POLYCONVEXIFICATION OF THE MULTI-LABEL OPTICAL FLOW PROBLEM

N. Papadakis, A. Baeza, P. Gargallo

Image Group - Barcelona Media
Avenida Diagonal 177
08017 Barcelona, Spain

V. Caselles

DTIC - Universitat Pompeu-Fabra
C/Roc Boronat 138
08018, Barcelona, Spain

ABSTRACT

In this paper the problem of optical flow and occlusion mask estimation is aborded. To that end, we consider a multi-label representation of the optical flow and we define an energy that models the problem. The convexification of the energy and its minimization with an iterative algorithm are studied. Our algorithm is implemented in GPU, since each pixel can be processed in parallel. From our experiments, the relation between the quality of the results obtained and computing time seems to be very promising.

Index Terms— optical flow, occlusion, total variation

1. INTRODUCTION

In the literature, the problem of optic flow is classically based on the well-known Horn and Schunck model [1]. This approach consists in derivating the continuous luminance conservation equation and leads to fast, as they can be parallelized, and high quality results [2, 3]. These methods assume that the mass of luminance is advected in time by the optical flow and are completed with a spatial regularization of the vector field. Nevertheless, in the boundaries of motion layers, this last assumption is not valid. Moreover, the basic resolution induced by these approaches is limited to the search of small displacements. To be able to recover fast motions, a multi-resolution strategy should be used [4]. Pyramids representing the different scales of the images are then created and the motion is recursively computed from the upper level to the lower one. At each step, the problem is linearized around the projection of the upper solution. In this case, the large displacements of small structures cannot be estimated, since the small objects disappear at the upper level of the pyramids.

Another possibility to tackle the optical flow problem is to consider a discrete set of possible values for the optical flow. This representation leads to a multi-label problem: for each pixel of the image, one label (corresponding to one discrete optical flow value of the predefined set) has to be estimated. In this case, representations based on Markov Random Fields and optimized through Graph-Cuts, Dynamic Programming or Believe Propagation have been proposed [5, 6, 7]. The discrete representation of the optical flow allows using the luminance conservation assumption with a non derived formulation. The model is then valid in all the non occluded areas of the image. The estimation of large displacement of small object is also technically possible. Nevertheless, such kind of approach can be very slow when the number of pixels and the number of possible labels become too high.

We here propose to rely on the recent work of Pock et al. [8], where the multi-label disparity estimation problem is formulated in a convex way. With this formulation, the minimization of the associated energy can be done in parallel, so that very fast algorithms can be obtained. We therefore propose to extend this approach to the vectorial case and tackle the problem of multi-label optical flow. We also adapt the formulation in order to include occlusion mask between the two images. The formulation is variational, so that the energy is minimized through an iterative process to solve the Euler-Lagrange equations associated to the problem. More precisely, the Euler-Lagrange equation is a non linear equation of partial derivatives that can be expressed in terms of local interactions between pixels when discretized. For that reason, this approach allows using GPU implementation, since at each step, each pixel can be processed in parallel.

Let us now briefly describe the organization of the paper. In section 2, basic variational models and the associated methods of convexification are described. In section 3, the application of such a theory to the computation of optical flow and occlusion mask is described. Some results are finally given in section 4.

2. VARIATIONAL FORMULATION BASED ON ENERGY CONVEXIFICATION

We here focus on multi-label problems for image processing. Let Ω be the image domain. We aim at estimating a variable $u : \Omega \mapsto \Gamma$ which takes its values in a predefined discrete set containing N ordered elements: $\Gamma = \{a = u_0 < \cdots < u_{N-1} = b\}$. This section is then dedicated to the minimization of the following class of functionals:

$$J(u) = \int_\Omega |\nabla u(x)| dx + \int_\Omega \rho(x, u(x)) dx, \tag{1}$$

where ρ is a given positive data function and $\rho(x, u(x))$ represents the cost of assigning the value $u(x)$ to the pixel x.

We only assume that ρ is a bounded function that can be non linear with respect to u. The first term $\int_\Omega |\nabla u|$ is a spatial regularization of the unknown on the image domain. More precisely, it measures the integral of the perimeters of the level sets of u (see [9] for more details). Such a term, introduced by [10], is known as the total variation of u.

This regularization is general and has been applied to a lot of image processing problems as restoration [10], depth estimation [8], 3D reconstruction [11] or optical flow [12].

2.1. Convexification of the multi-label problem

We now briefly recall the convexification technique of Pock et al. [8]. The idea is to write the non-linearities of the functional $J(u)$ in a convex way, by introducing an auxiliary variable $\phi : \Omega \times \Gamma \mapsto \{0, 1\}$ that represents the different values of u. Let

$$\phi(x, s) = H(u(x) - s), \tag{2}$$

where H is the Heaviside function ($H(r) = 1$ if $r \geq 0$, and 0 otherwise). The unknown u can then be recovered from ϕ as

$$u(x) = u_0 + \int_\Gamma \phi(x, s)ds. \tag{3}$$

We can now rewrite the functional (1) as a function of ϕ instead of u.

$$J(\phi) = \int_\Omega \int_\Gamma |\nabla_x \phi(x, s)|dsdx + \int_\Omega \int_\Gamma \rho(x, s)|\phi_s(x, s)|dsdx. \tag{4}$$

In effect, by using the partial derivatives of ϕ,

$$\nabla_x \phi(x, s) = \delta(u(x)-s)\nabla u(x), \quad \phi_s(x, s) = -\delta(u(x)-s), \tag{5}$$

one can check that

$$\begin{aligned} \int_\Omega \int_\Gamma |\nabla_x \phi(x, s)|dsdx &= \int_\Omega |\nabla u(x)|dx, \\ \int_\Omega \int_\Gamma \rho(x, s)|\phi_s(x, s)|dsdx &= \int_\Omega \rho(x, u(x))dx. \end{aligned} \tag{6}$$

The functional (4) is convex in ϕ. To find the global minimum of (1), one can find a ϕ minimizing (4), and then recover u from ϕ. Care must be taken to ensure that it is possible to compute u from ϕ. In particular, ϕ must be binary and decreasing with s (i.e. $\phi_s < 0$). Unfortunately, the set of such functions ϕ is not convex.

To recover convexity, the function ϕ should be allowed to take values on $[0, 1]$. We introduce a convex set of admissible functions,

$$\mathcal{A} = \{\phi : 0 \leq \phi \leq 1,\ \phi(x, a) = 1,\ \phi(x, b) = 0,\ \phi_s \leq 0\}, \tag{7}$$

and the convex problem

$$\min_{\phi \in \mathcal{A}} J(\phi). \tag{8}$$

Let ϕ^* be its solution. Following [8], we have that any level set of ϕ^* is also a solution because $J(\mathbf{1}_{\phi^* \geq \alpha}) \leq J(\phi^*)$, for all $\alpha \in (0, 1]$. Therefore, a global mimimum of (1) can be recovered as a "cut" of ϕ. Given a threshold $\alpha \in (0, 1]$,

$$u(x) = \max\{s \in \Gamma : \phi^*(x, s) \geq \alpha\}. \tag{9}$$

is a global minimum of (1).

2.2. Numerical optimization

The previous minimization problem (8) can be written as a primal-dual problem:

$$\min_{\phi \in \mathcal{A}} \max_{Z \in \mathcal{C}} \int_\Omega \int_\Gamma \nabla_{x,s}\phi(x, s) \cdot Zdsdx. \tag{10}$$

where $\mathcal{C}$ is the set of vectorial fields $Z = (z_1, z_2, z_3)$, with z_i, $i = 1, 2, 3$ defined on $\Omega \times \Gamma$ such that

$$z_1^2(x, s) + z_2^2(x, s) \leq 1 \text{ and } |z_3(x, s)| \leq \rho(x, s), \forall (x, s) \in \Omega \times \Gamma \tag{11}$$

More details can be found in [8]. The basic ideal is to consider an alternate optimization using a proximal point algorithm [13, 14, 8]. A gradient ascent strategy is used on Z, while u is updated with a gradient descent method. At each step the variables are projected onto the convex sets $\mathcal{C}$ and $\mathcal{A}$. Derivating problem (10) with respect to ϕ or Z leads to simple expressions that can be considered independently for each pixel and implemented in parallel.

3. A VARIATIONAL MODEL FOR OPTICAL FLOW AND OCCLUSION MASK ESTIMATION

We aim at estimating the optical flow between two consecutive images I_1 and I_2. More precisely, we seek for the optical flow $w(x) = (u(x), v(x))$ going from image I_1 to image I_2 by minimizing the following energy:

$$J(u, v) = \int_\Omega |\nabla u(x)|dx + \int_\Omega |\nabla v(x)|dx + \frac{1}{\lambda} \int_\Omega \rho(x, u(x), v(x)) \tag{12}$$

We assume that $u(x)$ (resp. $v(x)$) takes its values in the interval $[a, b]$ (resp. $[c, d]$). The field $w(x) = (u(x), v(x))$ then represents the displacement field of pixels x of image I_1. The cost function ρ can then be defined as:

$$\rho(x, u(x), v(x)) = \frac{1}{\lambda}|I_1(x) - I_2(x + w(x))|. \tag{13}$$

Up to now, we did not include any occlusion mask. Observe also that the conservation of the luminance hypothesis is included in a non derived way. In optical flow literature, a linearized version is classically considered $\rho(x, w) = (I_t + w \cdot \nabla I)^2)$, to simplify the resolution. This kind of linearization is nevertheless only valid if the image gradient ∇I is smooth enough and is not able to deal properly with occlusions. Our data term is then more general but non linear and non convex with respect to w.

Following the calculus of section 2, we now write the non linearities of problem (12) in a convex way, by introducing additional variables that allow using optimization algorithms such as primal-dual ones.

3.1. Polyconvexification of the multi-label optical flow problem

Let $\phi(x, s) = H(u(x) - s)$ and $\psi(x, t) = H(v(x) - t)$ be two auxiliary functions, where H still denotes the Heaviside function and $s, t \in \mathbb{R}$. Note that ϕ (resp. ψ) is independent of t (resp. s). The unknown variables u and v can be recovered from ϕ and ψ with the integration (3). The problem can then be considered as a problem defined on binary variables ϕ and ψ. For simplicity, we assume that $b-a=d-c$. Following section 2, the class of admissible functions $\mathcal{A}_F$ is defined as

$$\begin{aligned}\mathcal{A}_F = \{(\phi,\psi) : 0 \le \phi,\psi \le 1, \phi_s \le 0, \psi_t \le 0,\\ \phi(x,a) = \psi(x,c) = 1, \phi(x,b) = \psi(x,d) = 0\}.\end{aligned} \tag{14}$$

Let us now compute the partial derivatives of these functions:

$$\begin{aligned}\nabla_x\phi(x,s) &= \delta(u(x)-s)\nabla u(x), \quad \phi_s(x,s) = -\delta(u(x)-s),\\ \nabla_x\psi(x,t) &= \delta(v(x)-t)\nabla v(x), \quad \psi_t(x,t) = -\delta(v(x)-t).\end{aligned} \tag{15}$$

Following section 2, the regularization terms of energy (12) can be written as

$$\begin{aligned}\int_\Omega |\nabla u(x)|dx &= \frac{1}{d-c}\int_\Omega\int_a^b\int_c^d |\nabla_x\phi(x,s)|dtdsdx,\\ \int_\Omega |\nabla v(x)|dx &= \frac{1}{b-a}\int_\Omega\int_a^b\int_c^d |\nabla_x\psi(x,t)|dtdsdx.\end{aligned} \tag{16}$$

The data term is finally rewritten as

$$\int_\Omega \rho(x,u(x),v(x))dx = \int_\Omega\int_a^b\int_c^d \rho(x,s,t)\phi_s(x,s)\psi_t(x,t)dtdsdx. \tag{17}$$

We end up with the following problem

$$\min_{(\phi,\psi)\in\mathcal{A}_F} \mathcal{F}(\phi,\psi), \tag{18}$$

with

$$\begin{aligned}\mathcal{F}(\phi,\psi) &= \frac{1}{\lambda}\int_\Omega\int_a^b\int_c^d \rho(x,s,t)\phi_s(x,s)\psi_t(x,t)dtdsdx\\ &+\int_\Omega\int_a^b\int_c^d |\nabla_x\phi(x,s)|dtdsdx + \int_\Omega\int_a^b\int_c^d |\nabla_x\psi(x,t)|dtdsdx,\end{aligned} \tag{19}$$

where the factor $b-a=d-c$ has been absorbed by the parameter λ. The energy $\mathcal{F}(\phi,\psi)$ is not convex in variables ϕ and ψ, but polyconvex [15]. The reason is that $\rho \ge 0$ and $\phi_s(x,s)\psi_t(x,t)$ is nothing else than the determinant of the Jacobian matrix in (s,t) as:

$$\begin{pmatrix} \phi_s & \phi_t \\ \psi_s & \psi_t \end{pmatrix} = \begin{pmatrix} \phi_s & 0 \\ 0 & \psi_t \end{pmatrix}.$$

Polyconvex functionals are quasiconvex. Quasiconvexity is the right extension of the notion of convexity for vector valued functions. Under assumptions, it guarantees the existence of minimizers and the well-posedness (in a certain sense) of the energies. Even if the numerical analysis of such functionals is not so well-developed, we use an alternate minimization scheme to solve (18). Assuming that ψ (resp. ϕ) is known, the energy is minimized with respect to ϕ (resp. ψ). In both cases, we can linearize with respect to the optimized variable. For instance, when optimizing with respect to ϕ, we have the problem

$$\min_\phi (d-c)\int_\Omega\int_a^b |\nabla_x\phi(x,s)|dsdx + \frac{1}{\lambda}\int_\Omega\int_a^b \Psi(x,s)\phi_s(x,s)dsdx, \tag{20}$$

where the factor Ψ can be computed as:

$$\Psi(x,s) = \int_c^d \rho(x,s,t)\psi_t(x,t)dtdx, \tag{21}$$

and the numerical optimization method of section 2.2 can be used to minimize (20).

3.2. Occlusion mask

In order to include the estimation of the occlusion mask, we redefine the data term as:

$$\rho(x,u(x),v(x),m(x)) = \frac{m(x)}{\lambda}|I_1(x) - I_2(x+w(x))| + \gamma(1-m(x)). \tag{22}$$

where the binary function $m(x)$ represents the occlusion mask and takes the value 1 when the pixel x is visible in both images and the value 0 otherwise. In this second case, an occlusion cost $\gamma > 0$ is paid. The final functional is obtained by incorporating this term in the energy (12) and adding a regularization term on the occlusion mask $\mu\int_\Omega |\nabla m|$, with $\mu \ge 0$. The optimization of such a functional with respect to m is also done following a primal-dual strategy, by relaxing the binary mask onto the convex interval [0,1] and cutting with respect to a threshold $\alpha \in (0,1]$.

4. EXPERIMENTS

In this section, the method is applied to data available on the Middlebury website (http://vision.middlebury.edu/flow/data). The regularization parameter λ was set to 20, the occlusion parameter γ to 80 and the regularization parameter of the occlusion mask is $\mu = 10$. The method has been implemented in GPU using CUDA. On Figure 1, we show results obtained on the RubberWhale images attesting that the visual quality of the estimated optical flow and mask are fine.

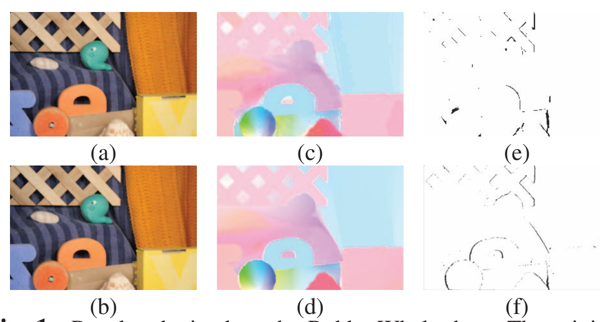

Fig. 1. Results obtained on the RubberWhale data. The original images (a) and (b) are used to compute the optical flow (c) and the occlusion mask (e). The ground truth is given in images (d) and (f).

On table 1, we also give the associated errors and the needed computational time. In this table, the Root Mean Square Error (RMSE) between estimation w and ground truth w_t is given. The mean and the maximal values of the groundtruth motion norm are also given. The error is thus inferior to 10% of the maximum amplitude of the norm. We also give N_u and N_v, the number of possible values for the motion components u and v. For such a rough discretization, the results are quite interesting. The method is fast but not real time, as 10 alternate iterations of ϕ and ψ are needed in practice, and the multiple computations of relation (21) are time consuming. We finally illustrate the application of the process to more data in figure 2.

Images	RMSE $\|w - w_t\|$	Mean $\|w_t\|$	Max $\|w_t\|$	N_u	N_v	Time
RubberWhale	0.36	1.24	4.6	31	31	5s
Dimetrodon	0.22	1.96	4.7	51	51	20s
Grove2	0.46	3.09	5	25	25	15s
Grove3	1.6	3.91	18.6	28	28	17s
Hydrangea	0.46	3.48	11.1	23	23	8s
Urban2	1.3	8.39	221.18	45	30	15s

Table 1. Comparison of the quantitative results on the Middlebury optical flow benchmark datasets. The RMS Error $|w - w_t|$ between estimation w and ground truth w_t is given for the 6 studied examples.

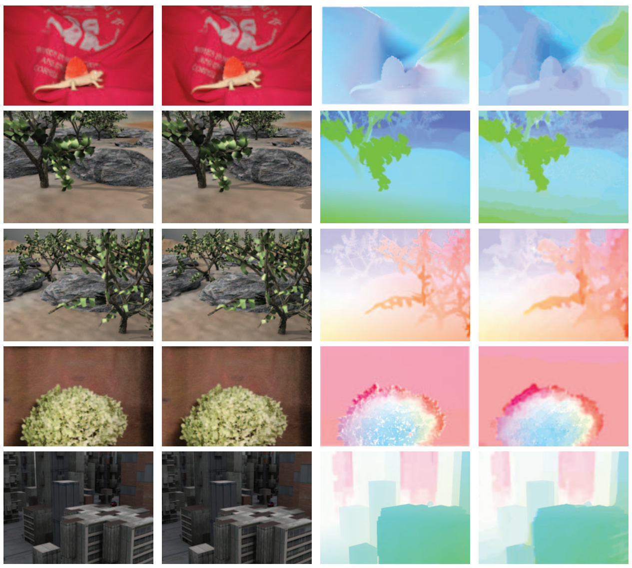

Fig. 2. Results obtained on the Middlebury data Dimetrodon, Grove2, Grove3, Hydrangea and Urban2. The original images are shown in the first and second columns. The true optical flow in the third column is then compared to the estimated one in the forth

5. CONCLUSIONS AND PERSPECTIVES

In this paper, we proposed and implemented a variational model for the estimation of optical flow. The proposed energy can be minimized through standard and simple numerical methods and can be implemented in parallel on the GPU. Up to our knowledge, this is the first variational method that solves the optical flow problem using its polyconvexification. The quality of the results with respect to the required computational time is interesting. As a drawback, the process needs a large number of possible labels to reach better accuracy in the results. As the convexification implies increasing the dimension of the unknown variable, the method is currently limited to rough estimations. Indeed, the more labels, the more memory space is needed to store the auxiliary functions. To circumvent this limitation, either we develop a multi-label refinement algorithm (a kind of multi-resolution strategy on labels), or we use a narrow band type method in order to reduce the computations to a band around the solution as in [16]. This will be the subject of a future work.

Acknowledgements

This work was partially funded by Mediapro through the Spanish project CENIT-2007-1012 i3media and by the Centro para el Desarrollo Tecnológico Industrial (CDTI), within the Ingenio 2010 initiative. V. Caselles and P. Gargallo acknowlege partial support by IP project "2020 3D Media: Spacial Sound and Vision", Financed by EC. P. Gargallo and N. Papadakis acknowledge support from the Torres Quevedo program of the Ministerio de Educación y Ciencia in Spain. V. Caselles also acknowledges partial support by MICINN project, reference MTM2009-08171, by GRC reference 2009 SGR 773 and by "ICREA Acadèmia" prize for excellence in research, both funded by the Generalitat de Catalunya.

6. REFERENCES

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

[2] M. Black and P. Anandan, "The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields," in *IEEE Int. Conf. Comp. Vis (ICCV' 96)*, 1996, vol. 63, pp. 75–104.

[3] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, "High accuracy optical flow estimation based on a theory for warping," in *Europ. Conf. on Comput. Vis. (ECCV'04)*. 2004.

[4] E. Mémin, P. Pérez, and D. Machecourt, "Dense estimation and object-based segmentation of the optical flow with robust techniques," *Technical report INRIA*, no. 2836, 1996.

[5] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. on Pat. Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, 2001.

[6] P.F. Felzenszwalb and D. P. Huttenlocher, "Efficient belief propagation for early vision," in *IEEE Conf. Comp. Vis. Pat. Recogn. (CVPR'04)*, 2004.

[7] B. Glocker, N. Paragios, N. Komodakis, G. Tziritas, and N. Navab, "Optical flow estimation with uncertainties through dynamic mrfs," in *IEEE Conf. Comp. Vis. Pat. Recogn (CVPR'08)*, 2008.

[8] T. Pock, T. Schoenemann, D. Cremers, and H. Bischof, "A convex formulation of continuous multi-label problems," in *Europ. Conf. Comp. Vis. (ECCV'08)*, 2008.

[9] L. Ambrosio, N. Fusco, and D. Pallara, "Functions of Bounded Variation and Free Discontinuity Problems," in *Oxford Math. Monogr., The Claredon Press, Oxford University Press*, 2000.

[10] L.I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, no. 1-4, pp. 259–268, 1992.

[11] K. Kolev, M. Klodt, Thomas T. Brox, and D. Cremers, "Continuous global optimization in multiview 3d reconstruction," *Int. J. Comput. Vision*, vol. 84, no. 1, pp. 80–96, 2009.

[12] C. Zach, T. Pock, and H. Bischof, "A duality based approach for realtime TV-L^1 optical flow," in *Annual Symp. of the German Assoc. for Pat. Recogn. (DAGM 2007)*, 2007, pp. 214–223.

[13] R. T. Rockafellar, "Augmented lagrangians and applications of the proximal point algorithm in convex programming," *Mathematics of operations research*, vol. 1, no. 2, pp. 97–116, 1976.

[14] M. Zhu and T. Chan, "An efficient primal-dual hybrid gradient algorithm for total variation image restoration," *Computational and Applied Math Report 08-34, UCLA*, 2008.

[15] J.M. Ball, "Convexity conditions and existence theorems in nonlinear elasticity," *Archive Rational Mechanics Analysis*, vol. 63, pp. 337–403, 1977.

[16] A. Baeza, V. Caselles, P. Gargallo, and N. Papadakis, "A narrow band method for the convex formulation of discrete multi-label problems," Preprint, available at `http://www.dtic.upf.edu/~vcaselles/BCGP_MMS09.pdf`.