Stereo Matching Using Population-Based MCMC

Wonsik Kim & Joonyoung Park & Kyoung Mu Lee

Department of EECS, ASRI, Seoul National University

DM research Lab., LG Electronics Inc.

Presented by
Pravee Kruachottikul

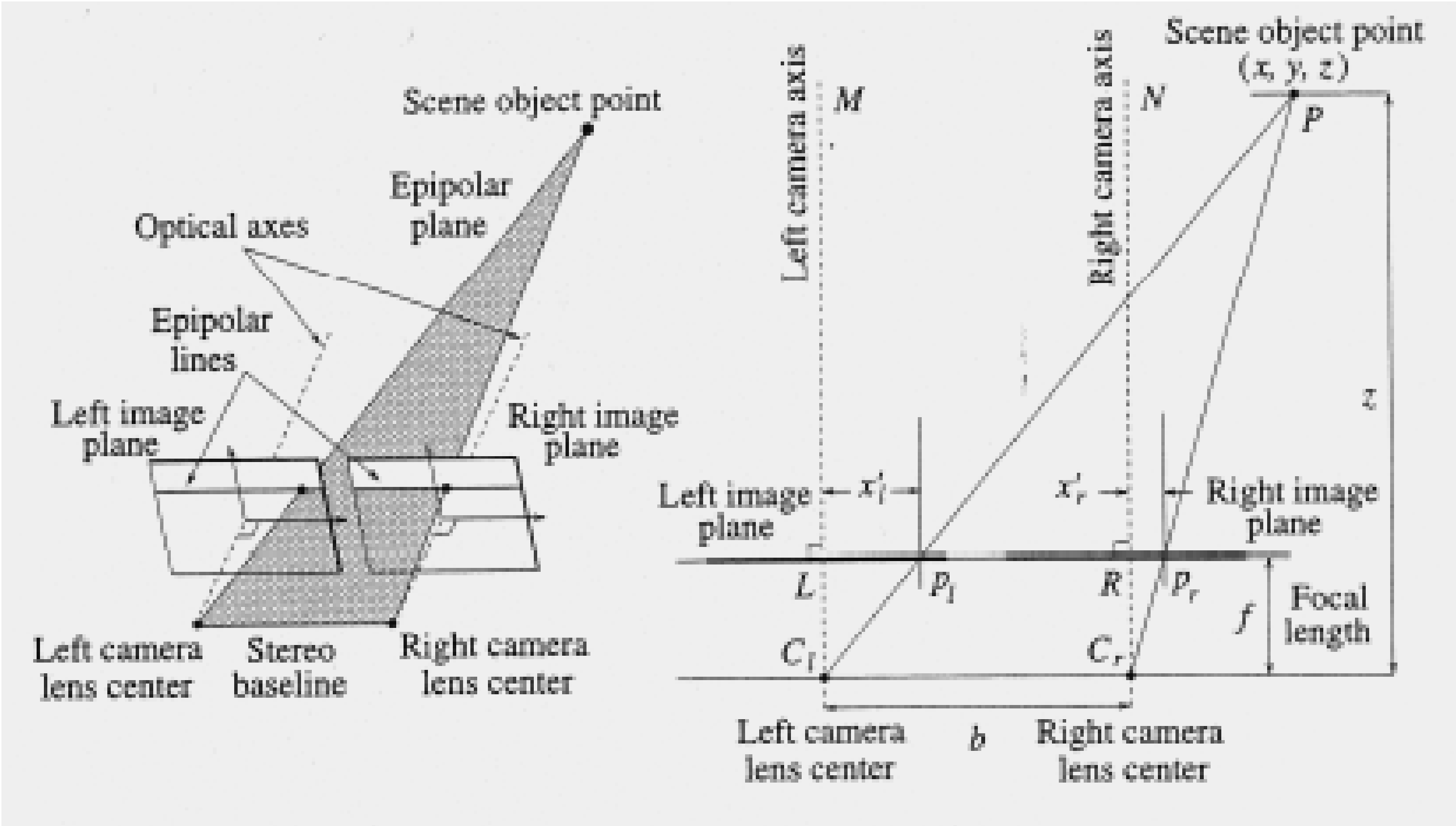
Topics

1. Objectives
2. Problems of Previous Works
3. Proposed Concepts
4. Proposed Algorithms
5. Experimental Results
6. Limitations
7. Conclusions

Objectives

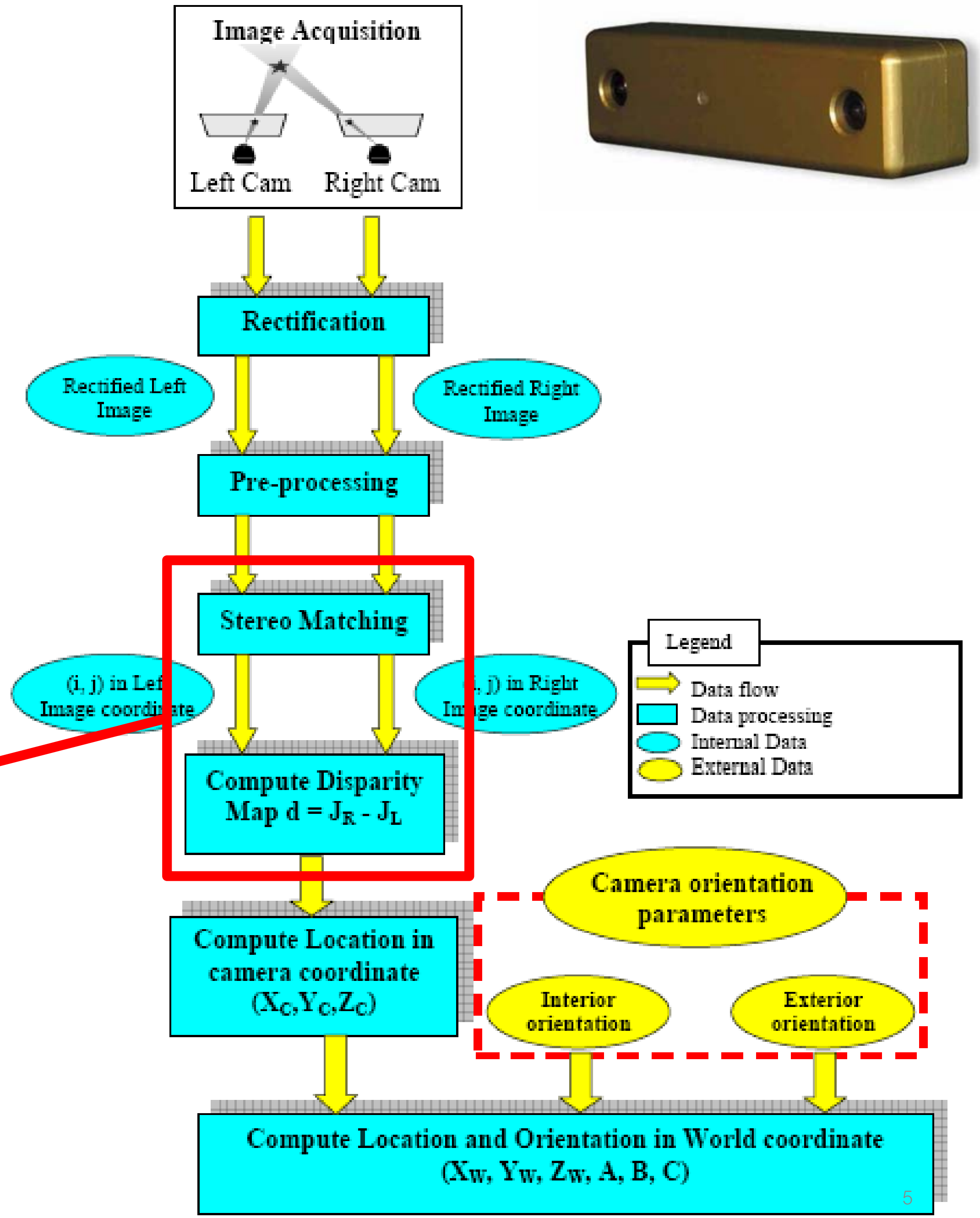
1. To propose a new stereo matching method using the population-based Markov Chain MonteCarlo (Pop-MCMC), which belongs to the sampling-based methods.

2. Comparison of performance of POP-MCMC with other methods such as SA, SWC, and BP.



- The goal of stereo matching is to determine disparities (the displacement between the positions of the two points, Z).

$$z = \frac{bf}{(x_l' - x_r')}$$



If we obtain an accurate disparity map, we can recover the correct 3-D scene information.

Stereo Vision Paradigm and Location computing algorithm

Recall Stereo Matching Algorithm

It can be classified into two approaches.

1. **Local Approach:**
   - Disparities are determined by comparing the intensity values in local windows by some measures such as SAD, SSD, and Birchfield-Tomasi measure.
   - Although they are fast, they have difficulties in obtaining an accurate disparity map due to some intrinsic difficulties including the noise and choosing appropriate window size problems. → Better to use global approach

2. **Global Approach:**
   - Exchanging data between multiple samples.
   - Apply the smoothness constraint that reflects the smoothly varying surface assumption.

Problems of Previous Works

- Originally, Monte-Carlo (MC) method is used since it's the most primitive sampling based method (it generates samples from a given target distribution or to integrate functions in high dimensional space for energy minimization to solve the stereo matching problem).

  But there is a problem to apply to vision applications as an optimizer, which takes infinitely long time.

- Thus, to resolve these problems, Markov Chain Monte Carlo (MCMC) methods had been tried (a new sample is drawn from the previous sample with a local transition probability, based on the Markov chain).

  But there is also a problem since most MCMC methods allow only local moves in a large solution space, it still takes very long time to reach the global optimum.

Problems of Previous Works

- To overcome the limitations of MCMC methods as an optimizer, Swendsen-Wang Cuts, SWC, was proposed.
- Advantages over the previous methods are :
  - Bigger local moves are possible than in previous methods while maintaining the detailed balance.
  - SWC uses Simulated Annealing (SA) to find the global optimum.
- Although SWC allows bigger local moves, a very slow annealing process is needed to approach the global optimum with probability 1. **This is an apparent drawback of SWC.**
- Therefore, we need a faster annealing process for real vision applications. However, fast annealing does not always guarantee the global optimum and the samples are often trapped in local optima.

The goal of Pop-MCMC

1. To overcome the drawbacks of SWC for stereo matching problem

2. To obtain a lower energy state faster than other sampling methods such as SWC (Optimization).

Done by

- Perform Global move (two or more samples are drawn at the same time and samples can exchange information with each other) so the mixing rate of drawn samples becomes faster (for optimization, faster mixing rate means shorter time for the samples to approach the global optimum).

Proposed Concepts

- It uses multiple chains in parallel and produces multiple samples at a time. It enables global moves by exchanging data between samples, leads to faster mixing rate.

- To apply to the stereo matching problem, we design two effective 2-D mutation and crossover (Genetic Algorithm) moves among multiple chains to explore a high dimensional state space efficiently.

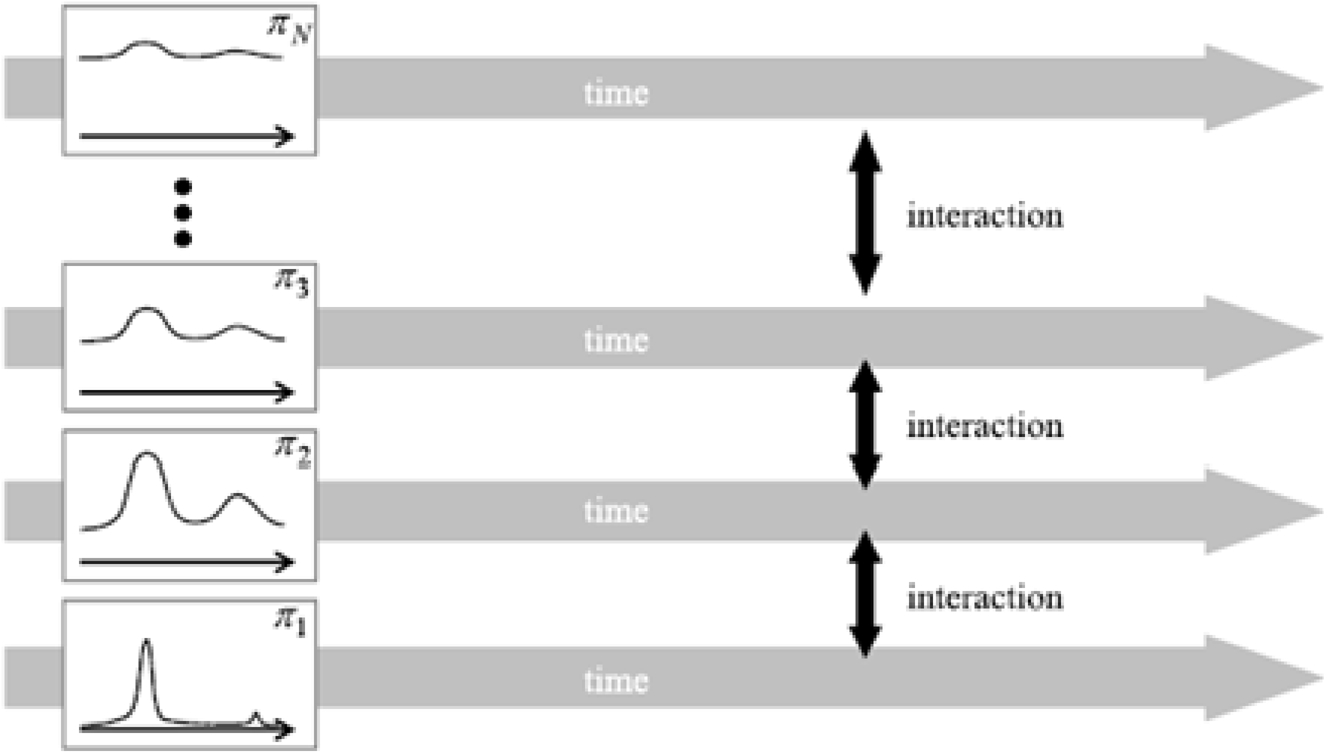
Proposed Concepts

- Pop-MCMC generates multiple chains in parallel with different temperatures, and exchanges information among them to accelerate the mixing rate (the Parallel Tempering , PT)

- PT aims to overcome the problems of single process MCMC using a Metropolis-Hastings update, which has low mixing rate. The basic idea of PT is to simulate multiple replicas of the original system in parallel at a series of different temperatures, and swap the configurations with a Metropolis-Hastings criterion.

Proposed Concepts

- The target distribution of the ith chain is $\pi_i(\mathbf{X}) = \pi(\mathbf{X})^{\frac{1}{T_i}}$,
- The idea of PT is illustrated below



Proposed Algorithm

1. Segment-Based Stereo Energy Model : to improve the accuracy of the disparity map. It also reduces running time since the number of nodes is much smaller than pixel-based model. And, mean-shift algorithm is used for the segmentation.

2. Pop-MCMC : to apply to Stereo Matching to find disparities for creating high dimensional state space efficiently.

Segment-Based Stereo Energy Model

1. Each segment is defined as a node : $v \in V$ and each node is assigned a label $X_i \in \{1, 2, \ldots, L\}$ (# of possible labels is L)
2. Neighboring nodes 's' and 't' are connected with edges : $\langle s, t \rangle \in E$
3. Constructing a graph G = (V,E)
4. The energy function is defined by

$$E(\mathbf{X}) = \sum_{v \in V} C_{\text{SEG}}(f_v) + \sum_{\langle s,t \rangle \in N} \beta_{s,t} \mathbf{1}(f_s \neq f_t),$$

where

- X = current state of every segment
- f_v = An estimated plane for each segment
- $C_{SEG}(f_v)$ = A matching cost

$$C_{\text{SEG}}(f_v) = \sum_{(x,y) \in V} C(x, y, f_v(x, y)),$$

Segment-Based Stereo Energy Model

- $\beta_{s,t}$ = A penalty for different neighboring nodes of s and t

$$\beta_{s,t} = \gamma \cdot BL(s,t) \cdot S(s,t),$$

where

- $C(x,y,f_v(x,y))$ = The Birchfield-Tomasi cost
- $BL(s,t)$ = The shared border length
- $S(s,t)$ = The mean color similarity defined by

$$S(s,t) = \frac{1}{2}\left(1 - \min\left(1, \frac{|R_{V_s} - R_{V_t}| + |G_{V_s} - G_{V_t}| + |B_{V_s} - B_{V_t}|}{255}\right)\right) + \frac{1}{2},$$

where

- R_{Vs}, G_{Vs}, and B_{Vs} are ave. intensity values of segment Vs (bet. 0 - 255)
- Mean color similarity value is between ½ and 1 (when 2 neighboring segments have similar intensities, it is closer to 1)

- By varying γ, we can control the relative effect of matching and smoothness cost.

Segment-Based Stereo Energy Model

5. For each pixel, we calculate the initial disparity by using SAD and WTA, using these initial disparities, we fit a plane for each segment. Thus, the eq. of a plane in 3D-space is as

$$d(x, y) = c_1 x + c_2 y + c_3,$$

where

- x and y are the coordinates of a pixel
- d(x,y) is its disparity

Segment-Based Stereo Energy Model

6. Construct the algebraic equation for each segment

$$\mathbf{A}\,[c_1, c_2, c_3]^{\mathrm{T}} = \mathbf{B},$$

where

- The ith row of the matrix A is the coordinates $[x_i, y_i, 1]$ of the ith pixel
- The ith row of the matrix B is the disparity $d(x_i, y_i)$ of that pixel

7. We can find c1,c2,c3 from a least squares solution of above eq.
8. Outlier disparities are initially detected and removed by a disparity crosschecking method.
9. Once we find the plane parameters, we can further identify more outlier disparities that are not close to the fitted plane.

Segment-Based Stereo Energy Model

10. For those outlier disparities pixels, we re-estimate the correct disparities by confining the search range to be small near the fitted plane.
11. The least squares method is repeated to update parameters c1,c2,c3 based on the modified disparities.
12. Next, above plane fitting process is repeated for each segment and newly found planes are added to a list.
13. After that, each segment is assigned to a plane in the list that has lowest C_{SEG} value.
14. Then, we group the segments assigned to the same plane.
15. And, for each group, the above plane fitting is repeated to improve the plane accuracy.
16. At last, we'll have the final list of the planes to use.

Pop-MCMC

1. Given a target probability distribution $\pi(\mathbf{X}) \propto \exp\{-E(\mathbf{X})\},$
2. Aim is to find the state X where the probability is maximized.
3. In Pop-MCMC, we draw multiple samples from multiple chains at the same time with respect to the following distributions.

$$\pi_i(\mathbf{X}_i) = \pi(\mathbf{X}_i)^{\frac{1}{T_i}} \propto \exp\left\{-\frac{E(\mathbf{X}_i)}{T_i}\right\},$$

where

- T_i is the temperature of ith chain

4. Each sample from each chain is chromosome, which interacts with each other helps perform global moves.

Algorithm Flow Chart

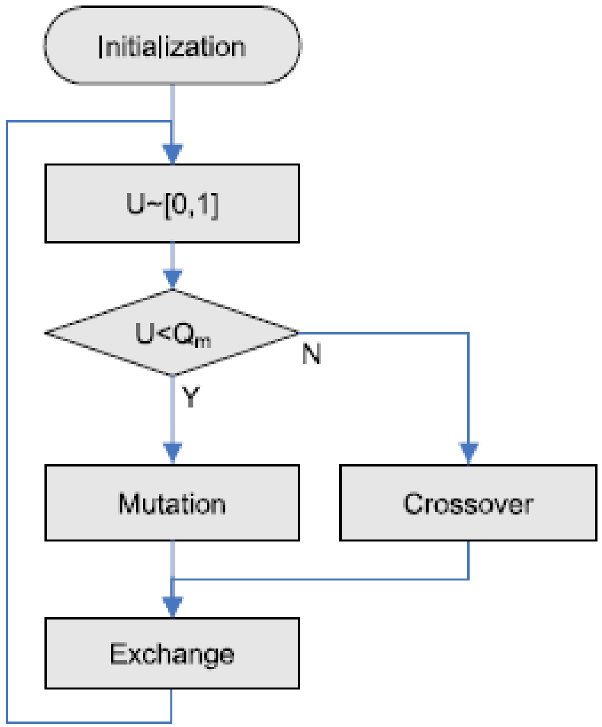


Fig. 2 The overall flow chart of the proposed Pop-MCMC algorithm applied to stereo matching

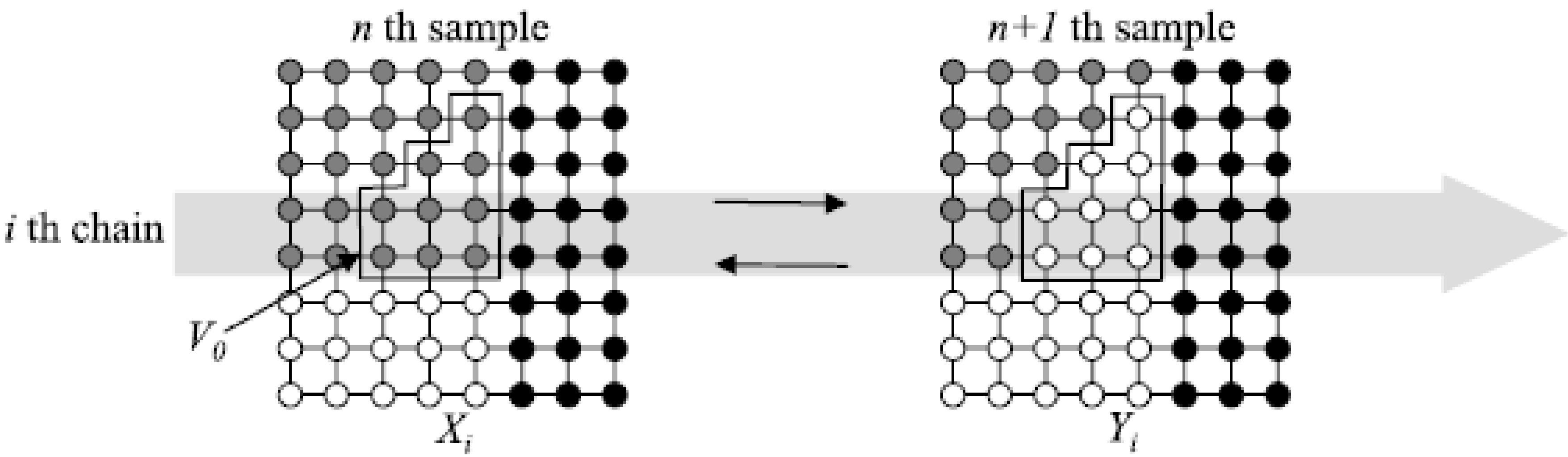
- U = random number between 0-1
- Q_m = mutation rate

<u>Note:</u>

- By varying Qm, we can control the rates between the global move (crossover) and local move (mutation).
- If a large number of chains are used, Qm is usually set to a small value for faster convergence.

- The three moves are repeatedly performed and samples are generated at each iteration.
- Proper Qm can be chosen due to the given problem, model, or number of chains. For ex, if a large number of chains are used, Qm is usually set to a small value for faster convergence.

Mutation Move



Mutation Move

- If the labels of two neighboring nodes s and t are different, the edge connecting two nodes is removed.
- If the labels are the same, we determine whether the edge is retained or not with the probability q_e.
- If there exists external field, we consider it in designing the probability q_e.

$$q_e = 1 - \exp\left(-\frac{K_i \cdot S(s,t)}{\frac{C_{SEG}(f_{v1})}{N(v_1)} + \frac{C_{SEG}(f_{v2})}{N(v_2)} + 2}\right)$$

where

- v_1 and v_2 represent neighboring nodes
- N(v) = number of the pixels in the node (segment) v,
- K_i = a weighting factor for the chosen ith chain.

Mutation Move

- This process is repeated for all edge e = s, t ∈ E.
- Next, nodes connected by remaining edges are considered as a cluster. Thus, one cluster Vo is randomly selected.

Mutation Move

- The new label l' for the selected cluster Vo is proposed as

$$q(l' \mid V_0, \mathbf{X}_i) = \exp\left[-\left\{ \frac{\sum_{v \in V_0} C_{\mathrm{SEG}}(f_v = l')}{\sum_{v \in V_0} N(v)} + 1 - \prod_{\langle v_1, v_2 \rangle \in N, v_1 \in V_0, v_2 \notin V_0} \mathbf{1}(l' = f_{v_2}) \right\} \right],$$

where

- l' is the newly proposed label for Vo.
- X_i is the current state of selected ith chain.

Mutation Move

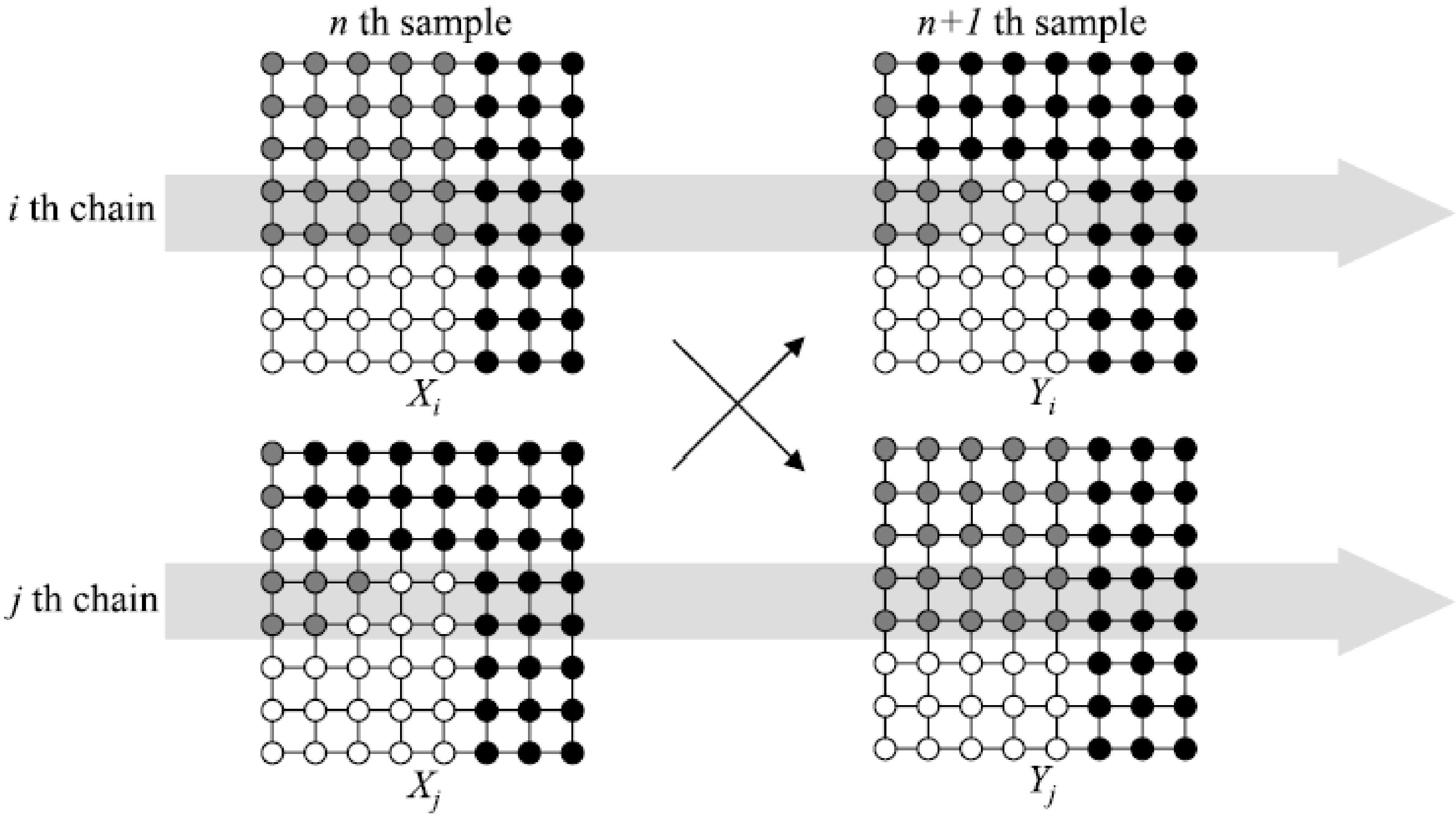
- Based on Metropolis-Hastings, The acceptance probability is

$$\alpha = \min(1, \gamma_m)$$
$$= \min\left(1, \frac{\pi_i(\mathbf{Y}_i)}{\pi_i(\mathbf{X}_i)} \cdot \frac{T(\mathbf{Y}_i \to \mathbf{X}_i)}{T(\mathbf{X}_i \to \mathbf{Y}_i)}\right)$$
$$= \min\left(1, \exp\left\{\frac{E(\mathbf{X}_i) - E(\mathbf{Y}_i)}{T_i}\right\}\right.$$
$$\left. \cdot \frac{q(V_0 \mid \mathbf{Y}_i) q(l \mid V_0, \mathbf{Y}_i)}{q(V_0 \mid \mathbf{X}_i) q(l' \mid V_0, \mathbf{X}_i)}\right),$$

where

- Y_i = the proposed state of the ith chain
- $Q(Vo \mid X_i)$ = the probability for selecting cluster Vo when current state is Xi.

Exchange Move



Exchange Move

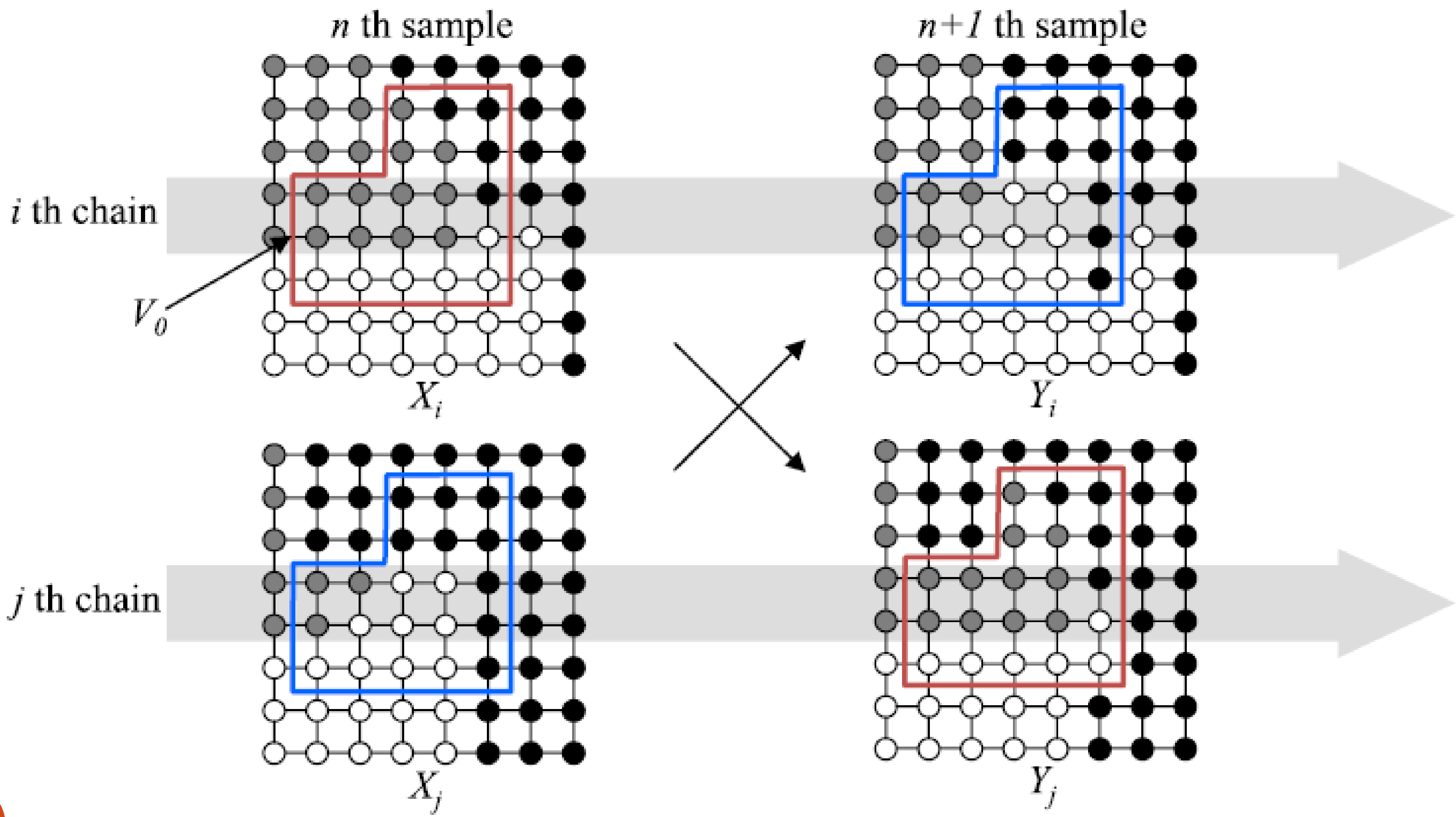
- We choose two chains and propose to exchange the chromosomes of two chains.
- The proposal is accepted or not by the Metropolis-Hastings rule.

$$\begin{aligned}\alpha &= \min(1, \gamma_e) \\ &= \min\left(1, \frac{\pi_i(\mathbf{X}_j)\pi_j(\mathbf{X}_i)}{\pi_i(\mathbf{X}_i)\pi_j(\mathbf{X}_j)}\right) \\ &= \min\left(1, \exp\left[\{E(\mathbf{X}_i) - E(\mathbf{X}_j)\} \cdot \left(\frac{1}{T_i} - \frac{1}{T_j}\right)\right]\right),\end{aligned}$$

where

- X_i and T_i are the current state and temperature of the ith chain.

Crossover Move



Crossover Move

- Select 2 chains randomly and construct a cluster Vo similar to Mutation Move except
  1. q_e is constant, not adaptively determined with the matching costs or the intensities of the input image.
  2. When calculate q_e, no need to check whether the labels of the nodes are the same or not, so Vo can have nodes with different labels.
- Compared with the mutation move (requires the identifying and removing processes of all the edges connecting the nodes with different labels), the selecting scheme and calculation of the acceptance probability of Vo in the crossover move is much simpler.
- Enables high efficiency in computation, and freedom in constructing of Vo helps to achieve faster convergence.

Crossover Move

- The acceptance probability

$$
\begin{aligned}
\alpha &= \min(1, \gamma_c) \\
&= \min\left(1, \frac{\pi_i(\mathbf{Y}_i)\pi_j(\mathbf{Y}_j)}{\pi_i(\mathbf{X}_i)\pi_j(\mathbf{X}_j)} \cdot \frac{q(\mathbf{X}_i, \mathbf{X}_j | \mathbf{Y}_i, \mathbf{Y}_j)}{q(\mathbf{Y}_i, \mathbf{Y}_j | \mathbf{X}_i, \mathbf{X}_j)}\right) \\
&= \min\left(1, \frac{\pi_i(\mathbf{Y}_i)\pi_j(\mathbf{Y}_j)}{\pi_i(\mathbf{X}_i)\pi_j(\mathbf{X}_j)}\right) \\
&= \min\left(1, \exp\left[\frac{E(\mathbf{X}_i) - E(\mathbf{Y}_i)}{T_i} + \frac{E(\mathbf{X}_j) - E(\mathbf{Y}_j)}{T_j}\right]\right),
\end{aligned}
$$

Algorithm 1 Proposed Pop-MCMC algorithm

```
(Initialize)
Initialize the population X_{1:N} by Winner-Takes-All manner with data cost.
Set the temperatures T_1 < T_2 < ··· < T_N.
repeat
  if U ~ [0, 1] < Q_m then
    for i = 1 to N do
      (Mutation)
      Select a random node v in ith chain.
      Draw a cluster from a node v with SWC-2.
      Propose a new label for the cluster and determine whether accept it or not with Metropolis-Hastings rule.
    end for
  else
    for i = 1 to ⌊N/5⌋ do
      (Crossover)
      Select two random chains and a random node v.
      Draw a cluster from node v with modified SWC-2.
      Determine whether swap the cluster or not with Metropolis-Hastings rule.
    end for
  end if
  for i = N − 1 to 1 do
    (Exchange)
    Perform the exchange move onto ith and i + 1th chains with Metropolis-Hastings rule.
  end for
until The algorithm converges.
```

Experimental Results

- Implemented the proposed algorithm on a 2.8GHz Pentium IV PC platform.
- Comparing the performance with other conventional methods as SWC-2, SA, BP, and Graph Cuts.
- Illustrate the effects of each move, temperature parameter, the number of chains.
- Tested the algorithm on several benchmark images in the Middlebury datasets (*http://vision.middlebury.edu/stereo*)
- Using a segment based energy model for testing.
- Repeat 10 times on each test stereo image pair, then AVE. and SD of the resulting energies were used to compare.

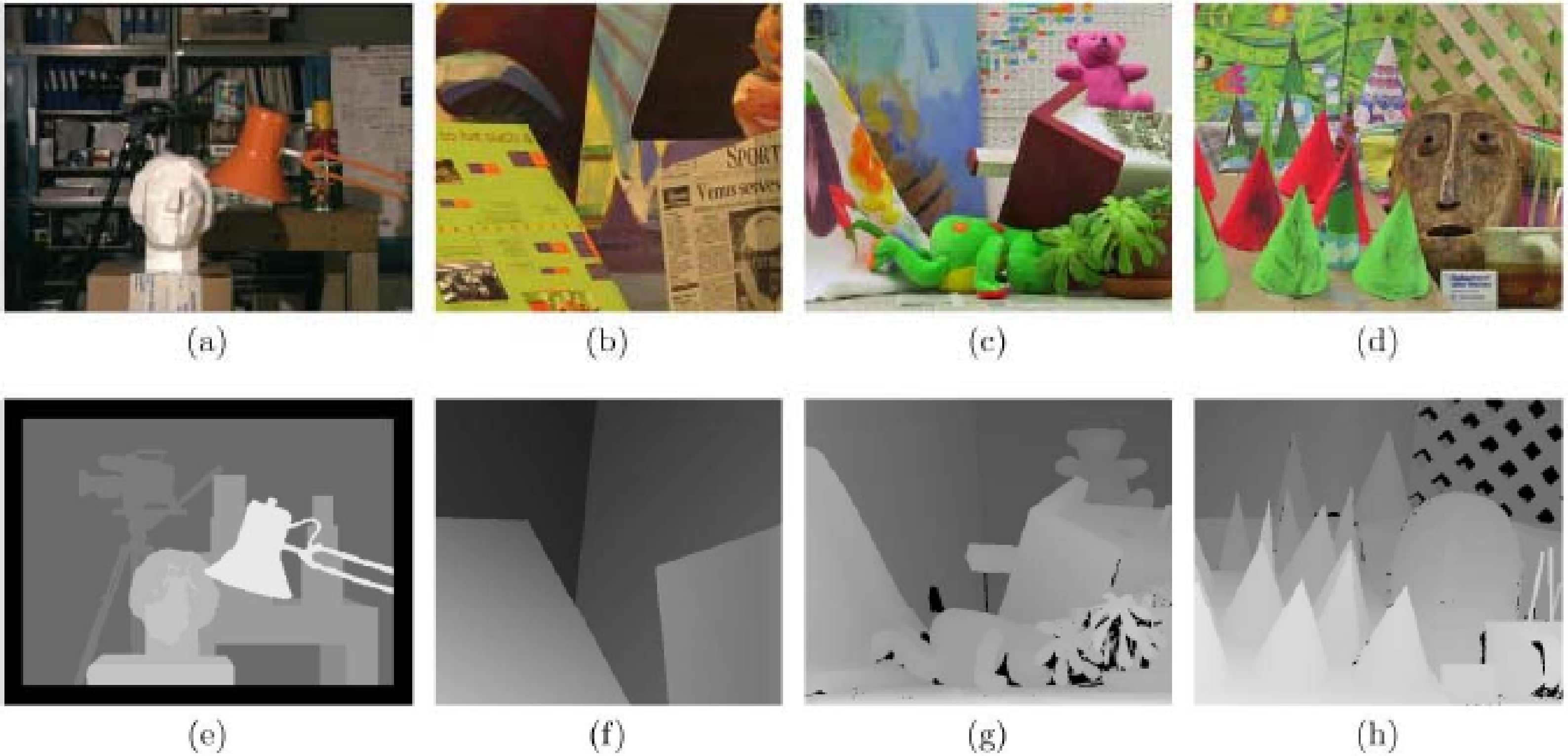
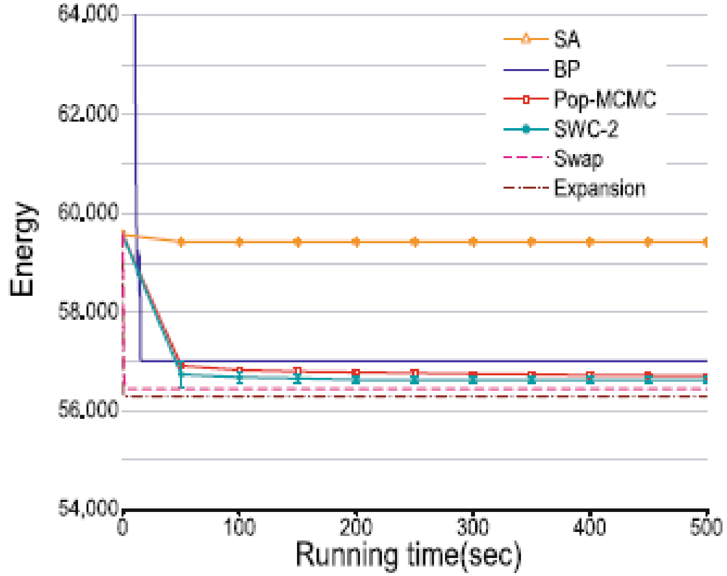
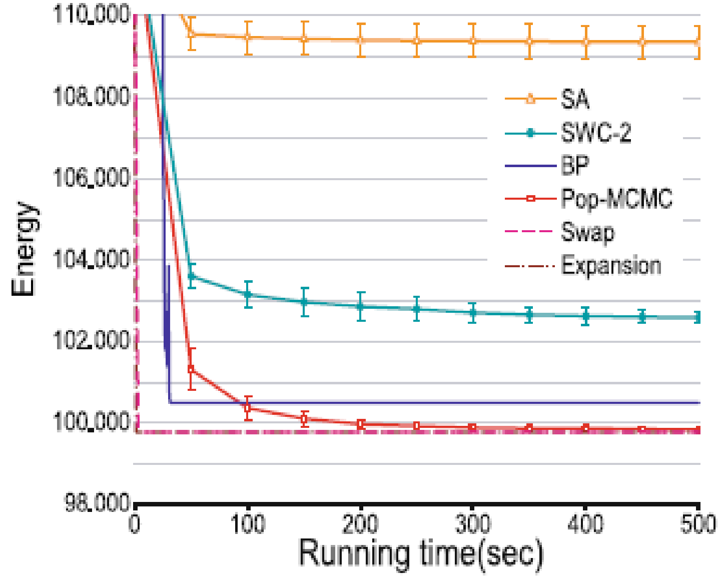


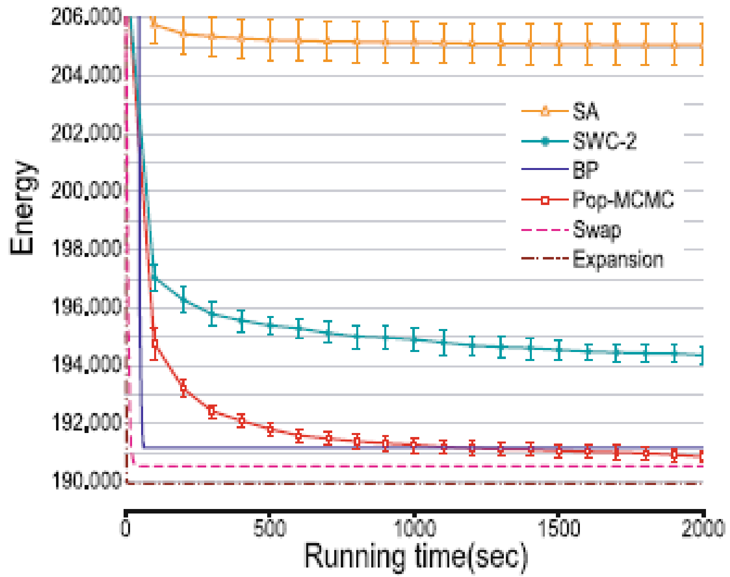
Fig. 6 Test stereo images: (**a**)–(**d**) reference images, (**e**)–(**h**) ground truth disparity maps. (**a**, **e**) Tsukuba, (**b**, **f**) Venus, (**c**, **g**) Teddy, and (**d**, **h**) Cones



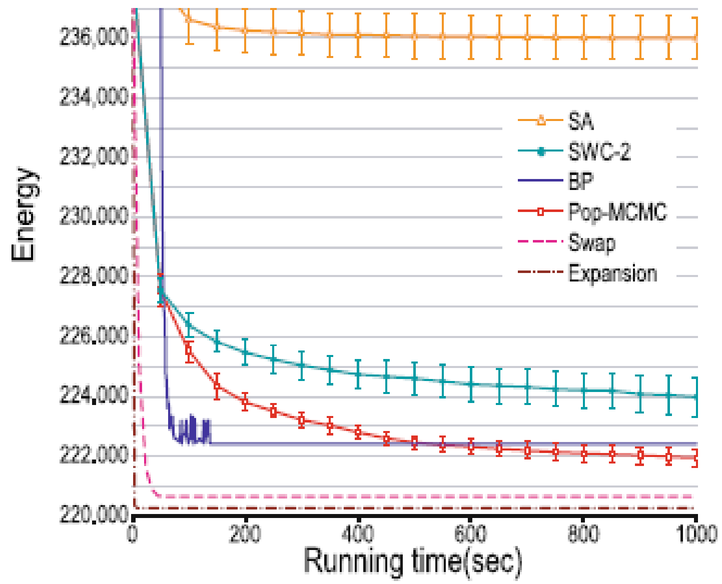
(a) Tsukuba



(b) Venus



(a) Teddy



(b) Cones

Table 1 The error rates for each test image (http://vision.middlebury.edu/stereo). For the sampling-based methods, we denote the average and standard deviation for ten trials. *nonocc*, *all*, and *disc* represent the error rate within non-occluded region, the whole image, and the vicinity of discontinuity, respectively

Method	Tsukuba			Venus			Teddy			Cones		
	nonocc	*all*	*disc*	*nonocc*	*all*	*disc*	*nonocc*	*all*	*disc*	*nonocc*	*all*	*disc*
Pop-MCMC	3.35 (±0.42)	3.88 (±0.42)	10.3 (±0.92)	0.22 (±0.01)	0.35 (±0.02)	2.89 (±0.17)	12.0 (±0.56)	17.9 (±0.69)	21.7 (±0.63)	13.3 (±0.37)	19.2 (±0.54)	23.7 (±0.57)
SWC	3.69 (±1.24)	4.28 (±1.23)	10.4 (±1.21)	0.9 (±0.27)	1.1 (±0.26)	5.57 (±0.11)	11.6 (±0.72)	17.8 (±0.86)	22.2 (±0.99)	13.5 (±0.72)	20.3 (±0.97)	23.4 (±0.76)
SA	3.5 (±0.28)	4.09 (±0.3)	9.58 (±0.48)	0.94 (±0.16)	1.34 (±0.21)	7.67 (±0.78)	14.8 (±0.61)	21.4 (±0.59)	24.3 (±0.79)	15.6 (±0.69)	22.9 (±0.78)	25.1 (±0.43)
BP	3.12	3.76	10.5	0.21	0.34	2.81	10.5	16.5	20.4	12.9	19.2	23.3
α-expansion	4.12	4.73	12.2	0.21	0.34	2.81	10.9	12.4	19.1	12.5	18.6	23.1
α-β-swap	2.56	3.09	9.15	0.21	0.34	2.81	10.5	12.0	19.7	13.0	19.0	23.6

Fig. 8 Results of the proposed algorithm: the disparity maps of (**a**) Tsukuba, (**b**) Venus, (**c**) Teddy, and (**d**) Cones

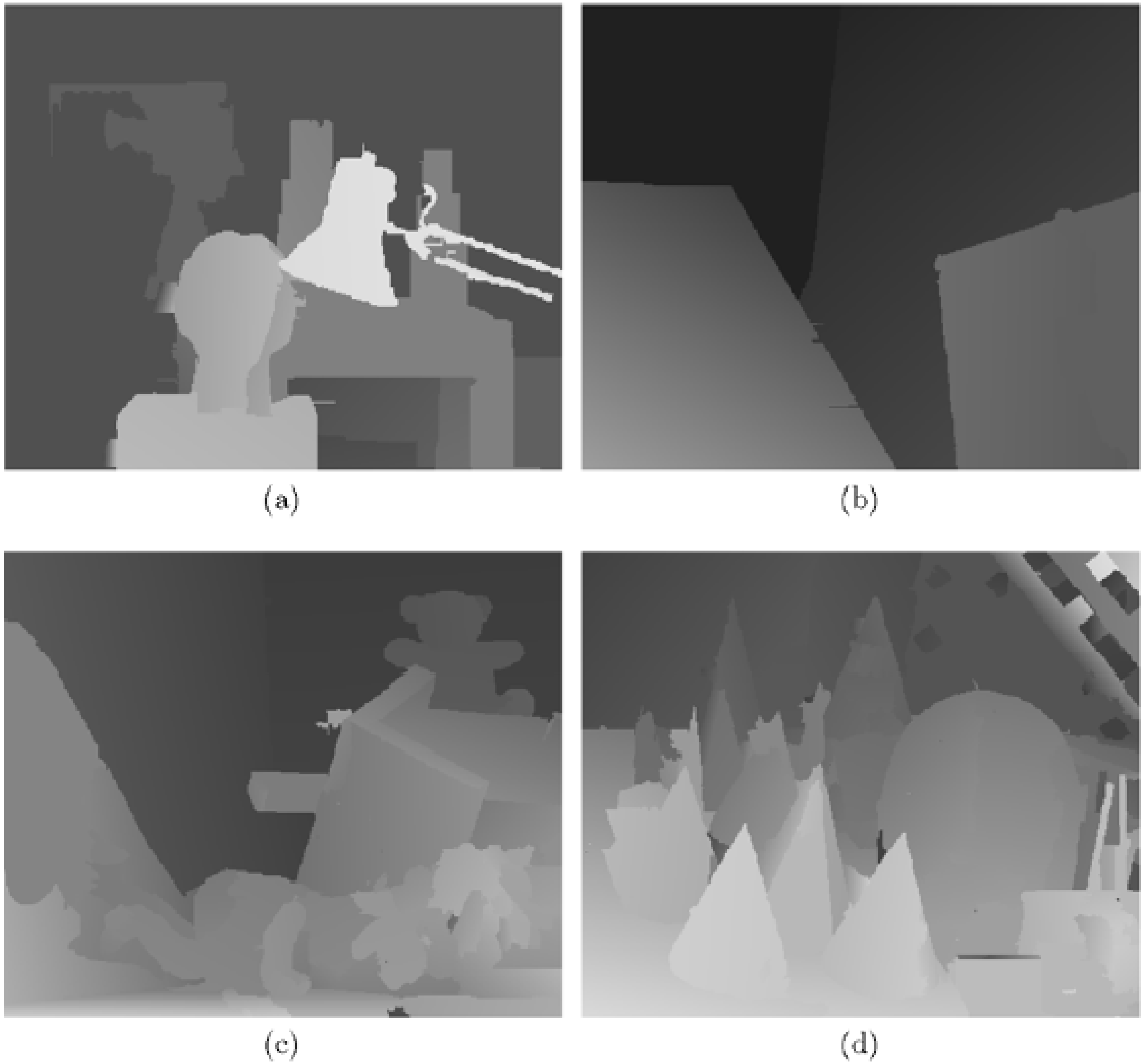
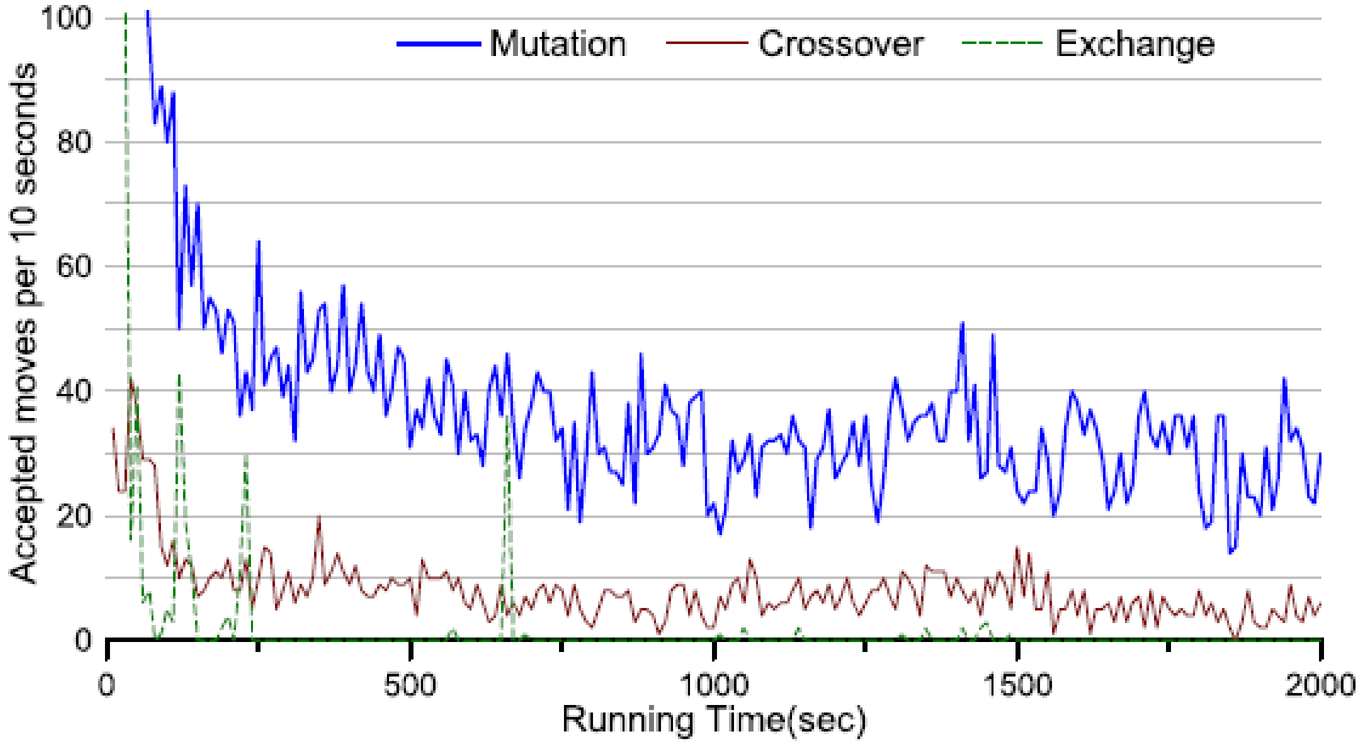


Fig. 9 The statistics of each move while running the algorithm on the Venus images. Initially, all the three moves are quite active, and then tend to decrease as time goes on. While mutation and crossover moves consistently occur, exchange move occurs occasionally



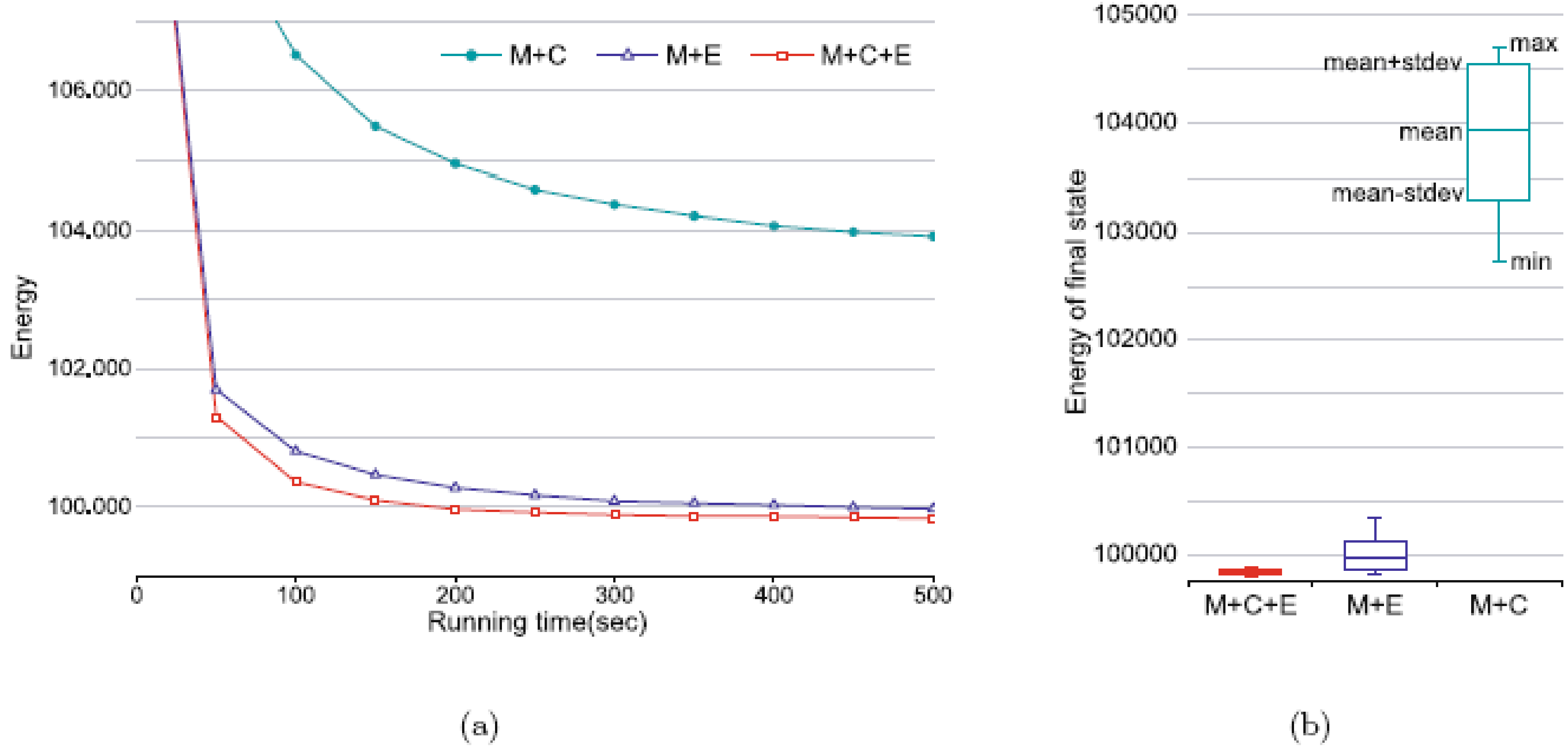
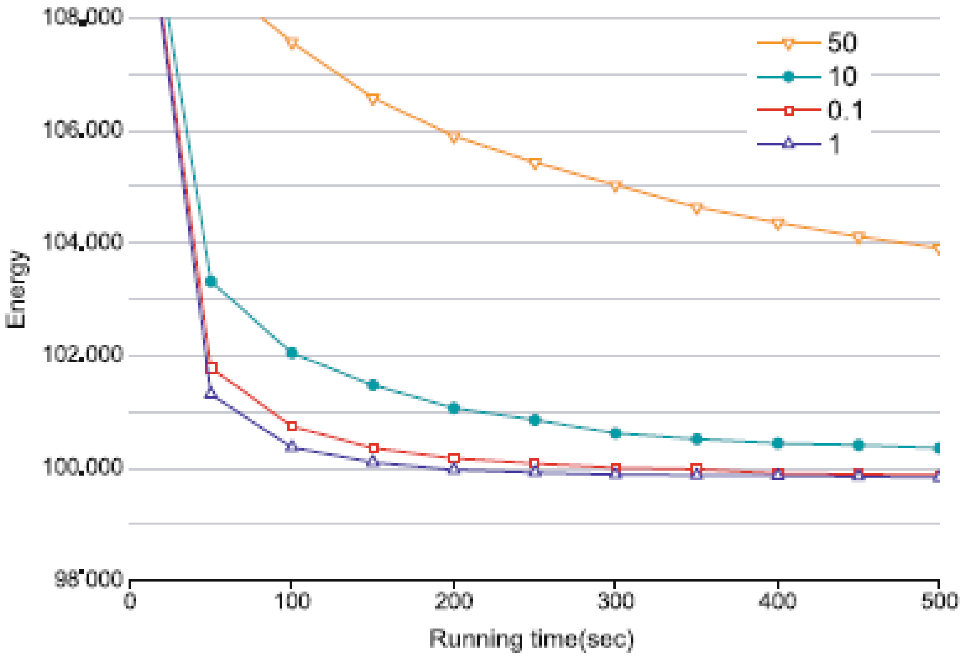
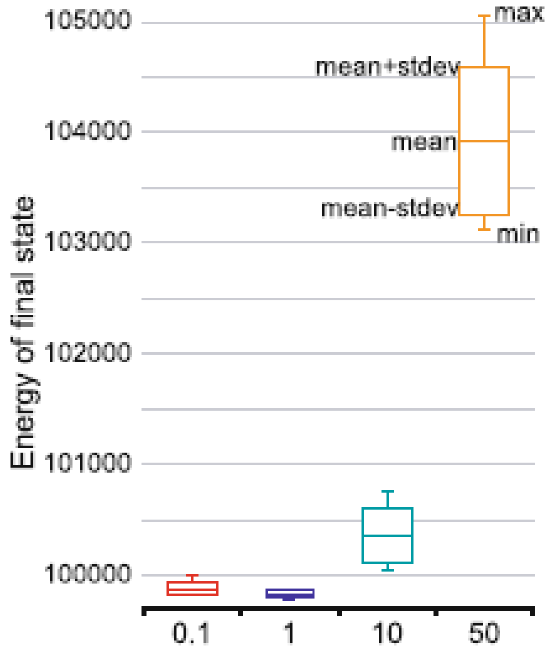


Fig. 10 The performance of the Pop-MCMC for different combinations of moves: (**a**) Energy curves, (**b**) boxplots of the final states



(a)



(b)

Fig. 11 The performance of the Pop-MCMC for different max temperature values: (**a**) Energy curves, (**b**) boxplots of the final states

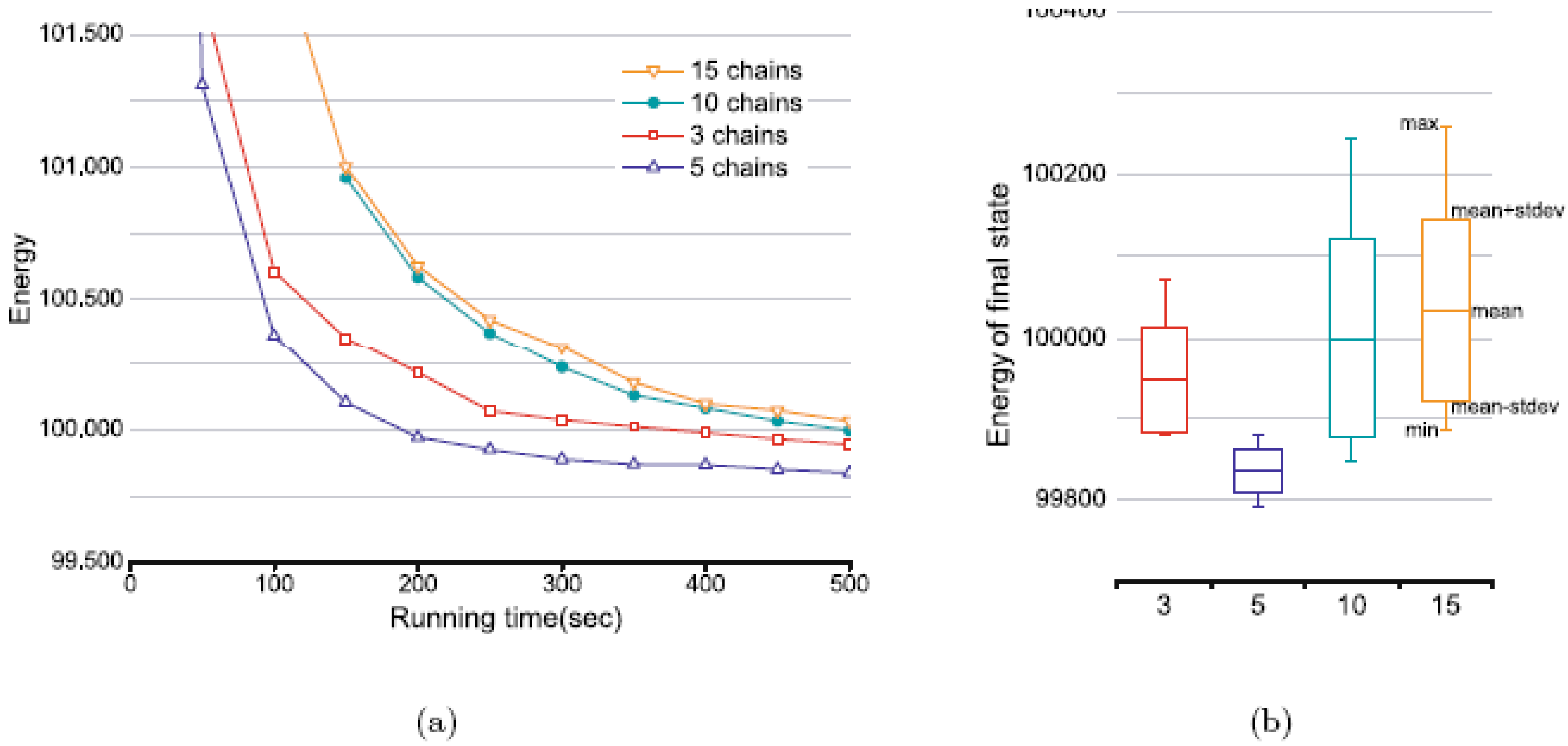


Fig. 12 The performance of the Pop-MCMC for different number of chains: (**a**) Energy curves, (**b**) boxplots of the final states

Limitations

- When objects are piecewise planar, the results are quite good. However, for the cases of Teddy and Cones that include objects with curved surfaces, the performance seems not satisfactory.

- For a front-parallel plane, a non-segment based energy model can be better than the segment-based energy model due to the smaller number of labels.

- Since occlude or visibility was not considered in our stereo model, the error rates at the vicinity of discontinuity were relatively large.

Conclusions

The proposed algorithm gives

1. Much faster convergence rate than conventional sampling-based methods including SA (Simulated Annealing) and SWC (Swendsen-Wang Cuts).

2. Consistently lower energy solutions than BP (Belief Propagation).

Thanks a lot for your attention

Question ..!!!