MOTION ESTIMATION FROM NOISY IMAGE DATA

Ikhlas M. Abdelqader, Sarah A. Rajala
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, the problem of motion estimation is formalised as a problem in nonlinear optimisation. The algorithm is based on modeling the displacement fields as Markov Random Fields. The Markov Random Fields-Gibbs distribution equivalence is used to convert the problem into one of finding an appropriate energy function that describes the motion fields. Mean field annealing, a technique for finding the global minima in nonconvex optimisation problems, is used to minimise the Hamiltonian. The estimated displacement vector fields are accurate, even for scenes containing noise or intensity discontinuities.

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 image frames. The objective is to determine what information in the image plane is necessary to characterise the movement of the objects in the scene. Unfortunately, the estimation of dense displacement vector fields from image sequences is an ill-posed problem. This is primarily due to the fact that the data contained in the 2-D projection of a scene does not itself provide sufficient information to completely determine the displacement vector field. The objective of this research is to produce dense displacement vectors that are accurate in noise presence and near discontinuities. In previous work [2], a new algorithm for motion estimation from image sequences was developed. The algorithm models the displacement field as a Markov random field since the displacement value of a certain vector depends on the displacement values of its neighboring pixels. A Hamiltonian is developed that describes the displacement field and any constraints imposed on it. Mean field annealing is then used to find the optimum displacement field. In this paper, the algorithm is extended to account for noise present in the data, hence it produces dense displacement fields that are insensitive to noise.

2 The Hamiltonian

The algorithm developed in this paper seeks an estimate of the following vector for each pixel in the image,

$$\mathbf{f_i} = [d_{xi} \quad d_{yi} \quad I_{t1i} \quad I_{t2i}]^t$$

where the components of this vector are as follows,

- d_{xi} - is the horisontal component of the displacement vector of pixel i.
- d_{yi} - is the vertical component of the displacement vector of pixel i.
- I_{t1i} - is the true intensity value of pixel i in frame $t1$.
- I_{t2i} - is the true intensity value of pixel i in frame $t2$.

The noisy data will be denoted by g_{t1} and g_{t2}, for frame $t1$ and frame $t2$. That is we assume that the system is linear and subjected to additive noise,

$$\mathbf{g} = \mathbf{I} + \mathbf{n}$$

where $\mathbf{I}$ is the true underlying signal, $\mathbf{g}$ is the measurement or the noisy data and $\mathbf{n}$ is the random noise. The noise signal is assumed additive and to have a Gaussian distribution. The mean field vector μ will be given by:

$$\mu_i = [\,\mu_{xi} \quad \mu_{yi} \quad \mu_{I_{t1i}} \quad \mu_{I_{t2i}}\,]^t$$

where μ_{xi} and μ_{yi} are the mean field parameters for the horisontal and the vertical components of the displacement field, respectively. The parameters $\mu_{I_{t1i}}$ and $\mu_{I_{t2i}}$ are the mean field parameters for the true intensity value of pixel i in both frames.

The proposed Hamiltonian consists of three parts, the smoothness Hamiltonian, H_s , the intensity gradient Hamiltonian, H_g, and the noise Hamiltonian H_n.

$$H(\mathbf{f}) = H_s(\mathbf{f}) + H_g(\mathbf{f}) + H_n(\mathbf{f}) \tag{1}$$

The first part of the energy function H_s, states that the displacement vector is expected to vary smoothly over position. In other words, pixels close to each other and within the same object tend to have the same displacement, i.e. a piece-wise constant displacement field. An energy function which enforces this piece-wise constancy can be represented as:

$$H_s(\mathbf{f}) = \lim_{\tau \to 0} \sum_i \sum_{j \in \aleph_i} \frac{-\alpha}{\sqrt{2\pi\tau}} e^{-\|\mathbf{d}_i - \mathbf{d}_j\|^2/2\tau} \tag{2}$$

where α is a weighting factor, $\mathbf{d}_i$ and $\mathbf{d}_j$ are the displacement vectors for pixel i and j respectively, such that $\mathbf{d}_i = [d_{xi} \quad d_{yi}]^t$, where d_{xi} is the horizontal component of the displacement vector of pixel i, and d_{yi} is the vertical component of the displacement of pixel i. $\aleph_i$ is the neighborhood of pixel i (denoting the image as $N^2 \times 1$ vector), the maximum size for a neighborhood is the nearest 8-pixels. Oriented smoothnes is implemented in the sense that a pixel j does not belong to the neighborhood of pixel i if its intensity value is larger than a certain threshold value. The $\| . \|$ designates the norm of the difference between the two displacement vectors. Taking the limit of H_s in equation (2) as τ approaches zero would make the energy function approximates a differentiable dirac delta function. $H_s(\mathbf{d})$ reaches its minimum when $\mathbf{d}_i = \mathbf{d}_j$. The second part of the Hamiltonian function $H_g(\mathbf{d})$, models the intensity variation between the two frames. One possible model is to assume that the intensity is preserved under motion leading us to choose

$$H_g(\mathbf{f}) = \lim_{\tau \to 0} \sum_i \frac{-\beta}{\sqrt{2\pi\tau}} e^{-(I_{t_1}[i] - I_{t_2}[i+d_i])^2/2\tau} \tag{3}$$

where $I_{t_1}[i]$ 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 H_g. The third term in the Hamiltonian is the noise Hamiltonian H_n. It is the term responsible for reducing the effect of the noise on the displacement estimates. This term is,

$$H_n(\mathbf{f}) = \sum_i \frac{1}{2\sigma_n^2} [(I_{t1i} - g_{t1i})^2 + (I_{t2i} - g_{t2i})^2] \tag{4}$$

where σ_n^2 is the variance of the noise, assuming Gaussian, zero mean noise.

3 MFA for Motion Estimation

Use of MFA implies the existence of a mean field Hamiltonian H_o [4]. For a continuous-space MRF, one useful mean field Hamiltonian $H_o(\mathbf{f}, \mu)$ is given by

$$H_o(\mathbf{f}, \mu) = \sum_i \| \mathbf{f}_i - \mu_i \|^2 \tag{5}$$

and μ_i is the mean field parameter for pixel i or it is the mean value of the displacement of pixel i. H_o has a minimum of $\mu = \mathbf{f}$. The mean field Gibbs distribution is then,

$$p_o(\mathbf{f}, \mu) = \frac{1}{Z_o} e^{-H_o(\mathbf{f}, \mu)/T} \tag{6}$$

The next task is to determine μ such that $H_o(\mathbf{f}, \mu)$ best approximates $H(\mathbf{f})$. MFA minimizes the expected value of the difference of the two energy functions.

The expected value of the mean field Hamiltonian, $H_o(\mathbf{f}, \mu)$, will have the following form:

$$< H_o(\mathbf{f}, \mu) >= (2T)N^2 \tag{7}$$

Hence, the algorithm is seeking a solution to the following equation:

$$\nabla_\mu < H(\mathbf{f}) > = 0 \tag{8}$$

at a given temperature.

The evaluation of the expected value of $H(\mathbf{f})$ proceeds as follows:

$$< H(\mathbf{f}) > = \int_{R^{N^2}} H(\mathbf{f}) \; p_o(\mathbf{f}, \mu) \, d\mathbf{f} \tag{9}$$

The expected value of each term of the Hamiltonian is evaluated separately in Appendix B. The expected value of $H_s(\mathbf{f})$:

$$< H_s(\mathbf{f}) > = \; - \sum_i \sum_{j \in \aleph_i} \frac{\alpha}{2\pi T} e^{-[(\mu_{xi} - \mu_{xj})^2 + (\mu_{yi} - \mu_{yj})^2]/2T} \tag{10}$$

The expected value of $H_g(\mathbf{f})$ is:

$$< H_g(\mathbf{f}) >= \sum_i C_i' \tag{11}$$

$$\cdot e^{-(\frac{d\mu I}{dt} + \frac{d\mu I_{t2}}{dx} \mu_x + \frac{d\mu I_{t2}}{dy} \mu_y)^2 / [2T + T((\frac{d\mu I_{t2}}{dx})^2 + (\frac{d\mu I_{t2}}{dy})^2)]}$$

where

$$C_i' = \frac{-\beta}{\sqrt{\pi[2T + T(\frac{d\mu I_{t2}}{dx})^2 + T(\frac{d\mu I_{t2}}{dy})^2]}} \tag{12}$$

The expected value of H_n is evaluated in two steps, Let

$$H_{n1} = \frac{1}{2\sigma_n^2} \sum_i (I_{t1i} - g_{t1i})^2 \tag{13}$$

and

$$H_{n2} = \frac{1}{2\sigma_n^2} \sum_i (I_{t2i} - g_{2ti})^2 \tag{14}$$

The expected value of H_{n1} is,

$$< H_{n1} > = \frac{1}{2\sigma_n^2} \left[\frac{N^2(T/2)}{\sqrt{\pi T}} + \sum_i (\mu_{I_{t1i}} - g_{t1i})^2 \right] \quad (15)$$

H_{n2} has a similar form.,

Using the gradient descent and an annealing schedule to minimise the expected value of the total Hamiltonian with respect to μ, the optimum displacement field is obtained.

4 Experimental Results

To demonstrate the performance of the algorithm it is implemented on several types of sequences. A displacement vector of two components d_x and d_y is estimated for each pixel in the image. To show the accuracy of the estimates, the first frame of the sequence is used with the estimates of the displacement field to reconstruct the second frame of the sequence. This reconstructed frame will be denoted as the "displaced frame".

In Figure 1 the first frame of a sequence entitled SALESMAN is shown. Each image in this sequence is $288 \times 360 \times 8$. The displacement field was estimated. In order to better visualise the results, the estimated displacement field for a window in the SALESMAN sequence is zoomed up. The displacement field for this implementation is shown in Figure 2. The displacement field shown in Figure 2 is used reconstruct the second frame of the sequence. This is shown in Figure 3. For comparison, the same area of the window over the hand from the original frames are zoomed up in Figures 4 and 5.

In all cases, 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.

5 Conclusions

A new algorithm for motion estimation is developed by casting motion estimation as an energy minimisation problem. A Hamiltonian modeling the displacement field was formulated. The noise Hamiltonian was included to reduce the estimates sensitivity to noise presence. The MFA as an optimisation algorithm is used to reach for an optimum estimate of the displacement vector field. Motion estimation from image sequences using MFA produces dense displacement estimates that are correct in the presence of noise and near ares of discontinuities. It is

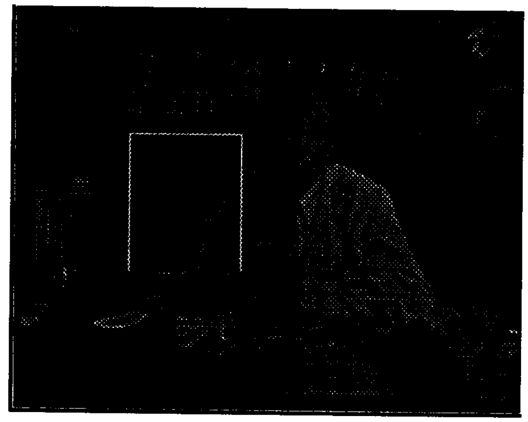

Figure 1: Frame 1 of the SALESMAN sequence with a window showing the location for which the displacement field is estimated.

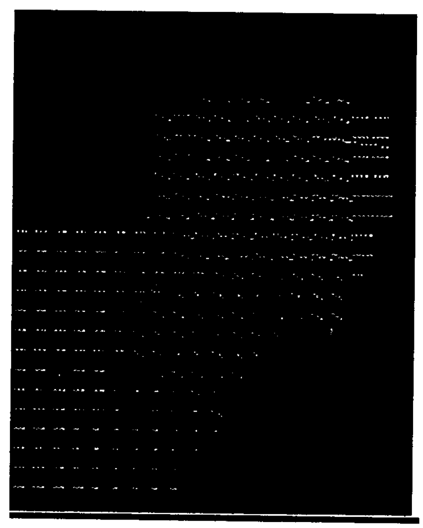

Figure 2: The displacement map for the window area over the arm as shown in Figure 1 for the Salesman sequence.

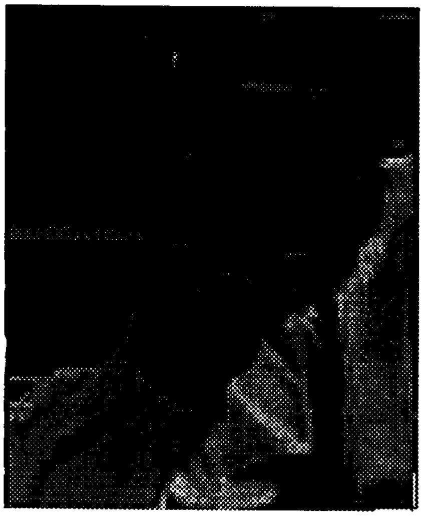

Figure 3: The displaced frame, frame 1 of the Salesman sequence moved by the estimated displacement field shown in Figure 2.

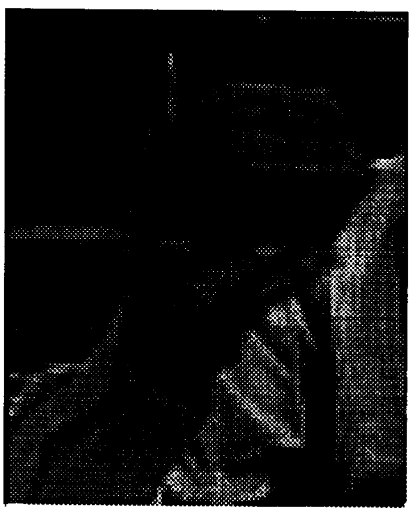

Figure 4: A zoom in over the arm window area from the first frame of the sequence as shown in Figure 1.

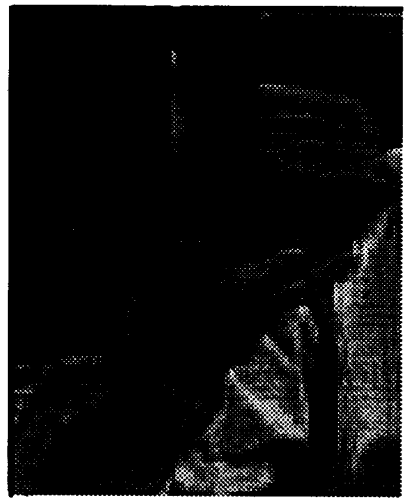

Figure 5: A zoom in over the arm window area from the second frame of the sequence.

an algorithm that is independent of initial displacement values and hence no prior information is needed.

References

[1] Abdelqader, I., "Motion Estimation From Image Sequences Using Mean Field Annealing," *Ph.D. Dissertation,* Electrical Engineering Department, North Carolina State University, Raleigh, August, 1992.

[2] Abdelqader, I., Rajala, S., Snyder, W., and Bilbro, G., "Energy Minimization Approach To Motion Estimation," *Journal Of Signal Processing,* September, 1992.

[3] Bilbro, G., and Snyder, W.E., "Range Image Restoration using Mean-Field Annealing," *Advances in Neural Network Information Processing Systems I,* Edited by Touretsky, D.S., Morgan Kaufmann Publishers, San Mateo, CA,, Pages 594-601, 1988.

5*Journal of the Optical Society of America A,* Volume 6, Number 12, Pages 1901-1912,

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