A Variational Algorithm For Motion Compensated Inpainting

François Lauze, Mads Nielsen
The IT University of Copenhagen
Rued Langgaards Vej 7
2300 Copenhagen S, Denmark
francois,malte@itu.dk

Abstract

A novel variational algorithm is developed for video inpainting. Within a Bayesian framework, using standard maximum a posteriori to variational formulation rationale, we derive a minimum energy formulation for the estimation of a reconstructed sequence as well as motion recovery. From the Euler-Lagrange Equations, we propose a full multiresolution algorithm in order to compute a good local minimizer for our energy and discuss its numerical implementation. Experimental results for synthetic as well as real sequences are presented.

1 Introduction

Since the birth, more than a century ago, of the cinema, and the apparition of video in the 50's, a huge amount of recorded material has been accumulated. Many of these suffer from degradation, because of recording problems, bad manipulation, or simply the "patina of time". Among these degradations, *blotches* – i.e, following Kokaram ([7] "regions of high contrast that appear at random position in the frame (...)". In these regions, the original data is usually entirely lost.

With the increasing computational power of modern computers, more and more sophisticated algorithms are developed for inpainting of still frames as well as video sequences. They can be roughly divided into stochastic and deterministic methods

In this article we are interested in deterministic algorithms for video sequence inpainting. To our knowledge, not so many deterministic algorithms exist for this problem. Chanas, in its Ph.D. dissertation ([5]) introduced an energy based method using explicitly motion. Our work is based upon his, providing a Bayesian interpretation to it. The "dictionary" between Bayesian and variational framework ([8]) allows to define a generic class of energy methods for image inpainting and motion recovery. Based upon a very recent development in motion recovery, we instanciate this generic class into a novel energy and present a a fully multiscale minimization algorithm for it.

The paper is organized as follow: in the next section we introduce a Bayesian framework for joint inpainting and apparent motion recovery. Using standard rationale we deduce a continuous energy minimization formulation for the posterior probability. We show how naturally the work of Chanas fits in this setting. Then, in section 3, we introduce a novel variational formulation for simultaneous motion recovery and motion

BMVC 2004 doi:10.5244/C.18.80

compensated inpainting, built upon a very recent work of Brox et al. for optical flow computation. From the Calculus of Variations, we deduce a coupled system of PDEs for optical flow recovery and motion compensated inpainting. Section 4 presents the optimization strategy we have used. Section 5 is devoted to some of the numerical aspects. In section 6 we show some results obtained by our algorithm, both in situations where ground truth is known and more realistic situations. We conclude in section 7.

2 A Bayesian Framework for Image Inpainting and Motion Recovery

We introduce first some notations used in the sequel. The spatial digital domain of a sequence will be denoted D_s, $D_s = \{0,\dots,M\} \times \{0,\dots,N\}$, the temporal domain will be denoted $D_t = \{0,\dots,T\}$ and we will also use $D_t^- = \{0,\dots,T-1\}$. A sequence u is a function $D := D_s \times D_t \to \mathbb{R}$. A motion field $\vec{v}$ is a 2D vector valued function $D_s \times D_t^- \to \mathbb{R}^2$, $\vec{v}(x,y,t)$ describing the motion of a pixel at location (x,y) between time t and time $t+1$. The degraded sequence will be denoted u_0, and $\Omega \subset D$ will denote the missing data locus. We will also keep the same notations, with obvious modifications, for the continuous spatio-temporal domain formulation, with this time the assumption that D and Ω are open subsets of $\mathbb{R}^2 \times \mathbb{R}$.

Given the degraded sequence u_0, we assume that the missing data locus Ω is known. Our goal is to reconstruct a sequence u on D and to compute a motion field $\vec{v}$ for that sequence. We introduce therefore the conditional $p(u,\vec{v}|u_0,\Omega)$. Using Bayes' Theorem, we can write

$$p(u,\vec{v}|u_0,\Omega) \propto p(u_0|u,\vec{v},\Omega)\, p(u,\vec{v}|\Omega). \tag{1}$$

We assume the independence of u_0 and $\vec{v}$ so that the likelihood $p(u_0|u,\vec{v},\Omega) = p(u_0|u,\Omega)$. We also clearly assume the independence of $(u,\vec{v})$ and Ω and the prior is therefore

$$p(u,\vec{v}|\Omega) = p(u,\vec{v}) = p(u|\vec{v})\, p(\vec{v}).$$

We will now decompose this prior. For that, let's imagine the following common situation of a person watching a video on a TV set. He/she pushes the "pause" button of the VCR's remote control. There could be motion blur due to large motion and finite aperture times at image acquisition, but she/he expects to see a meaningful still image displayed on the TV screen. When the "play" button is pushed again, the animation resumes, and our person expects to see an animation which is coherent with the content of the still, at least for a sufficiently small amount of time, this coherence corresponding of course to the *apparent motion* of the sequence. This leads us to assume that $p(u|\vec{v})$ has the form $p(u_s,u_t|\vec{v})$ where u_s and u_t denote local spatial (still frame) and temporal (animation) distributions for u, and we factor it as

$$p(u_s,u_t|\vec{v}) = p(u_t|u_s,\vec{v})\, p(u_s|\vec{v}). \tag{2}$$

As we see, the first term of this factorization is a direct translation of this expected temporal coherence. For sake of simplicity, we assume the independence of u_s and $\vec{v}$. This is not necessarily true, as motion edges and image edges have a tendency to be correlated, a fact exploited by several motion recovery algorithms, starting with the work of Nagel and Enkelmann (see for instance [9, 1]). In the other hand, many optical flow algorithms do not use this potential dependency and provide nevertheless very accurate motion estimations. Putting all these elements together, we finally obtain the following posterior distribution

$$p(u,\vec{v}|u_0,\Omega) \propto \underbrace{p(u_0|u,\Omega)}_{P_1}\,\underbrace{p(u_s)}_{P_2}\,\underbrace{p(u_t|u_s,\vec{v})}_{P_3}\,\underbrace{p(\vec{v})}_{P_4} \tag{3}$$

where (P_1) is the likelihood of u, (P_2) is the spatial prior for the sequence, (P_4) is the motion prior, and (P_3) is a coupling term that acts both as a temporal prior for the sequence and a likelihood for the motion – as is the gray level constancy assumption along apparent motion trajectories seen from either image or motion point of view. In the sequel, we choose the simple likelihood

$$P(u_0|u) = \begin{cases} 1 & \text{if } u = u_0 \text{ on } D\backslash\Omega \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Then we seek for the MAP (Maximum A Posteriori) of expression (3). In term of energy, taking $E_i = -\log(P_i)$, $i = 2,3,4$, we simply seek for a $(\mathbf{u},\vec{\mathbf{v}})$ satisfying

$$\begin{cases} (\mathbf{u},\vec{\mathbf{v}}) & = \underset{(u,\vec{v})}{\mathrm{Argmin}}\, E(u,\vec{v}) = E_2(u_s) + E_3(u_s,u_t,\vec{v}) + E_4(\vec{v}) \\ u = u_0 & \text{on } D\backslash\Omega. \end{cases} \tag{5}$$

From now, using the rationale [8], limiting expressions for the probability distribution involved give rise to a *continuous* energy formulation. The "game" is therefore to choose meaningful expressions for each of the E_i. If we replace the likelihood (4) by the assumption that outside Ω, u_0 is of the form $u+\eta$ for some Gaussian white noise η, so that $E_1 = -\log(P_1)$ is well defined, the reader can check that the following energy, proposed by Chanas et al. ([5], [6]), falls clearly in our framework:

$$E(u,\vec{v}) = \underbrace{\int_D \chi(u-u_0)^2 d\mu}_{E_1} + \underbrace{\lambda_1 \int_D \varphi_1(|\nabla u|^2) d\mu}_{E_2} + \underbrace{\lambda_2 \int_D \varphi_2(|\vec{v}\cdot\nabla u + u_t|^2) d\mu}_{E_3} + \underbrace{\lambda_3 \int_D \left(\varphi_3(|\nabla v_1|^2) + \varphi_3(|\nabla v_2|^2)\right) d\mu}_{E_4} \tag{6}$$

where μ is the Lebesgue measure on D, χ is the characteristic function of the set $D\backslash\Omega$, where the data is assumed to be known, ∇ is the *spatial* gradient operator, v_1 and v_2 are the x and y components of the motion vector field, $\varphi_i(x^2) = x^2$ or $\varphi_i(x^2) = \sqrt{x^2+\varepsilon^2}$ and the $\lambda_i > 0$ are some constant. The sequence term $\mathbf{u}$ of a minimizer for this energy will therefore be an inpainted version of u_0 inside Ω and a denoised version of u_0 outside.

3 The Proposed Energy

Brox, Brühn, Papenberg and Weickert have recently introduced a new variational algorithm for recovering optical flow in a sequence in [2]. Denoting by $\vec{V}$ the vector field

$(\vec{v},1)^T$, $\mathbf{p} = (x,y,t)$ a spatio-temporal location and $f^{\vec{V}}$ the shifted argument function $\mathbf{p} \mapsto f(\mathbf{p}+\vec{V})$, the energy they seek to minimize is the following

$$E(\vec{v}) = \int_D \varphi\left((u^{\vec{V}} - u)^2 + \gamma|\nabla u^{\vec{V}} - \nabla u|^2\right) d\mu + \alpha \int_D \varphi\left(|D\vec{V}|^2\right) d\mu \tag{7}$$

where $\varphi(s^2) = \sqrt{s^2 + 0.001^2}$ and $D\vec{V}$ is the spatio-temporal differential of $\vec{V}$. Here not only the gray level constancy assumption is enforced, but also the spatial gradient constancy. Multiresolution settings and a careful minimization of (7) provides one of the most accurate optical flow algorithms. For motion compensated inpainting, and especially for fine details reconstruction, the accuracy of the flow is extremely important. It seems therefore appealing to use this term to build our energy. Before doing so, we note that terms of the form $f^{\vec{V}} - f$ can be interpreted as *forward differences* semi-discretizations of $\frac{\partial f}{\partial \vec{V}}$. Using this remark we replace (7) by

$$E(\vec{v}) = \int_D \varphi\left(\left(\frac{\partial u}{\partial \vec{V}}\right)^2 + \gamma\left|\frac{\partial \nabla u}{\partial \vec{V}}\right|^2\right) d\mu + \alpha \int_D \varphi\left(|D\vec{V}|^2\right) d\mu. \tag{8}$$

We then propose a variational formulation, performing inpainting and optical flow recovery which is of the form (5). In this formulation, the likelihood term is the one given by (4), the spatial term is the one from (6) and we use the above mentioned optical flow energy. The problem is thus to find $(\mathbf{u},\vec{\mathbf{v}})$ minimizing the inpainting energy E_I

$$\begin{cases} E_I(u,\vec{v}) & = \underbrace{\frac{\lambda_1}{2}\int_D \varphi(|\nabla u|^2) d\mu}_{E_2} + \underbrace{\frac{\lambda_2}{2}\int_D \varphi\left(\left(\frac{\partial u}{\partial \vec{V}}\right)^2 + \gamma\left|\frac{\partial \nabla u}{\partial \vec{V}}\right|^2\right) d\mu}_{E_3} + \underbrace{\frac{\lambda_2}{2}\alpha\int_D \varphi\left(|D\vec{v}|^2\right) d\mu}_{E_4} \\ u = u_0 & \text{on } D\backslash\Omega \end{cases} \tag{9}$$

where φ is the function used by Brox et al. Using the Calculus of Variations, such a minimizer should satisfy the following partial differential equation <u>in Ω</u>

$$\frac{\partial E_I}{\partial u}(\mathbf{u},\vec{\mathbf{v}}) = 0 \tag{10}$$

and the following equation <u>in D</u>

$$\frac{\partial E_I}{\partial \vec{v}}(\mathbf{u},\vec{\mathbf{v}}) = 0 \tag{11}$$

We will not detail the derivative in (11) since it has been done in [2]. Using the calculus of variations, we obtain

$$\frac{\partial E_I(u,\vec{v})}{\partial u} = -\lambda_1 \operatorname{div}(A\nabla u) - \lambda_2 \operatorname{div}\left(B\frac{\partial u}{\partial \vec{V}}\vec{V}\right) + \lambda_2\gamma \operatorname{div}\begin{pmatrix} \operatorname{div}(B\frac{\partial u_x}{\partial \vec{V}}\vec{V}) \\ \operatorname{div}(B\frac{\partial u_y}{\partial \vec{V}}\vec{V}) \end{pmatrix}. \tag{12}$$

where $A = \varphi'(|\nabla u|^2)$ and $B = \varphi'\left(\left(\frac{\partial u}{\partial \vec{V}}\right)^2 + \gamma\left|\frac{\partial \nabla u}{\partial \vec{V}}\right|^2\right)$. Although the resulting equation is apparently complex, some simple comments on its right hand terms may help provide a better understanding of it.

- The first term of the right-hand-side of (12) corresponds to a non-linear purely spatial Total Variation (TV) diffusion, which attempts to reconstruct spatially coherent structures.

- The second term can be rewritten, omitting λ_2, as

$$-\operatorname{div}\left(B\frac{\partial u}{\partial \vec{V}}\vec{V}\right) = -\frac{\partial}{\partial \vec{V}}\left(B\frac{\partial u}{\partial \vec{V}}\right) - B\frac{\partial u}{\partial \vec{V}}\operatorname{div}(\vec{v}).$$

  The first part of the right-hand side is essentially what we expect for a non-linear diffusion along the flow lines. The second part accounts for the flow divergence (intensity creation or disparition).

- The last term is the image dual of the gradient constancy constraint along flow lines in (7). This term attempts to enforce this constraint on the inpainted image, enhancing local structures, so that they match along the flow lines, and in that sense, it "diffuses" spatial gradients temporally, allowing *better reconstruction of small details*. It can be rewritten, omitting $\lambda_2\gamma$, as

$$\operatorname{div}\begin{pmatrix}\operatorname{div}(B\frac{\partial u_x}{\partial \vec{V}}\vec{V})\\ \\ \operatorname{div}(B\frac{\partial u_y}{\partial \vec{V}}\vec{V})\end{pmatrix} = \operatorname{div}\left(\frac{\partial}{\partial \vec{V}}\left(B\frac{\partial \nabla u}{\partial \vec{V}}\right)\right) + \operatorname{div}\left(B\frac{\partial \nabla u}{\partial \vec{V}}\operatorname{div}(\vec{v})\right).$$

  making more apparent the diffusion of the gradient along the flow line. Here also a correction factor accounting for the flow divergence appears.

4 Resolution Algorithm

We present here an algorithm that aims at converging to a reasonable minimizer by solving iteratively for the image and the flow, in a *spatial* multiresolution setting. A main difference with the algorithm proposed by Chanas *et al.* in [5, 6] is that not only the optical flow is computed using multiresolution, but also the inpainting. We implicitly build pyramids for the image sequence and the flow. In the sequel we assume that we have $N+1$ resolution levels, level N being the coarsest, level 0 being the finest, original one. We do not require a $\frac{1}{2}$ spatial scaling factor between consecutive levels.

At coarsest resolution we inpaint the sequence by computing a minimizer of (6) with assumed zero-motion, i.e. that the flow lines are orthogonal to the spatial domain: u^N is computed as the solution of the following PDE

$$-\operatorname{div}\left(\varphi'(|\nabla u|^2)\right) - \lambda\partial_t\left(\varphi'(u_t^2)u_t\right) = 0, \quad u_{|_{D^N\setminus\Omega^N}} = u_0^N. \tag{13}$$

where D^N (resp. Ω^N, resp. u_0^N) is the sequence spatio-temporal domain (resp. degradation locus, resp. original data) at resolution level N. The spatial divergence term of this equation is the Total Variation (TV) inpainting of Chan and Shen ([3]), while the term $\partial_t\left(\varphi'(u_t^2)u_t\right)$ provides a simple form of motion adaptivity. If u_t is small, then our assumption of zero-motion is fulfilled, and we can diffuse pixel values along the time direction. If it is large, temporal diffusion is slowed down, and more weight is given to the spatial inpainting part.

Once we have a coarsest resolution inpainted sequence, we can start our multiresolution process. What we do is to modify the optical flow algorithm by running the inpainter after the interpolation of the lower resolution flow, and before updating the flow at that resolution level. Figure 1 illustrates a typical iteration and table 1 describes the sketchy resolution algorithm. At level k, the flow $\vec{v}^k$ is computed as a minimizer of (7), with only *one* resolution level, as an update of the intermediary flow field $\vec{v}_i^k$. Since the inpainted image was reconstructed using $\vec{v}_i^k$, this flow must close enough of a “good” minimizer of (7) at that resolution. At that level the inpainted sequence u^k is obtained by gradient descent of the corresponding Euler-Lagrange equation

$$\partial_\tau u^k = -\frac{\partial E_I}{\partial u}(u^k, \vec{v}_i^k), \quad u^k(-,-,\tau)_{|_{D^k \setminus \Omega^k}} = u_0^k. \tag{14}$$

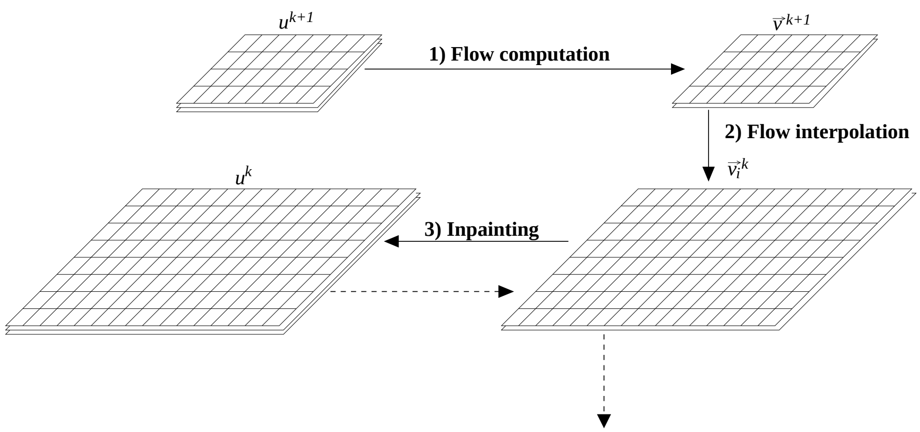

Figure 1: Multiresolution inpainting. From a sequence u^{k+1} at resolution level $k+1$, 1) the optical flow is computed, 2) it is interpolated at resolution level k and 3) used to inpaint the sequence at level k. This process is then iterated at lower pyramid levels.

1. Compute u^N by spatial TV + temporal motion adaptive diffusion.
2. For $k = N-1$ down to 1
   - Compute $\vec{v}^{k+1}$ from u^{k+1}
   - Compute intermediary flow $\vec{v}_i^k$ by interpolation of $\vec{v}^{k+1}$
   - Inpaint u^k using the $\vec{v}_i^k$
3. Compute $\vec{v}^0$ from u^0
4. Output $(u^0, \vec{v}^0)$.

Table 1: Inpainting algorithm

5 Numerics

In this section we briefly describe the numerics used for our inpainting PDE. The discretization of the optical flow equation has been discussed in details in [2] and we use the scheme proposed by Chan and Shen [4] for the discretization of the spatial term of equation (12). At coarsest resolution we use equation (13), which is solved directly, using a fixed point strategy to cope with nonlinearities. Its spatial term is discretized using the scheme mentioned above, the temporal part being also of that form, but simpler, since it is only one-dimensional.

We discuss now in greater details the schemes used for operators involving the optical flow field.

1. The first term to discretize is

$$\operatorname{div}\left(B\frac{\partial u}{\partial \vec{V}}\vec{V}\right) = \frac{\partial}{\partial \vec{V}}\left(B\frac{\partial u}{\partial \vec{V}}\right) + B\frac{\partial u}{\partial \vec{V}}\operatorname{div}(\vec{v}).$$

The left hand-side of this equation is not easily handled due to the apparition of cross-derivatives. We have therefore chosen to discretize the right-hand side. We describe first a scheme for terms of the form $\frac{\partial}{\partial \vec{V}}\left(B\frac{\partial f}{\partial \vec{V}}\right)$. Given a spatio-temporal location $\mathbf{p} = (x,y,t)$, $\mathbf{p}^+$ will denote the point *following* $\mathbf{p}$ along the flow line at $\mathbf{p}$, i.e. $\mathbf{p}^+ = \mathbf{p} + \vec{V}(\mathbf{p})$, while $\mathbf{p}^-$ will denote the point *preceding* $\mathbf{p}$ along the flow line at $\mathbf{p}$, i.e. $\mathbf{p} = \mathbf{p}^- + \vec{V}(\mathbf{p}^-)$, and notice that we have $(\mathbf{p}^+)^- = (\mathbf{p}^-)^+ = \mathbf{p}$. Then we introduce two finite difference operators

$$D^+_{\vec{V}} f(\mathbf{p}) = f(\mathbf{p}^+) - f(\mathbf{p}), \quad D^-_{\vec{V}} f(\mathbf{p}) = f(\mathbf{p}) - f(\mathbf{p}^-)$$

(when $\mathbf{p}^\pm$ is not on the image natural grid, the quantity $f(\mathbf{p}^\pm)$ is computed by spatial bilinear interpolation – its temporal coordinate is always on the temporal grid). Recalling that B is defined as

$$B = \varphi'\left(\left(\frac{\partial u}{\partial \vec{V}}\right)^2 + \gamma\left|\frac{\partial \nabla u}{\partial \vec{V}}\right|^2\right) \approx \varphi'\left(\left(u(\mathbf{p}^+) - u(\mathbf{p})\right)^2 + \gamma\left|\nabla u(\mathbf{p}^+) - \nabla u(\mathbf{p})\right|^2\right) = B(\mathbf{p},\mathbf{p}^+)$$

(observe that we are back to the formulation (7) used by Brox et. al.) the proposed discretization of $\frac{\partial}{\partial \vec{V}}\left(B\frac{\partial f}{\partial \vec{V}}\right)$ is

$$\begin{aligned} D^-_{\vec{V}}\left(B(\mathbf{p},\mathbf{p}^+) D^+_{\vec{V}} f(\mathbf{p})\right) = D^+_{\vec{V}}\left(B(\mathbf{p}^-,\mathbf{p}) D^-_{\vec{V}} f(\mathbf{p})\right) = \\ B(\mathbf{p},\mathbf{p}^+) f(\mathbf{p}^+) - \left(B(\mathbf{p},\mathbf{p}^+) + B(\mathbf{p}^-,\mathbf{p})\right) f(\mathbf{p}) + B(\mathbf{p}^-,\mathbf{p}) f(\mathbf{p}^-). \end{aligned} \tag{15}$$

2. The location $\mathbf{p}^+$ is directly available from the output of the flow algorithm, but the location $\mathbf{p}^-$ is not. This information is difficult to obtain, if possible at all (at occlusion/desocclusion, and other types of singularities of the flow field). In order to overcome this problem, we compute also the *backward in time* optical flow $\vec{v}_-$ from $u(-,t)$ to $u(-,t-1)$ and we estimate $\mathbf{p}_-$ as $\mathbf{p} + \vec{v}_-(\mathbf{p})$. From now we adopt the notation $\vec{v}_+$ for the forward-in-time flow and $\vec{v}_-$ for the backward-in-time flow. The corresponding B functions will be denoted B_+ and B_- respectively. We obtain therefore

$$\frac{\partial}{\partial \vec{V}}\left(B\frac{\partial f}{\partial \vec{V}}\right)(\mathbf{p}) \approx B_+(\mathbf{p}) f(\mathbf{p}^+) - \left(B_+(\mathbf{p}) + B_-(\mathbf{p})\right) f(\mathbf{p}) + B_-(\mathbf{p}) f(\mathbf{p}^-). \tag{16}$$

3. The term $B\frac{\partial u}{\partial \vec{V}}\,\mathrm{div}(\vec{v})$ must be handled carefully. A naive discretization of the form

$$\frac{1}{2}B_+(u(\mathbf{p}^+) - u(\mathbf{p}))\,\mathrm{div}(\vec{v}_+) - \frac{1}{2}B_-(u(\mathbf{p}) - u(\mathbf{p}^-))\,\mathrm{div}(\vec{v}_-)$$

where $\mathrm{div}(\vec{v}_\pm)$ is discretized by standard central differences, turned to be unstable (the "-" sign in the second term of the above formula comes from the fact that $\vec{v}_-$ is *backward in time*). This is not totally unexpected since it is a sort of "transport" term. We favor the best available information, i.e. the one for which the difference between the information at $\mathbf{p}$ and its two flow lines neighbors is minimal. This information is quantified by the function B. Indeed, $B_\pm$ provides an explicit measure of the match between the information at location $\mathbf{p}$ and $\mathbf{p}^\pm$, not only gray level variation, but also gradient variation, a low value of $B_\pm$ corresponding to a poor match, while a large value correspond to a good one[1]. We use either $\mathbf{p}^+$ or $\mathbf{p}^-$ according to the value of $B_\pm$:

$$B\frac{\partial u}{\partial \vec{V}}\,\mathrm{div}(\vec{v})(\mathbf{p}) \approx \begin{cases} B_+ D^+_{\vec{V}} u(\mathbf{p})\,\mathrm{div}(\vec{v}_+), & \text{if } B_+ > B_- \\ -B_- D^-_{\vec{V}} u(\mathbf{p})\,\mathrm{div}(\vec{v}_-), & \text{otherwise.} \end{cases} \tag{17}$$

4. Discretization of the higher order terms. The first one we treat is

$$\mathrm{div}\left(\frac{\partial}{\partial \vec{V}}\left(B\frac{\partial \nabla u}{\partial \vec{V}}\right)\right) = \partial_x \frac{\partial}{\partial \vec{V}}\left(B\frac{\partial u_x}{\partial \vec{V}}\right) + \partial_y \frac{\partial}{\partial \vec{V}}\left(B\frac{\partial u_y}{\partial \vec{V}}\right).$$

Standard central differences are used for the spatial derivatives, while the resulting flow line terms are discretized according to (16). Using an half-grid point scheme, as it is often the case for divergence operators, involves a large amount of interpolations and produces experimentally less accurate results. Central differences are also used for the third order term spatial divergence, using this time (17) for the terms $B\frac{\partial u_z}{\partial \vec{V}}\,\mathrm{div}(\vec{v})$, $z = x, y$.

6 Experimental Evaluation

We present here results on three sequences. The first one is the well known `Yosemite` sequence, very often used for optical flow evaluation. The sequence is artificial and the ground truth is known for the flow. It was degraded by removing large polygonal patches, three of them *overlapping consecutively* in time, on 6 of the 15 frames. The entire sequence has been treated in one run. Figure 2 shows frame 3 of the original sequence, the corresponding degraded frame and the reconstructed one. The second and third examples we present are taken from the companion CD ROM of Kokaram's book [7]. This second sequence, called `mobcal`, is a real one with 25 frames and with complex motion patterns. It is then artificially degraded to simulate blotches. (approximatively 1.22% of the image is degraded with blocthes of multiple size, which do not overlap in time). We show in figure 3 one frame from the original sequence, the corresponding degraded frame as well as the inpainted one. Then in figure 4 we look at two particular blotches and their reconstruction, the first one is almost perfect, while the second one is of a bit lesser quality. The problem for that one seems to come from an incorrect motion recovery.

[1] It can be used for occlusion detection, and probably for blotch detection. But as is, it lacks robustness for that purpose, as we constated it experimentally.

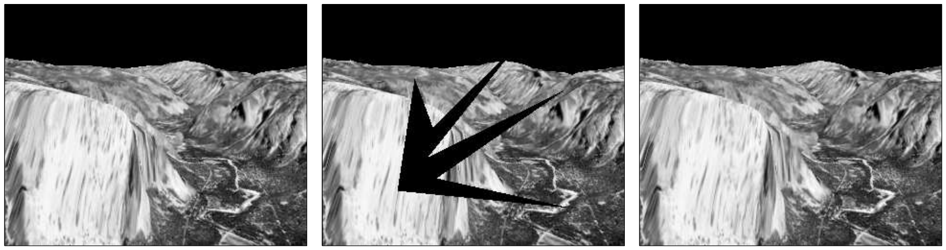

Figure 2: `Yosemite` sequence. (a) Left: Original frame 3. (b) Middle: corresponding degraded frame. (c) Right: restored frame: at its center, the restoration is slightly unsharp, but 3 consecutive blotches overlap here in the degraded sequence.

Figure 3: `Mobcal` sequence. (a) Left: Original frame 9. (b) Middle: frame with artificial blotches. (c) Right: inpainted frame.

The third sequence, called `Frankenstein`, is a real degraded one, with 64 frames, for which no ground truth is known. In our experimentations we used only a subsequence of 5 frames, with frame 3 presenting a blotch on Frankenstein's hair. In figure 5 we show this frame and its reconstruction. In figure 6 we show a close-up of the damaged hair of the character with the detected blotch and the reconstruction. Blotches were detected using the so called rank order detector, as described in [7], modified for the optcal flow algorithm. Fine texture details were very plausibly recreated.

7 Conclusion and Future Work

In this work we have introduced a new variational algorithm for blotch removal. Based on a Bayesian framework for joint image reconstruction and motion interpolation, and using a recent algorithm for optical flow recovery, we have proposed a novel energy formulation and derived a fully multi-scale algorithm for our energy minimization. We have described the discretization and showed the capabilities of our algorithm on some examples. We are currently working on an inpainting-denoising formulation as well as extensions for color and more generally vector valued image sequences. A longer term goal is to incorporate an automatic detection of the corrupted areas.

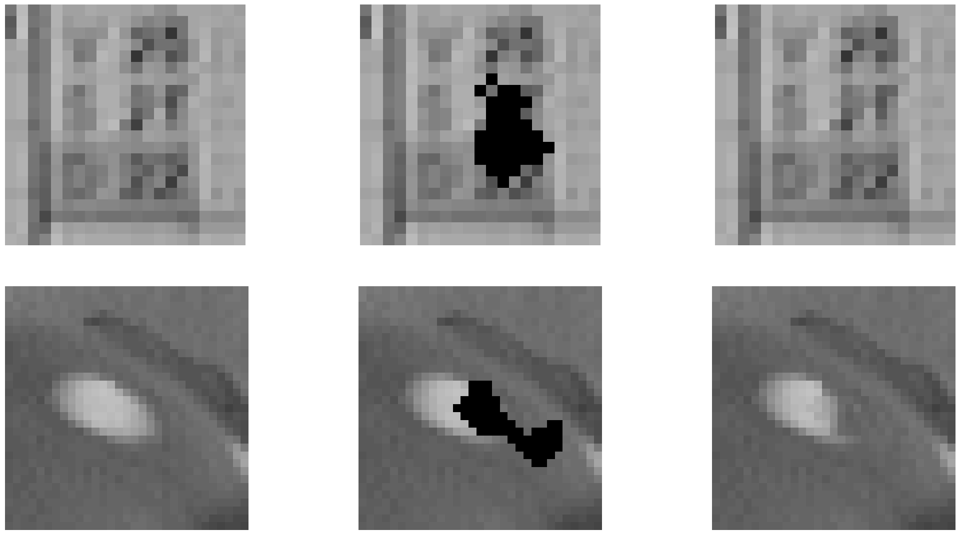

Figure 4: `mobcal` sequence, details of frame 9: (a) Top left: close-up at a digit inside the calendar. (b) Top center: blotch on that digit. (c) Top right: almost perfect reconstruction. (d) Bottom left: close-up at a white patch on the ball. (e) Bottom center: blotch. (f) Bottom right: partially failed reconstruction.

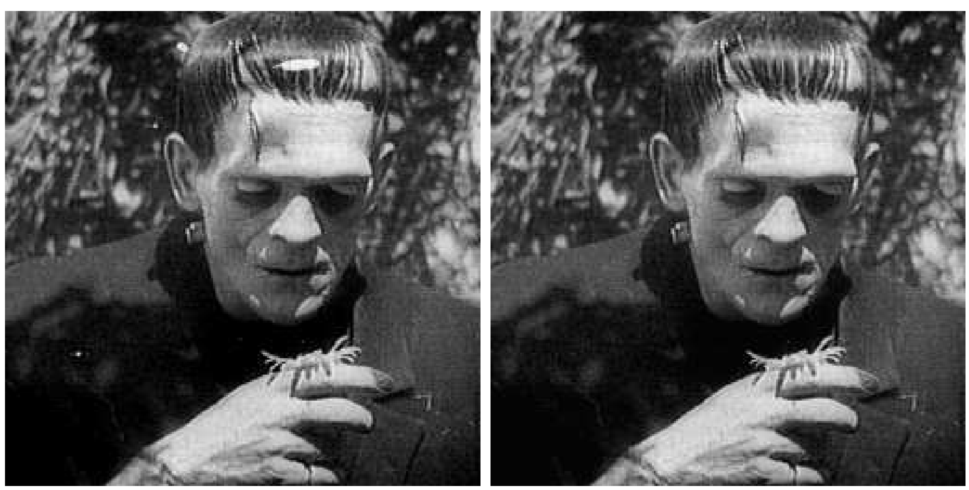

Figure 5: `Frankenstein` sequence. (a) Left: original damaged frame – observe the blotch on the hair and at the left of the head. (b) Right. Reconstructed frame by our algorithm.

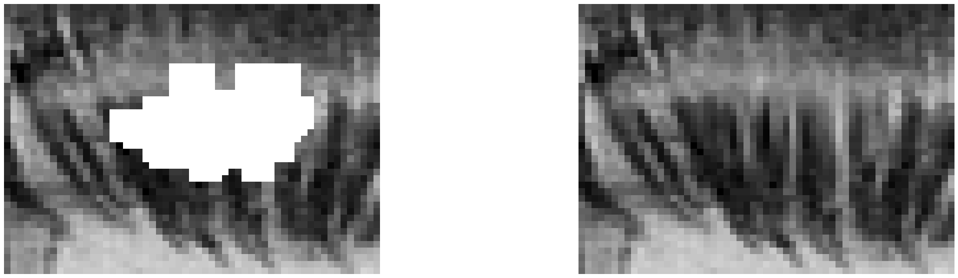

Figure 6: `Frankenstein` sequence, Hair close-up. Left: the detected blotch, which is larger that the actual degradation. Right: reconstruction, the hair texture is nicely recreated.

References

[1] L. Alvarez, R. Deriche, J. Weickert, and J. Sànchez. Dense disparity map estimation respecting image discontinuities: A PDE and scale-space based approach. *International Journal of Visual Communication and Image Representation, Special Issue on Partial Differential Equations in Image Processing, Computer Vision and Computer Graphics*, 2000.

[2] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In T. Pajdla and J. Matas, editors, *Proceedings of the 8th European Conference on Computer Vision*, Prague, Czech Republic, 2004. Springer–Verlag.

[3] T. Chan and J. Shen. Mathematical models for local deterministic inpaintings. Technical Report 00-11, Department of Mathematics, UCLA, Los Angeles, March 2000.

[4] T.F. Chan and J. Shen. Non-texture inpainting by curvature-driven diffusions (cdd). *J. Visual Comm. Image Rep.*, 12(4):436–449, 2001.

[5] Laurent Chanas. *Méthodes variationnelles pour la restauration de séquences d'images fortemenent dégradées. Application aux images infrarouges éblouies par laser*. PhD thesis, Université de Cergy-Pontoise, October 2001.

[6] J.P. Cocquerez, L. Chanas, and J. Blanc-Talon. Simultaneous inpainting and motion estimation of highly degraded video-sequences. In *Scandinavian Conference on Image Analysis*, pages 523–530. Springer-Verlag, 2003. LNCS, 2749.

[7] A. Kokaram. *Motion Picture Restoration: Digital Algorithms for Artefact Suppression in Degraded Motion Picture Film and Video*. Springer, 1998.

[8] D. Mumford. Bayesian rationale for the variational formulation. In *Geometry-Driven Diffusion In Computer Vision*, pages 135–146. Kluwer Academic Publishers, 1994. Computational Imaging And Vision.

[9] H.H. Nagel and W. Enkelmann. An investigation of smoothness constraint for the estimation of displacement vector fields from image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:565–593, 1986.