A Deterministic Approach for Stereo Disparity Calculation

Chienchung Chang[1] *and Shankar Chatterjee*[2]

[1] Qualcomm Incorporated, 10555 Sorrento Valley Rd. San Diego, CA 92121, USA
[2] Department of ECE, University of California, San Diego, La Jolla, CA 92093, USA

Abstract. In this work, we look at mean field annealing (MFA) from two different perspectives: information theory and statistical mechanics. An iterative, deterministic algorithm is developed to obtain the mean field solution for disparity calculation in stereo images.

1 Introduction

Recently, a deterministic version of the *simulated annealing* (SA) algorithm, called *mean field approximation* (MFA) [1], was utilized to approximate the SA algorithm efficiently and successfully in a variety of applications in early vision modules, such as image restoration [8], image segmentation [3], stereo [12], motion [11] surface reconstruction [4] etc.

In this paper, we apply the approximation in the stereo matching problem. We show that the optimal Bayes estimate of disparity is, in fact, equivalent to the mean field solution which minimizes the relative entropy between an approximated distribution and the given posterior distribution, if (i) the approximated distribution has a Gibbs form and (ii) the mass of distribution is concentrated near the mean as the temperature goes to zero. The approximated distribution can be appropriately tuned to behave as close to the posterior distribution as possible. Alternatively, from the angle of statistical mechanics, the system defined by the states of disparity variables can be viewed as isomorphic to that in magnetic materials, where the system energy is specified by the binary states of magnetic spins. According to the MRF model, the distribution of a specific disparity variable is determined by two factors: one due to the observed image data (external field) and the other due to its dependence (internal field) upon the neighboring disparity variables. We follow the mean field theory in the usual Ising model of magnetic spins [7] to modify Gibbs sampler [5] into an iterative, deterministic version.

2 An Information Theoretic Analysis of MFA

The optimal Bayes estimate of the disparity values at uniformly spaced grid points, given a pair of images, is the *maximum a posteriori* (MAP) estimate when a uniform cost function is assumed. To impose the prior constraints (*e.g.*, surface smoothness etc.), we can add energy terms in the objective energy (performance) functional and/or introduce an approximated distribution. The posterior energy functional of disparity map $\mathbf{d}$ given the stereo images, $\mathbf{f}_l$ and $\mathbf{f}_r$, can usually be formulated in the form [2]:

$$U_p(\mathbf{d}|\mathbf{f}_l,\mathbf{f}_r) = \alpha(T)\sum_{i=1}^{M}|\mathbf{g}_l(\mathbf{x}_i) - \mathbf{g}_r[\mathbf{x}_i + (d_{\mathbf{x}_i},0)]|^2 + \sum_{i=1}^{M}\sum_{\mathbf{x}_j\in N_{\mathbf{x}_i}}(d_{\mathbf{x}_i} - d_{\mathbf{x}_j})^2 \qquad (1)$$

where $\mathbf{g}_l$ and $\mathbf{g}_r$ represent the vectors of matching primitives extracted from intensity images; $N_{\mathbf{x}_i}$ is the neighborhood of $\mathbf{x}_i$ in the image plane Ω and is defined through a neighborhood structure with radius r, $\aleph_r = \{N_{\mathbf{x}}, \mathbf{x} \in \Omega\}$, $N_{\mathbf{x}} \triangleq \{\mathbf{y}, |\mathbf{y}-\mathbf{x}|^2 \le r\}$; $M = |\Omega|$ is the number of pixels in the discretized image plane; and the disparity map is defined as $\mathbf{d} \triangleq \{d_{\mathbf{x}}, \mathbf{x} \in \Omega\}$. The first term represents the photometric constraint and the second term describes the surface smoothness. If the disparity is modelled as an MRF given the image data, the posterior distribution of disparity is given as

$$P(\mathbf{d}|\mathbf{f}_l,\mathbf{f}_r) = \frac{1}{Z_p}\exp\left[-\frac{U_p(\mathbf{d}|\mathbf{f}_l,\mathbf{f}_r)}{T}\right] \tag{2}$$

where Z_p and T are the normalization and temperature constants respectively. The MAP estimate of disparity map is the minimizer of the corresponding posterior energy functional $U_p(\mathbf{d}|\mathbf{f}_l,\mathbf{f}_r)$. It is desirable to describe the above equation by a simpler parametric form. If the approximated distribution is P_a, which is dependent on adjustable parameters represented by vector $\bar{\mathbf{d}} = \{\bar{d}_{\mathbf{x}}, \mathbf{x} \in \Omega\}$ and has the Gibbs form:

$$P_a(\mathbf{d}|\bar{\mathbf{d}}) = \frac{1}{Z_a}\exp\left[-\frac{U_a(\mathbf{d}|\bar{\mathbf{d}})}{T}\right], \quad U_a(\mathbf{d}|\bar{\mathbf{d}}) = \sum_{i=1}^{M}(d_{\mathbf{x}_i} - \bar{d}_{\mathbf{x}_i})^2 \tag{3}$$

where Z_a is the partition function and $U_a(\mathbf{d}|\bar{\mathbf{d}})$ is the associated energy functional. For the specific U_a, the approximated distribution is Gaussian. In information theory, *relative entropy* is an effective measure of how well one distribution is approximated by another [9]. Alternative names in common use for this quantity are discrimination, Kullback-Liebler number, direct divergence and cross entropy. The relative entropy of a measurement $\mathbf{d}$ with distribution P_a relative to the distribution P is defined as

$$S_r(\bar{\mathbf{d}}) \triangleq \int P_a(\mathbf{d}|\bar{\mathbf{d}})\log\frac{P_a(\mathbf{d}|\bar{\mathbf{d}})}{P(\mathbf{d}|\mathbf{f}_l,\mathbf{f}_r)}d\mathbf{d} \tag{4}$$

where $P(\mathbf{d}|\mathbf{f}_l,\mathbf{f}_r)$ is referred to as reference distribution. Kullback's principle of minimum relative entropy [9] states that, of the approximated distributions P_a with the given Gibbs form, one should choose the one with the least relative entropy. If $\bar{\mathbf{d}}$ is chosen as the mean of the disparity field $\mathbf{d}$, the optimal mean field solution is apparently the minimizer of relative entropy measure. After some algebraic manipulations, we can get

$$S_r(\bar{\mathbf{d}}) = \frac{1}{T}(F_a - F_p + E(U_p) - E(U_a)) \tag{5}$$

where the expectations, $E(\cdot)$, are defined with respect to the approximated distribution P_a. $F_a \triangleq -T\log Z_a$, $F_p \triangleq -T\log Z_p$ are called free energy. In statistical mechanics [10], the difference between the average energy and the free energy scaled by temperature is equal to entropy, or $F = E - TS$. From the *divergence inequality* in information theory, the relative entropy is always non-negative [6] $S_r(\bar{\mathbf{d}}) \geq 0$, with the equality holding if and only if $P_a \equiv P$. And since temperature is positive,

$$F_p \leq F_a + E(U_p) - E(U_a) \tag{6}$$

which is known as *Perierls's inequality* [1]. The MFA solution, realized as the minimizer of relative entropy, can be alternatively represented as the parameter $\bar{\mathbf{d}}$ yielding the tightest bound in (6). In other words, we have

$$\min_{\bar{\mathbf{d}}} S_r(\bar{\mathbf{d}}) = \min_{\bar{\mathbf{d}}}[F_a + E(U_p) - E(U_a)] \tag{7}$$

since F_p in (5) is not a functional of the parameter $\bar{\mathbf{d}}$ at all. The choice of U_a relies on a prior knowledge of the distribution of the solution. Gibbs measure provides us with the flexibility in defining the approximated distribution P_a as it depends solely on the energy function U_a, which in turn can be expressed as the sum of clique potentials [5]. Next we discuss an example of U_a which is both useful and interesting. For the energy function given in (3), the corresponding approximated distribution is Gaussian and the adjustable parameters are, in fact, the mean values of disparity field. As the temperature (variance) approaches zero, it will be conformed to the mean value with probability one. Since the disparity values at lattice points are assumed to be independent Gaussian random variables, both the free energy and expected approximate energy can be obtained as:

$$F_a = -T \log Z_a = -\frac{MT}{2}\log(\pi T), \quad E(U_a) = \sum_{i=1}^{M} E(d_{\mathbf{x}_i} - \bar{d}_{\mathbf{x}_i})^2 = \frac{MT}{2} \tag{8}$$

The mean posterior energy can be written as:

$$E(U_p) = \alpha(T)\sum_{i=1}^{M} E\left(|\mathbf{g}_l(\mathbf{x}_i) - \mathbf{g}_r[\mathbf{x}_i + (d_{\mathbf{x}_i}, 0)]|^2\right) + \sum_{i=1}^{M}\sum_{\mathbf{x}_j \in N_{\mathbf{x}_i}} E\left[(d_{\mathbf{x}_i} - d_{\mathbf{x}_j})^2\right] \tag{9}$$

The second term in the right hand side (RHS) can be rewritten as:

$$\sum_{i=1}^{M}\sum_{\mathbf{x}_j \in N_{\mathbf{x}_i}} E\left[(d_{\mathbf{x}_i} - d_{\mathbf{x}_j})^2\right] = \sum_{i=1}^{M}\sum_{\mathbf{x}_j \in N_{\mathbf{x}_i}} [T + (\bar{d}_{\mathbf{x}_i} - \bar{d}_{\mathbf{x}_j})^2] \tag{10}$$

On the other hand, if the first term at RHS of (9) can be approximated by (the validity of approximation will be discussed later)

$$\alpha(T)\sum_{i=1}^{M} E\left(|\mathbf{g}_l(\mathbf{x}_i) - \mathbf{g}_r[\mathbf{x}_i + (d_{\mathbf{x}_i}, 0)]|^2\right) \approx \alpha(T)\sum_{i=1}^{M} \left|\mathbf{g}_l(\mathbf{x}_i) - \mathbf{g}_r[\mathbf{x}_i + (\bar{d}_{\mathbf{x}_i}, 0)]\right|^2 \tag{11}$$

then, by combining (10) and (11), the upper bound in Peierls's inequality becomes

$$F_a + E(U_p) - E(U_a) \propto \alpha(T)\sum_{i=1}^{M} \left|\mathbf{g}_l(\mathbf{x}_i) - \mathbf{g}_r[\mathbf{x}_i + (\bar{d}_{\mathbf{x}_i}, 0)]\right|^2 + \sum_{i=1}^{M}\sum_{\mathbf{x}_j \in N_{\mathbf{x}_i}} (\bar{d}_{\mathbf{x}_i} - \bar{d}_{\mathbf{x}_j})^2 \tag{12}$$

It is interesting to note that the format of the above functional of mean disparity function, $\bar{\mathbf{d}}$, is identical to that of the posterior energy functional, $U_p(\mathbf{d}|\mathbf{f}_l, \mathbf{f}_r)$ up to a constant. Hence, it is inferred that the MAP estimate of disparity function is, in fact, equivalent to the mean field solution minimizing the relative entropy between the posterior and approximated Gaussian distributions. Regarding the approximation in (11), as the temperature $T \to 0$, all the mass of $P_a(\mathbf{d}|\bar{\mathbf{d}})$ will be concentrated at mean vector $\mathbf{d} = \bar{\mathbf{d}}$ and (11) holds exactly. At least, in the low temperature conditions, the MFA solution coincides with the MAP solution.

3 MFA Based on Statistical Mechanics

When a system possesses a large interaction degree of freedom, the equilibrium can be attained through the mean field [10]. It serves as a general model to preview a complicated physical system. In our case, each pixel is updated by the expected (mean) value given the mean values of its neighbors [7].

With Gibbs sampler [2], we visit each site $\mathbf{x}_i$ and update the associated site variable $d_{\mathbf{x}_i}$ with a sample from the local characteristics

$$P(d_{\mathbf{x}_i}|d_{\mathbf{y}}, \forall \mathbf{y} \neq \mathbf{x}_i, \mathbf{f}_l, \mathbf{f}_r) = \frac{1}{Z_i}\exp\left[-\frac{1}{T}U_i(d_{\mathbf{x}_i})\right] \tag{13}$$

where the *marginal* energy function $U_i(d_{\mathbf{x}_i})$ is derived from (1) and (2). If the system is fully specified by the interactions of site (disparity) variables and the given data, the uncertainty of each variable is, in fact, defined by the local characteristics. In a magnetic material, each of the spins is influenced by the magnetic field at its location. This magnetic field consists of any external field imposed by the experimenter, plus an internal field due to other spins. During the annealing process, the mean contribution of each spin to the internal field is considered. The first term in (1) can be interpreted as the external field due to the given image data and the second term as internal field contributed by other disparity variables. SA with Gibbs sampler simulate the system with the samples obtained from the embedded stochastic rules, while MFA tries to depict the system with the mean of each system variable.

In summary, the MFA version of Gibbs sampler can then be stated as:

1. Start with any initial mean disparity $\bar{\mathbf{d}}_0$ and a relative high initial temperature.
2. Visit a site $\mathbf{x}_i$ and calculate the marginal energy function contributed by given image data and the mean disparity in the neighborhood $N_{\mathbf{x}_i}$ as

$$\bar{U}_i(d_{\mathbf{x}_i}) = \alpha\,|\mathbf{g}_l(\mathbf{x}_i) - \mathbf{g}_r[\mathbf{x}_i + (d_{\mathbf{x}_i}, 0)]|^2 + \sum_{\mathbf{y}\in N_{\mathbf{x}_i}} (d_{\mathbf{x}_i} - \bar{d}_{\mathbf{y}})^2 \qquad (14)$$

3. Calculate the mean disparity $\bar{d}_{\mathbf{x}_i}$ as

$$\bar{d}_{\mathbf{x}_i} = \sum_{d_{\mathbf{x}_i}\in R_D} d_{\mathbf{x}_i} P(d_{\mathbf{x}_i}|\bar{d}_{\mathbf{y}}, \forall \mathbf{y} \neq \mathbf{x}_i, \mathbf{f}_l, \mathbf{f}_r) = \sum_{d_{\mathbf{x}_i}\in R_D} d_{\mathbf{x}_i} \frac{\exp\left[-\bar{U}_i(d_{\mathbf{x}_i})/T\right]}{Z_i} \qquad (15)$$

4. Update in accordance with steps **2** and **3** until a steady state is reached at the current temperature, T.
5. Lower the temperature according to a schedule and repeat the steps **2**, **3** and **4** until there are few changes.

Consequently, MFA consists of a sequence of iterative, deterministic relaxations in approximating the SR. It converts a hard optimization problems into a sequence of easier ones.

4 Experimental Results

We have used a wide range of image examples to demonstrate that SR can be closely approximated by MFA. Due to the space limitation, we only provide an image example: Pentagon (256×256). The matching primitives used in the experiments are intensity, directional intensity gradients (along horizontal and vertical directions), *i.e.*, $\mathbf{g}_s(x, y) = \left(f_s(x, y), \frac{\partial f_s}{\partial x}, \frac{\partial f_s}{\partial y}\right), \quad \forall (x, y) \in \Omega,\ s = l, r$. We try to minimize the functional $U_p(\bar{\mathbf{d}}|\mathbf{f}_l, \mathbf{f}_r)$ by deterministic relaxation at each temperature and use the result at current temperature as the initial state for the relaxation at the next lower temperature. The initial temperature is set as 5.0 and the annealing schedule used is where the temperature is reduced 50% relative to the previous one. The neighborhood system $\aleph_2$ is used in describing surface smoothness. The computer simulation results are shown in Fig 1. One could compare the result with those obtained by SA algorithm using Gibbs sampler.

In MFA version of SA with Gibbs sampler, we follow the algorithm presented in Section 3. The initial temperature and the annealing schedule are identical to those in above. The results are also shown in Fig 1. When they are compared with the previous results, we can see that the MFA from both approaches yield roughly the same mean field solution and they approximate the MAP solution closely.

5 Conclusion

In this paper, we have discussed, for stereo matching problem, two general approaches of MFA which provide good approximation to the optimal disparity estimate. The underlying models can be easily modified and applied to the other computer vision problems, such as image restoration, surface reconstruction and optical flow computation etc. As the Gaussian distribution is the most natural distribution of an unknown variable given both mean and variance [9], it is nice to see that the mean values of these independent variables that minimize the relative entropy between the assumed Gaussian and the posterior distribution is equivalent to the optimal Bayes estimate in MAP sense.

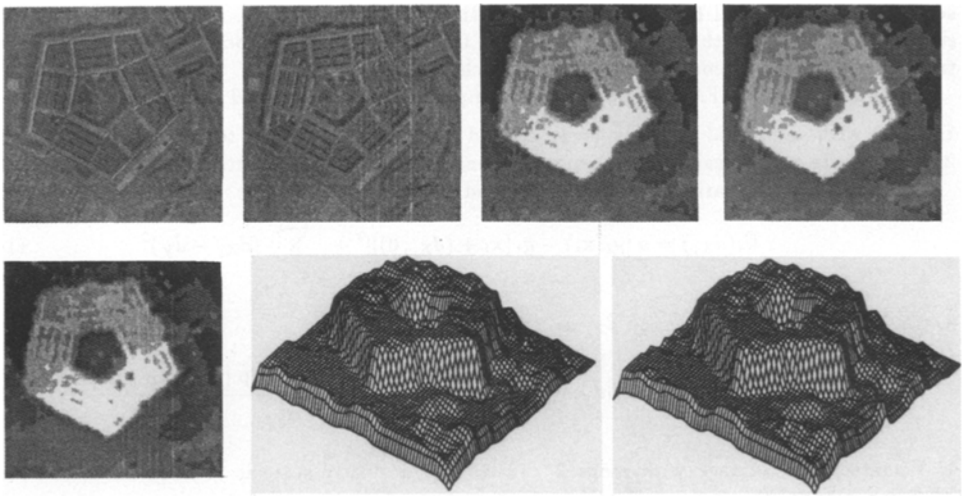

Fig. 1. Upper row (left to right): the left and right images of Pentagon stereo pair, the mean field result based on information theoretic approach, and the result using SA. Bottom row (left to right): the result using deterministic Gibbs sampler, the three dimensional (3-D) surface corresponding to information theoretic MFA, and the 3-D surface corresponding to deterministic Gibbs sampler.

References

1. G.L. Bilbro, W.E. Snyder, and R.C. Mann. Mean-field approximation minimizes relative entropy. *Jour. of Optical Soci. America*, Vol-8(No.2):290–294, Feb. 1991.
2. C. Chang. *Area-Based Methods for Stereo Vision: the Computational Aspects and Their Applications*. PhD thesis, University of California, San Diego, 1991. Dept. of ECE.
3. C. Chang and S. Chatterjee. A hybrid approach toward model-based texture segmentation. *Pattern Recognition*, 1990. Accepted for publication.
4. D. Geiger and F. Girosi. Mean field theory for surface reconstruction. In *Proc. DARPA Image Understanding Workshop*, pages 617–630, Palo Alto, CA, May 1989.
5. S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. on Patt. Anal. & Mach. Intel.*, Nov. 1984.
6. R.M. Gray. *Entropy and Information Theory*. Springer-Verlag, New York, NY, 1990.
7. J. Hertz, A. Krogh, and R.G. Palmer. *Introduction to The Theory of Neural Computation*. Addison Wesley, Reading, MA, 1991.
8. H.P. Hiriyannaiah, G.L. Bilbro, W.E. Snyder, and R.C. Mann. Restoration of piecewise-constant images by mean field annealing. *Jour. of Optical Soci. America*, Vol-6(No.12):1901–1912, Dec. 1989.
9. S. Kullback. *Information Theory and Statistics*. John Wiley & Sons, New York, NY, 1959.
10. G. Parisi. *Statistical Field Theory*. Addison Wesley, Reading, MA, 1988.
11. A. Yuille. Generalized deformable models, statistical physics and matching problems. *Neural Computation*, Vol.2(No.1):1–24, 1990.
12. A. Yuille, D. Geiger, and H. Bulthoff. Stereo integration, mean field theory and psychophysics. In *Proc. 1st European Conf. on Comp. Vision*, pages 73–88, Antibes, France, April 1990.

This article was processed using the LaTeX macro package with ECCV92 style