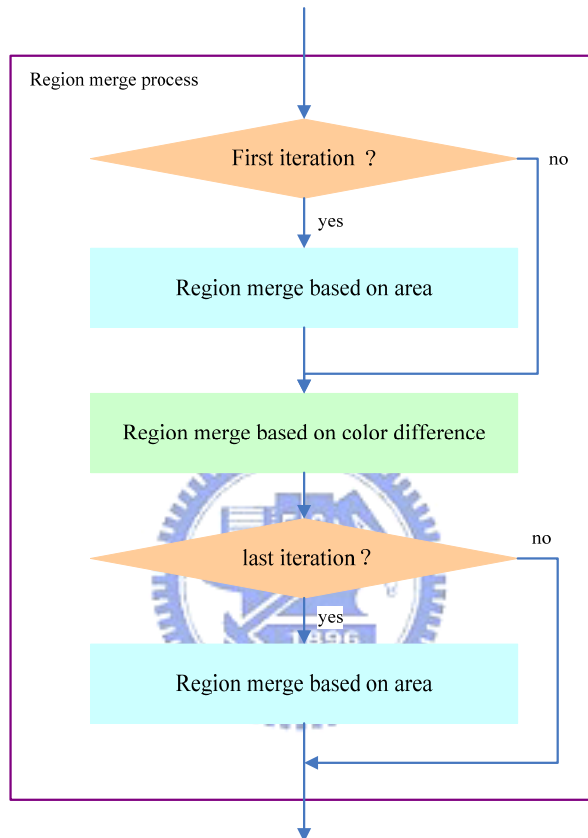


Fig. 3.8. Region merge process in four-window based method.

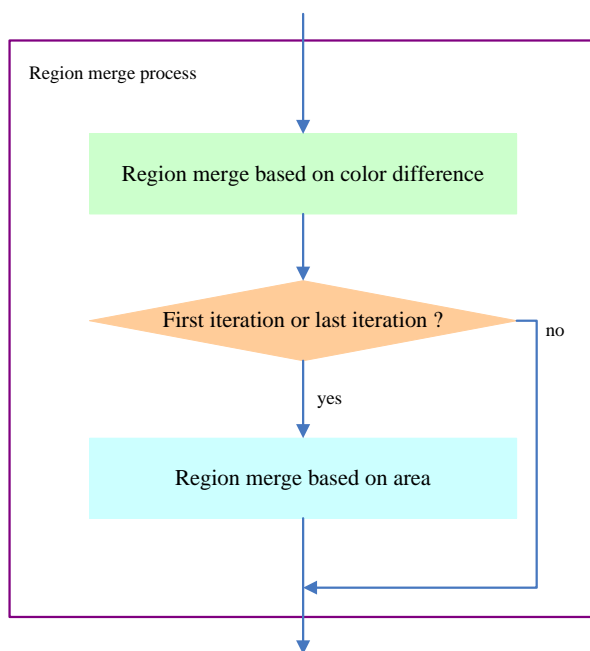


Fig. 3.9. Region merge process in one-window based method.

3.4.4. Convergence of Local Belief Aggregation

Similar to BP, there is no theoretical proof to guarantee the convergence of the proposed local belief aggregation method. However, we suspect that LBA is likely to achieve convergence in practice. We use empirical result to demonstrate LBA's convergence trend. For the purpose of minimizing the energy term E_{cw} , it is reasonable that the convergence of the energy term E_{cw} could imply the convergence of the proposed local belief aggregation algorithm. Fig. 3.10 shows the E_{cw} curve for the test image 253055 from Berkeley Segmentation Database [27]. As we can see, after several iterations of estimating posterior probability by LBA, the energy term E_{cw} oscillates around the value 0.136. Hence, we believe that the LBA tends to converge as the E_{cw} energy term is converging.

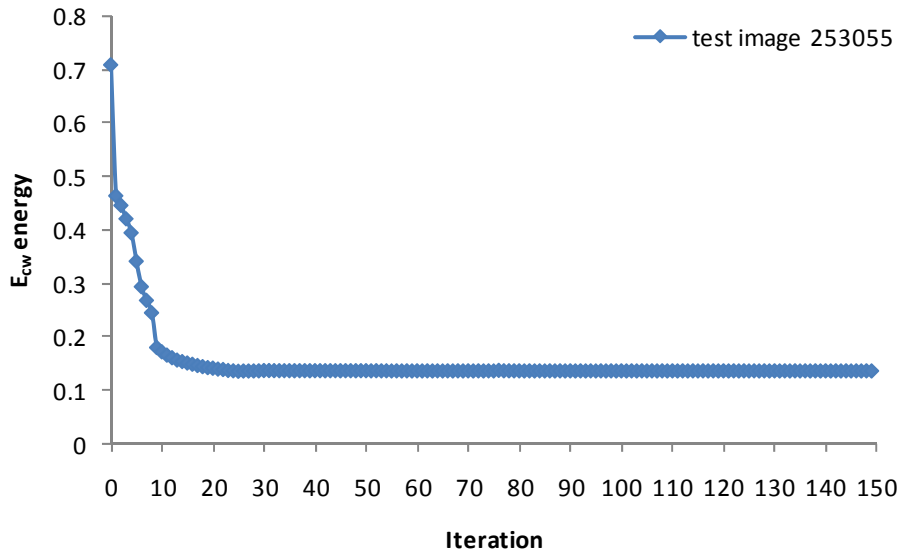


Fig. 3.10. E_{cw} convergence of four-window based LBA

In addition to the E_{cw} curve, another empirical method to demonstrate the convergence trend is to check the results of the segmentation image subjectively. If the result of the segmentation images is convergence, than we believe that the LBA also tends to converge. For test images we use, this convergence tend is guaranteed so far.

3.5. MRF Model Comparison

We suspect that the MRF model using (3.12), (3.13) and (3.24), (3.25) is not suitable for the proposed algorithm for two reasons. First, two-value model provides less information for a labeling problem such as the image segmentation problem. An energy distribution proportional to color difference can give a Gaussian-like measure and thus is considered to be a more suitable measure. Second, in the proposed LBA algorithm, the region merge process will merge two neighbor segment regions with

average color distance smaller than a predefined threshold. If the region merge threshold is selected larger than or equal to the visible color difference threshold th_v , then the input of the unit step function in (3.13) and (3.25) will always be smaller or equal to zero. From (3.19), this will cause the results of unit step function that appears in (3.13) and (3.25) to always be equal to zero. Thus the concept of unit step function is considered not suitable for the region merge process of the proposed LBA. Fig. 3.11 shows the LBA results using (3.12), (3.13) and (3.24), (3.25) respectively. The region merge threshold is chosen to be 2 in order to be less than the threshold for visible color th_v and we run the LBA for 5 iterations. All two images perform bad segmentation results, both methods cannot successfully segment out giraffes of the test image 253055; however, Fig. 3.11 (b) is better than Fig. 3.11 (a) among the two. We suspect that this is because MRF model used in the Fig. 3.11 (b) has include the color distance values which is better than the two-value MRF model used in Fig. 3.11(a). From the above reasons and the empirical experiment, (3.27) and (3.28) is adopted as the energy of the likelihood and the prior respectively. The segmentation results of the selected MRF model will be demonstrate in the next section and throughout the thesis.

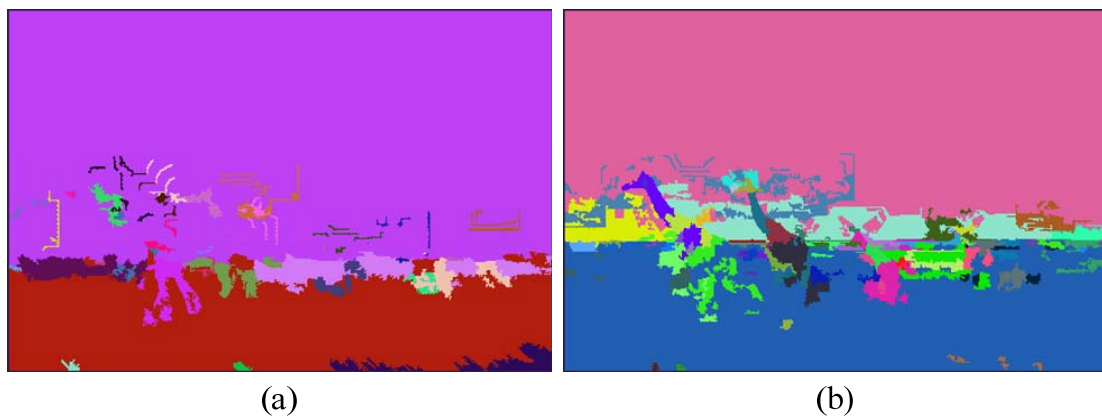


Fig. 3.11. Segmentation results with different MRF model using LBA algorithm. (a) MRF model using (3.12) and (3.13). (b) MRF model using (3.24) and (3.25).

3.6. Complexity Analysis and Comparison

We suspect that one-window based reliable message aggregation is likely to achieve less computational complexity than the four-window based method when executes one LBA iteration. To examine the assumption, the empirical results are used to demonstrate the computational complexity of above two methods. With the same parameter settings, we run LBA on both methods for 50 iterations and then calculate average counts of each computational element. Parameter settings used in this experiment is shown in Table 3.1. Since there is no restriction on the number of the reliable message for one-window based method, the corresponding blank is vacant. Table 3.2 demonstrates the average execution counts of both two methods. As we can see from the table, each computational element of one-window based LBA reduces from 27% ~ 86% than four-window based LBA does. Hence, we believe that computational complexity of one-window based LBA is much less than four-window based LBA. Fig. 3.12 demonstrates the reduced rate of each computational element in graph. Note that it should be noticed that our codes have not been optimized yet. Thus the execution counts in Table 3.2 are just for reference. However, we believe that the execution counts for both methods will be much improved if the codes have been optimized and the conclusion will still be the same as we concluded above.

Table 3.1. Parameter settings of image 253055 for LBA.

	Local search window	Region merge threshold	LBA iteration	Number of reliable message	Area merge threshold
1-window based	5×5	5	50	X	20/20
4-window based	5×5	5	50	6	20/20

Table 3.2. Average execution counts of image 253055 for LBA.

	Load/store	Add/sub	Div	Mul	Comp	Sqrt	Exp
1-window-based	3.15E+08	7.37E+07	2.62E+07	3.64E+07	4.74E+07	7.51E+06	7.49E+06
4-window-based	6.50E+08	1.30E+08	3.58E+07	7.07E+07	3.43E+08	2.18E+07	1.35E+07
Reduced rate	52%	43%	27%	49%	86%	66%	44%

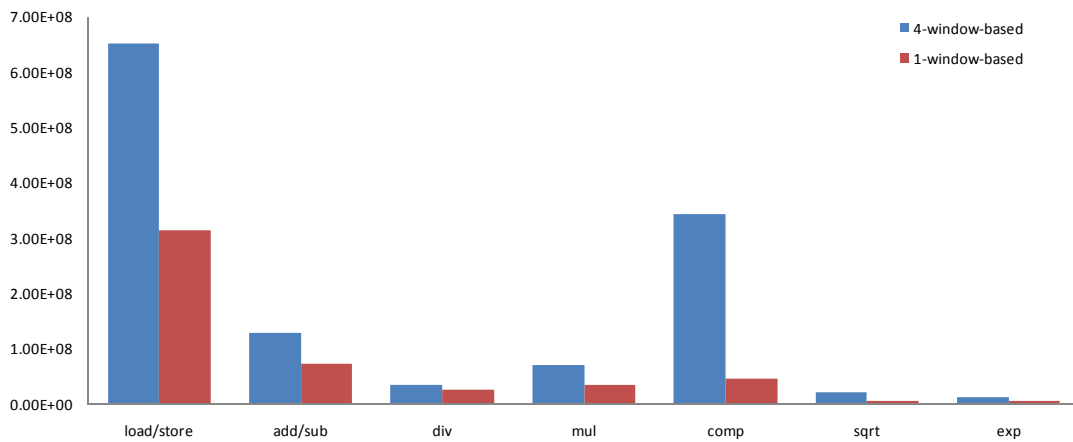


Fig. 3.12. Comparison of each computational unit between different methods used in the proposed LBA.

The segmentation results of both methods in the same parameter settings mentioned above are also presented in Fig. 3.13. As we can see, four-window based method segments out more details as shown in the lag of the first and second giraffe. Although four-window based method restricts the number of reliable messages to be aggregated from the neighbor nodes, it can obtain more neighbor segment labels than one-window based method when using the same size of the local search window. Thus it is possible for four-window based method to aggregate more reliable message than the other based on the same parameters does and hence can segment out a more detail one than one-window based method.

It should also be noticed that the parameter settings used in here may not be the best one of all. Therefore, a quantitative and qualitative evaluation of these two methods will be compared in Chapter 4.

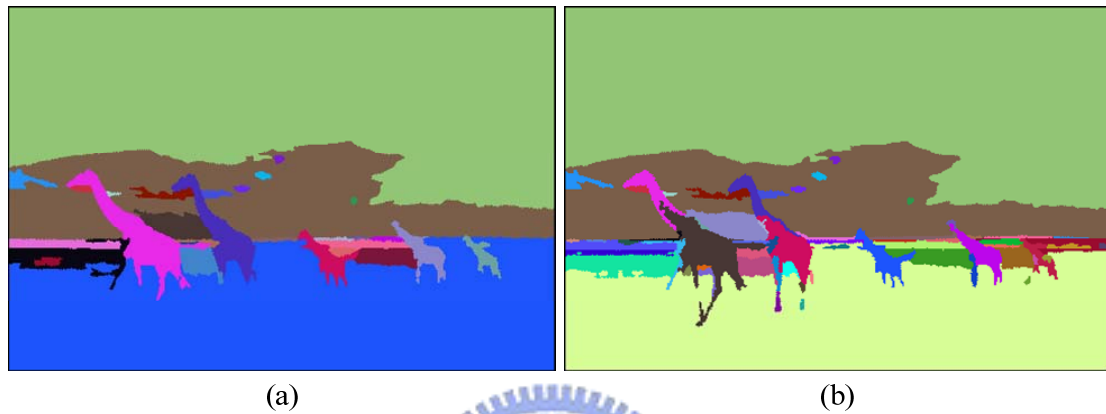
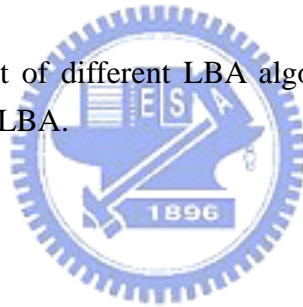


Fig. 3.13. Segmentation result of different LBA algorithms. (a) One-window based LBA. (b) Four-window based LBA.



Chapter 4. Experimental Results and Analysis

4.1. Introduction

In this chapter, we present the experimental results of the proposed LBA algorithm on test images. The test images are used from the Berkeley Segmentation Database. E_{cw} graph [23] is adopted for objective quantitative performance evaluation. E_{cw} is the weighted sum of E_{intra} and E_{inter} which measures the intra-region visual error and inter-region visual error respectively. According to the definition of these two terms (3.16) - (3.19), higher E_{intra} and E_{inter} value implies that the segment result is under-segmentation and over-segmentation respectively. Thus, to achieve better segmentation result, E_{cw} should be minimized. For this reason, the curve that is closer to the origin is considered as a better segmentation algorithm [23]. However, current existing unsupervised evaluation methods are far from perfect in evaluating segmentation algorithms [25], thus the result segmentation images are also provided for subjective evaluation. In addition to the performance of the proposed four-window based LBA method, we included the performance of mean-shift and watershed methods for comparison. The source codes of mean shift and watershed are provided from [6] and [15] respectively.

4.2. Segmentation Results

4.2.1. Chen's Quantitative Evaluation Metrics

Chen and Wang [23] proposed an unsupervised evaluation method based on the concept of visible color difference to mimic the way human perceives the color difference for color image segmentation. This method, which is abbreviated as E_{cw} , defines two measurements called intra-region visual error and inter-region visual error to evaluate the quality of color segmentation.

Intra-region visual error which is abbreviated as E_{intra} is used to evaluate the degree of under-segmentation. In each segment region, color information of each pixel should be as close as the average color of the segment region. Thus, intra-region visual error defined as (3.16) is used to measure the total number of pixels in each segment region with visual color difference larger than a pre-defined visual threshold away from the average color. Intuitively, the value of intra-region visual error should be as small as possible in a properly segmented region. Higher intra-region visual error implies the higher probability of missing boundaries in a segment region, thus the higher probability that under-segmentation results it has. In this evaluation method, a CIE $L^*a^*b^*$ color space is adopted and the color difference (3.6) which is larger than the value 6 is considered as visible as shown in Table 4.1 [28]. Hence the visual threshold is chosen to be 6 throughout this evaluation method.

Table 4.1. Effect of color difference in CIE L*a*b* color space for human's visual perception [28].

ΔE_{ab}^*	Effect
< 3	Hardly perceptible
$3 < 6$	Perceptible, but acceptable
> 6	Not acceptable

Inter-region visual error which is abbreviated as E_{inter} is used to evaluate the degree of over-segmentation. Intuitively, color difference between two neighbor segment regions should be as far as possible. Thus inter-region visual error defined as (3.17) is used to measure the total number of pixels with joint length between two neighbor segment regions in which average color of these two neighbor regions is smaller than a visual threshold. Higher inter-region visual error implies higher probability that false boundaries is made between the neighbor regions, thus the higher probability that over-segmentation results it has.

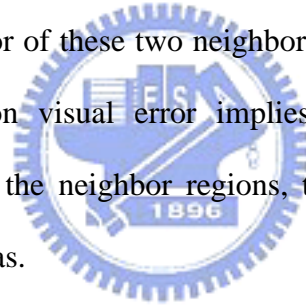


Fig. 4.1 [23] shows the E_{cw} graph which combines the intra-region visual error and the inter-region visual error to give a quantitative segmentation evaluation. As described above, higher E_{intra} value and higher E_{inter} value represents the under-segmentation and over-segmentation respectively. Thus, for a better segmentation result, the E_{cw} graph should be as close to origin as possible. In addition to the E_{cw} graph, E_{cw} which is defined as a weighted sum of E_{intra} and E_{inter} is also served as a suitable criterion for quality evaluation. Again, the smaller the value of E_{cw} , the better quality of segmentation image it has.

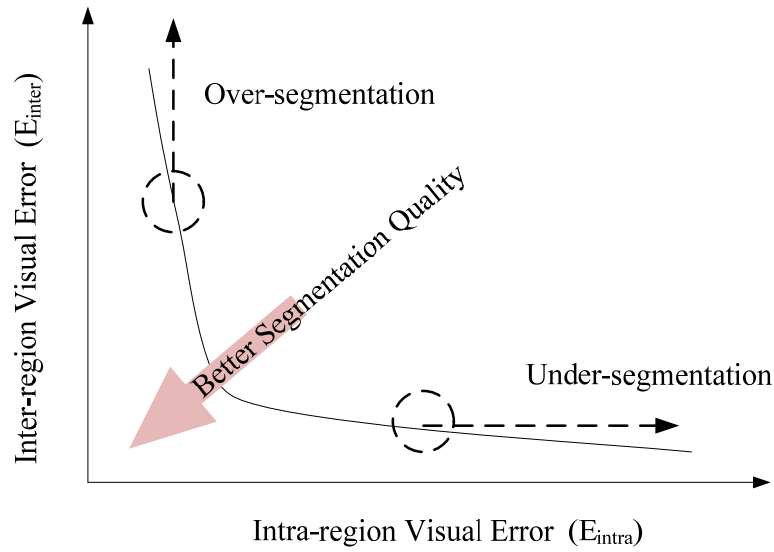


Fig. 4.1. The plot of intra-region visual error v.s. inter-region visual error [23].

4.2.2. Comparison between One-window Based LBA and Four-window Based LBA

In this section, we compare the quality of the segmentation image of the proposed one-window based and four-window based LBA using quantitative and qualitative evaluation. Fig. 4.2 – Fig. 4.4 shows the E_{cw} curves of the two proposed algorithm. For each test image, parameters are set to be the same except for the region merge threshold. The region merge threshold is adjusted from the range of 2 to 9 in integer to plot the E_{cw} curve. From the figures, the curves of two methods is almost closed to each other, however the curves of four-window based method is closer to the origin than one-window based method is. Thus from the definition of the E_{cw} curves, four-window based LBA performs segmentation results better than the one-window based method does. We suspect that this is because in the same constraint, four-window based method can reach more nodes than one-window based method does and thus it is possible to obtain more segment labels and aggregate more reliable messages than one-window

based method does.

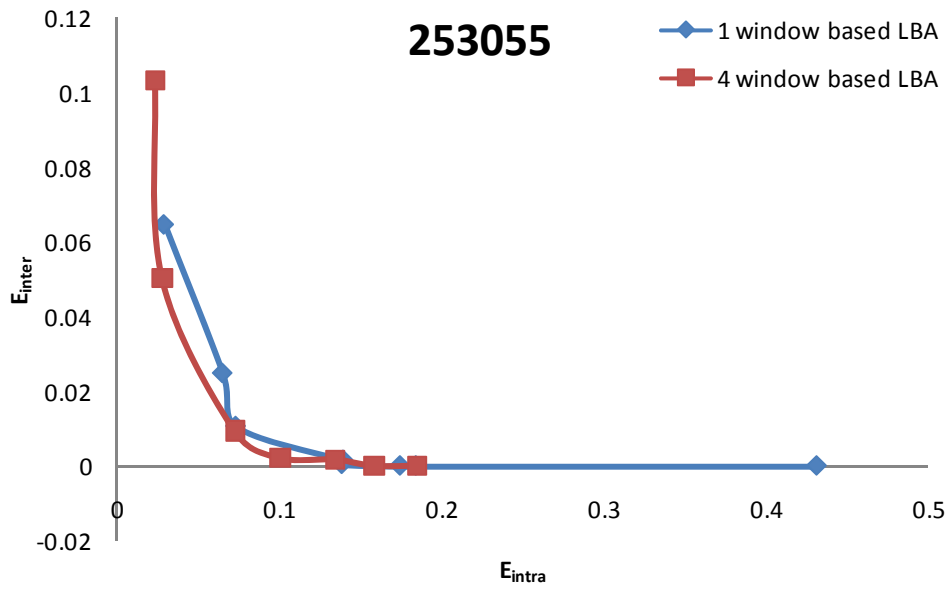


Fig. 4.2. E_{cw} curve on test image 253055.

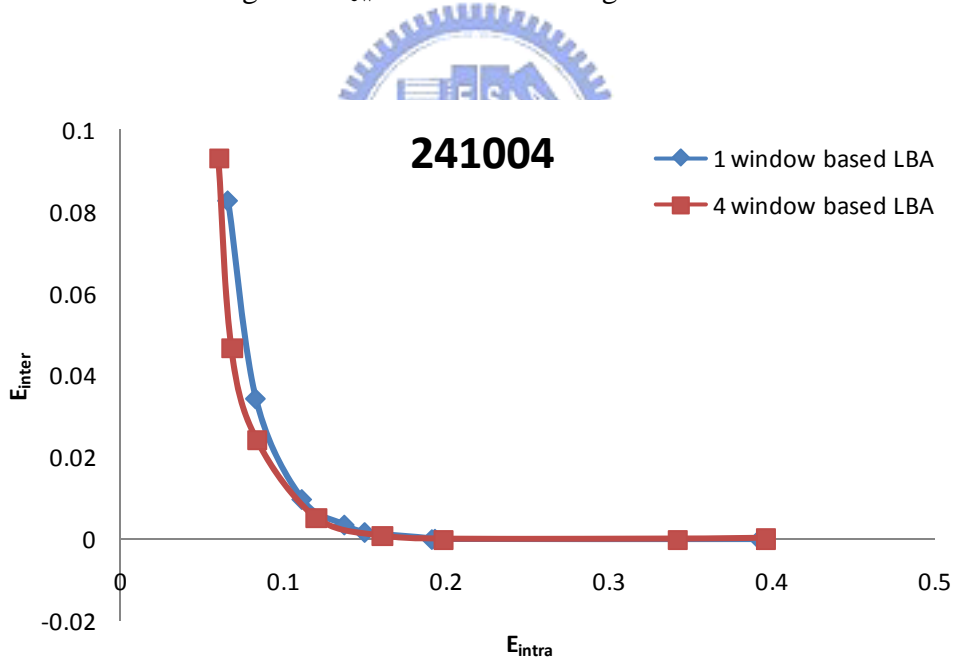


Fig. 4.3. E_{cw} curve on test image 241004.

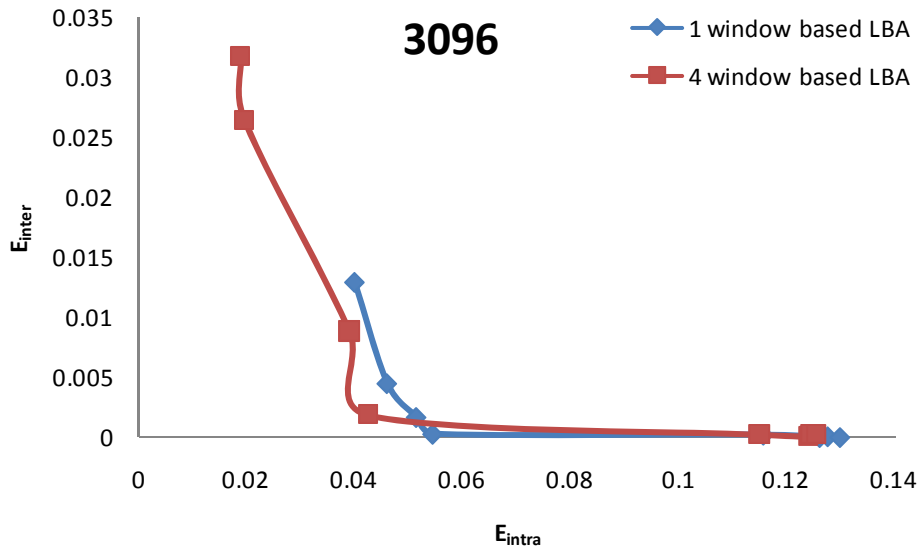


Fig. 4.4. E_{cw} curve on test image 3096.

In addition to the quantitative evaluation demonstrated above, the segmentation images are also provided for qualitative evaluation. Fig. 4.5 shows the segmentation results of the proposed algorithm with one-window based and four-window based method after 150 iterations. The parameter settings are all the same throughout the experiments except for the size of local search window ranged from 5×5 to 7×7 , threshold of region merge process ranged from 6 to 7 and the number of reliable messages aggregate from neighbors. The segmentation results of these parameter settings are selected based on subjective preference. Note that the number of reliable messages aggregate from neighbors are different based on the intrinsic difference between two methods. In four-window based method, this number is determined by a predefined threshold which is selected to be 6 in the experiments. In one-window based method, this number is determined by the size of the local search window.

In the test image 253055 as shown in Fig. 4.5 (a), there are four segment regions that

should be partitioned. They are five giraffes, grassland, blue sky and clouds. Subjectively speaking, both methods achieve comparable segmentation results. It is hard to tell the big difference between two methods. However, the lag of the first giraffe is segmented out in four-window based method as shown in Fig. 4.5 (c) whereas the one-window based method isn't. To segment out the lag of the first giraffe in one-window based LBA, one can try to adjust the threshold of the region merge process to the lower one, but this will result in a more over segmentation result than the one with higher threshold.

In the test image 241004 as shown in Fig. 4.5 (d), there are five major parts that should be partitioned. They are sky, the rocks, grassland, light mountains in the far side and dark mountains in the near side. Note that it depends on human evaluator to decide how many segment regions of the rocks and mountains should be partitioned. Subjectively speaking, both methods achieve comparable segmentation results. However, the rocks on the left side and right side of the image are over-segmented in one-window based method as shown in Fig. 4.5 (e) compared with the results using four-window based method as shown in Fig. 4.5 (f).

In the test image 3096 as shown in Fig. 4.5 (d), there are two major parts that should be partitioned, the clouds and the airplane. Both methods produce similar segmentation results. It is hard to tell the difference between them.

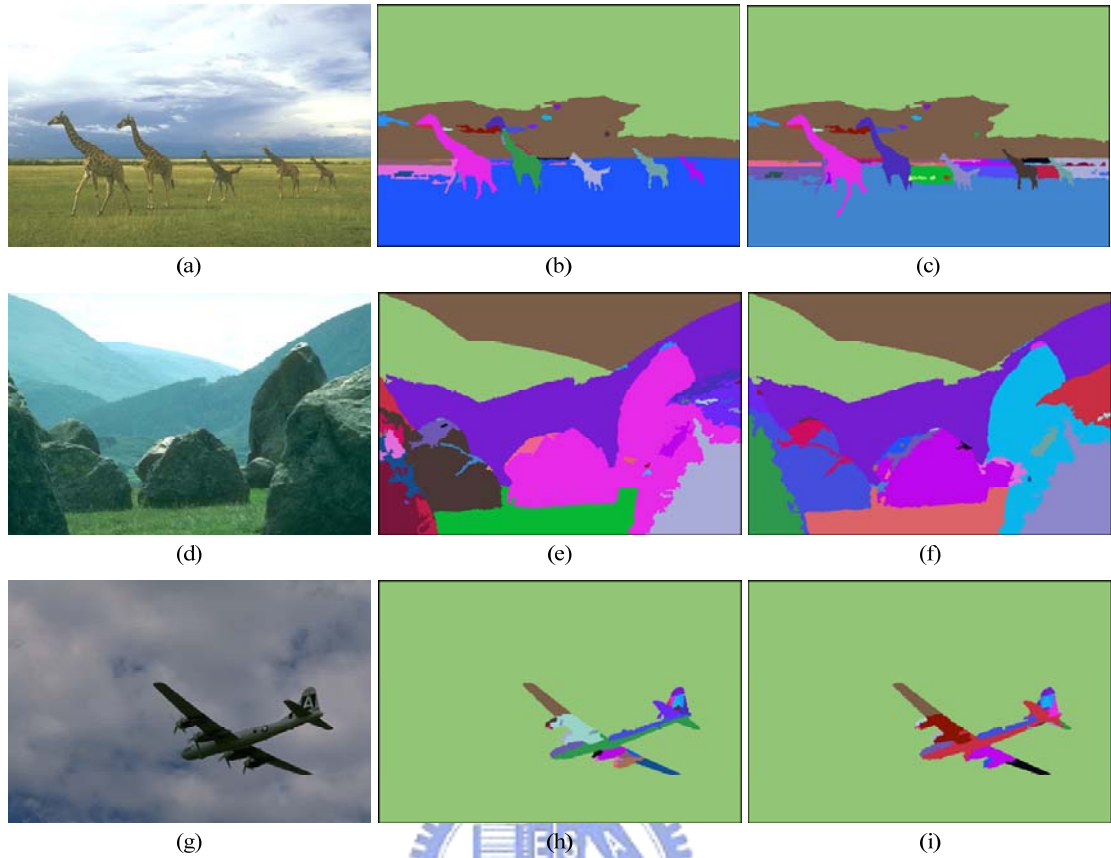


Fig. 4.5. Segmentation results with the proposed two methods. (a) Test image 253055. (d) Test image 241004. (g) Test image 3096. (b)(e)(h) Segmentation results using one-window based LBA. (c)(f)(i) Segmentation results using four-window based LBA.

4.2.3. Comparison between Different Segmentation

Algorithms

In this section, we compare the segmentation results of the proposed four-window based LBA with mean shift and watershed algorithm. Fig. 4.6 – Fig. 4.8 shows the E_{cw} curves of the proposed algorithm, mean shift, and watershed. From the E_{cw} curves, the proposed algorithm performs better than the rest. This is because our formulation tends to minimize E_{cw} . However, [25] has pointed out that E_{cw} may not always match

human perception. Another issue with E_{cw} is that although we have tried to find parameter settings that pushed the curves as close to the origin as possible, better parameter settings may still exist. Thus E_{cw} curve here is only a reference quantitative evaluation metric. Qualitative evaluation is still necessary to compare the segmentation results of different algorithms.

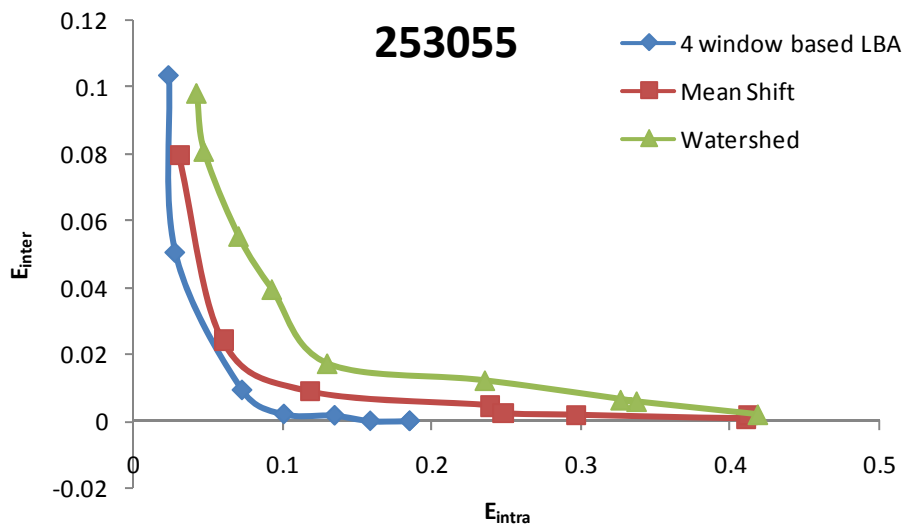


Fig. 4.6. E_{cw} curve of different segmentation algorithms on test image 253055.

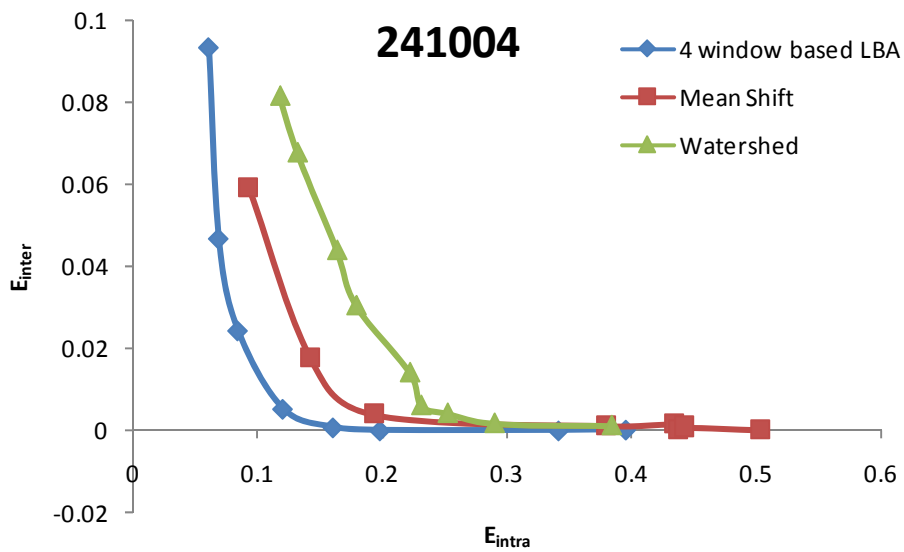


Fig. 4.7. E_{cw} curve of different segmentation algorithms on test image 241004.

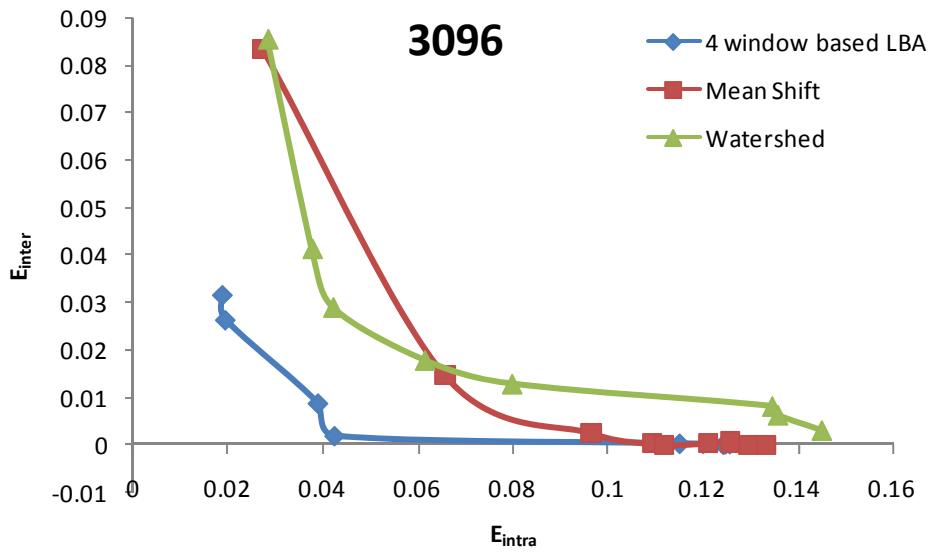


Fig. 4.8. E_{cw} curve of different segmentation algorithms on test image 3096.

Fig. 4.9 shows the segmentation images of the proposed four-window based LBA algorithm after 150 iterations and other algorithms. The parameter settings are selected based on subjective preference. In other words, the segmentation images are selected from the ones we felt of having the best quality. Hence, the parameter settings between algorithms are different. Subjectively speaking, the proposed four-window based LBA algorithm achieved comparable segmentation image quality to mean shift. One can hardly tell if the segmentation images were the result of applying the proposed algorithm or mean shift. The watershed algorithm, on the other hand, achieved the worst segmentation image quality of the three. Although it seems to perform well in test image 3096 as shown in Fig. 4.9 (i), but the mountain in the left side of test image 241004 as shown in Fig. 4.9 (f) cannot be perfectly segmented out. Besides, watershed performs poor segmentation results in which giraffes are miss-segmented on test image 253055 as shown in Fig. 4.9 (c). This is due to watershed algorithm's

being sensitive to noise and needs a pre-processing step to de-noise. The watershed segmentation results will much depend on the de-noise effect.

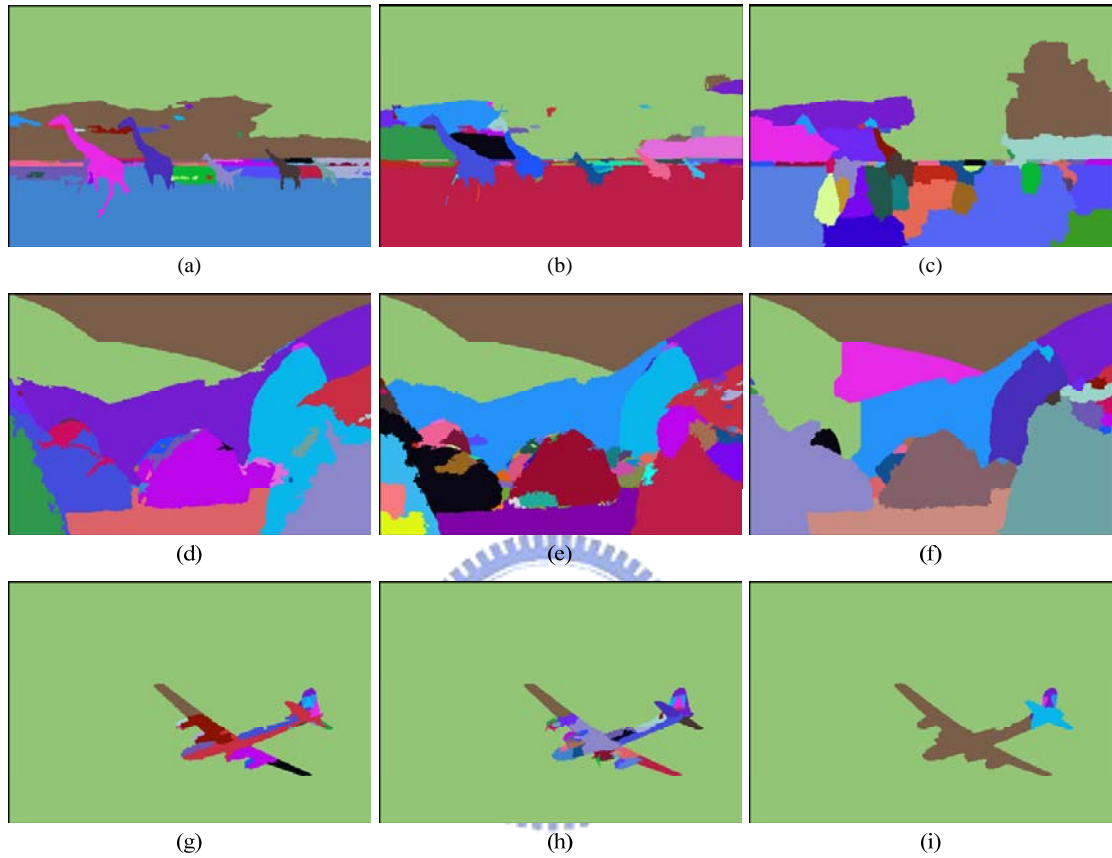


Fig. 4.9. Segmentation results with different algorithms. (a)(d)(g) Segmentation results using proposed four-window based LBA. (b)(e)(h) Segmentation results using mean shift algorithm. (c)(f)(i) Segmentation results using watershed algorithm.

Chapter 5. Conclusion and Future Works

5.1. Conclusion

In this thesis, we formulated a MRF model based on the concept of intra and inter region constraint for color image segmentation. Furthermore, we proposed a local belief aggregation (LBA) algorithm to find the best segmentation image based on our MRF formulation. Compared with the original BP algorithm, the usage of memory storage is much reduced while using the proposed LBA algorithm to find the MAP solution of the proposed MRF model.

Experimental results demonstrated that the proposed color image segmentation algorithm using four-window based LBA achieved segmentation image quality comparable to the mean shift algorithm. To further reduce the computational complexity of the four-window based LBA, a one-window based LBA is presented. Experimental results demonstrated that one-window based LBA can reduce 27% ~ 86% computational complexity than four-window based LBA with comparable segmentation results to the one-window based LBA.

5.2. Future work

There are two issues remained in the LBA estimation. First, the LBA may sometimes incorrectly discard the segment states that should not be discarded. This may results in incorrect final segmentation image. Besides, this can also results in more LBA iterations to reach to the convergence. The other issue is computational speed of LBA

estimation algorithm which still remains slow. Therefore, we will be working on remedying the incorrect segment state discarding issue and optimizing the speed of the LBA estimation algorithm in the future.



Reference

- [1] A. Klaus, M. Sormann, and K. Karner, "Segment-based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure," in *Proceedings of IEEE International Conference on Pattern Recognition*, vol. 3, pp. 15-18, 2006.
- [2] J. R. Jimenez-Alaniz, V. Medina-Banuelos, and O. Yanez-Suarez, "Data-driven Brain MRI Segmentation Supported on Edge Confidence and a Priori Tissue Information," *IEEE Transactions on Medical Imaging*, vol. 25, no. 1, pp. 74-83, January 2006.
- [3] S. Y. Chien, S. Y. Ma, and L. G. Chen, "Efficient Moving Object Segmentation Algorithm Using Background Registration Technique," *IEEE Transaction on Circuits and Systems for Video Technology*, vol. 12, no. 7, pp. 577-586, July 2002.
- [4] S. Y. Chien, Y. W. Huang, and L. G. Chen, "Predictive Watershed: A Fast Watershed Algorithm for Video Segmentation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 5, pp. 453-461, May 2003.
- [5] L. Vincent and P. Soille, "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583-598, June 1991.
- [6] D. Comaniciu and P. Meer, "Mean Shift: A Robust Approach toward Feature Space Analysis," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603-619, May 2002.
- [7] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol.6, no. 6, pp. 721-741, 1984.

- [8] H. Deng and D. A. Clausi, "Unsupervised Image Segmentation Using a Simple MRF Model with a New Implementation Scheme," *Pattern Recognition*, vol. 37, no. 12, pp. 2323-2335, December 2004.
- [9] J. Wu, and A. C. S. Chung, "A Segmentation Model Using Compound Markov Random Fields Based on a Boundary Model," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 241-252, January 2007.
- [10] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers, CA, 1988.
- [11] J. Sun, N. N. Zheng, and H. Y. Shum, "Stereo Matching Using Belief Propagation," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 25, no. 7, pp. 787-800, July 2003.
- [12] J. Coughlan, and H. Shen, "Dynamic Quantization for Belief Propagation in Sparse Spaces," *Computer Vision and Image Understanding*, vol. 106, no.1, pp. 47-58, April 2007.
- [13] J. Fairfield, "Toboggan Contrast Enhancement for Contrast Segmentation," in *Proceedings of IEEE International Conference on Pattern Recognition*, vol. 1, pp. 712-716, June 1990.
- [14] E. N. Mortensen and W. A. Barrett, "Toboggan-based Intelligent Scissors with a Four-parameter Edge Model," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 452-458, June 1999.
- [15] Y. C. Lin, Y. P. Tsai, Y. P. Hung, and Z. C. Shih, "Comparison between Immersion-based and Toboggan-based Watershed Image Segmentation," *IEEE Transaction on Image Processing*, vol. 15, no. 3, pp. 632-640, March 2006.
- [16] Stan Z. Li, *Markov Random Field Modeling in Image Analysis*, Springer-Verlag New York, Inc., Secaucus, NJ, 2001.

- [17] V. Andersen, "Smoothing 3D Meshes using Markov Random Fields," Master Thesis, Multimedia Technology at IT-University in Copenhagen, Denmark, 2007.
- [18] J. M. Hammersley and P. Clifford, *Markov Fields on Finite Graphs and Lattices*, Unpublished, 1971.
- [19] J. Besag, "On the Statistical Analysis of Dirty Pictures," *Journal of the Royal Statistical Society, Series B*, vol. 48, no. 3, pp. 259-302, 1986.
- [20] Y. Boykov, O. Veksler, and R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222-1239, November 2001.
- [21] M. I. Jordan, *Learning in Graphical Models*, MIT Press, 1998.
- [22] R. M. Haralick and L. G. Shapiro, "Survey: Image Segmentation Techniques," *Computer Vision, Graphics and Image Processing*, vol. 29, pp. 100-132, 1985.
- [23] H. C. Chen and S. J. Wang, "The Use of Visible Color Difference in the Quantitative Evaluation of Color Image Segmentation," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 593-596, May 2004.
- [24] H. C. Chen, "A Study of Image Analysis Techniques Based on Luminance/Color Contrast," Ph.D. Dissertation, Department of Electronic Engineering, National Chiao-Tung University, Hsinchu, Taiwan, June 2006.
- [25] H. Zhang, J. E. Fritts, and S. A. Goldman, "Image Segmentation Evaluation: A Survey of Unsupervised Methods," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 260-280, May 2008.
- [26] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear Total Variation Based Noise Removal Algorithms," *Physica D*, vol. 60, pp. 259-268, 1992.
- [27] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A Database of Human Segmented

Natural Images and Its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics,” in *Proceedings of IEEE International Conference on Computer Vision*, vol. 2, pp. 416-423, July 2001.

- [28] J. Y. Hardeberg, “Acquisition and Reproduction of Colour Images: Colorimetric and Multispectral Approaches,” Ph.D. Dissertation, Ecole Nationale Supérieure des Télécommunications, Paris, France, February 1999.



Appendix

In this section, we discuss the relationship between the number of reliable message number and the iteration of convergence using four-window based LBA. In the experiments, the parameter settings are all the same except the number of the reliable message number. We evaluate the segmentation results to subjectively determine the convergence iteration of the test image 253055. The segmentation image that remains unchanged in the contiguous iterations or oscillates between iterations is determined as the convergence of the respective image. Furthermore, the E_{cw} energy curve is also used to objectively determine the convergence iteration. Same as in the subjective method, the iteration in the curve that has trend to converge or oscillate is determined as the convergence iteration.



Fig. A.1 shows the relationship between the convergence iteration counts and the reliable message numbers where the convergence iteration counts is determined subjectively. From the figure, we could see that the 4 reliable messages and 8 reliable messages need about 104 iterations to converge. The others only need 33 to 59 iterations to converge. Fig. A.2 shows the relationship objectively using the E_{cw} energy plot of reliable message number with value 4 and 9. From the figure, we could see that the convergence iteration of reliable message number with value 4 and 9 is about 105 and 44 respectively which meets the subjectively evaluation results.

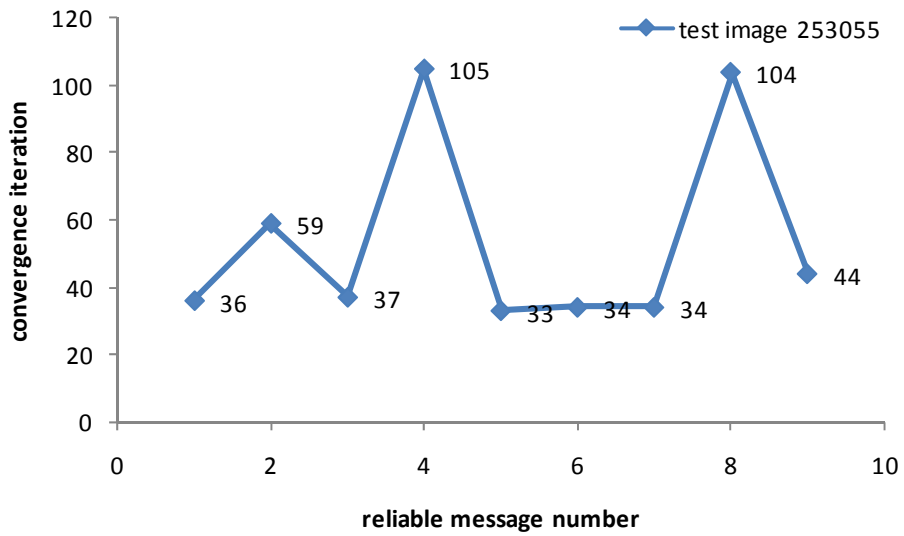


Fig. A.1. Subjectively determine the relation between reliable message number and convergence iteration in four-window based LBA with test image 253055.

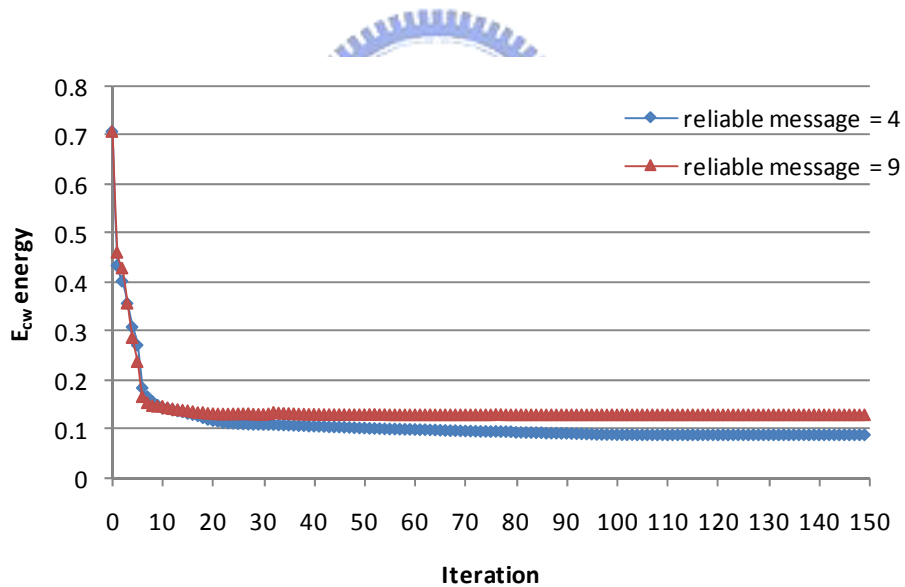


Fig. A.2. Objectively determine the relation between reliable message number and convergence iteration in four-window based LBA with test image 253055.

We suspect that the difference iterations of these two parts come from the different initial seeds produced after the first iteration of the proposed LBA algorithm. Here, we take the 4 reliable messages which needs 105 iterations to converge and 9 reliable

messages which needs 44 iterations to converge for example to discuss the probable reason. The constraints are the same during the initial guess, thus the initial segmentation images are the same. After the first iteration of the LBA, the intermediate segmentation result becomes an image with initial seeds due to the region merge process. The quality of the initial seeds will affect the final segmentation results and hence affect the iterations to converge. Fig. A.3 and Fig. A.4 show the intermediate results of 4 reliable messages and 9 reliable messages at different iterations respectively. In Fig. A.3 (a), there are two segment regions being segmented out on the right down side of the sky. Then after the first iteration, the LBA algorithm iteratively segment out the probable segment regions using proposed MRF models. The sky is segmented out in three parts as shown in Fig. A.3 (f). The blue sky is colored in green, the dark cloud in yellow and the white cloud in dark blue. Most of the iterations are spent to progressive segment out the dark cloud in yellow color as shown in Fig. A.3 (d) - Fig. A.3 (f). Five giraffes and grassland have been segmented out already. The reason of the much iteration used to segment out the yellow color which represents the dark cloud in the example is due to the local search window metric we used in the LBA algorithm. The local search window metric only considers the candidate segment states appeared in the window.

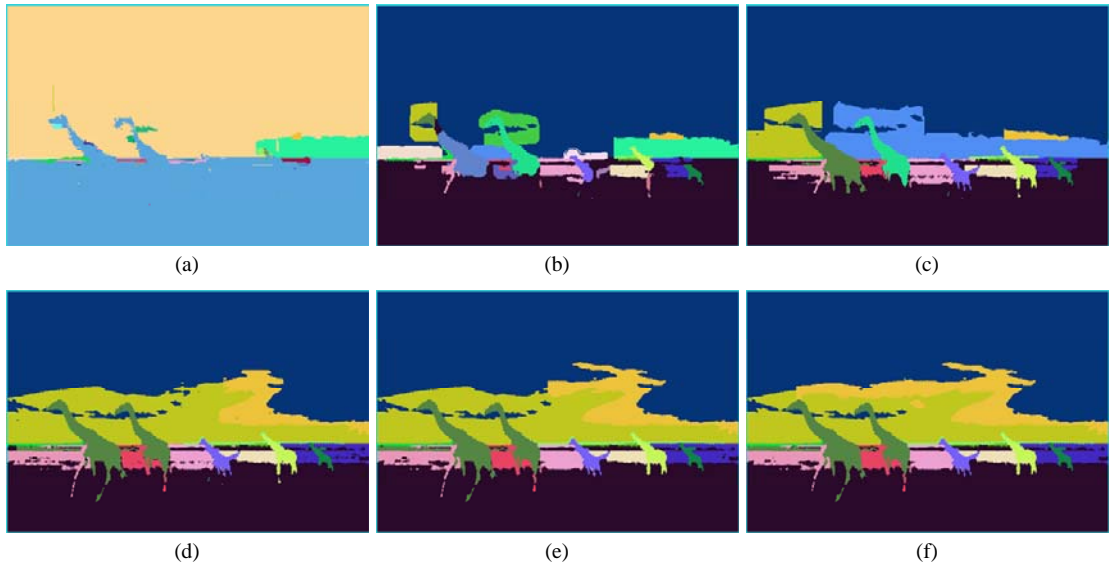


Fig. A.3. Intermediate results on 253055 at different iterations using 4 reliable messages. (a) Iteration 1. (b) Iteration 5. (c) Iteration 10. (d) Iteration 25. (e) Iteration 40. (f) Iteration 80.

In contrast with the one using 4 reliable messages, there is only one segment region being segmented out on the right down side of the sky in the first iteration of LBA using 9 reliable messages as shown in Fig. A.4 (a). Without the need of spending iterations to segment out the dark cloud in yellow color as shown in Fig. A.3 (d) – Fig. A.3 (f), LBA using 9 reliable messages converge faster than the one using 4 reliable messages.

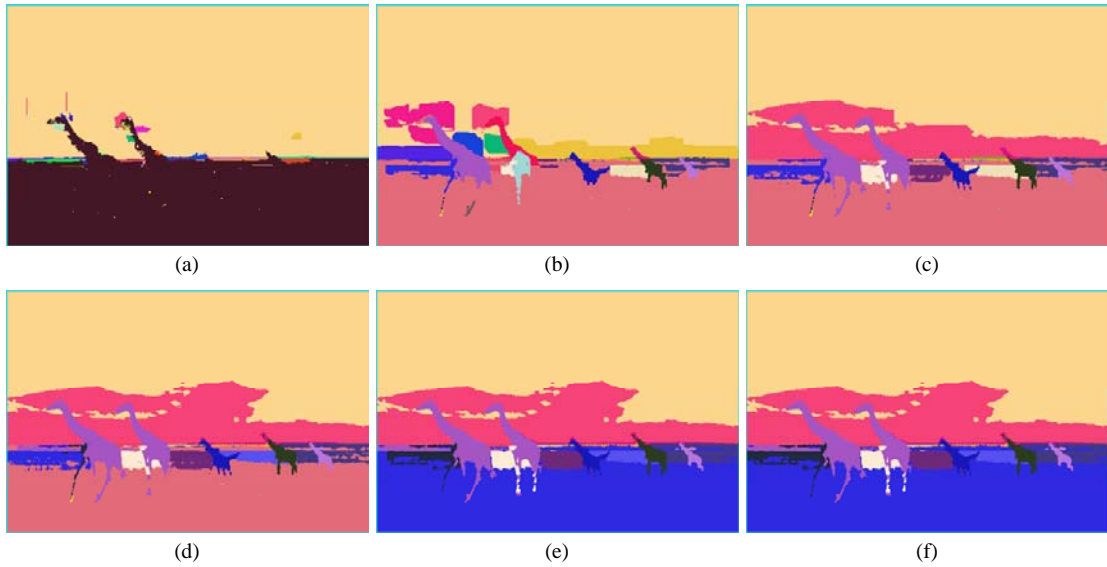


Fig. A.4. Intermediate results on 253055 at different iterations using 9 reliable message number. (a) Iteration 1. (b) Iteration 5. (c) Iteration 10. (d) Iteration 25. (e) Iteration 40. (f) Iteration 44.

Fig. A.5 demonstrates the final segmentation results with different reliable message number using LBA. Both Fig. A.5 (d) and Fig. A.5 (h) segment out the sky part in three segment regions and thus require more iteration to converge as illustrated in Fig. A.1. From the segmentation results, it seems that except for the one using 1 reliable message in Fig. A.5 (a), the others using reliable message number larger than 2 can segment out the sky, grassland and five giraffes successfully for test image 253055. We suspect this with two reasons. First, in the four-window based method, the restriction of the number of reliable message is performed on each side individually when aggregate the reliable message from neighbors. Thus, if we assign the restriction number to 2 for example, there are 8 probable candidate segment states to be calculated for the belief in the worst case. Second, after the first iteration of LBA, most of the redundant segment states are merged and only a few segment states that are important exist. Thus, for the small amount of reliable message such as 2, the LBA algorithm may also work well for the above two reasons.

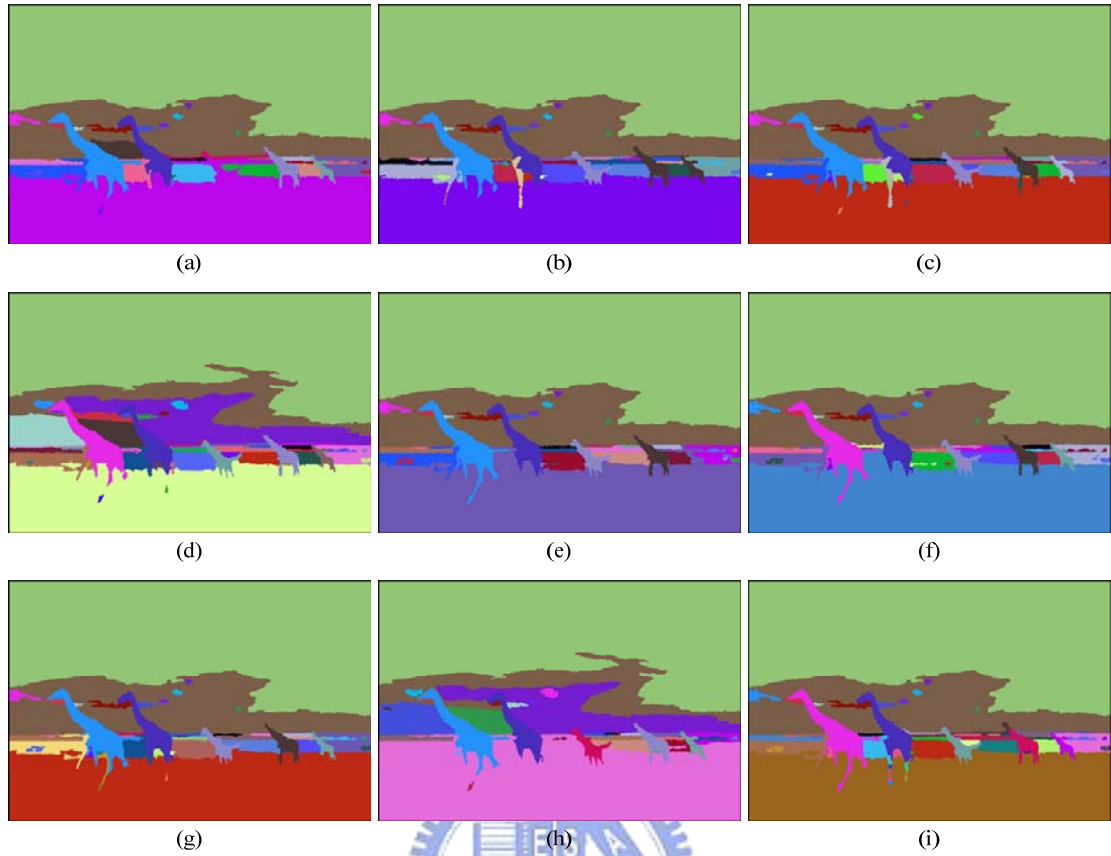


Fig. A.5. Segmentation results with different reliable message number using proposed 4 window based LBA. (a) 1 reliable message. (b) 2 reliable messages. (c) 3 reliable messages. (d) 4 reliable messages. (e) 5 reliable messages. (f) 6 reliable messages. (g) 7 reliable messages. (h) 8 reliable messages. (i) 9 reliable messages.

作者簡歷

姓名：詹景竹

籍貫：高雄市

學歷：

國立高雄師範大學附屬高級中學 (民國 88 年 09 月 ~ 民國 91 年 06 月)

國立交通大學電機與控制工程學系 學士 (民國 91 年 09 月 ~ 民國 95 年 06 月)

國立交通大學電子所系統組 碩士 (民國 95 年 09 月 ~ 民國 97 年 09 月)

著作：

國內會議

- 
- [1] Jing-Chu Chan, Nelson Yen-Chung Chang, Yu-Cheng Tseng and Tian-Sheuan Chang, “Local Belief Aggregation for MRF-based Color Image Segmentation,” in *IPPR Conference on Computer Vision, Graphics and Image Processing*, August 2008.

國際會議

- [2] Jing-Chu Chan, Nelson Yen-Chung Chang, Tian-Sheuan Chang, “ISID: In-order Scan and Indexed Diffusion Segmentation Algorithm for Stereo Vision,” in *Proceedings of IEEE International Symposium on Circuits and Systems*, pp. 3478-3481, May 2008.