DISPARITY MAP RESTORATION BY INTEGRATION OF CONFIDENCE IN MARKOV RANDOM FIELDS MODELS

V. Murino, U. Castellani and A. Fusiello

Dipartimento Scientifico e Tecnologico, University of Verona
Strada Le Grazie 15, 37134 Verona, Italy
{murino,castellani,fusiello}@sci.univr.it

ABSTRACT

This paper proposes some Markov Random Field (MRF) models for restoration of stereo disparity maps. The main aspect is the use of confidence maps provided by the Symmetric Multiple Windows (SMW) stereo algorithm to guide the restoration process. The SMW algorithm is an adaptive, multiple-window scheme using left-right consistency to compute disparity and its associated confidence in presence of occlusions. The MRF approach allows to combine in a single functional all the available information: observed data with its confidence, noise, and a-priori hypotheses. Optimal estimates of the disparity are obtained by minimizing an energy functional using simulated annealing. Results with a real stereo pair show the improvement obtained by the restoration using a MRF approach integrating confidence data.

1. INTRODUCTION

Three-dimensional (3D) reconstruction is a fundamental issue in Computer Vision, and in this context, structure from stereo algorithms plays a major role. The process of stereo reconstruction aims at recovering the 3D scene structure from a pair of images by matching *conjugate points*, thereby finding the *disparity* [1]. Matching errors generate outliers and marked degradations of the disparity map, which makes a restoration process necessary. In this paper, the Symmetric Multiple Windows (SMW) stereo algorithm [2] is used, which performs a correlation between windows in the two images, and provides a confidence or reliability value of the disparity. Our contribution is the proposal of MRF models for the restoration of the disparity map integrating confidence data as well.

Many work about the image restoration or reconstruction by using MRFs has been done on both intensity (optical) [3] and range images [4]. Moreover, different approaches were proposed aimed at integrating additional information in the MRF model, like, for instance, edges to guide the line processing [5], or confidence data to guide the restoration and the segmentation of underwater acoustic images [6].

Regarding stereo algorithms, a large literature addresses the correspondence problem (see [1] for a survey). In general this algorithms select the best match according to a similarity measure that can be also used as a reliability (confidence) indicator of the goodness of the disparity. Apart from dynamic stereo frameworks [7], confidence information is utilized only to evaluate *a-posteriori* the quality of the disparity image, but they are not used to improve the disparity estimate.

This work proposes different models for the integration of the confidence data in a MRF framework, testing the real possibilities to use such information actively.

The rest of the paper is organized as follow. In Section 2, the stereo process is described. In Section 3, the MRF basic concepts are reported. Actual MRF models are detailed in Section 4, results are shown in Section 5 and conclusions are drawn in Section 6.

2. THE STEREO PROCESS

The disparity image is generated using the *Symmetric Multi-Window* (*SMW*) stereo algorithm [2], whose main features are summarized below.

Multiple windows. For each pixel the correlation is performed with nine different windows and the disparity with the smallest SSD *(Sum of Squared Differences)* error value is retrived. The idea is that points within a window close to a surface discontinuity come from two different planes, therefore a single "average" disparity cannot be assigned to the whole window without making a gross error. The multiple windows approach can be regarded as a robust technique able to fit a constant disparity model to a piecewise-constant surface.

Left-right consistency. Occlusions create points that do not belong to any conjugate pairs. Different conjugate pairs are obtained, in general, by swapping the reference images. This means that some points are involved in more than one conjugate pairs. As each point in one image can match at most one point in the other image, such pairs are discarded and the points marked as *occluded*.

Uncertainty map. Area-based algorithms are likely to fail not only in occluded regions, but also in poorly-textured regions, where signal-to-noise ratio (SNR) is low. It is therefore essential to assign to each disparity an *uncertainty* value inversely proportional to the SNR. The key observation is that disparity estimation is more sensitive to the window's shape where the SNR is low. Consequently, we define the uncertainty as the variance of the disparity measures obtained with nine asymmetric windows; occluded points are assigned infinite variance. Experimental results show [2] that such uncertainty measure is consistent, i.e., it grows as the SNR decreases.

3. MRF BASICS AND RESTORATION

A MRF is defined on a finite lattice I of elements i called sites. Let us define a family of random variables $F=\{F_i=f_i, i\in I\}$, and let us suppose that each variable may assume values taken from a discrete and finite set (e.g., grey level set). Every site i is associated to a random variable F_i (being f_i its realization), and, owing to the Markov property, the conditional probability $P(f_i \mid f_{I-\{i\}})$ depends only on the value on the neighboring set of i, N_i [8]. The Hammersley-Clifford theorem establishes the Markov-Gibbs equivalence between MRFs and Gibbs Random Fields [9], so the probability distribution takes the following form:

$$P\left(f\right) = Z^{-1} \cdot e^{-\beta \cdot U(f)} \tag{1}$$

where Z is a normalization factor called partition function, β is a parameter called temperature and $U(f)$ is the energy function, which can be written as a sum of local energy potentials dependent only on the cliques $c\in C$ (local configurations) relative to the neighboring system [8]:

$$U\left(f\right) = \sum_{c\in C} V_c\left(f\right) \tag{2}$$

Given the observation g, the problem is solved computing the estimate f according to a Maximum A-Posteriori (MAP) probability criterion. Since the posterior probability is still of the Gibbs type, we have to minimize $U(f \mid g) = U(g \mid f) + U(f)$,where $U(g \mid f)$ is the *observation model* and $U(f)$ is the *a-priori model* [8]. When the MRF model is applied to image processing, the observation model describes the noise that degrade the image, and the a-priori model describes the a-priori information like the smoothness of the surfaces composing the scene objects.

To deal with the stereo problem the scene is modeled as composed by a set of planes located at different distances to the observer, so that each disparity value corresponds to a plane in the scene. Therefore, the a-priori model is piecewise constant [9].

The observation model is harder to define because the disparity map is not formed by a physical process. We assume a simple noise model based on different local costs, even if a more articulated noise model could be devised to get more accurate reconstruction.

We use the confidence map on the basis of the principle that reliable points are more likely to keep the observed values, while uncertain points are more likely to assume a value forced by the a-priori hypothesis. The minimization of the global energy function is performed by Simulated Annealing using the Metropolis sampler [10][9].

In order to define the MRF model, we introduce a random field F to estimate the disparity map, a random field G to model the observed disparity and a random field S to model the confidence. It is worth noting that in the image S, dark gray levels are associated to reliable disparity estimates and vice versa.

4. CONFIDENCE MRF MODEL

4.1. MRF with a binary confidance image (model A)

In order to obtain a more consistent information, the confidence image is binarized applying a deterministic criterion which assumes that a reliable point in an uncertain region becomes uncertain too, and vice versa. A global threshold is set up to build a simple binary image, and an hysteresis thresholding method is subsequently applied.

Given the observed disparity image g (the realization of the field G) and the confidence image s (the realization of field S), we search for the estimated disparity f (the realization of the field F) that maximizes the posteriori probability $P(f|g,s)$, minimizing the following energy functional:

$$U\left(f \mid g,s\right) = \sum_{i\in I}\left[V_1\left(s_i\right)\cdot T_i^O + T_i^P\right] \tag{3}$$

where $V_1(s_i) \cdot T^O$ is the observation model and T^P is the a-priori model, computed to estimate the energy cost of each configuration. The observation model is defined as:

$$T_i^O = \begin{cases} 0 & if \quad f_i = g_i \\ k_1 \quad if & f_i \neq g_i \wedge \exists g_j : f_i = g_j, j\in N_i \\ k_2 & otherwise \end{cases} \tag{4}$$

where k_1 and k_2 ($k_1 < k_2$) are parameters that depend on the noise that corrupt the disparity image. In this way, the observation constraint takes a local significance on the neighboring set of i. When noise is low, high values for k_1 and k_2 should be set in order to favour estimates close to local observations. The potential $V_1(s_i)$ introduces the confidence information and is defined as:

$$V_1\left(s_i\right) = \begin{cases} c_1 & if \quad s_i = 0 \\ c_2 & if \quad s_i = 255 \end{cases} \tag{5}$$

where c_1 and c_2 ($c_1 > c_2$) are constant parameter chosen heuristically. In the a-priori term, we impose that all pixels assume the same value in a region. The a-priori term is defined as:

$$T_i^P = \sum_{j \in N_i} V_2(s_i) \cdot V(f_i, f_j) \cdot k_p \tag{6}$$

where

$$V(f_i, f_j) = \begin{cases} 1 & if \quad f_i \neq f_j \\ 0 & if \quad f_i = f_j \end{cases} \tag{7}$$

and k_p is a parameter chosen heuristically. The potential confidence dependent $V_2(s_i)$ is defined as:

$$V_2(s_i) = \begin{cases} c_3 & if \quad s_i = 0 \\ c_4 & if \quad s_i = 255 \end{cases} \tag{8}$$

where c_3 and c_4 ($c_3 < c_4$) are constant parameter chosen heuristically.

4.2. MRF with coupled restoration of disparity and confidence map (model B)

In this model the confidence map is actively included as a coupled field associated to the disparity map. Therefore, restoration of the confidence and disparity fields is simultaneously carried out in a *cooperative* way. We have to define an observation model and a-priori model for the confidence image to be included in the energy functional. The energy function is defined as:

$$U(f, s \mid g, gs) =$$

$$\sum_{i \in I} \left[V_1(s_i) \cdot T_i^O + T_i^P \right] + \sum_{i \in I} \left[TS_i^O(s_i, gs_i) + TS_i^P \right] \tag{9}$$

where s_i and gs_i are a generic estimate and observed confidence image pixel, respectively. The first sum concerns to the disparity restoration described by Eq. (3). The second sum concerns the restoration of the confidence field as a generic binary image [9] in witch the observation model TS_i^O is a sensor model for binary surface (actually, derived by the binary symmetric channel theory) [9], and the a-priori model TS_i^P is a piecewise constant model.

4.3. MRF with split confidence images (model C)

The *SMW* algorithm establishes that when a point violates the left-right consistency constraint (an occlusion), its uncertainty is set to infinity. The other points in the confidence image come from the variance of the disparity values measured by the nine correlation windows. Confidence data are split in two contributions, hence, they are distinguished in the restoration model. Two kinds of random fields are considered: the occlusions *so*, and the confidence *sc*. The energy functional is defined as:

$$\begin{aligned} U(f \mid g, so, sc) = \sum_{i \in I} \Big[& V_1(so_i, sc_i) \cdot T_i^O \\ & + \sum_{j \in N_i} V_2(sc_i, \quad sc_j) \cdot V(f_i, f_j) \cdot k_p \Big] \end{aligned} \tag{10}$$

where T_i^O and $V(f_i, f_j) \cdot k_p$ refer to Eqs. (4) and (6), respectively. The potential $V_1(so_i, sc_i)$ is defined as:

$$V_1(so_i, sc_i) =$$

$$\begin{cases} \left[\left(1 - \frac{sc_i}{255}\right) \cdot (c_1 - c_2) \right] + c_2 & if \quad so_i = 0 \\ c_{OCL} & if \quad so_i = 255 \end{cases} \tag{11}$$

where so_i is the *i*-th occlusion point and sc_i is the *i*-th confidence point, c_1 and c_2 ($c_1 < c_2$) are the parameters defined in section 4.1 and c_{OCL} is a new coefficient defined to insert a different contribution derived from the occlusions. The value of the coefficient c_{OCL} depends on the heuristic disparity assignment criterion set in the *SMW* algorithm for the occlusion points (a reasonable relation is $c_{OCL} < c_2 < c_1$). The potential $V_2(sc_i, sc_j)$ is defined as

$$V_2(sc_i, sc_j) = \left[\frac{sc_i}{255} \cdot \left(1 - \frac{sc_j}{255} \right) \cdot (c_4 - c_3) \right] + c_3 \tag{12}$$

where c_3, c_4 ($c_4 > c_3$) are the parameters defined in section 4.1.

It is worth pointing out the different contributions of the two images in the restoration model: the occlusion map *so* gives a binary confidence information between reliable and unreliable points, the confidence map *sc* introduces a continuous reliability information defined by the [c_2, c_1] interval in the observation model and the [c_3, c_4] interval in the a-priori model. Moreover, in the observation model, the constant c_{OCL} introduces a specific cost paid by the occlusions points.

5. ANALYSIS OF RESULTS

The proposed restoration models are applied to the stereo images[1] in Fig.1 (a and b). From the SMW algorithm we obtain the disparity and confidence images shown in Fig. 1 (c and d). Using our MRF models we get the results shown in Fig. 1(e, f and g). In order to evaluate these results, we compared them with the ground truth image using the mean absolute error (MAE) and the percentage of wrong pixels (Percentage Error, PE). As a zero-order restoration, we included in the comparison the output of median and Gaussian filters (Fig.2) showing the better MRF models performance. Results are shown in Table 1.

[1] Stereo pair and ground truth are courtesy of Dr. Y. Otha, University of Tsukuba, Japan.

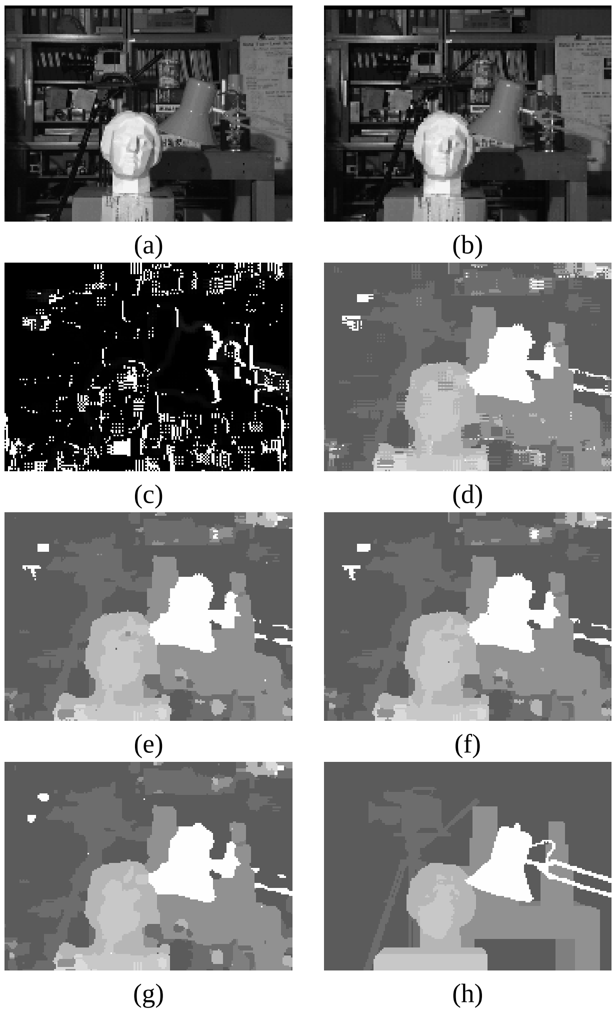

Fig. 1. Stereo images (a and b) , SMW confidence (c) and disparity images (d).Restored disparity map with Model A (e), with Model B (f), with Model C (g) and ground truth (h).

6. CONCLUSIONS

This paper constitutes a preliminar work aimed at proposing and testing the capabilities to directly take into account for confidence information (estimates' uncertainty) in several MRF model. The results proves the utility of using confidence data to improve disparity estimates, showing the capacity to remove spurious points, while preserving object contours. Future work will be devoted to directly model the SMW stereo process in an MRF framework.

7. REFERENCES

[1] U. R. Dhond and J. K. Aggarwal, "Structure from stereo – a review," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 19, no. 6, pp. 1489–1510, December 1989.

Fig. 2. Details of the output of Gaussian filter (left), median filter (center), and MRF model C (right).

Resulting restored image from	MAE	PE
Initial Disparity	10.74	24.98%
Median Filter	10.70	49.53%
Gaussian Filter	10.79	36.29%
MRF model A	9.93	23.41%
MRF model B	9.95	23.51%
MRF model C	9.63	23.14%

Table 1. Performance values measured in the different restoration models.

[2] A. Fusiello, V. Roberto, and E. Trucco, "Symmetric stereo with multiple windowing," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 14, no. 8, pp. 1053–1066, December 2000.

[3] H. Derin and H. Elliot, "Modeling and segmentation of noisy and textured images using gibbs random fields," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 9, no. 1, pp. 39–54, 1987.

[4] G.S Nadabar A.K. Jain, "Range image segmentation using mrf models," in *Markov Random Fields Theory and Application*, A.K. Jain and R. Chellappa, Eds., pp. 542–572. Academic Press, 1993.

[5] E. Gamble and T. Poggio, "Visual integration and detection of discontinuities: the key role of intensity edge," A.I. Memo 970, Massachusetts Institute of Technology, 1987.

[6] V. Murino, A. Trucco, and C.S. Regazzoni, "A probabilistic approach to the coupled reconstruction and restoration of underwater acoustic images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 9–22, January 1998.

[7] L. Matthies, T. Kanade, and R. Szelisky, "Kalman filter based algorithms for estimating depth from image sequences," *International Journal of Computer Vision*, vol. 3, pp. 209–236, 1989.

[8] S. Geman and D. Geman, "Stochastic relaxation, gibbs distribution, and bayesian restoration of images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.

[9] J. K. Marroquine, *Probabilistic Solution of Inverse Problem*, Ph.D. thesis, Massachusetts Institute of Technology, 1985.

[10] C.D. Gellat Jr S. Kirkpatrik and M.P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4, pp. 671–680, 1983.