BAYESIAN ESTIMATION OF DISCONTINUOUS MOTION IN IMAGES USING SIMULATED ANNEALING

Janusz Konrad[†] *and Eric Dubois*

INRS-Télécommunications
3 Place du Commerce, Verdun
Québec, Canada, H3E 1H6

Abstract

In this paper we study the Bayesian estimation of motion in image sequences based on stochastic relaxation algorithms. Unlike previous work, where motion smoothness over the complete image was assumed, here we reformulate the Bayesian approach by allowing motion discontinuities. The motion-image relationship is based on the widely used assumption of constant image intensity along motion trajectories. For the image acquisition process we use white Gaussian noise and we conclude that the displaced pel differences can be modeled by independent Gaussian random variables. The displacement field is modeled by a 2D vector Markov random field such that it encourages smooth motion but allows occasional discontinuities at motion boundaries. These discontinuities, called *line elements*, are modeled by a coupled 2D Markov random field favouring absence of line elements, slightly penalizing straight lines and corners, and severely penalizing line ends, intersections and double lines. The maximum *a posteriori* probability estimation criterion is derived from the above models. The optimization problem involves several thousands of unknowns, and is solved by simulated annealing based on the Gibbs sampler. Results of the estimation algorithm applied to television sequences with natural motion are presented. The motion fields obtained by application of the above model are compared with the fields estimated using the model disregarding discontinuities.

1. INTRODUCTION

Estimation of motion (optical flow) in 2D time-varying images has been proved, like many other early vision tasks, to be an ill-posed problem [1]. Since usually the unknown motion field is related to the observed image sequence through a *structural model*, such as the assumption of constant image intensity along motion trajectories, there are infinitely many motion fields consistent with the observed images. In order to resolve this ambiguity and find a unique solution, various *regularization* methods have been proposed. These methods impose constraints, such as smoothness of the motion field, in the form of a stabilizing functional. Motion estimation techniques developed by Horn and Schunck [2], Hildreth [3], and Nagel [4] are examples of regularization.

A different approach to solving this non-uniqueness problem is through Bayesian estimation. Such a Bayesian formulation has been developed by Konrad and Dubois for 2D motion estimation at single resolution level [5] and also via a hierarchical approach [6]. They used the maximum *a posteriori* probability (MAP) estimation criterion, which resulted in minimization of a certain cost functional. This cost functional had the same form as in the regularization approach, and in fact regularization methods can be considered as special cases of Bayesian estimation, with particular choice of the random field model for the motion.

If the cost functional derived from the *structural model* and the displacement field model is quadratic with respect to the motion vectors, the necessary conditions for optimality can be easily established, resulting in a large system of linear equations. Such a system is usually solved using either deterministic relaxation methods (Gauss-Seidel, Jacobi) [2], [4] or the Chebyshev method [7]. There is no guarantee, however, that the global optimum will be attained. For this reason, *stochastic relaxation* algorithms, such as the Metropolis algorithm [8] or Gibbs sampler [9], have been proposed. Incorporated into *simulated annealing* [10], these methods provide convergence (under certain conditions) to the global minimum, and have been used to solve various ill-posed problems e.g., image reconstruction [9], [11], segmentation of moving planar surfaces [12], stereo matching [11], [13].

In this paper we extend the stochastic approach to motion estimation by incorporating motion discontinuities into the model. As in [5] we use a 2D vector Markov random field to model a displacement field, but we also introduce a coupled Markov random field called a *line process* (proposed in [9] to model intensity discontinuities for image restoration) to model motion discontinuities. These two models result in a piecewise-smooth description of motion with occasional discontinuities at the motion boundaries, which is expected to improve estimates in the occlusion areas. We derive the cost functional for MAP estimation of motion fields based on the above models, and minimize it using inhomogeneous simulated annealing implemented via the Gibbs sampler. We present some results of this algorithm applied to television sequences with natural motion, and compare them with motion fields obtained from the model disregarding discontinuities.

This work was supported by the Natural Sciences and Engineering Research Council of Canada under Strategic Grant G-1357

[†] Also with McGill University, Montreal, on leave from the Technical University of Szczecin, Poland

2. PROBLEM FORMULATION

2.1 Terminology

The images considered in this paper are time-varying, hence they are functions of 3 coordinates (horizontal, vertical and temporal). Let u denote the true underlying image defined as an illuminance pattern in some image plane acquired from the observed scene via an ideal optical system. Let g be an observed image, which is obtained from u by some general transformation (filtering, sampling, quantization etc.). We consider g to be a sample from a random field (multidimensional stochastic process) G. The image g is assumed to be sampled on a lattice Λ_g in R^3. Such a lattice is a collection of sites in R^3 uniquely described by a sampling matrix [14]. We investigate here orthogonal lattices only, hence the sampling matrices are strictly diagonal. Each field of the image sequence contains M_g picture elements and consecutive fields are spaced by T sec. Let $(\mathbf{x}_i, t)$ denote the i-th pel in the image field at time t. The numbering order of the picture elements ($i = 1, ..., M_g$) in a given field is arbitrary (e.g., horizontal scan).

Let us assume that the unknown (true) displacement field $\boldsymbol{d}$ is a sample from a random field $\mathbf{D}$, and that it is defined over continuous spatio-temporal coordinates $(\mathbf{x}, t)$. The displacement vector $\boldsymbol{d}(\mathbf{x}, t)$ is defined as follows: the image point $\mathbf{x} - \boldsymbol{d}(\mathbf{x}, t)$ at time $t - T$ moves to position $\mathbf{x}$ at time t. We denote an estimate of the true displacement field $\boldsymbol{d}$ for a given image sequence by $\widehat{\mathbf{d}}$, while $\mathbf{d}$ denotes any sample field from $\mathbf{D}$. We assume that $\widehat{\mathbf{d}}$ and $\mathbf{d}$ are defined over an orthogonal lattice $\Lambda_{\mathbf{d}}$ in R^3, and that they comprise $M_{\mathbf{d}} = M_1 \times M_2$ vectors (M_1,M_2 denote the horizontal and vertical displacement field sizes, respectively). Clearly, $\mathbf{d}$ is an array of vectors defined over the lattice $\Lambda_{\mathbf{d}}$. In general Λ_g and $\Lambda_{\mathbf{d}}$ can be different, however here we assume that they are equal (then the number of displacement vectors $M_{\mathbf{d}}$ is equal to M_g).

Let $\boldsymbol{l}$ be the unknown underlying discontinuity field of the true displacement field $\boldsymbol{d}$. $\boldsymbol{l}$ is defined over continuous spatio-temporal coordinates $(\mathbf{x}, t)$, and can be understood as an indicator function (e.g., binary) for each $(\mathbf{x}, t)$. Both $\boldsymbol{d}$ and $\boldsymbol{l}$ are unobservable. In this paper we will model the true discontinuity field $\boldsymbol{l}$ by a binary random field L. A sample field from L will be denoted by l, while for an estimate of $\boldsymbol{l}$ we will use $\widehat{l}$. The random field L is also called a *line process* and its individual samples are called *line elements*. The random field L is defined over a union of shifted lattices $\Psi_l = \psi_h \cup \psi_v$ [14], where ψ_h and ψ_v are shifted orthogonal lattices for horizontal and vertical line elements respectively:

$$\psi_h = \Lambda_{\mathbf{d}} + [0, T_v^{\mathbf{d}}/2, 0]^T$$
$$\psi_v = \Lambda_{\mathbf{d}} + [T_h^{\mathbf{d}}/2, 0, 0]^T.$$

$T_h^{\mathbf{d}}$ and $T_v^{\mathbf{d}}$ are horizontal and vertical distances between displacement vectors, and superscript T denotes a transposition. Consequently there are $M_l = M_1 \times (M_2 - 1) + (M_1 - 1) \times M_2$ horizontal and vertical line elements.

Let the subscript t denote the restriction of a random field or of its realization to time t. Then, $\mathbf{d}_t$ is a sample field of $\mathbf{D}_t$ ($\mathbf{D}$ at time t), while $\mathbf{d}(\mathbf{x}_i, t)$ is a single displacement vector at spatial location $\mathbf{x}_i$ and time t.

It is assumed, for computational convenience, that the random fields G_t, $\mathbf{D}_t$ and L_t are defined over discrete state spaces $\mathcal{S}_g = (\mathcal{S}'_g)^{M_g}$, $\mathcal{S}_{\mathbf{d}} = (\mathcal{S}'_{\mathbf{d}})^{M_{\mathbf{d}}}$ and $\mathcal{S}_l = (\mathcal{S}'_l)^{M_l}$, respectively, where $(\mathcal{S}')^M$ denotes the M-fold Cartesian product of $\mathcal{S}'$. $\mathcal{S}'_g$, $\mathcal{S}'_{\mathbf{d}}$ and $\mathcal{S}'_l$ are single pel, single displacement vector and single line element state-spaces. The state spaces $\mathcal{S}'_g$ and $\mathcal{S}'_{\mathbf{d}}$ correspond to a sufficiently fine quantization of the underlying continuous image intensities and displacements, so that their characteristic properties are preserved. The state space of the line elements $\mathcal{S}'_l$ is assumed binary, corresponding to "off" and "on" states of the motion discontinuities. Hence, the probability measure used throughout the paper is a discrete probability distribution rather than a continuous density.

2.2 Estimation criterion

We want to find the true displacement field $\boldsymbol{d}(\mathbf{x}, t)$ corresponding to an underlying time-varying image $u(\mathbf{x}, t)$ on the basis of the observations $g(\mathbf{x}, t)$. Since the line field $\boldsymbol{l}(\mathbf{x}, t)$ describing motion discontinuities will be used in the displacement model, it has to be estimated, too. Although in general the number of observed image fields could be arbitrary, in this paper we address the estimation of $(\boldsymbol{d}_t, \boldsymbol{l}_t)$ from two images at times $t - T$ and t.

In order to determine the "best" or "most likely" displacement field $\widehat{\mathbf{d}}_t^* \in \mathcal{S}_{\mathbf{d}}$ and line field $\widehat{l}_t^* \in \mathcal{S}_l$ given the observations g_{t-T}, g_t, we have to find such a pair $(\widehat{\mathbf{d}}_t^*, \widehat{l}_t^*)$ which will satisfy the following relationship:

$$P(\mathbf{D}_t = \widehat{\mathbf{d}}_t^*, L_t = \widehat{l}_t^* | G_{t-T} = g_{t-T}, G_t = g_t) \geq$$
$$P(\mathbf{D}_t = \widehat{\mathbf{d}}_t, L_t = \widehat{l}_t | G_{t-T} = g_{t-T}, G_t = g_t)$$
$$\forall\, \widehat{\mathbf{d}}_t \in \mathcal{S}_{\mathbf{d}},\ \widehat{l}_t \in \mathcal{S}_l,$$

where P is a probability measure. To obtain the posterior distribution involved in the above relationship, Bayes rule for discrete random variables can be applied as follows (to simplify the notation the sample fields $\mathbf{d}_t, l_t, g_{t-T}, g_t$ are omitted in the probability expressions):

$$P(\mathbf{D}_t, L_t | G_{t-T}, G_t) = \frac{P(G_t | \mathbf{D}_t, L_t, G_{t-T}) \cdot P(\mathbf{D}_t, L_t | G_{t-T})}{P(G_t | G_{t-T})}. \quad (1)$$

In previous work [5], [6] it was assumed that the random field $\mathbf{D}$ at time t is statistically independent of the observation G at time $t - T$ i.e., that the knowledge of a single image field provides no information about the motion. Consequently, since no line process was involved in the displacement field model, the *a priori* probability $P(\mathbf{D}_t, L_t | G_{t-T})$ was simplified to $P(\mathbf{D}_t)$. That was a coarse approximation in the model, because the knowledge of intensity pattern only in one image may be helpful in displacement vector computations e.g., a uniform intensity area (still or moving) should not contribute to a discontinuity in a displacement field. In this work we still assume direct independence between $\mathbf{D}_t$ and G_{t-T}. However, the observation G_{t-T} is assumed to affect the displacement process $\mathbf{D}_t$ indirectly through the line process L_t.

Note that since the probability in the denominator of (1) is not a function of the displacement process $\mathbf{D}_t$ or of the line process L_t, it can be ignored when maximizing $P(\mathbf{D}_t, L_t | G_{t-T}, G_t)$ with respect to $(\widehat{\mathbf{d}}_t, \widehat{l}_t)$, and the MAP es-

timate of the pair $(\boldsymbol{d}_t, \boldsymbol{l}_t)$ is the solution to the following optimization problem:

$$\max_{(\hat{\mathbf{d}}_t, \hat{l}_t)} \Big[P(G_t = g_t | \mathbf{D}_t = \hat{\mathbf{d}}_t, L_t = \hat{l}_t, G_{t-T} = g_{t-T}) \cdot P(\mathbf{D}_t = \hat{\mathbf{d}}_t, L_t = \hat{l}_t | G_{t-T} = g_{t-T}) \Big]. \quad (2)$$

2.3 Models

2.3.1 Structural model

As it has been stated at the beginning, the displacement fields $\mathbf{d}_t$ are unobservable. Hence, in order to compute the motion vectors given some observed images, certain relationship or a "structural" model between those vectors and image intensity values must be assumed. Such a model is crucial to any motion estimation algorithm. It is assumed here that over the time interval $[t-T, t]$ the image intensity of the true underlying image u along the true motion trajectory $\boldsymbol{d}$ is constant. Thus the following holds:

$$u(\mathbf{x} - \boldsymbol{d}(\mathbf{x}, t), t - T) = u(\mathbf{x}, t). \quad (3)$$

Also other models e.g., allowing linear variation of intensity, have been devised [15], [16], but such approaches will not be pursued here.

2.3.2 Observation model

First let us consider a simplified case. Assume that the observed image g is related to the true underlying image u via additive white Gaussian noise n:

$$g(\mathbf{x}, t) = u(\mathbf{x}, t) + n(\mathbf{x}, t), \quad (4)$$

and that g, u, n are temporarily defined over continuous $\mathbf{x}$. Then, incorporating the observation model (4) in the structural model (3) we obtain the following relationship:

$$g(\mathbf{x}, t) - g(\mathbf{x} - \boldsymbol{d}(\mathbf{x}, t), t - T) = n(\mathbf{x}, t) - n(\mathbf{x} - \boldsymbol{d}(\mathbf{x}, t), t - T). \quad (5)$$

The two difference terms in (5) are called the displaced pel difference (DPD) and the displaced noise difference, respectively. Since the noise is white Gaussian, the displaced noise differences are independent Gaussian random variables (as the sum of two independent identically distributed Gaussian random variables).

In reality this simplified case does not hold because the true underlying image u undergoes various transformations and is subjected to noise influence before it becomes the observed image g. The transformations include γ-correction, spatio-temporal filtering, sampling, quantization, while the noise sources incorporate the image sensor noise, distortion due to aliasing, and quantization noise.

Because of the extreme complexity of the theoretical derivation of the DPD model, we have tried to establish experimentally the type of probability distribution reflecting the characteristic properties of displaced pel differences. We have computed a number of histograms and autocorrelation functions of displaced pel differences obtained by some existing motion estimation techniques. In particular we applied reliable methods of Horn and Schunck [2] and of Paquin and Dubois [17] to the test image used in this paper as well as to other data. The histograms exhibited shapes similar to that of a Gaussian distribution with "closeness" and variance depending on the image material and motion estimation technique used. Estimates of the autocorrelation functions very closely approximated a Dirac impulse, indicating near independence between the displaced pel differences.

Based on the above simplified derivation and on the experimental results, we will use independent, identically distributed discrete random variables drawn from the Gaussian distribution with variance σ^2, to model the displaced pel differences (5). Due to the independence of DPDs, the conditional likelihood from equation (1) can be expressed as

$$P(G_t = g_t | \mathbf{D}_t = \hat{\mathbf{d}}_t, L_t = \hat{l}_t, G_{t-T} = g_{t-T}) = \prod_{i=1}^{M_\mathrm{d}} p_{n_d}(g(\mathbf{x}_i, t) - \tilde{g}(\mathbf{x}_i - \hat{\mathbf{d}}(\mathbf{x}_i, t), t - T)) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{M_\mathrm{d}} \cdot e^{-U_g(g_t | \hat{\mathbf{d}}_t, g_{t-T})/(2\sigma^2)}, \quad (6)$$

where p_{n_d} is a Gaussian distribution with variance σ^2, and $\tilde{g}$ denotes an intensity value at spatial location $(\mathbf{x}_i - \hat{\mathbf{d}}(\mathbf{x}_i, t), t - T) \notin \Lambda_g$ obtained by some interpolation method (e.g., bilinear, biquadratic, bicubic). Consequently the energy U_g is defined as follows:

$$U_g(g_t | \hat{\mathbf{d}}_t, g_{t-T}) = \sum_{i=1}^{M_\mathrm{d}} [g(\mathbf{x}_i, t) - \tilde{g}(\mathbf{x}_i - \hat{\mathbf{d}}(\mathbf{x}_i, t), t - T)]^2. \quad (7)$$

Note that U_g expresses simply a pel matching problem, and solved on its own turns out to be ill-posed.

2.3.3 Displacement field model

In most scenes the motion can be attributed to the movements of rigid or almost rigid bodies. After projection onto the image plane, the three-dimensional motion induces an optical flow which consists of patches of similar (orientation and length) vectors, with discontinuities at the motion boundaries. In previous work [5], [6] it was assumed that motion fields are smooth functions of spatial position $\mathbf{x}$ (fixed t), and a 2D vector Markov random field (VMRF) $\mathbf{D}_t$ was used to model the motion fields $\mathbf{d}_t$. This assumption, however, is violated at the boundaries of moving objects. With this simple model the motion vectors in the vicinity of a motion boundary become oversmoothed. For example, if two vectors are positioned on the opposite sides of a motion boundary, the 2D VMRF model will enforce smoothness regardless of the fact that the two vectors may belong to two differently moving objects. In other words, smoothness is enforced homogeneously across the whole displacement field. In order to reduce this effect, we assume that the discontinuities l_t of a motion field $\mathbf{d}_t$ exist and that they can be modeled by another (coupled) MRF L_t, also called a *line process*. With such a two-layer model, the same two vectors can be decoupled by introducing a line element between them. In this way the smoothness is enforced (spatially) inhomogeneously. The idea of modeling the discontinuities by a coupled MRF has been introduced by Geman and Geman [9] for image restoration, and subsequently used for boundary detection [18] and segmentation of moving planar surfaces [12]. A deterministic version of line elements has been used in optical flow computation [19] via non-stochastic methods.

Let X denote a random field with samples χ. It has been shown [20] that given a neighbourhood system $\mathcal{N}$, X is a MRF

with respect to $\mathcal{N}$ if and only if its joint distribution is a *Gibbs* distribution relative to $\mathcal{N}$. Thus the spatial properties of a MRF are uniquely characterized by parameters of the Gibbs distribution

$$\pi(\chi) = \frac{1}{Z} e^{-U(\chi)/\beta}, \qquad U(\chi) = \sum_{c \in \mathcal{C}} V(\chi, c), \tag{8}$$

where U is an *energy function*, c is a *clique*, $\mathcal{C}$ is a set of cliques and β, Z are constants. A clique is a set of sites such that for a defined neighbourhood system every two sites belonging to this set are neighbours. $V(\chi, c)$ is called a *potential function* over clique c, and depends only on those elements from χ which belong to the clique c. Z is called a *partition function*, and is a normalizing constant such that π is a probability measure.

We assume that the pair $\chi = (\mathbf{d}_t, l_t)$ is a realization of the 2D MRF pair $X = (\mathbf{D}_t, L_t)$, which is characterized by the following Gibbs distribution:

$$P(\mathbf{D}_t, L_t) = \pi(\mathbf{d}_t, l_t) = \frac{1}{Z} e^{-U(\mathbf{d}_t, l_t)/\beta}. \tag{9}$$

The energy function $U(\mathbf{d}_t, l_t)$ consists of two non-negative energy terms $U(\mathbf{d}_t|l_t)$ and $U(l_t)$ [9]. Consequently $P(\mathbf{D}_t|L_t)$ and $P(L_t)$ are Gibbsian, and $\mathbf{D}_t$ given L_t as well as L_t by itself are Markovian. Applying the Bayes rule to the conditional probability $P(\mathbf{D}_t, L_t|G_{t-T})$ it follows that:

$$\begin{aligned} P(\mathbf{D}_t, L_t|G_{t-T}) =& P(\mathbf{D}_t|L_t, G_{t-T}) \cdot P(L_t|G_{t-T}) = \\ & P(\mathbf{D}_t|L_t) \cdot P(L_t|G_{t-T}). \end{aligned} \tag{10}$$

Note that the assumption that there is no direct dependence between D_t and G_{t-T} was used in the last line above. Later in this section the conditional probability $P(L_t|G_{t-T})$ will be defined in such a way that it will equal $P(L_t) \cdot e^{-\sum V_{l_1}}$, with non-negative potential V_{l_1}. Since $P(L_t)$ is Gibbsian and V_{l_1} is non-negative, $P(L_t|G_{t-T})$ is Gibbsian as well. Consequently $P(\mathbf{D}_t, L_t|G_{t-T})$ is also Gibbsian.

Let $P(\mathbf{D}_t|L_t)$ be defined as follows:

$$P(\mathbf{D}_t|L_t) = \pi(\mathbf{d}_t|l_t) = \frac{1}{Z_{\mathbf{d}}} e^{-U_{\mathbf{d}}(\mathbf{d}_t|l_t)/\beta_{\mathbf{d}}}, \tag{11}$$

where $Z_{\mathbf{d}}$,$\beta_{\mathbf{d}}$ have the same meaning as Z,β in (8), while the conditional energy $U_{\mathbf{d}}(\mathbf{d}_t|l_t)$ is defined as:

$$U_{\mathbf{d}}(\mathbf{d}_t|l_t) = \sum_{c_{\mathbf{d}} = \{\mathbf{x}_i, \mathbf{x}_j\} \in \mathcal{C}_{\mathbf{d}}} V_{\mathbf{d}}(\mathbf{d}_t, c_{\mathbf{d}}) \cdot [1 - l(<\mathbf{x}_i, \mathbf{x}_j>, t)]. \tag{12}$$

$c_{\mathbf{d}}$ is a vector clique, while $\mathcal{C}_{\mathbf{d}}$ is a set of all vector cliques defined over $\Lambda_{\mathbf{d}}$. In this paper we consider only two-element vector cliques. $V_{\mathbf{d}}$ is a (non-negative) potential function, and $(<\mathbf{x}_i, \mathbf{x}_j>, t) \in \Psi_l$ denotes a site of line element located between vector sites $\mathbf{x}_i$ and $\mathbf{x}_j$ which belong to $\Lambda_{\mathbf{d}}$.

The energy function (12) can be understood as follows. With every vector clique there is associated a value (cost) $V_{\mathbf{d}}(\mathbf{d}_t, c_{\mathbf{d}})$ which should increase if a field sample locally departs from the assumed *a priori* model of the field. This model is characterized by $\beta_{\mathbf{d}}$ and $V_{\mathbf{d}}$. If, however, the line element, separating the displacement vectors from clique $c_{\mathbf{d}}$ is "turned on" ($l(<\mathbf{x}_i, \mathbf{x}_j>, t) = 1$), there is no cost associated with the clique $c_{\mathbf{d}}$. In this way there is no penalty for introducing an abrupt change in length or orientation of a displacement vector.

The ability to zero the cost associated with vector cliques by inserting a line element must be penalized, however. Otherwise a line field with all elements "on" would give the zero energy $U_{\mathbf{d}}$. This penalty is provided by the line field model which is based on a binary MRF L_t, and described by the Gibbs probability distribution

$$\begin{aligned} P(L_t|G_{t-T}) =& \pi(l_t|g_{t-T}) = \frac{1}{Z_l} e^{-U_l(l_t|g_{t-T})/\beta_l}, \\ & U_l(l_t|g_{t-T}) = \sum_{c_l \in \mathcal{C}_l} V_l(l_t, g_{t-T}, c_l), \end{aligned} \tag{13}$$

where U_l is a line energy function, c_l is a line clique and $\mathcal{C}_l$ is a set of all line cliques defined over Ψ_l. The line potential function V_l provides a penalty associated with introduction of a line element, and Z_l, β_l are constants as before.

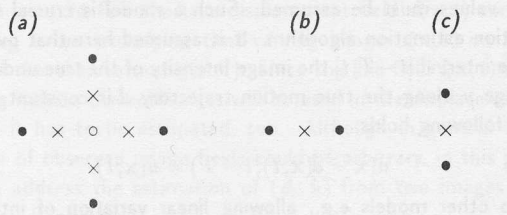


Fig. 1 First-order neighbourhood system $\mathcal{N}_{\mathbf{d}}^1$ for vector field $\mathbf{d}_t$ defined over $\Lambda_{\mathbf{d}}$ (a), and associated horizontal (b) and vertical (c) cliques ($\circ$ – center vector site, $\bullet$ – vector site, $\times$ – line site).

To uniquely specify the *a priori* models (and hence distributions) we need to define neighbourhood systems $\mathcal{N}_{\mathbf{d}}$, $\mathcal{N}_l$, and cliques $c_{\mathbf{d}}$, c_l, as well as the potential functions $V_{\mathbf{d}}$, V_l. In this work we use the first-order neighbourhood system $\mathcal{N}_{\mathbf{d}}$ depicted in Fig. 1.a, which consists of 2-element horizontal vector cliques (Fig. 1.b):

$$\mathcal{C}_h = \left\{ c_{\mathbf{d}} = \left\{ \mathbf{x}_i, \mathbf{x}_j \right\} : \mathbf{x}_i - \mathbf{x}_j = [T_h^{\mathbf{d}}, 0] \right\},$$

and 2-element vertical vector cliques (Fig. 1.c):

$$\mathcal{C}_v = \left\{ c_{\mathbf{d}} = \left\{ \mathbf{x}_i, \mathbf{x}_j \right\} : \mathbf{x}_i - \mathbf{x}_j = [0, T_v^{\mathbf{d}}] \right\},$$

where $(\mathbf{x}_i, t), (\mathbf{x}_j, t) \in \Lambda_{\mathbf{d}}$. The set of cliques $\mathcal{C}_{\mathbf{d}}$ is defined as the union of horizontal and vertical sets:

$$\mathcal{C}_{\mathbf{d}} = \mathcal{C}_h \cup \mathcal{C}_v.$$

Single-vector cliques, as a degenerate case, are not considered. Note that every displacement vector has 4 vector neighbours and 4 line neighbours. We define the potential function over $c_{\mathbf{d}}$ as follows:

$$V_{\mathbf{d}}(\mathbf{d}_t, c_{\mathbf{d}}) = \|\mathbf{d}(\mathbf{x}_i, t) - \mathbf{d}(\mathbf{x}_j, t)\|^2, \qquad c_{\mathbf{d}} = \{\mathbf{x}_i, \mathbf{x}_j\}, \tag{14}$$

where $\|\cdot\|$ is a norm in R^2 e.g., L^2 (note that the summation in (12) is now over i and j). This particular potential captures the smoothness of the displacement field process $\mathbf{D}_t$; for $\mathbf{d}(\mathbf{x}_i, t) = \mathbf{d}(\mathbf{x}_j, t)$ the potential is zero, and the probability of such a configuration is high, while any deviation from this equality causes a smooth reduction in the probability of such an arrangement.

The neighbourhood system for the "dual" structure Ψ_l is shown in Fig. 2. Note that since the union of shifted lattices $\Psi_l = \psi_h \cup \psi_v$ identifies positions of horizontal and vertical line elements, two neighbourhood systems are defined

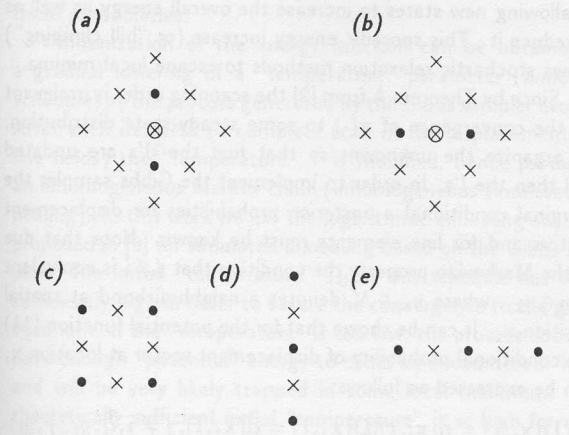


Fig. 2 Neighbourhood system $\mathcal{N}_l^1$ for line field l_t defined over Ψ_l: neighbourhood of horizontal (a) and vertical (b) line element, and associated four-element (c) and two-element (d),(e) cliques ($\otimes$ – center line site, $\times$ – line site, $\bullet$ – vector site).

(Figs. 2.a,2.b). Every line element has 8 line neighbours and 2 vector neighbours. We define four-element line cliques as proposed in [9] (Fig. 2.c), and also two-element vertical cliques of horizontal elements (Fig. 2.d) and two-element horizontal cliques of vertical elements (Fig. 2.e) as suggested in [11]. A simple analytical expression for potential V_l is not possible, hence it is tabulated as a cost associated with each configuration of a clique. Possible configurations (up to a rotation) and related costs are shown in Fig. 3.a for the four-element clique and in Fig. 3.b for the two-element clique. Note that such potentials encourage absence of line elements (V_{l_4}=0.0), slightly penalize straight lines (V_{l_4}=0.4) and corners (V_{l_4}=0.8), and more heavily penalize ends of lines (V_{l_4}=1.2) and intersections (V_{l_4}=1.2, V_{l_4}=1.2). The "double edges" and sharp turns are discouraged by the high penalty (V_{l_2}=3.2) associated with two-element cliques.

So far we have defined the line model based only on the relationship between the line elements and the displacement vectors. Note, however, that the *a priori* probability of line process (13) is conditioned on the observations. It means that we should take into account the image information g_{t-T} when computing the line samples l_t. If independence between L_t and G_{t-T} were claimed, then only $P(L_t)=\pi(l_t)$ could have been used, and motion discontinuities would have been inferred only from the displacement field $\mathbf{d}_t$. We had tried this approach and encountered some difficulties. Hence, based on the following reasoning (similar to that presented in [19]), we will use the observations g_{t-T} in the line model as well. Note, that in general a 3D scene giving rise to a motion discontinuity will also contribute to an intensity edge. Only under specific circumstances will a motion discontinuity not correspond to an edge of intensity. Hence, we assume that an introduction of a line element should coincide with intensity edge and we use the following potential function for one-element clique:

$$V_{l_1}(l_t, g_{t-T}, c_l) = \begin{cases} \frac{\alpha}{(\nabla_v g_{t-T})^2} \cdot l_h(<\mathbf{x}_i, \mathbf{x}_j>, t) & \text{for hor. } c_l = \{\mathbf{x}_i, \mathbf{x}_j\} \\ \frac{\alpha}{(\nabla_h g_{t-T})^2} \cdot l_v(<\mathbf{x}_i, \mathbf{x}_j>, t) & \text{for ver. } c_l = \{\mathbf{x}_i, \mathbf{x}_j\}, \end{cases}$$

where l_h, l_v are horizontal and vertical line elements, ∇_h, ∇_v are horizontal and vertical components of the spatial gradient at position $(<\mathbf{x}_i, \mathbf{x}_j>, t)$, and α is a constant. Note that potential V_{l_1} is non-negative hence $P(L_t|G_{t-T})$ from (10) is Gibbsian like $P(L_t)$. The above potential introduces a penalty only if a line element is "on" and the appropriate gradient is relatively small. For example with α=10.0 a vertical element at a position with horizontal gradient ∇_h=5.0 will cause a penalty (energy) of 0.4 i.e., equivalent to the smallest penalty of a non-zero line element (two in-line elements).

The line field potential function can be expressed now as:

$$V_l(l_t, g_{t-T}, c_l) = V_{l_4}(l_t, c_l) + V_{l_2}(l_t, c_l) + V_{l_1}(l_t, g_{t-T}, c_l),$$

where V_{l_4} and V_{l_2} are tabulated in Fig. 3, and V_{l_1} is given above.

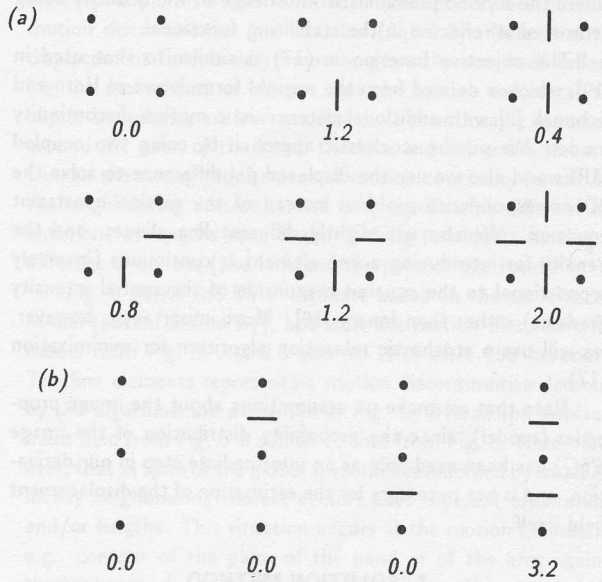


Fig. 3 Costs V_{l_4}, V_{l_2} associated with various configurations (up to rotation) of the four-element (a) and two-element (b) cliques. ($\bullet$ – vector site, — – line element "turned on").

2.4 A posteriori probability and cost functional

Combining the conditional likelihood $P(G_t = g_t|\mathbf{D}_t = \widehat{\mathbf{d}}_t, L_t = \widehat{l}_t, G_{t-T} = g_{t-T})$ from (6), the displacement *a priori* probability $P(\mathbf{D}_t = \widehat{\mathbf{d}}_t|L_t = \widehat{l}_t)$ from (11) and the line *a priori* probability $P(L_t = \widehat{l}_t|G_{t-T} = g_{t-T})$ from (13) via (1) we obtain the following Gibbs form of the *a posteriori* probability

$$P(\mathbf{D}_t = \widehat{\mathbf{d}}_t, L_t = \widehat{l}_t|G_{t-T} = g_{t-T}, G_t = g_t) = \frac{1}{P(G_t|G_{t-T})} \cdot \frac{1}{Z} e^{-U(\widehat{\mathbf{d}}_t, \widehat{l}_t, g_{t-T}, g_t)} \quad (15)$$

where Z is a new normalizing constant, and the new energy function $U(\hat{\mathbf{d}}_t, \hat{l}_t, g_{t-T}, g_t)$ is

$$U(\hat{\mathbf{d}}_t, \hat{l}_t, g_{t-T}, g_t) = \lambda_1 \cdot U_g(g_t|\hat{\mathbf{d}}_t, g_{t-T}) + \lambda_2 \cdot U_{\mathbf{d}}(\hat{\mathbf{d}}_t|\hat{l}_t) + \lambda_3 \cdot U_l(\hat{l}_t|g_{t-T}). \quad (16)$$

The conditional energies in the above relationship are defined in (7), (12) and (13) respectively, and $\lambda_1 = 1/(2\sigma^2)$, $\lambda_2 = 1/\beta_{\mathbf{d}}$, $\lambda_3 = 1/\beta_l$.

Having shown that the posterior distribution (1) is Gibbsian, it follows that MAP estimation can be achieved by means of the following minimization

$$\min_{\{\hat{\mathbf{d}}_t, \hat{l}_t\}} \lambda_1 \cdot U_g(g_t|\hat{\mathbf{d}}_t, g_{t-T}) + \lambda_2 \cdot U_{\mathbf{d}}(\hat{\mathbf{d}}_t|\hat{l}_t) + \lambda_3 \cdot U_l(\hat{l}_t|g_{t-T}) \quad (17)$$

Note that the functional under minimization is in a regularized form, where $U_g(g_t|\hat{\mathbf{d}}_t, g_{t-T})$ describes the original ill-posed problem, $\lambda_2 \cdot U_{\mathbf{d}}(\hat{\mathbf{d}}_t|\hat{l}_t) + \lambda_3 \cdot U_l(\hat{l}_t|g_{t-T})$ is a stabilizing functional and $1/\lambda_1$ is a regularization parameter [1]. This shows that regularization is a special case of Bayesian estimation, where the *a priori* probabilistic knowledge of the quantity being estimated is reflected in the stabilizing functional.

The objective function in (17) is similar to that used in [19], which is derived from the original formulation of Horn and Schunck [2] with additional deterministic motion discontinuity model. We pursue stochastic approach by using two coupled MRFs and also we use the displaced pel difference to solve the 2D correspondence problem instead of the motion constraint equation. We also use slightly different line cliques, and the penalty for introducing a line element is continuous (inversely proportional to the squared magnitude of the spatial intensity gradient) rather than binary [19]. Most importantly, however, we will use a stochastic relaxation algorithm for minimization (17).

Note that we make no assumptions about the image properties (model), since the probability distribution of the image $P(G)$ has been used only as an intermediate step in our derivation, and is not necessary for the estimation of the displacement field itself.

3. SOLUTION METHOD

The minimization (17) involves $2M_1M_2 + M_1(M_2-1) + (M_1-1)M_2$ unknowns, thus exceeding 10^6 for a 512 by 512 image. A method which is able to find the global optimum for such a complex problem is simulated annealing [10]. It is based on the idea of chemical annealing used for determining the low energy states of a material by a gradual lowering of temperature.

Geman and Geman [9] showed that for a MRF, if one starts with any initial configuration and keeps updating all sites $\mathbf{x}_i$, visiting one at a time, according to a pre-computed Gibbsian transition matrix, then the joint distribution will converge to the Gibbs distribution $\pi(\cdot)$ as $t \to \infty$, regardless of the initial state. The above procedure, called the *Gibbs sampler*, generates samples of a MRF such that after infinitely long evolution, the states of the samples are distributed according to that Gibbs distribution. The Gibbs sampler, along with the Metropolis algorithm [8], are examples of *stochastic relaxation*, a class of methods seeking the equilibrium state of a stochastic process by allowing new states to increase the overall energy as well as to reduce it. This sporadic energy increase (or "hill climbing") allows stochastic relaxation methods to escape local minima.

Since by Theorem A from [9] the scanning order is irrelevant for the convergence of $\pi(\cdot)$ to some steady-state distribution, we organize the unknowns so that first the $\mathbf{d}$'s are updated and then the l's. In order to implement the Gibbs sampler the marginal conditional *a posteriori* probabilities for displacement vectors and for line elements must be known. Note that due to the Markovian property the condition that $j \neq i$ is equivalent to $j \in \eta_{\mathbf{x}_i}$, where $\eta_{\mathbf{x}_i} \in \mathcal{N}$ denotes a neighbourhood at spatial position $\mathbf{x}_i$. It can be shown that for the potential function (14) the conditional probability of displacement vector at location $\mathbf{x}_i$ can be expressed as follows:

$$P(\mathbf{D}(\mathbf{x}_i,t) = \hat{\mathbf{d}}(\mathbf{x}_i,t) | \mathbf{D}(\mathbf{x}_j,t) = \hat{\mathbf{d}}(\mathbf{x}_j,t), j \neq i, \hat{l}_t, g_{t-T}, g_t) = \frac{e^{-U_{\mathbf{d}}^i(\hat{\mathbf{d}}(\mathbf{x}_i,t), \hat{l}_t, g_{t-T}, g_t)}}{\sum_{\boldsymbol{z} \in \mathcal{S}'_{\mathbf{d}}} e^{-U_{\mathbf{d}}^i(\boldsymbol{z}, \hat{l}_t, g_{t-T}, g_t)}}, \quad (\mathbf{x}_i,t), (\mathbf{x}_j,t) \in \Lambda_{\mathbf{d}} \quad (18)$$

where the local displacement energy function $U_{\mathbf{d}}^i$ is defined as

$$U_{\mathbf{d}}^i(\boldsymbol{z}, \hat{l}_t, g_{t-T}, g_t) = \lambda_1 \cdot [g(\mathbf{x}_i,t) - \tilde{g}(\mathbf{x}_i - \boldsymbol{z}, t-T)]^2 + \lambda_2 \cdot \sum_{j:\, \mathbf{x}_j \in \eta_{\mathbf{x}_i}} \|\boldsymbol{z} - \hat{\mathbf{d}}(\mathbf{x}_j,t)\|^2 \cdot [1 - \hat{l}(<\mathbf{x}_i, \mathbf{x}_j>, t)].$$

Due to the non-quadratic (data-dependent) form of the energy $U_{\mathbf{d}}^i$ it is not possible to use a standard variate generation algorithm followed by a simple transformation. To generate samples from the probability distribution (18) the conditional probability for each candidate $\hat{\mathbf{d}}(\mathbf{x}_i,t) \in \mathcal{S}'_{\mathbf{d}}$ must be computed. Since $\hat{\mathbf{d}}(\mathbf{x}_i,t)$ is a 2D vector, a 2D array of probabilities is obtained. Accumulating this probability array in one direction (e.g., horizontal), and then sampling according to this accumulated 1D (vertical) distribution, and after that sampling according to the appropriate row (horizontal) distribution will generate a needed sample vector.

Similarly it can be demonstrated that for the line potential function $V_l(l_t, g_{t-T}, c_l)$ the conditional probability of a line element at spatio-temporal location $(\mathbf{y}_i,t) \in \Psi_l$ can be expressed as follows

$$P(L(\mathbf{y}_i,t) = \hat{l}(\mathbf{y}_i,t) | L(\mathbf{y}_j,t) = \hat{l}(\mathbf{y}_j,t), j \neq i, \hat{\mathbf{d}}_t, g_{t-T}) = \frac{e^{-U_l^i(\hat{l}(\mathbf{y}_i,t), \hat{\mathbf{d}}_t, g_{t-T})}}{\sum_{z \in \mathcal{S}'_l} e^{-U_l^i(z, \hat{\mathbf{d}}_t, g_{t-T})}}, \quad (\mathbf{y}_i,t), (\mathbf{y}_j,t) \in \Psi_l \quad (19)$$

where the local line energy function U_l^i is defined as

$$U_l^i(z, \hat{\mathbf{d}}_t, g_{t-T}) = \lambda_2 \cdot \sum_{\substack{c_{\mathbf{d}} = \{\mathbf{x}_m, \mathbf{x}_n\}: \\ <\mathbf{x}_m, \mathbf{x}_n> = \mathbf{y}_i}} \|\hat{\mathbf{d}}(\mathbf{x}_m,t) - \hat{\mathbf{d}}(\mathbf{x}_n,t)\|^2 \cdot [1-z] + \lambda_3 \cdot \sum_{c_l:\, \mathbf{y}_i \in c_l} V_l(z, g_{t-T}, c_l).$$

Samples from the probability distribution (19) are obtained by computing the probabilities associated with possible states of

a line element (0 or 1), and generating a variate according to these probabilities.

Minimization of the energy function can be obtained by a gradual lowering of a "temperature" parameter (*annealing schedule*) as the process generated by the Gibbs sampler evolves. After each iteration (a complete scan of the displacement and line fields) the "temperature" T is modified, hence producing an inhomogeneous Markov chain (inhomogeneous simulated annealing). In this work we use the logarithmic annealing schedule proposed in [9] for simulated annealing based on the Gibbs sampler. The initial "temperature" T_0 of this schedule has to be sufficiently high in order to assure the convergence to the global optimum. If the "temperature" is too low, the process does not have enough "potential" energy to climb all encountered "hills" and will be very likely trapped in some local minimum. The theoretically sufficient initial "temperature" is so high for most of the problems that it would take prohibitively many iterations before the process converges. In practice one has to choose a lower T_0 experimentally, so that the result is not too severely degraded and the computation time is acceptable.

4. RESULTS

We tested the MAP estimation described above on image sequences containing natural motion. We used standard television signal with 2:1 interlace and a temporal spacing between the fields τ=1/60 s. The signal contained some inherent camera and aliasing noise. No pre-processing of the images was performed. As the observations g we used the luminance fields extracted from this signal and quantized to 8 bits (the effective luminance range was 40–200).

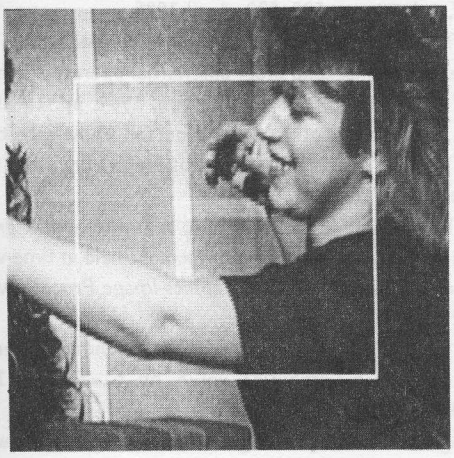


Fig. 4 Test image with natural motion.

Fig. 4 shows one frame from a sequence with such a natural motion. The window indicates that part of the image field (170 by 70 pixels) to which motion estimation was applied. For improved clarity the vector fields displayed in the sequel show only 25% of the estimates obtained (subsampling by 2 in each direction), unless otherwise indicated.

To perform the minimization (17) we must know the values of weights λ_1, λ_2 and λ_3. The constant λ_1 reflects the amount of noise expected in the observations g. The less confidence we have in the accuracy of the data, the smaller λ_1 should be (higher noise variance σ^2). The constants λ_2 and λ_3 control how well the estimate $(\widehat{\mathbf{d}}_t, \widehat{l}_t)$ matches the displacement and line models. At this point we have no means of estimating the λ's, but we recognize the importance of their values for the quality of the estimates. Since the ratios of λ's rather than absolute values matter in the total energy expressions, we defined λ_2 to be equal to 1.0. Then, we experimentally found the other two values which produce (subjectively) good estimates. The results presented below incorporate the motion field model based on the potential described by (14), and the line field model from Fig. 3. The constant α from the one-element line clique is set to 10.0 to significantly penalize introduction of line elements in the areas where magnitude of the corresponding gradient is much smaller than 10.0. The initial value of the "temperature" parameter T_0=1.0 in simulated annealing has been also identified by experimentation. Even fields of the TV sequence have been used as the data, hence the temporal spacing T is equal to 2τ.

Fig. 5 shows the MAP estimate of motion from the test image (Fig. 4) based on the displacement model presented in [5] with λ_1=0.01, after 200 iterations. This model does not take motion discontinuities into account (λ_3=0.0). Note that the displacement field is very smooth, generally very well reflecting the movements of objects, but the motion vectors at the object boundaries are oversmoothed. It is noticeable especially around the palm of the hand, the face and below the forearm. Due to the strong requirement of motion continuity (small λ_1) the neighbouring vectors cannot have significantly different orientations or lengths even if they belong to objects moving in different directions (the face and the palm of the hand).

Fig. 6 shows the MAP estimate based on the same vector model (potential and λ_1), but with the motion discontinuities model from Fig. 3, λ_3=0.3 and α=10.0, after 250 iterations. The line elements representing motion discontinuities detected by the algorithm are presented in Fig. 7. Overall the displacement field from Fig. 6 is similar to that from Fig. 5. Note, however, that in spite of the global smoothness enforced by small λ_1, locally neighbouring motion vectors have different orientations and/or lengths. This situation occurs at the motion boundaries e.g., contour of the palm of the hand or of the arm against the background, and in the occlusion areas e.g., contour of the face occluding the hand. To examine more precisely the motion fields with and without the discontinuity model, Fig. 8 shows sub-windows of the motion fields from Fig. 5 and Fig. 6 with no horizontal or vertical subsampling. Clearly the algorithm did not extract all of the motion boundaries present in this image (Fig. 7). The stronger ones were estimated correctly, however the less pronounced ones (the top edge of the forearm) were not. This effect can be explained by the very small intensity gradient between the object and the background, and also by relatively small motion in this area. Unfortunately a modification of λ_3 and α, so that the motion boundaries of such indistinguishable, slowly moving objects were visible, makes the algorithm too sensitive to variations in motion and intensity. This results in numerous motion boundaries in the areas where they are not present. Such a behaviour is very characteristic of many image processing algorithms, and cannot be easily avoided. We

are planning to investigate the possibility of estimating these crucial parameters.

In the example presented above, the inability to estimate the motion boundary around the forearm may have an impact on application of the motion estimates. If the line elements are supposed to guide an image segmentation process, the results will be quite disastrous. If, however, the vectors are going to be used in motion-compensated interpolation of image sequences (e.g., frame-rate conversion of TV images), the degradation will be unnoticeable. This is due to the fact that the background and object intensities are similar anyway, and if they are confused slightly (motion estimates are still reasonable) no major artifact will arise. The same confusion at an occlusion of two significantly intensity-different objects will result in quite unacceptable errors. There, however, the intensity gradient is large, hence our algorithm will have no difficulty with correct separation of the two motions.

5. CONCLUSIONS

This paper has presented an extension to previous results in Bayesian estimation of motion. The 2D VMRF model for displacement fields has been augmented with a binary 2D MRF model for motion discontinuities to deal with the very important problem of motion boundaries. Such extension permits the enforcement of smoothness constraint across the displacement field in an inhomogeneous fashion i.e., it breaks a "bond " between neighbouring vectors if there exists a line element between them. This new model has been shown to improve the quality of motion estimates at moving object boundaries and in occlusion areas. The line elements, responsible for the ability of motion vectors to depart from the *a priori* vector model (smoothness in this case), have been also associated with intensity gradients to improve reliability of the estimates.

REFERENCES

[1] T. Poggio and V. Torre, "Ill-posed problems and regularization analysis in early vision," MIT Artificial Intelligence Laboratory A.I. Memo 773, 1984.

[2] B.K.P. Horn and B.G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185–203, 1981.

[3] E.C. Hildreth, "Computations underlying the measurement of visual motion," *Artificial Intelligence*, vol. 23, pp. 309–354, 1984.

[4] H.-H. Nagel, "Displacement vectors derived from second-order intensity variations in image sequences," *Comput. Vision, Graphics Image Process.*, vol. 21, pp. 85–117, 1983.

[5] J. Konrad and E. Dubois, "Estimation of image motion fields: Bayesian formulation and stochastic solution," in *Proc. IEEE Int. Conf. on Acoust., Speech, and Signal Process. ICASSP'88*, 1988, pp. 1072–1074.

[6] J. Konrad and E. Dubois, "Multigrid Bayesian estimation of image motion fields using stochastic relaxation," in *Proc. IEEE Int. Conf. Computer Vision ICCV'88*, 1988, pp. 354–362.

[7] D. Lee, "Some computational aspects of low-level computer vision," *Proc. IEEE*, vol. 76, pp. 890–898, August 1988.

[8] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, H. Teller and E. Teller, "Equation of state calculations by fast computing machines," *J. Chem. Phys.*, vol. 21, pp. 1087–1092, June 1953.

[9] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. PAMI-6, pp. 721–741, November 1984.

[10] S. Kirkpatrick, C.D. Gelatt, Jr. and M.P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, pp. 671–680, May 1983.

[11] J.L. Marroquin, "Probabilistic solution of inverse problems," Ph.D. Thesis, MIT Dept. of Electr. Eng. and Comp. Science, 1985.

[12] D.W. Murray and B.F. Buxton, "Scene segmentation from visual motion using global optimization," *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. PAMI-9, pp. 220–228, March 1987.

[13] S.T. Barnard, "Stochastic stereo matching over scale," in *Proc. Image Understanding Workshop, IUS'88*, 1988, pp. 769–778.

[14] E. Dubois, "The sampling and reconstruction of time-varying imagery with application in video systems," *Proc. IEEE*, vol. 73, pp. 502–522, April 1985.

[15] O. Tretiak and L. Pastor, "Velocity estimation from image sequences with second order differential operators," in *Proc. IEEE Int. Conf. Pattern Recognition*, 1984, pp. 16–19.

[16] E.A. Krause, "Motion estimation for frame-rate conversion," Ph.D. Thesis, MIT Dept. of Electr. Eng. and Comp. Science, 1987.

[17] R. Paquin and E. Dubois, "A spatio-temporal gradient method for estimating the displacement field in time-varying imagery," *Comput. Vision, Graphics Image Process.*, vol. 21, pp. 205–221, 1983.

[18] D. Geman, "Stochastic model for boundary detection," *Image & Vision Computing*, vol. 5, pp. 61–65, May 1987.

[19] J. Hutchinson, Ch. Koch, J. Luo and C. Mead, "Computing motion using analog and binary resistive networks," *Computer*, vol. 21, pp. 52–63, March 1988.

[20] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Royal Statist. Soc.*, vol. 36, series B, pp. 192–236, 1974.

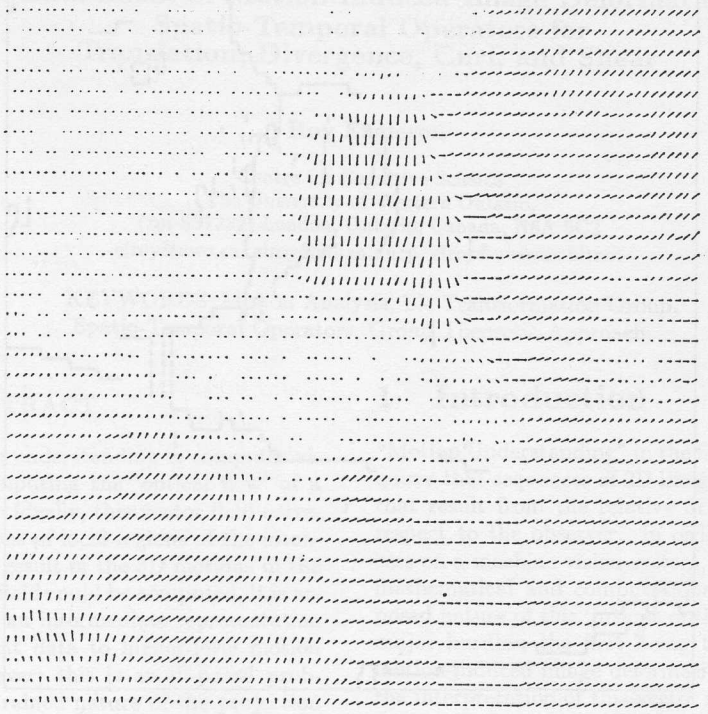


Fig. 5 MAP estimate of motion from the image in Fig. 4 without motion discontinuities model.

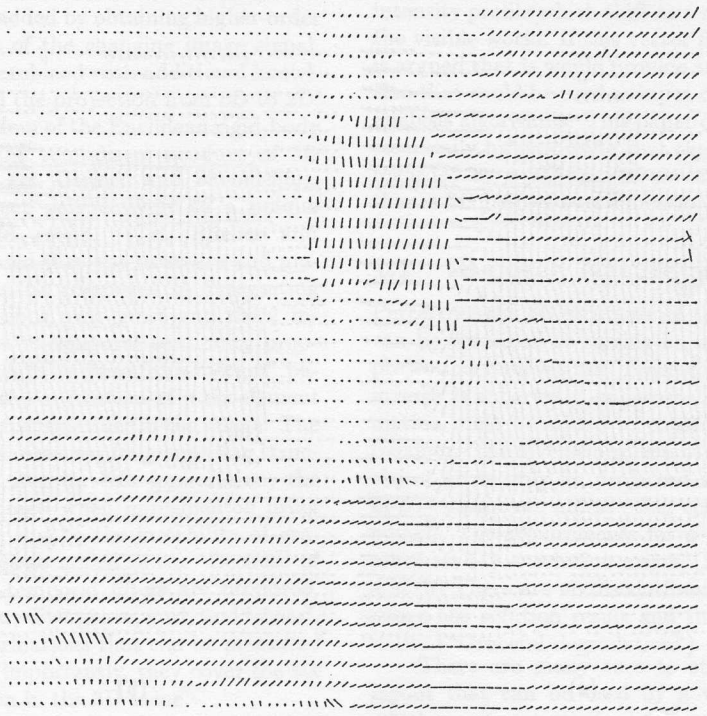


Fig. 6 MAP estimate of motion from the image in Fig. 4 with motion discontinuities model.

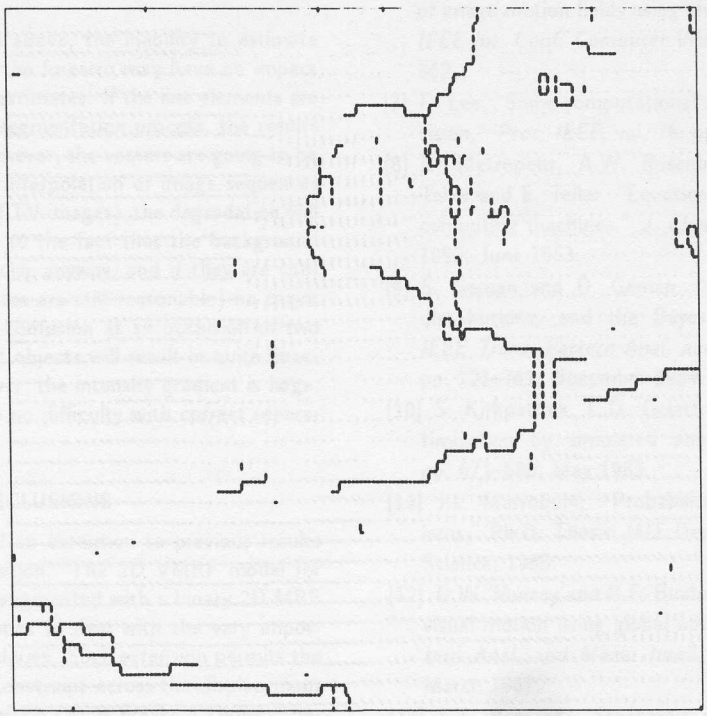
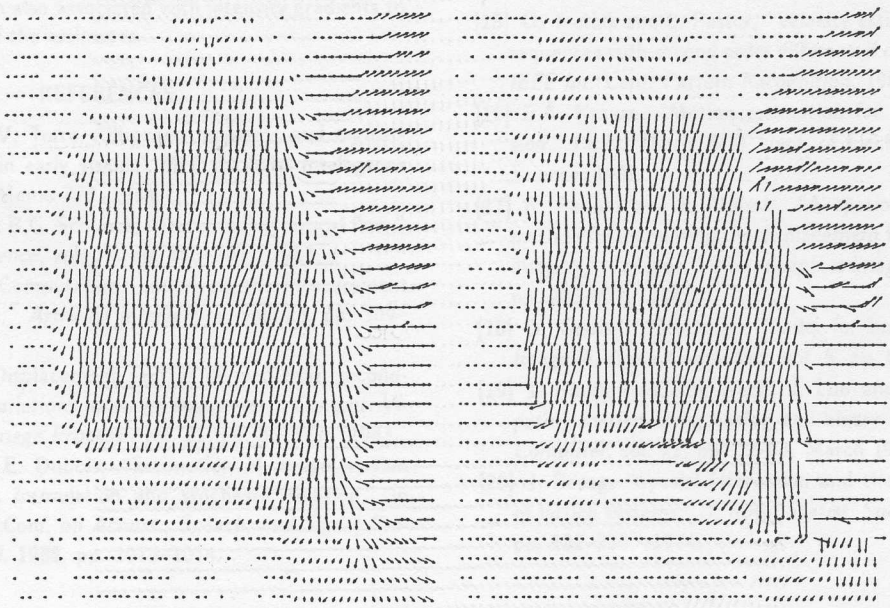


Fig. 7 MAP estimate of motion discontinuities from the image in Fig. 4 computed jointly with motion field in Fig. 6.



(a) (b)

Fig. 8 Enlarged sub-window of motion field from Fig. 5 (a) and from Fig. 6 (b); no horizontal or vertical subsampling of vectors.