Neurocomputing 220 (2017) 236–243

Contents lists available at ScienceDirect

Neurocomputing

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

Video interpolation using optical flow and Laplacian smoothness

Wenbin Li[a,b], Darren Cosker[b]

[a] Department of Computer Science, University College London, UK
[b] Centre for the Analysis of Motion, Entertainment Research and Applications (CAMERA), University of Bath, UK

ARTICLE INFO

Keywords:
Optical flow
Non-rigid deformation
Video interpolation
Laplacian smoothness

ABSTRACT

Non-rigid video interpolation is a common computer vision task. In this paper we present an optical flow approach which adopts a Laplacian Cotangent Mesh constraint to enhance the local smoothness. Similar to Li et al., our approach adopts a mesh to the image with a resolution up to one vertex per pixel and uses angle constraints to ensure sensible local deformations between image pairs. The Laplacian Mesh constraints are expressed wholly inside the optical flow optimization, and can be applied in a straightforward manner to a wide range of image tracking and registration problems. We evaluate our approach by testing on several benchmark datasets, including the Middlebury and Garg et al. datasets. In addition, we show application of our method for constructing 3D Morphable Facial Models from dynamic 3D data.

1. Introduction

Non-rigid video interpolation is a computer vision related problem that requires the tracking of non-rigid objects, calculation of dense image correspondences and the registration of image sequences containing highly non-rigid deformation. Existing algorithms to achieve this include model based tracking [1], dense patch identification and matching [2,3], group-wise image registration [4], space–time tracking [5–9] and optical flow [10–16]. All such models and the general dense tracking have been widely used in fields e.g. motion tracking [17,18], visualization [19,20], interaction [21] and Rotoscoping [22].

Optical flow is an attractive formulation as it provides a dense displacement field between image pairs. In most standard approaches, assumptions regarding gray value constancy between images and smoothness in motion between neighboring pixels are adopted [11,10]. Sun et al. [23] propose a different approach which overcomes these constraints by learning a probabilistic model for flow estimation. However, their approach requires training pre-calculated optical flow ground truths, which are difficult to obtain. In the general optical flow model, it is common to adopt a data term consisting of gray value and gradient constraints (e.g. Brox and Malik [11]) and an additional smoothness term. Nevertheless, most previous optical flow formulations only consider global smoothness and ignore formulations that preserve local image details.

Many optical flow techniques concentrate on problems where the scene movement is largely rigid in nature. However, there are many problem cases where we would like to calculate flow given highly non-rigid global and local image displacements over long image sequences. One recent problem highlighting this particular case is the alignment of 3D dynamic facial sequences containing highly non-rigid deformations [24,25]. The problem requires non-rigidly aligning a set of images to a reference – e.g. a neutral facial expression. Each image referred to as a UV map[1] is accompanied by a corresponding 3D mesh, and each mesh has a difference vertex topology. Once the UV maps are registered to a reference image (e.g. a neutral expression), vertex correspondence can be imposed. The technique is popular in 3D Morphable Model construction [24,26].

Beeler et al. [27] and Bradley et al. [25] adopt a slightly different approach to mesh correspondence. In their solutions, image displacement is calculated from camera views and then used to deform a reference mesh from an initial frame through a 3D sequence. The optical flow provides guides for adjusting pixel positions, and the mesh reduces artifacts by imposing a constraint to prevent faces on the mesh from becoming inverted or flipped. Further, mesh and image deformation research in graphics is an active area of research [28]. Such techniques provide flexible methods to invoke deformation while preserving some desired properties such as local geometric details. As such, it is also of interests to apply such solutions as smoothness constraints to optical flow calculation, which forms the central basis of our presented work. Li et al. [29] introduce a hybrid optical flow framework that takes into account a Laplacian mesh data term and a global smoothness term. However, their energy is highly nonlinear and hard to minimize.

[1] UV refers to the XY location of a pixel in the image. UV map is the graphical term for the *texture* for a 3D model. Each UV location maps to a 3D vertex on a corresponding mesh.

http://dx.doi.org/10.1016/j.neucom.2016.04.064
Received 17 March 2016; Accepted 5 April 2016
Available online 29 September 2016
0925-2312/

1.1. Contributions

In this paper we present an optical flow algorithm (*LCM-flow*) which adopts a smoothness term based on *Laplacian Cotangent Mesh Deformation*. Such deformation approaches have been widely used in graphics, particularly for preserving small details on deformable surface [30,31]. Such concept shows advantage in the non-rigid optical flow estimation [29]. Those energy is able to penalize local movements and preserve smooth global details. In our method, the proposed constraint on the local deformations is expressed in Laplacian coordinates encourage local regularity of the mesh whilst allowing globally non-rigid preservation.

Similar to Li et al. [29], our proposed algorithm applies a mesh to the image with a resolution up to one vertex per pixel. The Laplacian constraint is described in terms of a smoothness term, and can be applied in a straightforward manner to a number of optical flow approaches with the addition of our proposed minimization strategy. We evaluate our approach on the popular *Middlebury* dataset [32] as well as the publicly available non-rigid dataset proposed by Garg et al. [12]. We show our method to give high performance on *Middlebury* in terms of interpolation, and either outperform or show comparable accuracy against the leading publicly available non-rigid approaches when evaluated against Garg et al. In addition, we show an application of our optical flow approach for building dynamic 3D Morphable Models from dynamic 3D facial data, and outperform a current state of the art method.

The remainder of our paper is organized as follows: In Sections 2 and 3 our strategy for calculating optical flow displacements between image pairs is outlined. Section 4 shows an evaluation of *LCM-flow* on the *Middlebury* dataset and four other publicly available sequences of non-rigidly deforming objects [12].

2. Energy function definition

In this section the core energy function of our *Laplacian Cotangent Mesh* based Optical Flow approach is presented. In the formulation the algorithm considers a pair of consecutive frames in an image sequence. The current frame is denoted by $I_i(\mathbf{X})$ and its successor by $I_{i+1}(\mathbf{X})$, where $\mathbf{X} = (x, y)^T$ is a pixel location in the image domain Ω. We define the optical flow displacement between $I_i(\mathbf{X})$ and $I_{i+1}(\mathbf{X})$ as $\mathbf{w}_i = (\mathbf{u}, \mathbf{v})^T$. In the proposed optical flow estimation approach, the core energy function can be obtained from the following general formulation:

$$\mathbf{E}(\mathbf{w}) = \mathbf{E}_{Data}(\mathbf{w}) + \lambda\cdot\mathbf{E}_{Global}(\mathbf{w}) + \xi\cdot\mathbf{E}_{Lap}(\mathbf{w})$$

where $\mathbf{E}_{Data}(\mathbf{w})$ denotes a data term that contains both Gray Value and *Gradient Constancy* assumptions (Section 2.1) on pixel values between $I_i(\mathbf{X})$ and $I_{i+1}(\mathbf{X})$.

Two smoothness terms are also introduced into the formulation. Similar to [11,10], the first term $\mathbf{E}_{Global}(\mathbf{w})$ controls global flow smoothness. The second term represents our core contribution, i.e. a Laplacian Cotangent Mesh constraint $\mathbf{E}_{Lap}(\mathbf{w})$. In the following sections we next describe each term in detail, focusing on our Laplacian constraint in Section 2.3.

2.1. Data term definition

Following the standard optical flow assumption regarding *Gray Value Constancy*, we assume that the gray value of a pixel is not varied by its displacement through the entire image sequence. In addition, we also make a *Gradient Constancy* assumption which is engaged to provide additional stability in case the first assumption (*Gray Value Constancy*) is violated by changes in illumination. The data term of *LCM-flow* encoding these assumptions is therefore formulated as:

$$\mathbf{E}_{Data}(\mathbf{w}) = \sum_{\Omega} \Psi\left(I_{i+1}(\mathbf{X}+\mathbf{w}) - I_i(\mathbf{X})\right)^2 + \theta\cdot\Psi\left(\nabla I_{i+1}(\mathbf{X}+\mathbf{w}) - \nabla I_i(\mathbf{X})\right)^2$$

In order to deal with occlusions, we apply the increasing concave function $\Psi(s^2) = \sqrt{s^2 + \epsilon^2}$ with $\epsilon = 0.001$ [11] to solve this formation which enables $L1$ minimization. The remaining term $\nabla = (\partial_{xx}, \partial_{yy})^T$ is a spatial gradient and $\theta \in [0, 1]$ denotes weight that can be manually assigned with different values. In the experiments it is pre-defined as 0.5.

2.2. Global smoothness constraint

The first smoothness term of *LCM-flow* is a dense pixel based regularizer that penalizes global variation. The objective is to produce a globally smooth optical flow field (as in the data term, the robust function $\Psi(s^2)$ is again used):

$$\mathbf{E}_{Global}(\mathbf{w}) = \sum_{\Omega} \Psi\left(|\nabla\mathbf{u}|^2 + |\nabla\mathbf{v}|^2\right)$$

2.3. Laplacian Cotangent Mesh Smoothness Constraint

Global smoothness terms are widely used in optical flow formulation [32]. However, their definition means that local nonlinear variations between images – such as those in non-rigid motion – can be over smoothed. In order to improve optical flow estimation against the local complexity of non-rigid motion, a novel *Laplacian Cotangent Mesh* constraint is proposed in this section. The aim of this constraint is to account for non-rigid motion in scene deformation. This term is inspired by Laplacian mesh deformation research in graphics which aims to preserve local mesh smoothness under non-linear transformation [30]. Its use in computer vision research for optical flow estimation is introduced for the first time here. Although non-rigid motion is highly nonlinear, the movement of pixels in such deformations still often exhibit strong correlations in local regions. In order to represent this, we propose a quantitative *Cotangent Weight* based on a Laplacian framework and a differential representation. The scheme was originally presented by Meyer et al. [31] for mesh deformation.

We assume that the image is initially covered by a triangular mesh

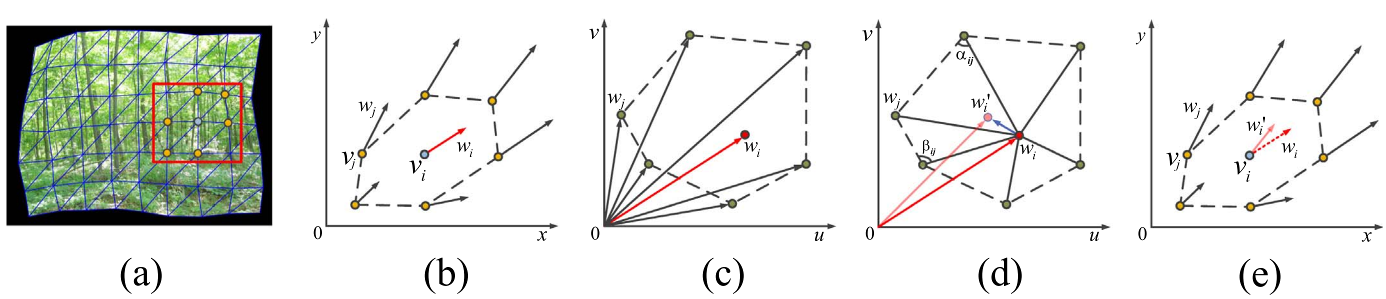

Fig. 1. Laplacian cotangent mesh constraint in optical flow vector space. (a) A mesh on a specific frame. (b) 1-Ring neighborhood based on vertices. (c) 1-Ring neighborhood based on endpoints of optical flow vectors. (d) $\delta(\mathbf{w}_i)$ (the blue vector) calculated by endpoints of optical flow vectors. (e) The modified optical flow vector $\mathbf{w}_{'i}$ based on $\mathbf{w}_i$ and $\delta(\mathbf{w}_i)$. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)