Bayesian Inference for Layer Representation with Mixed Markov Random Field

Ru-Xin Gao[1,2], Tian-Fu Wu[2], Song-Chun Zhu[3,2] and Nong Sang[1]

[1]IPRAI, Huazhong University of Science and Technology, China.
[2]Lotus Hill Research Institute, China.
[3]Departments of Statistics and Computer Science, UCLA.
rxgao.lhi@gmail.com, tfwu.lhi@gmail.com, sczhu@stat.ucla.edu, nsang@hust.edu.cn

Abstract. This paper presents a Bayesian inference algorithm for image layer representation [26], 2.1D sketch [6], with mixed Markov random field. 2.1D sketch is an very important problem in low-middle level vision with a synthesis of two goals: segmentation and 2.5D sketch, in other words, it is to consider 2D segmentation by incorporating occulision/depth explicitly to get the partial order of final segmented regions and contour completion in the same layer. The inference is based on Swendsen-Wang Cut (SWC) algorithm [4] where there are two types of nodes, instead of all nodes being the same type in traditional MRF model, in the graph representation: atomic regions and their open bonds desribed by address variables. These makes the problem a mixed random field. Therefore, two kinds of energies should be simultaneously minimized by maximizing a joint posterior probability: one is for region coloring/layering, the other is for the assignments of address variables. Given an image, its primal sketch is computed firstly, then some atomic regions can be obtained by completing some sketches into a closed contour. At the same time, T-junctions are detected and broken into terminators as the open bonds of atomic regions after being assigned the ownership between them and atomic regions. With this graph representation, the presented inference algorithm is performed and satisfactory results are shown in the experiments.

Key words: Layer Representation, 2.1D Sketch, Bayesian Inference, Contour Completion, Mixed Markov Field, Swendsen-Wang Cut, MCMC

1 Introduction

This paper presents a Bayesian inference algorithm for image layer representation [26], that is 2.1D sketch [6], with mixed Markov random field. The general goal of 2.1D sketch is to resume the occluded structure part of object and find the occlusion relation (partial order) of all objects in a scene. 2.1D sketch is a very important issue in low-middle level vision tasks and remains a challenging problem yet in the literature. Solving 2.1D sketch is a critical step for scene understanding in both still images and video, such as foreground/background separation, 2.5D sketch, motion analysis, etc.

Layer representation is firstly presented by Wang and Adelson for motion analysis [26]. Based on this idea, the concept of 2.1D sketch was then proposed firstly by Nitzberg and Mumford [6]. In their works, T-junctions detected on region's boundaries provide the cue information of occlusion relation, and an energy function is minimized to get 2.1D sketch with experiments on some simple images.

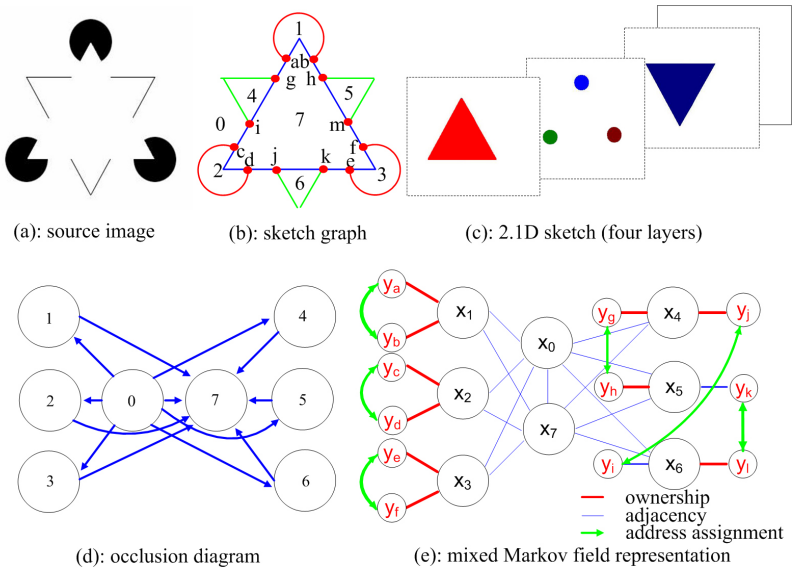

Fig. 1. Illustration of the 2.1D sketch. (a) is the original image, Kanizsa figure. (b) is the sketch graph computed by the primal sketch model after interactively labeling. there are 8 atomic regions and 12 terminators broken form T-junctions. (c) is the 2.1D sketch. There are 4 layer and the contour completion is performed using Elastica rules. (d) is the Hasse diagram for partial order relation, such as $< 7, 1 >$ means region 7 occludes region 1. (e) is the graph representation with mixed Markov random field. Each big circle denotes a vertex of atomic region, each blue bar denotes one terminator, each little circle denotes a vertex of open bound described as address variable. Each region may have two or more terminators. The blue line segments connect the two neighboring regions. The red two-way arrows connect two terminators and each terminator is assigned another terminator's address variable.

As Fig.1 shown, it is an illustration of the 2.1D sketch for the well known Kanizsa figure. The final layer representation with contour completion is shown in Fig.1(c). There are 8 atomic regions in total as shown in Fig.1(b), denoted by $0, 1, ..., 7$ in Fig.1(d) for showing the diagram of the partial order and $x_0, x_1, ..., x_7$ in Fig.1(e) for showing the the graph representation with mixed random field. There are 12 T-junctions in total and they are break into terminators a, b; c, d; e, f; g, h; i, j; k, m as shown in Fig.1(b). These terminators are called open

bonds described as address variables for corresponding atomic regions, denoted by $y_a, y_b; ...; y_k, y_m$ in Fig.1(e). Both the atomic region vertices and the corresponding address variables vertices yield the graph representation and address variables would reconfigure the graph structure. And it is these address variables that make the problem a mixed random field:

(1) The vertices are inhomogeneous with different degrees of connections and should be inferred from the images;

(2) The neighborhood of each vertex is no longer fixed but inferred as address variables, which yields and would reconfigure the graph structure.

The compatibility between any two terminators is a function defined on different cues such as appearance, geometric properties (eg. Elastic rules), etc. The compatibility function is then used as the weight of edge in the graph representation. For the simple case shown in Fig.1, Fig.2 shows the inferential computing procedure using Gibbs sampler with anneal simulating starting from initial temperature $T = 20$.

In the literature, there are some related works on 2.1D sketch. Stella [8] etc proposed a model for Figure-Ground segregation based on hierarchical Markov random field. They used clique potentials in MRF to encode local logical decision rules and demonstrated a system that automatically integrating sparse local relative depth cues arisen from T-junctions over long distance into a global ordering of relative depths. Eseddoglu and Riccardo March [7] proposed to segment image with depth but without detecting junctions. They described a technique in the variational formulation of minimizing Mumford-Shah's energy function that avoided explicit detection/connection of T-junction. These works did not consider the the problem of address variable explicitly.

In contrast, this paper formulates 2.1D sketch using mixed random field, and presents a Bayesian inference algorithm based on Swendsen-Wang Cut algorithm. According to the authors' knowledge, this is the first time to solve 2.1D sketch problem inferencing based on mixed random field.

Given an image, its primal sketch is firstly computed using the primal sketch model [2], then some atomic regions can be obtained by completing some sketches into a closed contour using some interactive operations for the time being as shown in Fig.3 and other results in experiments. From sketch to atomic region is not the main issue of this paper, so, currently we use some interactive operations, and this is also a common assumption in the literature. At the same time, T-junctions are detected using a method developed in the author's group recently [14] as shown in Fig.3 and other results in experiments. T-junctions are then broken into terminators as the open bonds or address variables of atomic regions after being assigned the ownership between them and atomic regions. Each address variable should be assigned with another one to make contour completion to get larger segmented regions consisting of some atomic regions. Then a graph representation can be obtained and consists of two kinds of vertices: atomic regions and its corresponding address variables. So there are two kinds of energies that should be minimized at the same time: region coloring/layering

and address variable assignment, by maximizing a joint posterior probability. Some results are shown in Section 2 and Section 3.

The remainder of the paper is organized as following. Section 2 formulates the problem under Bayesina framework and defines some used concepts and variables. Section 3 presents the inference algorithm based on Swendsen-Wang Cut algorithm with experiments in Section 4. Some conclusions and discussions are come at in Section 5.

2 Problem Formulation of the 2.1D Sketch

2.1 Graph Representation

Given an image defined on lattice Λ with occlusion relations among the objects in it, eg. Fig.1 and other examples in this paper, its sketch graph is computed using the primal sketch model [1, 2], followed by some manually interactive operations to get initial partition of the image domain with some atomic regions. Because the step from primal sketch to initial atomic regions are not the main issue of this paper and it is another research subject in the literature, we simplify this step by using manually interactive operations, and this is a common assumption in dealing with 2.1D sketch for the time being. These atomic regions are one kind of vertex in the graph. For each atomic region, there are some open bonds, described as address variables, which are obtained by breaking the T-junctions. Each T-junction is broken into a line segment and a terminator, and the terminator is described by address variable treated as the other kind of vertex in the graph. Address variables should be assigned to another one, inferring from the image. Some examples are shown in Fig.3, etc. With this graph representation, 2.1D sketch can be formulated as graph partition problem. Swendsen-Wang Cut(SWC) is a recently developed algorithm to solve general graph partition problem. In the following, some definitions of the problem domain are given, bayesian formulation and the inference algorithm based on SWC are derived.

2.2 Definition of the Problem Domain

In this section, we define the elements in the 2.1D sketch representation. As stated above, it is a mixed random field with two kinds of vertices in the graph representation: a set of connected regions and a set of terminators broken from T-junctions. The regions set are defined as:

$$\Omega_x = \{R_1, R_2, \cdots, R_n\} \tag{1}$$

the regions will be divided into an unknown number of K layers.

$$\Omega_l = \{1, 2, \cdots, K\}, K \in |\Omega_x| \tag{2}$$

each region $R_i, i \in [1, n]$ has a label for its layer, $l_{R_i} \in X$ defined below. The terminators are also called open-bonds to be completed after inference, which can be defined as:

$$\Omega_a = \{a, b, \cdots, N\} \tag{3}$$

on Ω_x we define the layer information:

$$X = (x_1, x_2, \cdots, x_n), x_i \in \Omega_l, \forall i \in \Omega_x \tag{4}$$

region R_i is said to occlude R_j if $x_i < x_j$, region R_i and R_j are in the same layer and can be connected into one region if $x_i = x_j$.

On Ω_a, a set of address variables are defined for the contour completion information of open bonds:

$$Y = \{y_a, y_b, \cdots, y_N\}, y_a \in \Omega_a, \forall a \in \Omega_a \tag{5}$$

this will tell to whom the terminator a is connected. A terminator connects to itself when the T-junction is not broken. An example is shown in Fig.1.

2.3 Bayesian Formulation

According to the above definitions, 2.1D sketch is then represented by a hidden vector variable W, which describes the world state for generating the image I.

$$W = (X, Y) \tag{6}$$

The solution vector W includes the following information in two graphical representation: (1) Partial orders represented by an Hasse diagram as shown in Fig.1 (d), where each arrow indicates an occlusion relation. The arrow between two nodes i, j often decide the ownership of the boundary (contour) between regions R_i and R_j. For example, if region 2 occludes region 0, then region 2 owns the red boundary and consequently, it owns the two terminators c and d. Similarly, if region 4 occludes region 0, then region 4 owns the green contour. Thus region 4 own terminator g and i. This is not a trivial problem. In some rare accidental cases, two regions may co-own the contour between them; and (2) Graph representation with mixed Markov random field as shown in Fig.1 (e).

In a Bayesian framework, we make inference about W from input image I over a solution space Ω:

$$W^* = (X, Y)^* = \textbf{arg max}\ p(X, Y|I) = \textbf{arg max}\ p(I|X, Y)p(X, Y) \tag{7}$$

or preserving multiple interpretations, each with a probability:

$$W_1, W_2, \cdots, W_k \sim p(I|X, Y)p(X, Y) \tag{8}$$

As we mentioned before, there are two kinds of energies that should be minimized simultaneously and described in the above joint probability distribution.

The prior probability $p(W)$ is a product of the following probabilities.

1. An exponential model for the number of layers $p(K) \propto \exp\{-\lambda_0 K\}$.

2. A model for the size of each "larger" region R, $A = |R|$, consisted of some atomic regions after contour completion in each layer. The prior encourage large and connected regions. We take a distribution $p(A) \propto \exp\{-\gamma A^c\}$, where $c = 0.9$ is a constant in this paper.

3. The prior for the layer of a region $p(l_R)$ is assumed to be uniform, and the prior for the address variable is also uniform.

In this paper, we adopt a mixture of Gaussians for generative region model,

$$J(x,y) = \prod_{v \in R} [\sum_{i=1}^{m} (\alpha_i G(I_v - \mu_i; \Sigma_i))]. \tag{9}$$

The image likelihood $p(I|X,Y)$ is defined as:

$$p(I|X,Y) \propto \exp\{-\sum_{i=1}^{K} \sum_{(x,y)\in D_i} |I(x,y) - J(x,y)| - \sum_{i=1}^{K} \sum_{m=1,n=1}^{N} T(m,n)\} \tag{10}$$

where K is the total number of layer, D_i is the domain of region in the layer i ,$I(x,y)$ is color value in RGB space for any $(x,y) \in \Lambda$. N is the total number of open bonds, $T(m,n)$ denote the link energy of terminator m and terminator n when the edge between the two terminators are 'on'.

3 Inference

The presented inference algorithm is based on Swendsen-Wang Cut algorithm [3, 4, 23] which generalized a well accepted cluster sampling algorithm for solving graph partition problem. Generally speaking, there are three steps in performing a SWC algorithm: (1) initialization of an adjacency graph by computing local discriminative probabilities for each edge based on image data and the prior; (2) graph clustering under a given partition by removing all edges between vertices of different labels, probabilistically turning on/off of each of the remaining edges according to its weight; (3) graph flipping of the label for a selected connected component with a probability driven by the posterior. Details of the SWC algorithm are referred to reference [3, 4]. In the following, how to design the discriminative probabilities as weights of edges is given firstly, then we describe the inference algorithm for 2.1D sketch.

3.1 Discriminative Probabilities on Edges

After being broken from T-junctions, terminators or open bonds need to find a match by testing the compatibilities on different cues such as appearance (eg. profile matching), geometric properties (eg. smoothness), etc. These cues define the discriminative probabilities as the weights of edges in the graph. By sampling these probabilities, we can get new configurations of the graph.

Given an adjacency graph, for each open bond of each atomic region (vertex in initial graph), we extract a patch from the terminator and compute a 3D histogram with 15-bin in each dimension then normalize the histogram. For each edge $e =< v_i, v_j >$, we assign a binary random variable $\mu_e \in \{on, off\}$ and define the discriminative probability on edge as following:

$$q_e = p(\mu)e^{-E_{terminator}} \tag{11}$$

$$E_{terminator} = \alpha_0 E_{profile} + \alpha_1 E_{elastica} \tag{12}$$

Where α_0 and α_1 are constant, they will be adjusted in experiments.

$$E_{profile} = (KL(p_i||p_j) + KL(p_j||p_i)/2 \tag{13}$$

$$E_{elastica} = \int_{\Gamma} (\nu + \alpha k^2) ds \tag{14}$$

Where ν and α are constant, Γ is the boundary. By minimizing the energy, the Elastica function considered the orientation consistency as well as the shortest pathway.

3.2 Algorithm

We summarize the presented algorithm in the following table.

Bayesian Inference for 2.1D Sketch

Input an image and set the parameters ($\alpha_0, \alpha_1, \nu, \alpha$, sweep number and temperature value for anneal simulating):
1. Computing primal sketch for input image and interactively completing the sketches into close contour to get atomic regions
2. T-junctions detection and obtaining the open bond for each atomic region.
3. Initialization of adjacency graph.
4. Computing the discriminative probabilities on edges.
5. Select a region and select a layer(label) to it as in SWC algorithm;
6. Sample the open bonds according to the edge probabilities in each layer;
7. Connect terminators with curves according to open bonds' link status and generate new regions if some terminators and line segments can be combined into contours;

Repeat 5-7 within given sweep number and at the same time reduce the temperature every several steps.

At each step, the acceptance probability is:

$$\alpha(A \to B) = \mathbf{arg}\min\{1, \frac{q(B \to A)}{q(A \to B)} \cdot \frac{p(B)}{p(A)}\} \tag{15}$$

Where A and B are the two different states in node flipping process. $q(B \rightarrow A)$ is the posterior probability from state B to state A and $q(A \rightarrow B)$ is from A to B, let π denote state A or B, then the likelihood is in equation 10.

the transition probability is:

$$q(A \rightarrow B) = q(A \rightarrow B)_{layer} * q(A \rightarrow B)_{connect} \tag{16}$$

and

$$q(A \rightarrow B)_{layer} = \frac{p(B \rightarrow A)}{p(A \rightarrow B)} = \frac{n_A + 1}{n_B + 1} \tag{17}$$

$$q(\pi) = \exp\{-\sum_{i=1}^{K} \sum_{(x,y) \in A_i} |I(x,y) - J(x,y)| - \sum_{i=1}^{K} \sum_{m=a,n=a}^{N} T(m,n)\} \tag{18}$$

where A_i is the terminator's patch domain. all pixels' profile can get from the primitive's photometric attributes. $I(x,y)$ is the profile getting from the primitive dictionary, $J(x,y)$ is the color value from a color model, here we use gaussian model. N is the total number of open bonds, $T(m,n)$ denote the link energy of terminator m and terminator n when the edge between the two terminators is 'on'.

Fig.3 is an experiments result for the algorithm. The image is divided into 3 layers: the foreground layer, middle layer and background layer. The gaps between terminators are filled by curves which come from minimizing the Elastica function. The area are filled with a kind of color, the initial boundaries condition are from the corresponding primitive patch , the primitives are selected from a primitive dictionary.

4 Experiments

The presented inference algorithm is tested on a variety of natural images.

As shown in Fig.3, the boundaries of the plate in the image are occluded by a chopstick and the chopstick is still occluded by another chopstick and the plate and chopstick's regions can be cut into several atomic regions. The algorithm connects these regions in each layer with some curves cued by the terminator's boundaries' information and completed with boundaries' profile using heat diffusion. The background sketchable also are filled with some curves and the area was completed using colors.

More results are shown in Fig. 4.

5 Discussion

This paper presents a Bayesian inference algorithm for image layer representation with mixed Markov random field. We conducted several experiments to demonstrate the algorithm by satisfactory results. The key contribution of this paper is that it is the first time to formulate the 2.1D sketch using mixed random field and present an inference algorithm to solve region coloring/layering and assignments of open bonds simultaneously.

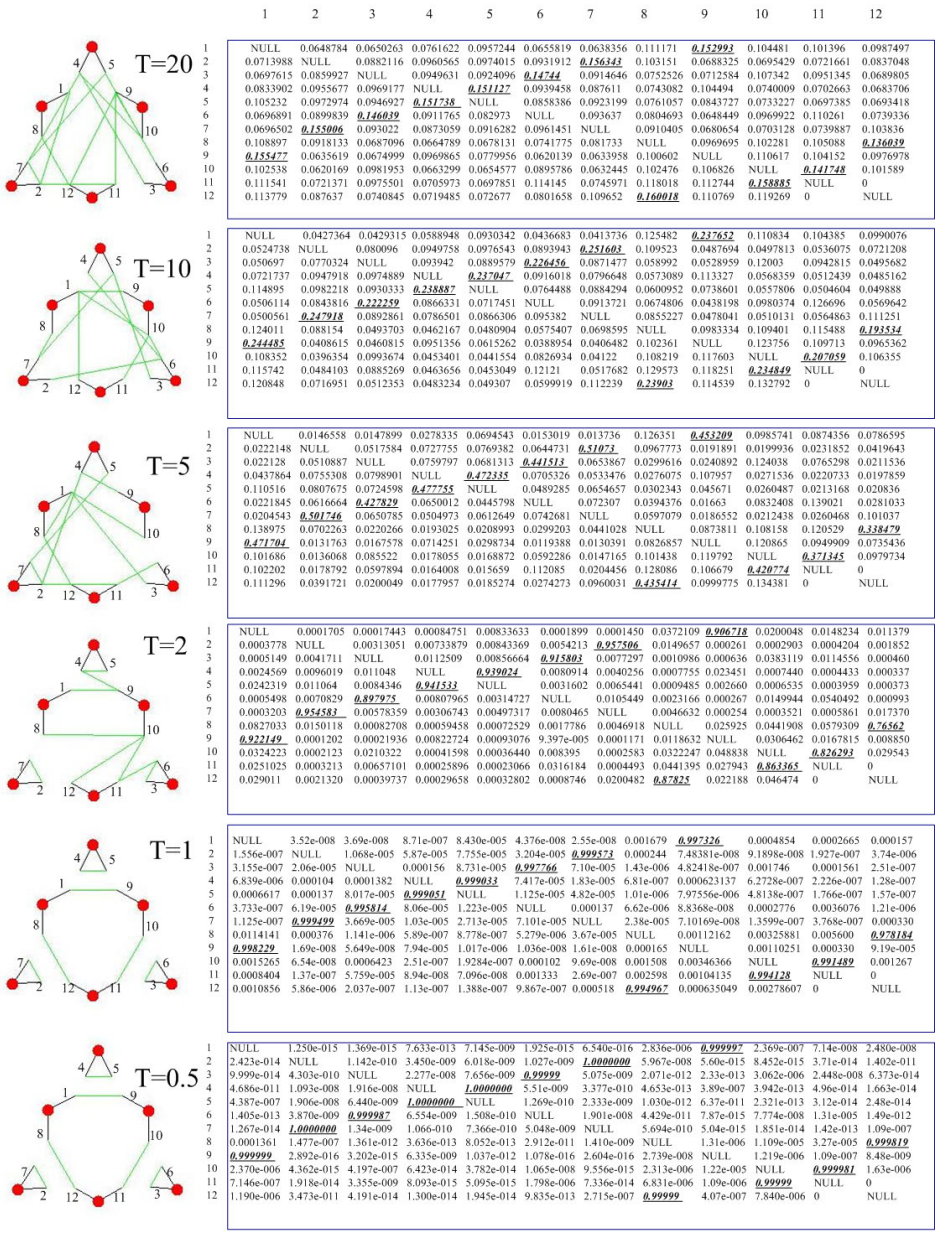

Fig. 2. Illustration of inferential Computing procedure. Each red node denotes a region, black line segment denotes terminator and the green line segment shows the inferred connection or assignment of address variable. Inference starts from a initial temperature $T = 20$, $(a) \sim (h)$ are the results in different temperatures. After $T = 1$, the result is right as in the figure.

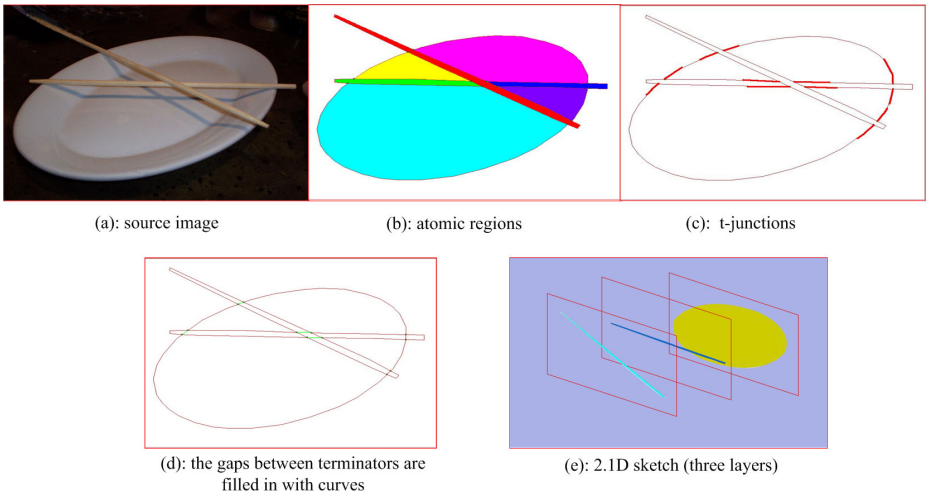

Fig. 3. One running example. There are three layers in the original image (a). (b) is atomic regions obtained by completing some sketched into a close contour; (c) shows the T-junction detection results based on the primal sketch.(d) shows the contour completion result after inference and (e)is the final layer representation.

6 Acknowledgements

This work is done when the authors are at the Lotus Hill Research Institute. The project is supported by : National 863 project Contact No. 2006AA01Z121, National Science Foundation China Contact No. 60672162, National Science Foundation China Contact No. 60673198, National 863 project contact No. 2006AA01Z339.

References

1. C.E.Guo, S.C.Zhu, and Y.N.Wu: Modeling visual patterns by integrating descriptive and generative models, *IJCV*, 53(1), 5-29,2003.
2. C.E.Guo, S.C.Zhu and Y.N. Wu: Primal sketch: integrating texture and structure, *Proc. Int'l Conf. on Computer Vision*, 2003.
3. A.Barbu and S.C.Zhu: Graph Partition by Swendsen-Wang Cuts.Proc. Int'l Conf. on Computer Vision,2003
4. A.Barbu and S.C.Zhu: Generalizing Swendsen-Wang to Sampling Arbitrary Posterior Probabilities.IEEE Trans. on PAMI,vol.27,2005 1239–1253
5. D.Marr: *Vision*, Freeman Publisher, 1983.
6. M.Nitzberg, D. Mumford and T. Shiota: *Filtering, Segmentation and Depth*, Springer Lecture Notes in Computer Science, **662**, 1993.
7. Selim Eseddoglu: Segment Image With Depth but Without Detecting Junction.Journal of Mathematical Imaging and Vision. vol. 18,2003

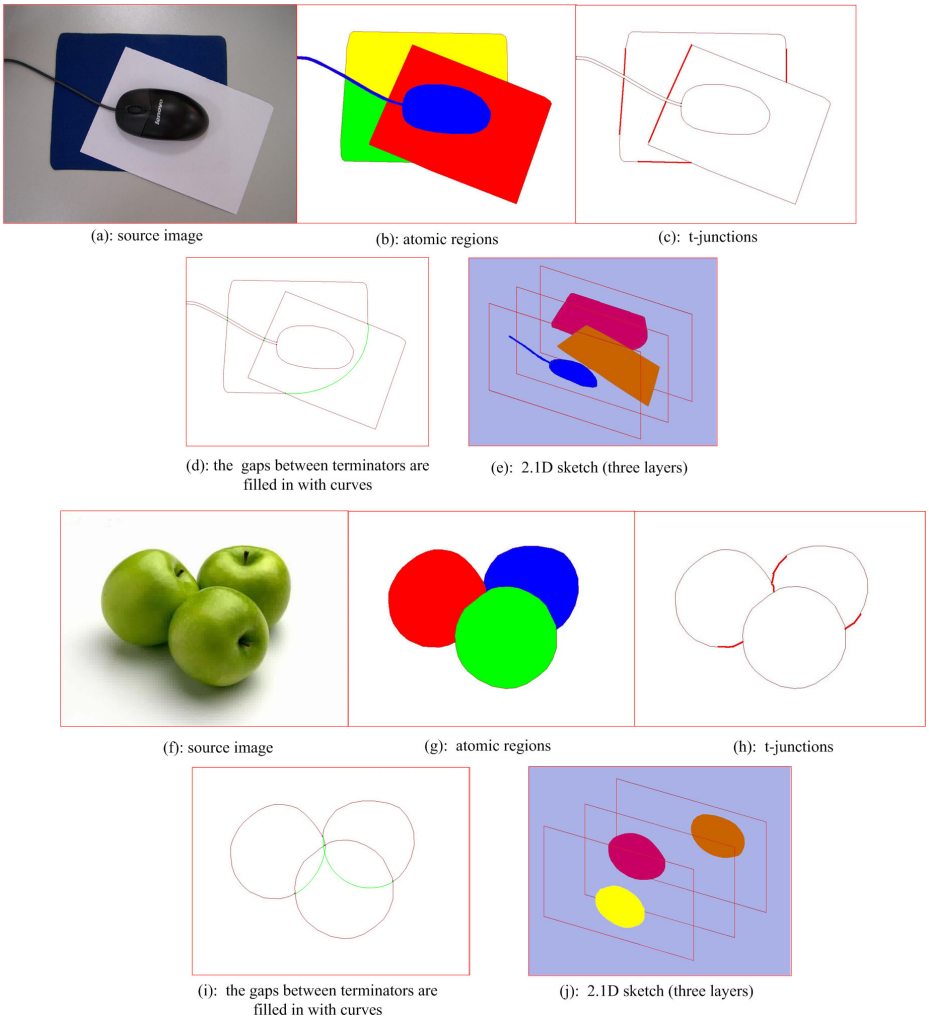

Fig. 4. More experiments

8. Stella X.Yu,T S.Lee and T Kanade: A Hierarchical Markov Random Field Model for Figure-Ground Segregation.EMMCVPR.2001 118–133
9. T.Chan and J.Shen: Mathematical Models for Local Nontexture Inpaintings.SIAM Journal of Applied Mathematics.vol. 62,2002 1019–1043
10. M.Bertalmio,G.Sapiro and C.Ballester: Image Inpainting.Computer, Graphics, SIGGRAPH 2000
11. L.Joyeux,O.Buisson and B.Besserer: Detection and Removal of Line Scratches in Motion Picture Films.Proceedings of CVPR'99, IEEE Int.Conf. on Computer Vision and Pattern Recognition, FortCollins.1999
12. oshi.S ,A.Srivastava and W.Mio: Hierarchical Organization of Shapes for Efficient Retrieval.In: Proc. ECCV 2004. 2004 570–591
13. Kumar, M.P.Torr and P.H.S.Zisserman: Obj Cut.In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.vol. 3,2005
14. Authors from the same group: Compositional boosting for computing hierarchical image structures. CVPR2007 (submitted).
15. B.B.Kimia,I.Frankel and A.M.Popescu: Euler spiral for shape completion.International journal of computer vision 54,2003 159–182
16. D.Mumford and J.Shah: Optimal approximations of piecewise smooth functions ans associated variatioanl problems.comm. in pure and appl.Math,vol. 42,1989 577–685
17. E.Saund: Perceptual organization of occluding contours generated by opaque surfaces.CVPR 1999.19999 624–630
18. H.Shum: Prior, Context and Interactive Computer Vision.The Microsoft Research Asia Technical Report.2006
19. B.K.P.Horn: The curve of least energy.ACM Transactions on Mathematical Software.vol. 9,1983 441–460
20. C.Ballester,M.Bertalmio and V.Caselles: Filling-In by Joint Interpolation of Vector Fields and Gray Levels.IEEE Transactions on Image Processing.vol. 10,2001 1200–1211
21. V.Kolmogorov and R.Zabih: What Energy Functions Can Be Minimized via Graph Cuts.IEEE Transactions on Pattern Analysis and Machine Intelligence.vol. 26,2004 147–159
22. Arthur Fridman: Mixed Markov models.Applied mathematics. PNAS, vol 100, no.14, 2003 8092–8096
23. Gilks,W.R.,Richardson,S.and Spiegelhalter: Markov Chain Monte Carlo In practive,Chapman and Hall.1996
24. Eric Saund: Perceptual organization of occluding contours generated by opaque surfaces. In Proceedings of the 1999 Conference on Computer Vision and Pattern Recognition, pages 624–630, 1999.
25. Davi Geiger, Krishnan Kumaran and Laxmi Parida: Visual organization for figure/ground separation. In Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pages 155–160, 1996
26. Edward A.Adelson and John Y.A.Wang: Representing Moving Images with Layers, IEEE Trans. on Image Processing, 3, pp. 625-638, Sept. 1994
27. Jingbin Wang, Erdan Gu and Margrit Betke: MosaicShape: Stochastic Region Grouping with Shape Prior ,Computer Vision and Pattern Recognition, pages 902–908 vol.1,2005.
28. Alexei A.Efros and William T.Freeman:Image Quilting for Texture Synthesis and Transfer, Proceedings of SIGGRAPH '01, Los Angeles, California, August, 2001.