Image and Vision Computing 69 (2018) 44–67

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Primal-dual optimization strategies in Huber-L^1 optical flow with temporal subspace constraints for non-rigid sequence registration ☆

Monica Hernandez

Robotics, Perception and Real Time Group, Aragon Institute on Engineering Research (I3A), University of Zaragoza, Spain

ARTICLE INFO

Article history:
Received 26 February 2016
Received in revised form 16 November 2017
Accepted 18 November 2017
Available online 23 November 2017

Keywords:
Convex optimization
Fenchel-Duality
Primal-dual
Preconditioning
Optical flow
Non-rigid sequence registration
Temporal consistency
Subspace constraints

ABSTRACT

This work studies the application of Fenchel-Duality principles to general convex optimization problems and their corresponding relaxed versions in the context of optical flow estimation. We derive the associated primal-dual optimization strategies in the problem of Huber-L^1 optical flow with temporal consistency for non-rigid sequence registration. Temporal consistency is imposed using a recently proposed approach that characterizes the optical flow using temporal subspace constraints, yielding solutions in a space spanned by a non-rigid orthogonal trajectory basis. The performance of the resulting optical flow methods has been studied in a framework for non-rigid sequence registration evaluation. In addition, we have compared the solution of the different methods in other challenging datasets. We have found that the strategies with the best outcome are among the ways of applying Fenchel-Duality principles that were not considered in previous works for the optical flow model with temporal subspace constraints. Indeed, our experiments have shown the simplest optimization strategy as the best performing one.

1. Introduction

Opticalflow or image registration in non-rigid sequences of images is among the most challenging and interesting problems in computer vision and medical image analysis. In computer vision, the estimation of the optical flow from sequences of non-rigid environments serves as input to other important applications related to non-rigid scene understanding (e.g. non-rigid structure from motion [25,71]). In medical image analysis, non-rigid image registration is crucial for the study of dynamic structures from medical images [40,67]. The most important challenges for solving the problem are how to properly deal with large displacements, deformations, and occlusions; how to include temporal consistency in the algorithms; how to work with changes of illumination in the scene; and how to adapt the methods to the particularities of the different medical image modalities [4,8,15,17,26,30,34,37,41,49,50,56-58,68,76,77,86]. In both disciplines, these challenges are approached by imposing models of flow that seek for meaningful solutions.

This work focuses on non-rigid sequence registration with temporal consistency preservation. Although consistency in the temporal dimension is a very desirable feature for this application, most optical flow methods yet focus on improving robustness under noise and varying illumination, or achieving spatial consistency, as reported in Volz et al. and Garg et al. [26,76]. From them, very few impose regularization on the temporal dimension. Regularization is mostly based on the spatio-temporal convolution or the differentiation of the flow [5,7,9,18,42,43,78,87]. In consequence, these methods are constrained to take advantage of the temporal consistency in two or at most few neighboring frames.

Some methods impose a regularizer in the temporal component of the flow [47,64,76,78]. Temporal regularizers have shown accurate results for smooth over time sequences. However, regularizing in the temporal dimension has adverse effects on the estimated flows in sequences including not only occlusions and other temporal discontinuities but also complex motions and large displacements. In addition, these proposals still allow considering few frames in the regularization.

The most effective approaches for including long temporal consistency in non-rigid sequence registration, characterize the time-varying flows as a linear combination of deformation bases. Some of these methods use spatio-temporal models for the deformation bases. The models are either given from spatio-temporal standard bases or

☆ This paper has been recommended for acceptance by Daniel Rueckert, PhD.
E-mail address: mhg@unizar.es.

https://doi.org/10.1016/j.imavis.2017.11.005
0262-8856/

learned from spatio-temporal training data [15,45,49,86]. However, the large dimensionality of the problem can extremely limit the accuracy of the solutions. Alternative methods use temporal models for the deformation bases. These methods are inspired by the dual shape-motion representation of non-rigid structure from motion [1,72]. Under the assumption that the time-varying flows at a given point in the image domain lie in a low-dimensional space, the model is given by standard temporal bases or learned from the temporal dimension of spatio-temporal training data [23,24,26]. With this approach, the dimensionality of the problem can be considerably reduced. As a result, the methods are able to consider long sequences in the optical flow computation.

Since the original contribution of Horn and Schunk for determining the optical flow [33], variational methods lead the techniques for solving all variants of the optical flow problem. Variational methods aim at the minimization of an energy functional. The energy is defined by the contribution of image and regularization terms. The image term measures the similarity of the reference and the moving images after registration. The regularization term allows constraining the solution to a desired space of flows. In the last decade, there has been a growing interest in regularizers based on the Total Variation norm [24,26,59,77,80,81,84]. These regularizers control the magnitude of the flow gradient while they can preserve discontinuities. This ability has led these methods to occupy top positions in optical flow benchmark studies [3,10,20] and evaluations [26].

The popularity of the Total Variation regularizer in computer vision problems has increased thanks to the availability of optimization methods for solving this difficult problem [13,14,21,28,52,62,75]. In particular, Chambolle and Pock proposed a fast first-order primal-dual algorithm for convex optimization and provided the equations for various computer vision problems including optical flow [13,53]. Their work recovered the principles of convex analysis and Fenchel-Duality and derived their primal-dual algorithm from the optimization of a saddle-point problem [61]. In addition, they showed linear convergence bounds for sufficiently smooth problems. Since Chambolle and Pock method, different primal-dual strategies have been proposed in the literature [26,52,80]. However, they have been partially formulated and evaluated in our application of interest.

The starting point of this article is the framework proposed by Garg et al. for non-rigid sequence registration with temporal subspace constraints [26]. In that work, the authors provided an extended version of *improved TV-L^1* optical flow algorithm [77] for *weighted Huber-L^1* variational formulation. The method reformulated the pairwise variational problem in order to deal with sequence registration. In addition, the time-varying flows were characterized using temporal subspace constraints on the whole sequence. Optimization was approached using quadratic relaxation methods and applying Fenchel-Duality principles to the resulting subproblems in two different fashions, yielding two primal-dual strategies for solving the problem. The authors noticed a third strategy arising from the parameterization of the optical flow in the temporal subspace as a hard constraint. The evaluation of the methods focused in showing the superiority of using the temporal subspace constraints. However, there are alternative ways of applying Fenchel-Duality principles to the variational problem that were not considered in that work, for example, the strategy associated with Chambolle and Pock method [13], or the preconditioned optimization method proposed in Pock and Chambolle [52]. We believe it would be of interest to derive, evaluate and compare the primal-dual optimization strategies associated with these alternatives.

Thus, the contribution of this article is to explore these different ways of applying Fenchel-Duality principles in general convex optimization problems, and to provide the derivation of the associated primal-dual optimization strategies in the problem of non-rigid sequence registration with temporal subspace constraints. We have derived seven different primal-dual optimization strategies in our application of interest. These strategies include the three methods in Garg et al. [26]. In addition, we have extended the non-rigid evaluation framework provided in Garg et al. [26] with images of different nature. The resulting optical flow methods have been evaluated in this framework, and compared in other challenging datasets [25,65].

The rest of the article is divided as follows. Section 2 describes the variational formulation of weighted Huber-L^1 optical flow with temporal subspace constraints. Section 3 reviews the foundations of Fenchel-Duality principles and their application to primal-dual optimization. Section 4 shows the derivation of the seven primal-dual optimization strategies. Section 5 gathers the most important implementation details. Results are presented in Section 6. Finally, Section 7 provides the most remarkable conclusions of our work.

2. Huber-L^1 optical flow in sequences of images

2.1. Variational formulation

Let $\Omega \subset \mathbb{R}^2$ be the domain of size $M \times N$ where the images are defined [1], $I_0 : \Omega \to \mathbb{R}$ the reference frame, and $I_f : \Omega \times \{1, \cdots, F\} \to \mathbb{R}$ the temporal sequence of images, respectively. The optical flow model used in this work aims at finding a sequence of time-varying vector fields $\mathbf{u}_f : \Omega \times \{1, \cdots F\} \to \mathbb{R}^2$ that estimates the motion existing between each frame of the sequence and the reference image. This problem can be formulated from the minimization of the general variational problem

$$E_{\text{total}}(\mathbf{u}) = E_{\text{reg}}(\mathbf{u}) + \alpha \sum_{f=1}^{F} E_{\text{img}}\left(I_f(\mathbf{x} + \mathbf{u}_f), I_0\right), \tag{1}$$

where E_{img} measures the similarity of the reference I_0 and the images I_f after warping with the transformation $\mathbf{x} + \mathbf{u}_f$, E_{reg} regularizes the space of solutions, and α balances the contribution of both terms to the total energy. Since the image similarity term is formulated with respect to a single image I_0, this model is suited to handle sequences from static viewpoints with minor changes across time.

In this work, we focus on weighted Huber-L^1 optical flow [81] extended to sequences of images

$$E_{\text{total}}(\mathbf{u}) = \sum_{f=1}^{F} \int_{\Omega} g(\nabla I_0) \|\nabla \mathbf{u}_f\|_{\epsilon} d\Omega + \alpha \sum_{f=1}^{F} \|I_f(\mathbf{x} + \mathbf{u}_f) - I_0\|_{L^1}. \tag{2}$$

Thus, the L^1 norm is used for measuring the image similarity, and Huber regularizer is used to control the smoothness of the flow. The expression of Huber regularizer is obtained replacing the L^1 norm in the expression of Total Variation (TV) regularizer by the Huber norm

$$\|x\|_{\epsilon} = \begin{cases} |x| - \epsilon/2, & \text{if} \quad |x| > \epsilon \\ |x|^2/(2\epsilon), & \text{if} \quad |x| \leq \epsilon. \end{cases} \tag{3}$$

The parameter $\epsilon > 0$ defines the tradeoff between the linear and the quadratic contributions to the norm yielding solutions in between TV and L^2 regularizers. The function used for weighting the Huber regularizer is an edge-preserving potential function. This function contributes to provide solutions with discontinuities located at the edges of the reference image. In this work, we use $g = \exp(-c|\nabla I_0|^2)$,

[1] In this work, we focus on 2D sequences of images. However, the framework can be extended to 3D with extra work on the derivations. The 3D case is especially interesting for medical imaging applications. In fact, we used the preconditioned optimization method 2 in Hernandez [31,32].