DISPARITY ESTIMATION USING COLOR COHERENCE AND STOCHASTIC DIFFUSION

Sang Hwa Lee, Nam Ik Cho*, and Jong-Il Park***

* School of Electrical Eng. and Computer Science, Seoul National Univ., Seoul, Korea.
** Division of Electrical and Computer Eng., Hanyang Univ., Seoul, Korea.
Emails: *lsh@ipl.snu.ac.kr, nicho@snu.ac.kr, jipark@hanyang.ac.kr*

ABSTRACT

This paper deals with disparity estimation based on Markov random field (MRF) models and color coherence. The disparity and line fields are explicitly modelled as MRFs, and are estimated by the stochastic diffusion. The potential functions are defined from the novel stochastic models between disparity and line fields. The color information is also utilized to model the textureless regions where the disparity estimation generally fails. And, the derived potential functions are minimized by the novel energy minimization method called stochastic diffusion. The stochastic diffusion diffuses the potential space using the probability distribution of neighboring fields, and searches for the optimal fields in the converged potential space. Some experiments show good performances of disparity estimation, which are compared with the other methods in a webpage.

1. INTRODUCTION

Recently, the main trend of correspondence (disparity and or motion) estimations is related to the energy minimization methods, which have two types of models, the likelihood and the prior models. In these approaches, the joint energy function of multiple fields is decomposed into two types of models by Maximum a posteriori (MAP) framework, and the prior models are defined by the MRFs. The optimal fields are estimated in the process of minimizing the energy function. The MAP-MRF estimations are suitable for the simultaneous and dense fields estimations such as correspondence, occlusion, line, and segmentation fields, since they utilize the prior information of the fields to reflect the interactions among the various fields[1, 2]. In addition, it shows good solutions in the ill-posed problems of computer vision while the standard regularization theory has limitations in the inverse problems. The important issues of the MAP-MRF framework are energy minimization scheme and MRF modelling of the fields. There are some different energy minimization methods based on the MAP-MRF framework, such as simulated annealing (SA) [5], mean field theory (MFT) [6], Bayesian diffusion [7], graph cut (GC)[8], belief propagation (BP) [9], and so on. These works deal with the energy minimization schemes defined by the MRF models and the disparity estimation.

The previous work of this paper has proposed the stochastic diffusion as an energy minimization method, and applied it to disparity estimation [10]. The stochastic diffusion considers the neighboring fields in MRF models with the joint probability density function (pdf). This is the main difference from the other energy minimization algorithms of MAP-MRF based estimators. The probability distribution of the neighborhood is introduced to the estimated fields in the form of probabilistic expectation. The each point, (i, j, d), has a potential value, and its potential is iteratively updated by the probabilistic expectation of the neighboring fields. The potential space defined by MRF models is iteratively diffused to a stable state, and the optimal fields are deterministically estimated by the winner-take-all manner at every pixel. As the expectation is iterated in the potential space, the effect of the neighboring fields is propagated to the farther positions, and the potential space seems to be stochastically diffused to a stable state. When the potential space is converged, the fields are determined at each pixel with the minimal potential value.

This paper focuses on the MRF modelling to estimate the disparity field in the MAP framework. It considers the energy minimization method and MRF modelling of the fields at the same time. The proposed algorithm combines with the line field to enforce the discontinuity of the disparity field at the image boundaries. The color coherence in the neighborhood are also exploited to improve the disparity estimates in the textureless regions. And, the stochastic diffusion is adopted in the energy minimization process. Though this paper is based on the stochastic diffusion of the previous works, the proposed method in the paper has some differences in the sense of MRF models of disparity and line fields, which is another important issue in the MAP-MRF framework.

This paper is organized as follows. Section 2 explains the basic formulation and the novel MRF models of the disparity and line fields. Section 3 shows some experimental results, and Section 4 concludes this paper.

0-7803-8554-3/04/$20.00 ©2004 IEEE.

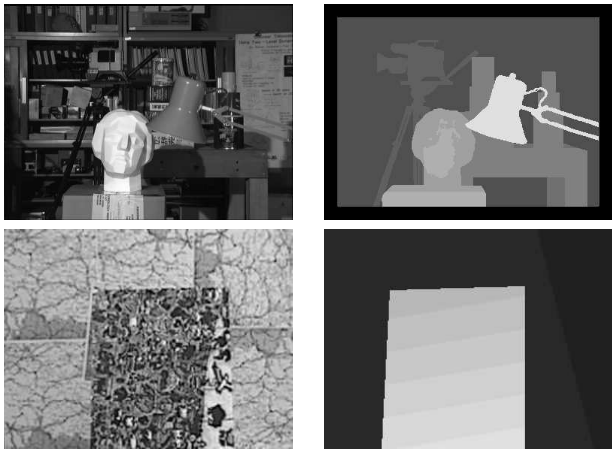


Fig. 1. Test images and the true disparity maps. (upper row) $Head$, disparity range:0-15, (lower) $Slanted$, 0-29.

2. STOCHASTIC MODELS

In the MAP-based estimation of disparity and line fields, the problem is to maximize the probability, $p(\mathbf{D}, \mathbf{L} \mid \mathbf{I}_r, \mathbf{I}_l)$. $\mathbf{I}_r, \mathbf{I}_l$ are stereo pair, and $\mathbf{D}$ and $\mathbf{L}$ are the disparity and line fields, respectively. The probability is generally decomposed into the marginal probabilities, and is transformed into a potential function to be minimized using the equivalence between MRF and Gibbs random field [5],

$$U(\mathbf{D}, \mathbf{L}|\mathbf{I}_r, \mathbf{I}_l) \propto U(\mathbf{I}_l|\mathbf{D}, \mathbf{L}, \mathbf{I}_r) + U(\mathbf{D}|\mathbf{L}, \mathbf{I}_r) + U(\mathbf{L}|\mathbf{I}_r). \quad (1)$$

The first term is the likelihood model which is related to the intensity difference. The second is the disparity field model which is based on MRF and plays an important role in regularization of the field. The third is the line field model to enforce the discontinuity of the disparity field.

In the next subsections, we derive the point-wise potential $U(d_{ij}, l_{ij})$ at the i-th row and j-th column in the image. The potential functions are defined by the MRF models of the fields and multiple cues such as color information.

2.1. Likelihood model

The likelihood model is the intensity-based error potential to describe how well the disparity field matches the points between two images. The usual block matching estimations are based on this measurement. Though the likelihood model is conditioned on the disparity and line fields, it is usually dependent only on the disparity field in real cases. The point-wise potential of the likelihood model is computed as the pixel matching algorithm,

$$U_o(\mathbf{I}_l(ij)|\mathbf{I}_r, d_{ij}) = \rho_i(\mathbf{I}_l(i,j) - \mathbf{I}_r(i, j + d_{ij})), \quad (2)$$

where, $\rho_i(\cdot)$ is a potential function such as the square function. Note that this paper assumes the epipolar geometry is

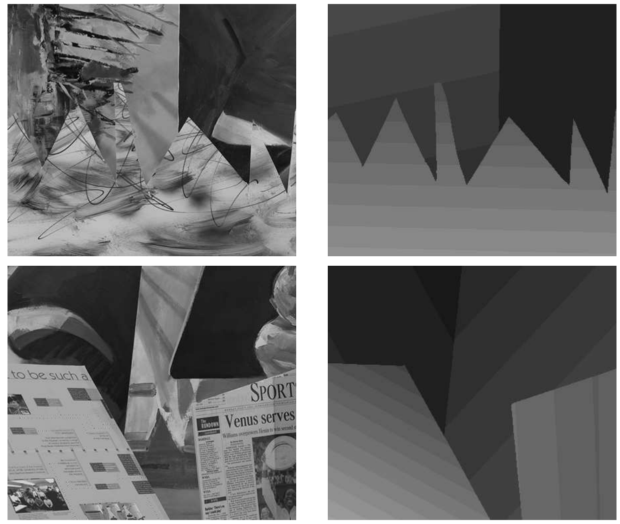


Fig. 2. Test images and true disparity maps. (upper row) $Sawtooth$, disparity range:0-19, (lower) $Venus$, 0-19.

1-D horizontal line.

2.2. Line field model

The line field enforces the discontinuity of the disparity field, and prevents the estimated disparity field from being blurred in the object boundaries. The MRF model of the line field is to isolate the adjacent field when the neighboring fields are likely to be discontinuous. That is to say, when the line field between the adjacent pixels, (ij) and (n), is *one*, the influence of the neighboring field is cut in the MRF models. Hence, the neighboring field is not related to the regularization process any more, and the estimated current field, d_{ij}, becomes independent of the neighboring field.

The line field model is based on the gradient of the intensity to examine the discontinuity of the disparity field. This results from the observation that the edges in the intensity are also likely to be edges in the disparity field. Therefore, the potential function of the line field should be defined to assign *one* at the high gradient. The higher the gradient is, the lower the potential value becomes to assign *one* to the line field. This principle implements the potential function of the line field to be inversely proportional to the gradient,

$$U(l_{ij}|\mathbf{I}_r) \propto \frac{C}{\nabla \mathbf{I}_r(ij, n)} \mathcal{L}(ij, n), \quad (3)$$

where, $\nabla \mathbf{I}_r(ij, n)$ is the gradient of the image, $\mathcal{L}(ij, n)$ is the line indicator between adjacent pixels, (ij) and (n), and a constant C is the penalty potential to determine the disparity discontinuity. The potential of the line field is a penalty to compensate for the potential of the disparity field at the discontinuous positions.

2.3. Disparity field model

The potential space of the disparity field is diffused by the stochastic model. The stochastic diffusion in the disparity field model is performed by iterative stochastic expectation using the probability distribution of the neighboring fields,

$$p(d_{ij}|\mathbf{L},\mathbf{I}_r) = \int_{\mathcal{N}\in\mathcal{S}} p(d_{ij}|\mathbf{d}_{\mathcal{N}},\mathbf{L},\mathbf{I}_r)p(\mathbf{d}_{\mathcal{N}}|\mathbf{L},\mathbf{I}_r)\, dd_n. \quad (4)$$

The potential space is diffused by propagating the effect of the neighborhood weighted by the probability distribution. And, the computational models of the disparity field, $U(d_{ij} \mid \mathbf{L},\mathbf{I}_r)$ are implemented to enforce both the smoothness of the field and the discontinuity in the boundary region. Even though the two purposes of the computational models have opposite properties, they can be simultaneously considered by the differences of the neighboring disparities and the line field. When the differences of neighboring disparities are large, the potential value of the disparity field is increased. On the other hand, the potential value is decreased when the differences are small. This mechanism regularizes the disparity distribution and guarantees the smoothness of the disparity field. In the paper, in addition, intensity information is utilized to improve the estimation in the textureless regions. The disparity estimation generally fails in the regions since the likelihood model is very noisy and the estimated fields are not unique. The proposed approach is based on the assumption that the pixels with similar intensity (or color) are likely to have the same disparity value. Therefore, we can introduce an additional term in the stochastic diffusion to consider the similarity of the intensities between adjacent pixels. The probability of the disparity field, $p(d_{ij}|d_{\mathcal{N}},\mathbf{L},\mathbf{I}_r)$, is implemented as

$$\prod_{n\in\mathcal{N}} \exp\left\{-\left[V_d(d_{ij},d_n)+W(d_{ij})\right]\left[1-\mathcal{L}(ij,n)\right]-U(l_{ij}|\mathbf{I}_r)\right\}, \quad (5)$$

where, $U(l_{ij}|\mathbf{I}_r)$ is $eq.$ (3), and $V_d(d_{ij},d_n)$ is a potential based on differences of the disparities. As described before in subsection 2.2, the line indicator isolates the neighborhood when the adjacent fields are discontinuous. The potential of the disparity field is replaced with that of line field model as a penalty. Note that the line field model in the paper is different from that of the previous work. This paper inserts the potential of the line field directly into that of the disparity field. Hence, the discontinuity of the disparity field is independently modelled for the each neighborhood while the previous work was modelled on the average for the neighborhood. This concept models the line field and disparity discontinuity more sharply. And, an additional term, $W(d_{ij})$, is based on the color coherence in the neighborhood,

$$W(d_{ij}) = \exp\left\{-\frac{(d_{ij}-\hat{d}_n)^2}{2\sigma_d^2}\right\}\cdot V_i(I_{ij},I_n), \quad (6)$$

where, $V_i(I_{ij},I_n)$ is the potential based on the difference between adjacent intensities. The term, $\exp\{-\frac{(d_{ij}-\hat{d}_n)^2}{2\sigma_d^2}\}$, prevents the current disparity, d_{ij}, from being assigned to the intermediately estimated neighboring disparity, $\hat{d}_n$, when the intensity difference, $V_i(I_{ij},I_n)$, is large. However, if the current disparity, d_{ij}, is different from the estimated neighboring disparities, the intensity difference effects little on the potential function, and the color coherence term has no meaning in the potential. That is to say, the color coherence makes an additional role in regularizing the disparity field only if $\hat{d}_n$ is similar to d_{ij}. By the additional term, the disparity fields with similar intensity (or color) are converged to the same value. Note that the pixels with similar intensities are not assigned to the same disparity directly. The coherence information is just exploited in the stochastic process as shown in $eq.$ (6), which is the main difference from the other color-based algorithms. Finally, the potential function of the disparity field in the stochastic diffusion is derived by substituting $eq.$ (4) with $eq.$ (5) and (6) and taking logarithmic function. We should make an emphasis that the probability of neighborhood fields, $p(\mathbf{d}_{\mathcal{N}}|\mathbf{L},\mathbf{I}_r)$ is the joint distribution which does not assume any probabilistic constraints such as the independence. The joint distribution of the neighborhood is implemented as,

$$p(\mathbf{d}_{\mathcal{N}}|\mathbf{L},\mathbf{I}_r) = \frac{\exp\left\{-\sum_{n\in\mathcal{N}} U_d(d_n|\mathbf{L},\mathbf{I}_r)\right\}}{\sum_{\mathcal{N}\in\mathcal{S}}\exp\left\{-\sum_{n\in\mathcal{N}} U_d(d_n|\mathbf{L},\mathbf{I}_r)\right\}}, \quad (7)$$

where, the joint pdf is proportional to the joint potential of the neighborhood.

2.4. Iterative procedure

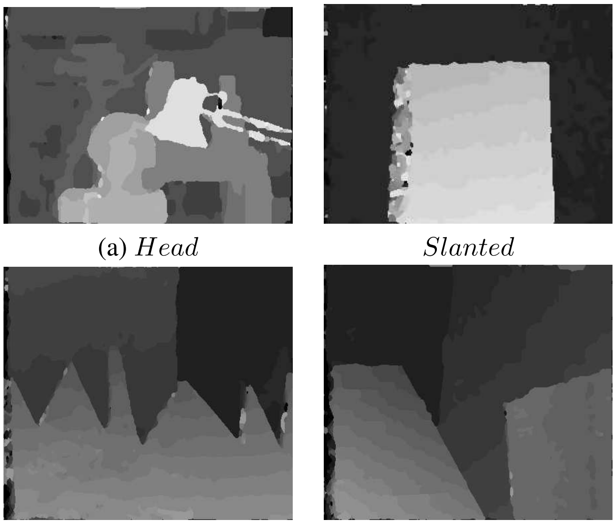
The 3D potential space is iteratively diffused to a stable state, and the fields with the minimum potential are selected as an optimal ones in each pixel. Since the line field is a binary value, the potentials of two values, 0 and 1, are examined at each pixel, respectively. The binary value with minimum potential is decided as the line field. The potential function of the disparity field as in $eq.$ (4) is iteratively diffused by the following relation,

$$U^{(k+1)} = \frac{1}{2}\{U_0 + \alpha U^{(k)} + \Delta U^{(k+1)}\}, \quad (8)$$

where, α is a weight, $\Delta U^{(k)}$ is the potential of the disparity field at the $k+1$ iterations, and $U^{(0)}$ is the likelihood model. The iterative relation is derived to preserve the relative weighting ratio between the potential terms since the priority of the potential terms depends on the relative weighting of the terms.

3. EXPERIMENTAL RESULTS

The experiments are performed with 4 stereo image pairs of grey levels as shown in Figure 1 and 2. These images are generally used since they have the true depth map and can be compared with the other methods in the website [11].



(a) *Head* *Slanted*

(c) *Sawtooth* (d) *Venus*

Fig. 3. Estimated disparity maps.

The neighborhood configuration used the plane configuration model since it showed the similar performances to those of the full neighborhood configurations from many experiments [10]. Since it's impossible to consider the whole configurations of the neighborhood in real implementations, we should select an optimal configurations to reflect the distribution as exactly as possible.

Figure 3 shows some experimental results of disparity maps. And, Table 1 shows some comparisons of error rates with the other methods that are based on MRF models with excellent performances. The error pixels have the absolute differences greater than 1 from true disparity values. The error rates are evaluated from the web site [11]. As shown in Table 1, the proposed method improves the previous works [10] for any images. The estimated disparity maps are improved at the textureless regions and the boundaries. These results prove that the color coherence and line field models are appropriately modelled in the stochastic process. In addition, the proposed method shows better performances for some images compared with the other best results. However, the proposed method should be optimized to show better performances for all the images. And, we should deal with the occlusion problem as shown in *Slanted* image in the future works.

4. CONCLUSIONS

Novel stochastic models to estimate disparity field are proposed in the paper. The disparity and line fields are estimated by the maximum a posteriori (MAP) framework and Markov random field (MRF) models. The disparity and line fields are explicitly modelled as MRFs, and are estimated in an energy optimization method called stochastic diffusion. The color coherence in the neighborhood is utilized to model the potential function of disparity field. Experimental results show the improved performance of the disparity estimation in the textureless regions. These results mean that the novel stochastic models including line field and color coherence are suitable for disparity estimation.

Table 1. The error rates (%) of disparity maps.

Image	[7]	[8]	[9]	[10]	Proposed
Head	6.49	1.86	1.15	3.95	2.88
Slanted	0.20	2.39	0.84	1.31	0.44
Sawtooth	1.45	0.42	0.98	2.45	2.01
Venus	4.00	1.69	1.00	2.45	1.38

5. REFERENCES

[1] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithm," *IJCV*, 47 (1/2/3), pp.7-42, 2002.

[2] A. Redert, E. Hendricks, and J. Biemond, "Correspondence estimation in image pairs," *IEEE Signal Processing Magazine*, Vol.16, No.3, pp. 29-46, 1999.

[3] Ze-Nian Li and Gongzhu Hu, "Analysis of disparity gradient based cooperative stereo," *IEEE Trans. on IP*,Vol.5, No.11, pp.1493-1506, Nov., 1996.

[4] C.L.Zitnick and T.Kanade, "A cooperative algorithm for stereo matching and occlusion detection," *IEEE Trans. on PAMI*, Vol.22, No.7, pp.675-684, July, 2000.

[5] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. on PAMI* , Vol.6, No.6, pp.721-741, 1984.

[6] J. Zhang and G. Hanauer, "The application of mean field theory to image motion estimation," *IEEE Trans. on IP*, Vol.4, No.1, pp.19-32. 1995.

[7] D.Scharstein and R. Szeliski, "Stereo matching with nonlinear diffusion," *IJCV*, Vol. 28, No.2, pp.155-174, 1998.

[8] Y.Boykov, O.Veksler, and R.Zahih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. on PAMI* , Vol.23, No.11, pp.1222-1239, Nov., 2001.

[9] J. Sun, H.Y. Shum, and N.N. Zheng, "Stereo Matching Using Belief Propagation," *ECCV-2002*, pp.510-524, 2002.

[10] S.H. Lee, Y. Kanatsugu, and J.I. Park, "Stochastic diffusion for stereo matching and line fields estimation," *IJCV*, vol. 47 (1/2/3) pp. 195-218, 2002.

[11] http://www.middlebury.edu/stereo