# A Markov Random Field Model and Method to Image Matching

Mohammed Ouali[1,3], Holger Lange[2], and Kheireddine Bouazza[1,4]
[1]LITIO Lab, University of Oran, Algeria
[2]R&D Group, General Dynamics Canada, Canada
[3]Department of Computer Science, Taif University, KSA
[4]Department of Computer Science, Umm Al Qura University, KSA

**Abstract:** *In this paper, the correspondence problem is solved by minimizing an energy functional using a stochastic approach. Our procedure generally follows Geman and Geman's Gibbs sampler for Markov Random Fields (MRF). We propose a transition generator to generate and explore states. The generator allows constraints such as epipolar, uniqueness, and order to be imposed. We also propose to embed occlusions in the model. The energy functional is designed to take into account resemblance, continuity, and number of occlusions. The disparity and occlusion maps as modeled by their energy functional, i.e., as a Gibbs-Boltzmann distribution, are viewed as a MRF where the matching solution is an optimal state.*

**Keywords:** *Disparity, MRF, image matching, stereo constraints, resemblance, epipolar geometry, uniqueness, continuity.*

*Received May 10, 2010; accepted January 3, 2011*

## 1. Introduction

Dense stereo matching allows extraction of depth physical properties from two images of the same scene. Depth physical properties may be used in applications such as robot navigation, obstacle avoidance, photogrammetry, teleoperation, and measurement in scanning electron microscopy to name a few. Depth is inferred from two or more views of the same scene. If the projections of the same scene point are known to correspond to each other, the horizontal disparity, respectively vertical, disparity can be computed as the difference of abscissas, respectively ordinates. Once the disparity map is established by matching the left and right image of the scene, depth may be computed using the disparity maps, the intrinsic cameras' parameters (focal distance, pixel size, CCD grid size, image coordinates of the center of projection), and the extrinsic cameras' parameters (mainly the computation of the rotation and translation matrices between the cameras). Cameras' parameters are determined through a calibration procedure [21].

Image matching relies on the choice of matching primitives and a matching algorithm. Primitives may be points, regions, or contours. Primitives should possess features that may be used for matching. Points possess the inherent gray level or brightness attribute. But they may also be characterized by features like Laplacian zero crossings, gradient local extrema, or even fancy photometric attributes such as the specular reflection coefficient, diffuse reflection coefficient, and the normal vector coefficient [15, 18]. The main features that may be used with regions are the center of mass of the region, the area of the region, the length of the perimeter and also some higher moments that measure symmetry and skewness. Contours are characterized by the length, the center of mass, orientation, and curvature. Contours may be used and produced by any edge detector or processed by polygonal approximation to yield an abstracted sketch of edges. The usage of a particular matching primitive depends on the quality of the desired disparity map. There are mainly two types of disparity maps: sparse and dense maps. Sparse maps only contain measurements in areas where primitives are present. Dense maps contain measurements almost everywhere. Hence, sparse maps are produced when contours and points of interest are used as matching primitives. Dense maps are used when points or regions are used as matching primitives.

Stereo matching is governed by a number of constraints. The most important are resemblance, epipolar, continuity, uniqueness, and order. The resemblance constraint stipulates that two primitives are matched only if they are similar. This constraint enforces the similarity of brightness or photometric attributes for points. This is particularly true in the case of Lambertian surfaces-for which the amount of reflected light is equally distributed for any point of view. For regions, it matches primitives with similar area, center of mass location, and dispersion. For contours, primitives that are matched should have the same length, orientation, and center of mass. The major issue with this constraint is the fact that in real imagery, primitives have seldom the exact same features and hard similarity measures do not perform

quite well. The epipolar constraint is the only constraint due to the stereo rig geometry. Its concept is simple: a physical point in the scene is projected onto the images by intersecting a projection line, going through the center of projection with the image plane; the physical point and the two centers of projection, one for each camera, form a plane that intersects both image planes-the epiplar plane; each intersection yields a line on the image-the epipolar line; a physical point in the scene determines two epipolar lines, one in each image-they are called conjugate epipolar lines. A point in the left image plane that is the result of the projection of a real physical point in the scene can only be matched with a point in the right image plane that belongs to the conjugate epipolar line. This has the advantage of reducing the search space from 2D to 1D. The continuity constraint is the result of the projection theorem. Physical surfaces are smooth and continuous and the perspective projection of a continuous surface is also continuous. Hence, the disparity should vary smoothly within a region. The uniqueness constraint translates the fact that a point in the left image should at most have one corresponding point in the right image. If the point is occluded in the right view or hidden by an object, there is obviously no corresponding point. The constraint of order is only valid for non transparent objects. A set of ordered points in one image is matched to a set in the other image with the same order.

Several algorithms have been proposed in the literature. These algorithms may be organized in approaches: spatial, frequency, and space-scale domain approaches. Spatial domain approaches are usually based on primitives extracted in the spatial domain (brightness, contours, regions). Frequency domain approaches use frequency and phase components. Scale-space domain approaches use multiscale decomposition. Among spatial approaches, we cite correlation [13, 19, 22]. It usually measures the resemblance peak between two images. As the images may have different levels and camera gain, a more robust operator is used: zero mean normalized cross correlation. Other forms of correlation include the localized correlation that uses the wavelets coefficients as matching primitives [16]. Similar spatial approaches are used with regions and contours. Frequency domain approaches translate the spatial shift (the disparity) into a phase modulation. The disparity is directly recovered as a phase difference. Other approaches use the phase as a matching primitive in the phase correlation. Cepstral correlation is another frequency-based method. Scale-space approaches use Gaussian pyramids and Burt's pyramids to create smaller versions of the original images. Matching is usually easier at low resolution and the computed disparity maps are propagated to higher resolutions. Previous works addressed energy minimization-based dense stereo matching and several methods were proposed. Among these we cite level set methods [8], and mean field annealing [2, 3, 5, 10, 11, 12, 20]. Stereo matching can always be cast as a problem of assigning each pixel a label. Such pixel labeling problems are naturally represented in terms of energy minimization. This framework can be justified in terms of maximum a priori of a Markov random field. A Gauss-Markov random field model can be used to compute the optimal Bayesian estimator. Global optimization methods define the matching as an energy or a cost function of the disparity field. This functional usually uses a correspondence term and a regularization term to enforce the spatial coherence of the disparity field. The form of these terms vary from one approach to another and many solvers based on Bayesian interpretation and Markov random field have been proposed [9]. Techniques for computing the optimal estimator range from global energy minimization [12] to mean field [6, 7, 14].

Stereo matching is an ill-posed problem [4] (it does not conform to the Hadamard mathematical paradigm for well-posedness) which solution is hidden by local minima and where the notion of gradient does not exist. This work aims at improving results by adding a global aspect to the stereo vision problem, the correlation having inherently a local aspect, and making this module cooperate with other early vision modules possible. For instance shape from shading and photometric stereo. In order to attain our objective, we formulate the problem as a Markov random field and we solve the functional minimization within a stochastic framework. We also propose a random transition generator that is compatible with all the stereo constraints as well as a disparity validity criterion that verifies the symmetric property of disparity maps-left and right disparities should be equal with opposite signs. The paper is organized as follows: section 2 deals with the interpretation of stereo constraints in the context of energy minimization and Markov random fields, where we lay down the states random generator; the energy minimization solver is presented in section 3; finally, section 4 discusses the experimental protocol and results in both synthetic and real imagery.

## 2. Stereo Matching

We are only interested in stereo matching general constraints: the resemblance of the primitives, the epipolar lines, the continuity of the disparities, the uniqueness, and the order. These constraints are expressed as energy functional and elementary transitions. We consider the correspondences in both directions (left to right and right to left), in order to apply the continuity criterion to the two images and the validity criterion given by the uniqueness constraint for the matching. We handle the disparities and occlusions of the two images in a consistent way. Images are noted $I_l(x,y)$ and $I_r(x,y)$, where

$x, y \in \{1 \cdots n\}$ denote columns and lines respectively, and $I_l$ and $I_r$ being the left-hand and right-hand side images. Two cameras and observe the same scene have a common field of vision. A point in the scene is projected on the two images in two points which have different coordinates due to the different points of view of the cameras. We express this shifting by horizontal and vertical disparities in the two directions (left with respect to right and right with respect to left): $dx^a(x,y) = x_b - x_a$ and $dy^a(x,y) = y_b - y_a$ where $(a,b) = \{(l,r),(r,l)\}$. The two cameras do not observe exactly the same physical points in the scene, because some points can be out of the common field of vision or hidden by objects. A point observed by a single camera is declared as an occlusion in the image where it appears. Occlusions are noted $O_l(x,y)$ and $O_r(x,y) \in \{0,1\}$, where 0 stands for an occlusion. Points that are seen by both cameras can be matched. Disparities can be calculated by this point matching. The disparities a depth can easily be computed by triangulation using the cameras and the stereo rig parameters.

## 2.1. Resemblance Constraint

This constraint allows the matching of two points only if they have the same gray level. In the case of lambertian surfaces, a point in the scene has the same gray level in both images. A point is a corner of a pixel for which we attribute the average of the gray levels of its four adjacent neighbors:

$$A(x,y) = \sum_{k=0}^{1} \sum_{l=0}^{1} I(x+k, y+l) \quad (1)$$

We also consider using redundant information on the scene by changing the lighting conditions. Each lighting condition provides a stereo image pair. Within the same stereo pair, corresponding points should have similar gray level values. We express this constraint by the square of the difference of the gray values associated with the matching candidates for each lighting condition *k*:

$$\left(A_k{}^l(x,y) - A_k{}^r\left(x + dx^l(x,y), y + dy^l(x,y)\right)\right)^2 \quad (2)$$

This term is weighted with the term representing occlusions, because if there is any occlusion there will be no matching. The energy associated with this constraint over the whole image is:

$$E_R = \sum_y \sum_x \left( A^l(x,y) - A^r\begin{pmatrix} x + dx^l(x,y), \\ y + dy^l(x,y) \end{pmatrix} \right)^2 O^l(x,y) \quad (3)$$

## 2.2. Epipolar Constraint

A point in the scene is projected onto a point in the image. All the points in the scene which belong to the line going through the point of the scene and the optical center are projected on the same point in the image. The projection of this line on the other image represents the epipolar line associated with the point. A point can only be matched with a point in the other image which lies on its epipolar line. The search space is reduced to the epipolar line. It is the only constraint that is due to stereoscopic geometry. It is expressed in the elementary transitions.

## 2.3. Continuity Constraint

We consider the physical surfaces as locally continuous. In this case, their Euclidean projection is continuous as well. The disparity varies a little, except on the edges. The reason is we cancel the continuity constraint at the edges because we cancel the constraint on occlusions; where no matching exists. This constraint is expressed by the energy functional given hereafter. Horizontal and vertical edges are respectively denoted by $c_x$ and $c_y$: $c_x^a(x,y), c_y^a(x,y) \in \{0,1\}$, where 0 stands for the presence of an edge.

$$\begin{aligned} E_C^a = & \\ & \sum_{x=1}^{n_x^a} \sum_{y=1}^{n_y^a} \left(\Delta dx_h^a(x,y)\right)^2 c_x^a(x,y) O^a(x,y) O^a(x+1,y) + \\ & \sum_{x=1}^{n_x^a} \sum_{y=1}^{n_y^a} \left(\Delta dx_v^a(x,y)\right)^2 c_y^a(x,y) O^a(x,y) O^a(x,y+1) + \\ & \sum_{x=1}^{n_x^a} \sum_{y=1}^{n_y^a} \left(\Delta dy_h^a(x,y)\right)^2 c_x^a(x,y) O^a(x,y) O^a(x+1,y) + \\ & \sum_{x=1}^{n_x^a} \sum_{y=1}^{n_y^a} \left(\Delta dy_v^a(x,y)\right)^2 c_y^a(x,y) O^a(x,y) O^a(x,y+1) \end{aligned} \quad (4)$$

## 2.4. Occlusions Counterweight

With occlusions all over the image, the energy is minimal (null), but this is not a viable solution. This is why we add a counterweight in the energy functional in order to limit the number of occlusions. For one image, this energy functional is:

$$E_O^a = \sum_x \sum_y \left(1 - O^a(x,y)\right) \quad (5)$$

## 2.5. Uniqueness Constraint

Every point can only be matched with at most one point from the other image. This constraint was introduced by Marr, and it is strictly verified in the case of the nontransparent objects. This constraint simplifies the computations and enables the application of the validity criterion for a given matching. A matching is valid if both disparities (left

to right and right to left) enable us to reach the point in the right hand side image from the point in the left hand side image and go back to the left hand side image initial point. This constraint is expressed in the elementary transitions. However, in case of foreshortening, corresponding surfaces do not exhibit the same number of pixels. As we are considering a point-based matching approach, we will end up with pixels that will not be matched at all. The algorithm will force these pixels to be occluded.

## 2.6. Energy Functional

The energy functional of the stereo matching problem is the weighted sum of the energy functionals derived from the stereo matching constraints:

$$E = \rho_R E_R + \rho_C (E_C^l + E_C^r) + \rho_O (E_O^l + E_O^r) \qquad (6)$$

The $\rho$ terms denote the weighting coefficients. It is always possible to set one of them to 1 and determine the others accordingly. This allows to balance between the functional terms or to give more importance to one term with respect to the others. For example, the cameras photometric/gain imbalance is an inherent degradation for any stereo rig. It is dealt properly by adjusting the weighting coefficients energy functional, to the detriment of the number of occlusions and the preservation of discontinuities. Similarly, for a continuous surface, we might want to increase the continuity weighting coefficient to obtain a smooth disparity map. This will result in matching pixels that do not have necessarily the same gray level and thus acts as a tolerance level.

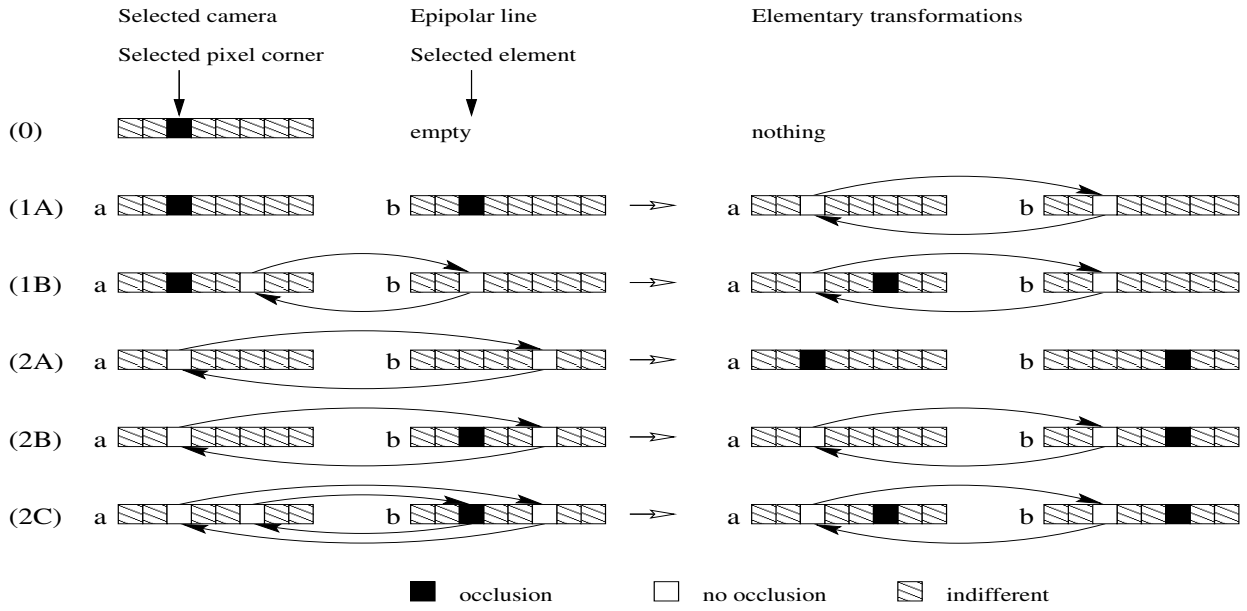



Figure 1. Elementary transitions.

## 2.7. Elementary Transitions

The random state generator is based on the definition of a set of elementary transitions with a decision rule for their application. Figure 1 illustrates the elementary transitions, these are the states before and after their application. Two conjugate epipolar lines are shown; before the transition has taken place and after the transition has been applied. Any complex transition could be obtained by combining these elementary transitions.

Left: two conjugate epipolar lines before transition takes place. *a* and *b* are the label of the epipolar lines that belong to images (*a*) and (*b*) respectively; (*a; b*) is in {(*l; r*); (*r; l*)}, where (*a*) is the label of the reference image. Right: the two conjugate epipolar lines after the transition. Depending on the selected point in the reference image epipolar line, the transitions {0*;* 1*A;* 1*B*} are possible if the point in the reference image is occluded, {2*A;* 2*B;* 2*C*} otherwise. The choice of a particular transition is fully determined by the point selected in the conjugate epipolar line.

### 2.7.1. Decision Rule

We select with equal randomization one of the two images and a first point. If this point has an occlusion, one of the transitions (0), (1A) or (1B) is applied. If its epipolar line is empty, (0) is applied; otherwise, a point in the epipolar line is selected with equal randomization for a potential matching. If the selected point has an occlusion, (1A) is applied; otherwise, we apply (1B). If the first point does not have an occlusion, we select with equal randomization either (2A) or {(2B), (2C)}. If the choice is {(2B), (2C)}, a point in the epipolar line is selected with equal randomization. If the selected point has an occlusion, (2B) is applied; otherwise (2C) is applied. We also

guarantee the same application probability for transitions (2A), (2B), and (2C) as (2A) is selected randomly, to balance the number of applied elementary transitions.

### 2.7.2. Transitions

The elementary transitions are depicted in Figure 1. The left side shows the matching in the current state. The right side shows the new matching corresponding to the new state. Depending on the current state, one of the elementary transitions is applied. The decision rule is outlined hereunder:

- *(0):* Do nothing (because of the empty epipolar line).
- *(1A):* Matches the first point with the selected point in its epipolar line.
- *(1B):* A new matching point between the first point and the selected point in its epipolar line; the former correspondent is occluded.
- *(2A):* Establishes two occlusions, one for the first point, and the other for its correspondent.
- *(2B):* Occludes the first point former correspondent and establishes a new matching with the selected point in the epipolar line.
- *(2C):* Establishes two occlusions, one for the first point former correspondent and the other for correspondent of the selected point in the epipolar line; a new matching between the first point and the selected epipolar line point is set.

The energy differential is computed locally. The configuration space corresponds to the configurations where the epipolar line and the uniqueness constraints as well as the validity criterion are respected.

## 3. Stochastic Optimization

Random walk processes associated with Gibbs-Boltzmann sampler is a stochastic approach for general nonlinear optimization problems. It is suitable for large size problems, where the global minimum is hidden by several local minima and where the notion of gradient has no signification. This approach is applied to a combinatorial discrete problem. Several conditions are required [1]: a description of all possible system configurations $x = \{x_i\}$, which are the generated states to be explored including the optimal states; an initialization $x^0 = \{x_i^0\}$, the initial state; a generator of random states $x \to x'$; an energy functional *J(x)* that describes the problem and which minimization is the goal of the procedure; a control parameter with its initial value $T_0$; and finally a decrease rule $T \to T'$.

### 3.1. Algorithm

The annealing is a well known framework in stochastic optimization. We are only interested in the particular breed that is based on Gibbs-Boltzmann distribution. The algorithm can be summarized as in Figure 2. Initially, the system is set to an initial state. The initial state can be chosen randomly. The energy *J(x)* is evaluated. A small change is made to the state vector $x$, corresponding to a state change. The dwelling in the new state is possible if the energy differential is negative. Otherwise, the probability to be in the new state is determined by the Gibbs-Boltzmann probability distribution. There are two parameters controlling the speed and quality of the convergence: the number of equilibrium plateaus and the number of iterations within an equilibrium plateau. Both numbers could be selected. However, selecting a low number of iterations might cause the system to step down to the next equilibrium plateau before reaching its equilibrium at that plateau. A high number of iterations will make the plateau last longer and may not affect the quality of the equilibrium. The equilibrium is reached when at least a certain proportion of the suggested transitions are accepted. Regarding the number of equilibrium plateaus, a small number results in an unstable solution while a high number of plateaus result in a long unnecessary computation, since no transitions are no longer accepted. This is why we choose to step down to the next equilibrium plateau as soon as the equilibrium is reached and we stop the annealing as soon as there are no suggested transitions accepted within few successive equilibrium plateaus.

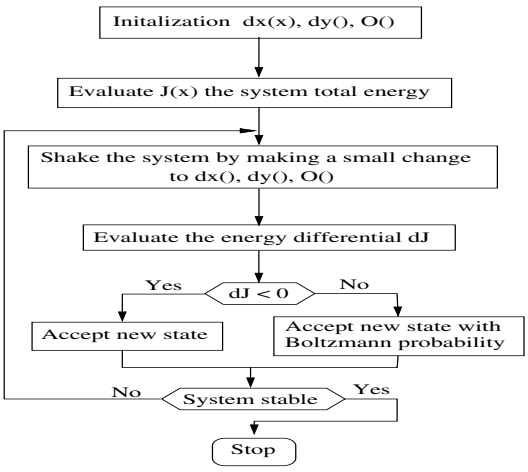



Figure 2. Block diagram.

Showing the steps performed from initialization to system equilibrium. The system state vector, namely the occlusions, horizontal, and vertical disparities are initialized. The initialization may be random or one can use a preliminary even an inaccurate result provided by a fast matching algorithm. The total energy of the system is evaluated. Then the system performs a small step in the configuration space. This destination is accepted as the new state if the system's energy decreases. From time to time, even new states increasing the system's energy are accepted. This

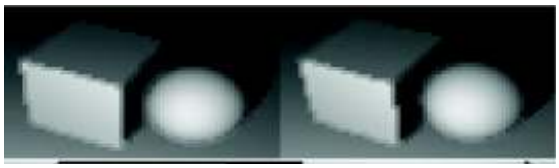
helps the system free itself if trapped in a local minimum valley. The procedure is repeated until no steps decreasing the energy are possible.

In practice, we compare a random number between 0 and 1 with the Gibbs-Boltzmann exponential $e^{-\Delta J(x)/t}$, and we accept the change if this number is smaller than this probability. The system reaches the thermal equilibrium in a given plateau after a certain number $M$ of elementary transitions. The number $M$ is set to 100 by some authors. This number should rather depend on the size of the definition domain of the variables. If we try 100 transitions on a variable, there is a difference in whether this variable can take one value out of 100 or one value out of 10000 values. In the first case each value is tried once in average, whereas. In the second case, only 1% of the values are used. We propose to define $M$ according to the size of the definition domain: $M=\tau$(SizeOfDomain). The parameter $\tau$ determines how many times the same value is used in average. The system is considered fairly unstable if more than 10% of $M$ attempts are accepted; otherwise, it is considered fairly stable. If the system is unstable, it is considered in local equilibrium after 10% of $M$ successful transitions. We get to the next equilibrium floor by stepping down the control parameter $T$. The procedure stops when no more transitions are accepted.

## 4. Experimental Results

A simple method is used to compute the weighting coefficients in the energy function. If a tolerance of 0.5 is allowed on the resemblance constraint, the matching is cancelled if the gray level difference is higher than this value: $0.5\times\rho_R \geq 2\times\rho_O$. With $\rho_O=1$, we have $\rho_R \geq 4$. Furthermore, we assume for a smooth surface that the local difference in disparities is not higher than 1. If the accumulated disparities differences in 4 directions for a pair of matched points are larger than the energy provided by 2 occlusions, the 2 occlusions are set: $2\times 4\times\rho_C \geq 2\times\rho_O$. With, $\rho_O=1$ we have $\rho_C \geq 0.25$. Finally, as we handle the two cases (left to right and right to left) altogether, we should set: $\rho_O=2$. This allows to balance between the functional terms or to give more importance to one term with respect to the others. The results obtained with the weighting gains given above are already better than the ones obtained with the initialization. Even better results could be obtained with thorough experimentations on the weighting gains. The starting control parameter $T_0$ is set to 2, which corresponds to the system's state where, virtually, all the elementary transitions are accepted. The parameter $\tau$ is set to 100. This parameter's value has proven to give good results. To provide an acceptable initial state, we considered using the Sum of Squared Differences-based correlation (SSD); we have used the same resemblance constraint, epipolar constraint, uniqueness constraint, and validity criterion. The correlation consists in sweeping the candidate point over the epipolar line and to establish a matching with the point that minimizes the resemblance constraint. This is done in both directions (left w.r.t. right and right w.r.t. left). Only valid matches are kept invalid matches are cancelled and occlusions are set instead.

We considered both synthesized and real images to evaluate our approach. Figure 3 shows the synthesized dataset and the ground-truth. The gray level images (top row) are generated using an in-house image synthesis system. The image synthesis system is developed for the purpose of verifying and analyzing an image analysis system. A geometric model of the objects and background is given along with the model and positioning of the cameras and light sources. The ground-truth (disparity maps) is also generated for comparisons against the computed disparity maps. The disparity maps generated by the image synthesis system and therefore corresponding to the solution are used to show the matching errors produced by our method. In the disparity maps shown in Figure 3-b, black pixels stand for occlusions while gray shades correspond to different depths. Dark areas are closer to the camera while lighter ones are farther.



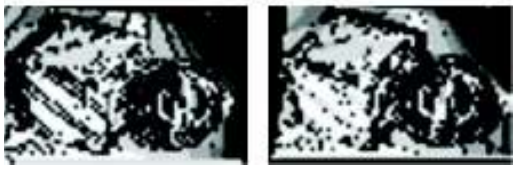
a) left and right gray level images.

b) left and right disparity maps (ground-truth).

Figure 3. Synthesized dataset.

In Figure 4, we show the initialization (using the SSD correlation-based matching), and the error against the ground-truth. The correlation produced a large number of occlusions, due to a large number of invalid matches. The disparity map contains too many errors. The error maps are computed as the difference between the produced disparity maps and the ground-truth. In the error maps, black pixels stand for areas where no measurements could have been done; dark areas represent large errors while lighter ones represent small or no errors.

a) Lleft and right disparity maps computed with SSD correlation.



b) Left and right error maps computed as the difference between the initialization and ground-truth.

Figure 4. Synthesized data. Initialization and error map computed against the ground-truth.

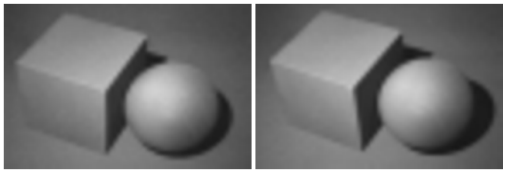
The MRF model results in synthesized imagery as shown in Figure 5. The correlation disparity map has been used as an initialization for the stochastic approach. This really does not influence the stochastic algorithm since the starting control parameter $T_0$ is high enough to allow all the transitions to be accepted at the beginning, thus moving the system far way from the initialization. Another way to initialize the system could have been by setting up occlusions everywhere and initializing all the disparities to 0. The results generated by the stochastic algorithm not only shows a small number of occlusions, but more importantly, occlusions are judiciously placed. The areas where disparity is not correctly produced are less important than for the correlation.

a) Left and right disparity maps; four lighting conditions stereo pairs were considered at the same time.

b) Left and right error maps computed as the difference between the disparity maps and ground-truth.

Figure 5. Synthesized data. Image matching using MRF and stochastic approach.

Experimental results on real imagery are shown in Figure 6. Similarly, we considered four lighting conditions during the image acquisition. The four lighting conditions stereo pairs were used at the same time to generate the disparity maps. First row shows one of the four stereo image pairs. Second row shows the disparity maps computed by the stochastic approach on the MRF.

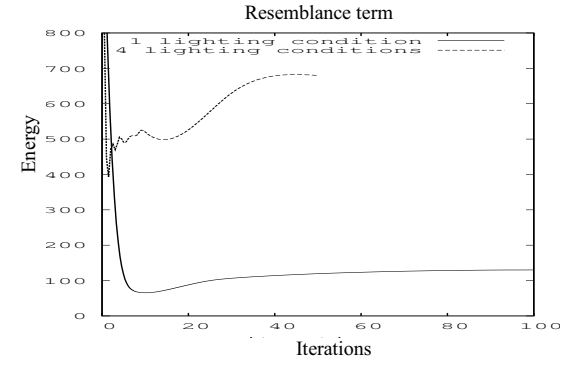


a) Gray level stereo pair.

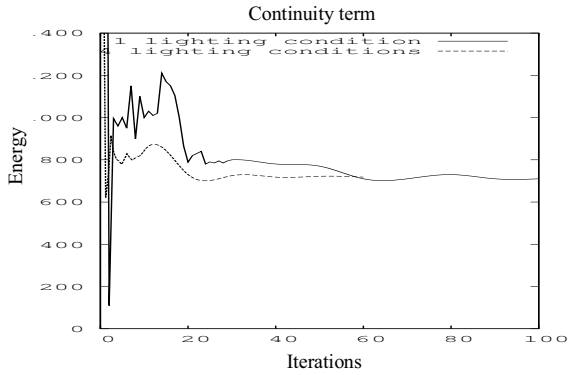
b) Left and right disparity maps.

Figure 6. Real imagery.

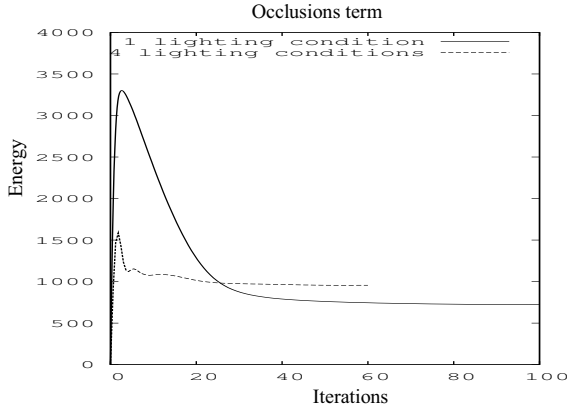
Finally, the evolution of the energy functional is shown in Figure 7. We note that the continuity and number of occlusions functional are decreasing and converging towards small values. However, the gray level functional somehow increases, although it reaches a stability plateau. This is due to the fact that corresponding pixels do not have necessarily the same gray level and also the continuity weight coefficient that favors local continuity over gray level resemblance.




a) Matching energy.




b) Continuity energy.




c) Number of occlusions energy.

Figure 7. Evolution of the energy functional components.

Note that both the number of occlusions and continuity energies decrease and converge towards a small value. The matching energy also converges to a stable value. This value is higher than the initial value as the matched pixels do not have necessarily the same gray level.

## 5. Conclusions

We tackled the dense stereo matching problem in a stochastic energy minimization framework. We proposed to formulate the problem as a random walk, moving the system from one state to another based on its energy and the Gibbs-Boltzmann sampler. We also proposed a random transitions generator that allows the system to only visit viable states, i.e., states that are compatible with the stereo constraints. As an initial state, we used the SSD correlation disparity maps. We showed that the stochastic framework improves significantly the disparity and occlusion maps. Experimental results are satisfactory on both synthesized and real imagery. Future perspectives of this work include the cooperation of this low level module with the stereo photometry module, where the shape of the objects is used as a constraint and the gray level for the resemblance constraint is replaced by the photometric attribute of the Phong/Cook-Torrance model [17, 18], taking into consideration the gray level changes as a function of the position of the cameras.

## References


[1] Aarts E. and Laarhoven P., *Simulated Annealing and Applications*, Kluwer Academic Publishers, Holland, 1987.

[2] Benameur W., "Computing the Initial Temperature of Simulated Annealing," *Computational Optimization and Applications*, vol. 29, no. 3, pp. 369-385, 2004.

[3] Ben-Ari R. and Sochen N., "A Geometric Approach for Regularization of the Data Term in Stereo Vision," *Journal of Mathematical Imaging and Vision*, vol. 31, no. 1, pp. 17-33, 2008.

[4] Bertero M., Poggio T., and Torre V., "Ill-Posed Problems in Early Vision," *in Proceedings of the IEEE*, pp. 869-889, 1988.

[5] Boykov Y. and Kolomogorov V., "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124-1127, 2004.

[6] Castellani U., Fusiello A., Gherardi R., and Murino V., "Automatic Selection of MRF Control Parameters by Reactive Tabu Search," *Image and Vision Computing*, vol. 25, no. 11, pp. 1824-1832, 2007.

[7] Corso J., Tu Z., and Yuille A., "MRF Labeling with a Graph-Shifts Algorithm," *in Proceedings of Lecture Notes in Computer Science*, Berlin, pp. 172-184, 2008.

[8] Faugeras O. and Keriven R., "Complete dense Stereo Vision using Level Set Methods," *in Proceedings of Lecture Notes in Computer Science*, pp. 379-398, 1998.

[9] Geman S. and Geman D., "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721-741, 1984.

[10] Huang J. and Liu H., "Stereo Vision using Microcanonical Mean-Field Annealing Neural Network," *Neural Computations*, vol. 8, no. 1, pp. 87-104, 1997.

[11] Huq S., Abidi B., and Abidi M., "Stereo-Based 3D Face Modeling using Annealing in Local Energy Minimization," *in Proceedings of 14th International Conference on Image Analysis and Processing*, Modena, pp. 265-272, 2007.

[12] Ishikawa H., "Exact Optimization for Markov Random Fields with Convex Priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1333-1336, 2003.

[13] Jia Y., Zhang X., Li M., and An L., "A Miniature Stereo Vision Machine (MSVM-III) for Dense Disparity Mapping," *in Proceedings of the 17th International Conference on Pattern Recognition*, pp. 728-731, 2004.

[14] Komodakis N., Tziritas G., and Paragios N., "Performance vs. Computational Efficiency for Optimizing Single and Dynamic MRFs: Setting the State of the Art with Primal-Dual Strategies," *Computer Vision and Image Understanding, Special Issue on Discrete Optimization in Computer Vision*, vol. 112, no. 1, pp. 14-29, 2008.

[15] Ne'ma B. and Ali H., "Multipurpose Code Generation using Fingerprint Images," *International Arab Journal of Information Technology*, vol. 6, no. 4, pp. 418-423, 2009.

[16] Radhakrishnan S., Subbarayan G., and Vikram K., "Wavelets Based Video Encoder using KCDS," *The International Arab Journal of Information Technology*, vol. 6, no. 3, pp. 245-249, 2009.

[17] Radziszewski M. and Alda W., "Family of Energy Conserving Glossy Reflection Models," *in Proceedings of Lecture Notes in Computer Science*, Berlin, pp. 46-55, 2008.

[18] Schlick C., "An Inexpensive BRDF Model for Physically-Based Rendering," *Computer Graphics Forum*, vol. 13, no. 3, pp. 233-246, 1994.

[19] Sunyoto H., Mark W., and Gavrila D., "A Comparative Study of Fast Dense Stereo Vision Algorithms," *Intelligent Vehicles Symposium, IEEE*, vol. 14, no. 17, pp. 319-324, 2004.

[20] Szeliski R., Zabih R., Scharstein D., Veksler O., Kolmogorov V., Agarwala A., Tappen M., and Rother C., "A Comparative Study of Energy Minimization Methods for Markov Random Fields," *in Proceedings of Lecture Notes in Computer Science*, pp. 16-29, 2006.

[21] Wang L., Yao H., and Cheng H., "Effective and Automatic Calibration using Concentric Circles," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 22, no. 7, pp. 1379-1401, 2008.

[22] Werda I., Chaouch H., Samet A., Ben-Ayed M., and Masmoudi N., "Optimal DSP Based Integer Motion Estimation Implementation for H.264/AVC Baseline Encoder," *The International Arab Journal of Information Technology*, vol. 7, no. 1, pp. 96-104, 2010.

**Mohammed Ouali** received his PhD in real-time, robotics, and control from Mines ParisTech-France in 1999. He has over 10 years of experience in developing embedded systems in aeronautics, automotive industry and CAD systems for VLSI reverse engineering. His main research interests are in computer vision and pattern recognition.

**Holger Lange** received his PhD in Real-time, Robotics, and Control from Mines ParisTech-France in 1995. He has over 15 years of experience in developing embedded as well as medical imagery systems. He is currently, a product director at Aperio Technologies Inc., Vista, USA. His main interests are in computer vision and embedded systems.

**Kheireddine Bouazza** is an associate professor with the Department of computer science, University of Oran, Algeria. He received his PhD in Automatic Control from the university of Nancy-France in 2004. His main interests are in time-delay systems and robust control.