

MOTION ESTIMATION OPTIMIZATION

Sarah A. Rajala, Ikhlas M. Abdelqader, and Griff L. Bilbro
Department of Electrical and Computer Engineering
North Carolina State University, Raleigh, NC 27695-7911
Wesley E. Snyder
Department of Radiology, Bowman Gray School of Medicine

ABSTRACT

In this paper, we cast motion estimation as a problem in energy minimization. This is achieved by modeling the displacement field as a Markov random field. The equivalence of a Markov random field and a Gibbs distribution is then used to convert the problem into one of defining an appropriate energy function that describes the motion and any constraints imposed on it. The energy function is then minimized using the Mean Field Annealing algorithm, a technique which finds the global or near global minima in nonconvex optimization problems. Analysis of the algorithm and experimental results are presented.

1 INTRODUCTION

Motion estimation is the process of describing the movement of an object in the three-dimensional physical world from its two-dimensional projection on a sequence of frames. Ideally, a motion estimation method should be general enough to be applicable to a wide range of images, and since all real world images contain noise, the algorithm should not be strongly sensitive to noise.

In this paper, motion estimation is cast as a problem in energy minimization. This is achieved by modeling the displacement field as a Markov random field (MRF) and formulating a Hamiltonian (also called an objective or energy function) for the displacement vector field. Geman and Geman [1], noted the equivalence of a Gibbs distribution and a Markov random field. This equivalence allowed them to find a maximum *a posteriori* (MAP) estimate by finding the global minimum of the Hamiltonian, instead of finding the most probable state in the *a posteriori* distribution. Geman and Geman then used stochastic simulated annealing (SSA) to find the global minimum. Bilbro and Snyder [2],[3] extended the original formulation of stochastic simulated annealing to arrive at the concept of mean field annealing (MFA). They replaced the stochastic search in SSA with a deterministic one and replaced the

minimization of the energy function H with a minimization of the difference between H and H_0 . H_0 is a simple approximation to H and is referred to as the mean field Hamiltonian.

2 MEAN FIELD ANNEALING

Mean field theory in physics states that the interactions of molecules or particles comprising the system can be approximated by an effective or mean value. This allows the use of an approximate function that is much simpler, but still resembles the actual interactions between the molecules. The idea behind MFA is to find a simpler energy function, or the Hamiltonian H_0 , that approximates the given Hamiltonian H . This approximate Hamiltonian depends on a parameter μ , where the vector μ represents the mean field or effect of the neighboring pixels on the pixel of interest. The task is then switched from minimization of the energy function H to minimization of the difference between H and H_0 . A standard gradient-descent can be used to reach the desired minimum. The probability measure $p(\mathbf{f})$ can be represented by the Gibbs distribution:

$$p(\mathbf{f}) = \frac{1}{Z} e^{-H(\mathbf{f})/T} \quad (1)$$

MFA approximates the Hamiltonian $H(\mathbf{f})$ by a simpler Hamiltonian H_0 . A quadratic function for H_0 has been shown to be suitable for many image processing problems [3]. That is,

$$H_0(\mathbf{f}, \mu) = \sum_{\mathbf{v}_i} (f_i - \mu_i)^2 \quad (2)$$

where μ_i is the mean field variable that represents the mean of the image measure \mathbf{f} at pixel i . Next, the mean field parameters μ_i need to be determined such that $H_0(\mathbf{f}, \mu)$ best approximates the actual $H(\mathbf{f})$. This is ac-

completed by using Jensen's inequality as follows:

$$\langle H(\mathbf{f}) \rangle_{\mu} = \int_{RN^2} H(\mathbf{f}) p_o(\mathbf{f}, \mu) d\mathbf{f} \quad (3)$$

where $\langle H(\mathbf{f}) \rangle_{\mu}$ is the mean of $H(\mathbf{f})$ with respect to μ and $p_o(\mathbf{f}, \mu)$ is the *a posteriori* mean field density. $p_o(\mathbf{f}, \mu)$ is given by

$$p_o(\mathbf{f}, \mu) = \frac{1}{Z_o} e^{-H_o(\mathbf{f}, \mu)/T} = \frac{1}{(\pi T)^{N^2/2}} e^{-H_o(\mathbf{f}, \mu)/T} \quad (4)$$

where N^2 is the size of the image and Z_o is the mean field partition (normalization) function. Defining,

$$Q = -\frac{1}{T} [H(\mathbf{f}) - H_o(\mathbf{f}, \mu)] \quad (5)$$

Then, by Jensen's inequality for convex functions ,

$$\ln \langle e^Q \rangle_{\mu} \geq \langle Q \rangle_{\mu} \quad (6)$$

Substituting in for the value of Q and multiplying by $-T$,

$$-T \ln \left[\int e^{-H(\mathbf{f})/T} d\mathbf{f} \right] \leq -T \ln \left[\int e^{-H_o(\mathbf{f}, \mu)/T} d\mathbf{f} \right] + \langle H(\mathbf{f}) - H_o(\mathbf{f}, \mu) \rangle_{\mu}$$

which can be rewritten as

$$F \leq F_o + \langle H(\mathbf{f}) - H_o(\mathbf{f}, \mu) \rangle_{\mu} \quad (7)$$

where F represents an upper bound on the free energy of the system with respect to $H(\mathbf{f})$ and F_o represents the free energy of the lattice with respect to the mean field Hamiltonian H_o . If the upper bound F in equation (7) is minimized with respect to the mean field μ , then H_o that best approximates approximates the given posterior Hamiltonian $H(\mathbf{f})$ can be obtained by solving the following equation with respect to μ

$$\nabla_{\mu} \{ \langle H(\mathbf{f}) - H_o(\mathbf{f}, \mu) \rangle_{\mu} + T \ln \left[\int e^{-H_o(\mathbf{f}, \mu)/T} d\mathbf{f} \right] \} = 0 \quad (8)$$

By simplifying, equation (8) results in:

$$\nabla_{\mu} \langle H(\mathbf{f}) \rangle_{\mu} = 0 \quad (9)$$

at a given temperature. In MFA a deterministic relaxation method is used to find μ that minimizes $\langle H(\mathbf{f}) \rangle$. At each T , any nonlinear minimization method may be used such as the gradient descent or conjugate gradient descent method.

3 THE HAMILTONIAN

The proposed energy function for motion estimation consists of two parts, the smoothness constraint, $H_s(\mathbf{d})$, and the gradient constraint, $H_g(\mathbf{d})$. Let $H(\mathbf{d})$ represent the total Hamiltonian energy function and \mathbf{d} represents the displacement field, then

$$H(\mathbf{d}) = H_s(\mathbf{d}) + H_g(\mathbf{d}) \quad (10)$$

The first part of the energy function $H_s(\mathbf{d})$, states that the displacement vector is expected to vary smoothly over position due to the inertia of the moving object. Thus one would expect to find smooth displacement fields in the image plane, as long as no edges are encountered. In other words, pixels close to each other and within the same object tend to have the same displacement, i.e. a piecewise constant displacement field. The energy function can be represented as:

$$H_s(\mathbf{d}) = - \sum_i \sum_{j \in N_i} \frac{\alpha}{\sqrt{2\pi\tau}} e^{-\|\mathbf{d}_i - \mathbf{d}_j\|^2 / 2\tau} \quad (11)$$

where α represents the weighting factor for this part of the energy function, N_i is the neighborhood of pixel i denoting the image as $N^2 \times 1$ vector, \mathbf{d}_i and \mathbf{d}_j are the displacement vectors for pixel i and j respectively, such that $\mathbf{d}_i = [d_{xi} \ d_{yi}]^T$, where d_{xi} is the x-component of the displacement of pixel i , and d_{yi} is the y-component of the displacement of pixel i . τ is a problem dependent tolerance parameter, which controls the accuracy of the results. From Equation (11), it is seen that $H_s(\mathbf{d})$ reaches its minimum when $\mathbf{d}_i = \mathbf{d}_j$ which satisfies the smoothness constraint. The $\|\cdot\|$ designates the norm of the difference of the two vectors. The second part of the Hamiltonian function $H_g(\mathbf{d})$, is concerned with the intensity variation over the displacement field. One possible model is to assume that the intensity is preserved under motion leading to:

$$H_g(\mathbf{d}) = - \sum_i \frac{\beta}{\sqrt{2\pi\tau}} e^{-(I_{t_1}[i] - I_{t_2}[i+d_i])^2 / 2\tau} \quad (12)$$

where $I_{t_1}[1]$ is the intensity at location i in frame t_1 , $I_{t_2}[i+d_i]$ is the intensity at location $i+d_i$ in frame t_2 , and β is the weighting coefficient for this part of the energy function.

4 MFA FOR MOTION ESTIMATION

For a continuous-space MRF, one useful mean field Hamiltonian $H_o(\mathbf{d}, \mu)$ is given by

$$H_o(\mathbf{d}, \mu) = \sum \|\mathbf{d}_i - \mu_i\|^2 \quad (13)$$

and μ_i is the mean field parameter for pixel i , that is $\mu_i = [\mu_{x_i} \ \mu_{y_i}]^T$.

As mentioned earlier, the task is to determine μ such that $H_o(\mathbf{d}, \mu)$ best approximates $H(\mathbf{d})$. The MFA estimates of one dimensional and two dimensional motion have been both developed, the two-dimensional motion is presented in the following section.

4.1 Two-Dimensional Motion

For estimates of 2-D motion, the mean field parameter is a two component vector $\mu_i = [\mu_{x_i} \ \mu_{y_i}]^T$. The expected value of the approximate Hamiltonian, H_o , will have the following form:

$$\langle H_o \rangle_\mu = (\pi T)^{3/2} N^2$$

The expected value of $H_s(\mathbf{d})$, which is given in equation (11), is

$$\langle H_s(\mathbf{d}) \rangle_\mu = \sum_i \sum_{j \in \mathbb{N}_i} \frac{-\alpha}{2\pi T(\tau + T)} e^{-[(\mu_{x_i} - \mu_{x_j})^2 + (\mu_{y_i} - \mu_{y_j})^2]/(2T + 2\tau)}$$

Differentiating $\langle H_s(\mathbf{d}) \rangle$ with respect to the mean field parameter μ_i yields:

$$\nabla_{\mu_{x_i}} \langle H_s \rangle = \sum_{j \in \mathbb{N}_i} \frac{\alpha(\mu_{x_i} - \mu_{x_j})}{2\pi T(\tau + T)^2} e^{-[(\mu_{x_i} - \mu_{x_j})^2 + (\mu_{y_i} - \mu_{y_j})^2]/(2T + 2\tau)} \quad (14)$$

and,

$$\nabla_{\mu_{y_i}} \langle H_s \rangle = \sum_{j \in \mathbb{N}_i} \frac{\alpha(\mu_{y_i} - \mu_{y_j})}{2\pi T(\tau + T)^2} e^{-[(\mu_{x_i} - \mu_{x_j})^2 + (\mu_{y_i} - \mu_{y_j})^2]/(2T + 2\tau)} \quad (15)$$

The expected value of $H_g(\mathbf{d})$, which is given in equation (12), is:

$$\langle H_g(\mathbf{d}) \rangle_\mu = -\frac{\sum_i \beta}{T\sqrt{2\pi\tau}} \left(\frac{1}{\left(\frac{dI_{t_2}}{dx}\right)\left(\frac{dI_{t_2}}{dy}\right)} \right) \left[\frac{\sqrt{2\tau T \left(\frac{dI_{t_2}}{dx}\right)^2}}{\sqrt{\left(\frac{dI_{t_2}}{dx}\right)^2 T + 2\tau}} \right] \cdot \left[\frac{\sqrt{2\tau + \left(\frac{dI_{t_2}}{dx}\right)^2 \left(\frac{dI_{t_2}}{dy}\right)^2 T^2}}{\sqrt{2\tau + \left(\frac{dI_{t_2}}{dx}\right)^2 T + \left(\frac{dI_{t_2}}{dy}\right)^2 T}} \right] e^{-\left(\frac{dI}{dt} + \frac{dI_{t_2}}{dx} \mu_{x_i} + \frac{dI_{t_2}}{dy} \mu_{y_i}\right)^2 / (2\tau + T \left[\left(\frac{dI_{t_2}}{dx}\right)^2 + \left(\frac{dI_{t_2}}{dy}\right)^2 \right])} \quad (16)$$

Taking the gradient of equation (16) with respect to μ_{x_i} and μ_{y_i} along with the gradient in equations (14), (15),

the gradient descent method can be used with a certain annealing schedule to obtain the optimum value of the mean field parameter which is the best estimate of the displacement vector field.

5 EXPERIMENTAL RESULTS

5.1 Two-Dimensional Motion

Both synthetic and real image sequences were used to evaluate the two-dimensional motion estimation algorithm. A sequence was generated in which a Gaussian sphere was moved diagonally toward the lower left at a distance of 3 pixels per frame. The estimated displacement field for the diagonally moving sphere is shown in Figure 2, while the sphere is shown in Figure 1. The ratio $\alpha/\beta = 5/2$ was the same as in the one-dimensional experiment. The results were good, however, to assess the real potential of the algorithm, it needs to be evaluated on real data. Figures 3 and 4 show two frames of a sequence entitled BOB. Each image in the sequence is 128x128x8. The estimated displacement field is shown in Figure 5. The displacement estimates are smooth over the spatial extent of the moving objects and the algorithm performs well even in the presence of the recording noise present in the real data. No pre- or post-filtering was performed on any of the images.

6 SUMMARY AND CONCLUSIONS

In this paper, we have developed a new algorithm for motion. The algorithm was the result of formalizing the motion estimation as a problem in energy minimization, modeling the displacement field as a Markov Random Field, and hence using mean field annealing as an optimization technique to solve it. The algorithm produced a dense displacement field that does not depend on a priori information or initial values of the displacement map. The method proposed here does not require pre-processing, nor does it need any post-filtering to smooth the displacement estimates.

REFERENCES

1. S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, No. 6, pp. 721-741, November 1984.

2. G. Bilbro, R. Mann, T.K. Miller, W.E. Snyder, D.E. Bout, and M. White, "Optimization by Mean-Field Annealing," *Advances in Neural Network Information Processing Systems I*, Edited by D.S. Touretzky, Morgan Kaufmann Publishers, San Mateo, CA, pp. 91-98, 1988.

3. W.E. Snyder, G.L. Bilbro, and V. den Bout, "New Techniques in Optimization," Technical Report Number NETR-89-12, Center of Communications and Signal Processing, North Carolina State University, 1988.

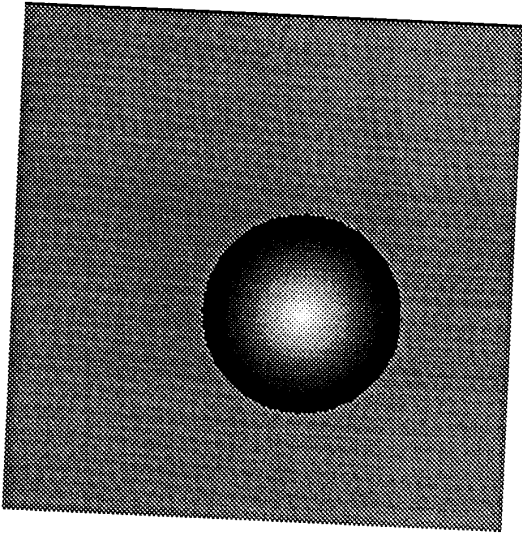


Figure 1: A synthetic image of a Gaussian circle.



Figure 3: First frame of a real world sequence, BOB.

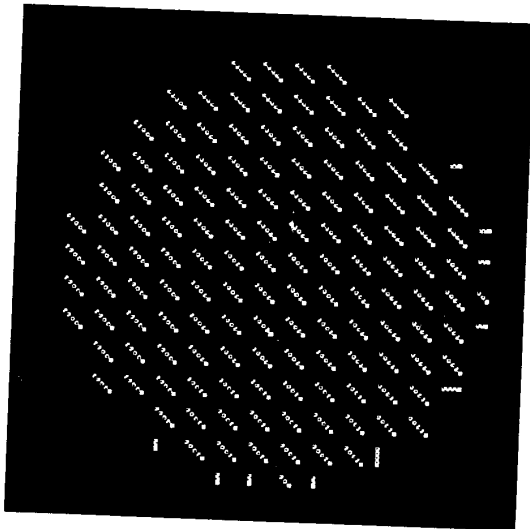


Figure 2: The displacement field for a diagonal motion ($\alpha = 5$ and $\beta = 2$).

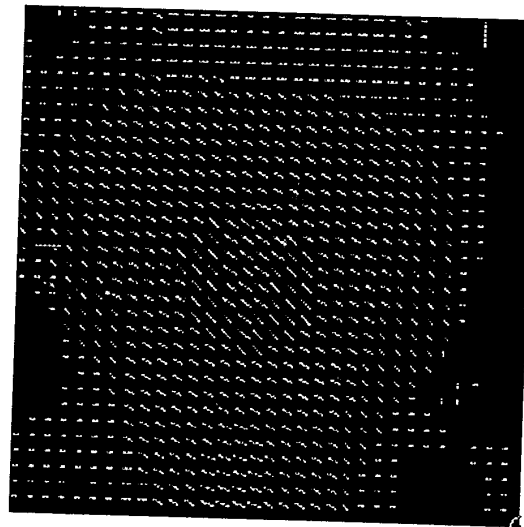


Figure 4: The displacement field from the two frames shown in Figure 3 and 4.