



CHALMERS

Chalmers Publication Library

On robust optical flow estimation on image sequences with differently exposed frames using primal–dual optimization

This document has been downloaded from Chalmers Publication Library (CPL). It is the author's version of a work that was accepted for publication in:

Image and Vision Computing (ISSN: 0262-8856)

Citation for the published paper:

Bengtsson, T. ; McKelvey, T. ; Lindström, K. (2017) "On robust optical flow estimation on image sequences with differently exposed frames using primal–dual optimization". Image and Vision Computing, vol. 57 pp. 78-88.

<http://dx.doi.org/10.1016/j.imavis.2016.11.003>

Downloaded from: <http://publications.lib.chalmers.se/publication/247723>

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)

On Robust Optical Flow Estimation on Image Sequences with Differently Exposed Frames using Primal-Dual Optimization

Tomas Bengtsson^{a,*}, Tomas McKelvey^a, Konstantin Lindström^b

^aChalmers University of Technology, Gothenburg, Sweden

^bVolvo Cars AB, Gothenburg, Sweden

Abstract

Optical flow methods are used to estimate pixelwise motion information based on consecutive frames in image sequences. The image sequences traditionally contain frames that are similarly exposed. However, many real-world scenes contain high dynamic range content that cannot be captured well with a single exposure setting. Such scenes result in certain image regions being over- or underexposed, which can negatively impact the quality of motion estimates in those regions. Motivated by this, we propose to capture high dynamic range scenes using different exposure settings every other frame. A framework for OF estimation on such image sequences is presented, that can straightforwardly integrate techniques from the state-of-the-art in conventional OF methods. Different aspects of robustness of OF methods are discussed, including estimation of large displacements and robustness to natural illumination changes that occur between the frames, and we demonstrate experimentally how to handle such challenging flow estimation scenarios. The flow estimation is formulated as an optimization problem whose solution is obtained using an efficient primal-dual method.

Keywords: optical flow estimation, high dynamic range, temporal coherency, primal-dual

*Corresponding author

Email address: tomasbengtsson83@gmail.com (Tomas Bengtsson)

1. Introduction

Optical flow (OF) methods are used to estimate apparent motion in image sequences [1, 2, 3]. They provide dense displacement fields that contain 2-dimensional motion vectors for each pixel location of a reference image. The produced low-level flow data is in turn used for higher-level computer vision tasks such as segmentation, tracking, motion analysis or image registration [4, 5, 6]. The literature on OF methods is focused almost exclusively on motion estimation for image sequences whose frames are taken with the same exposure settings. This is despite the fact that many real-world scenes contain high dynamic range (HDR) content that cannot be captured with a single exposure setting due to dynamic range limitations of camera sensors [7, 8]. In such a scenario, some regions of a given image will be over- or underexposed, which leads to an inability to estimate the motion of objects in those regions. For instance, lack of contrasts in the image data are reported as a bottleneck for the quality of resulting flow estimates for vehicle driver assistance applications [9]. In recent years, HDR functionality has been integrated into consumer cameras and other mobile devices. It captures the full dynamic range of the scene by taking and merging multiple images with varying exposure settings. If there is non-global motion between the images, reconstruction artifacts are typically avoided by using a HDR deghosting method [10]. Compensating for complex, local motion patterns is typically not attempted due to the difficulty of estimating the motion with sufficient accuracy [11]. However, there is recent work on estimating a HDR image based on dense OF motion data that provides promising results [12, 13].

In this paper, we propose a framework for optical flow estimation on image sequences with differently exposed frames, a setup devised particularly for HDR scenarios. Specifically, four frames are used, taken with two different exposure settings that are used every other frame. The intention is that all image regions should be properly exposed (non-saturated) for at least one of the exposure settings, such that the combined dynamic range is sufficiently high with respect to the imaged scene. The OF method is pursued mainly for its own purposes as an

enabler to motion analysis applications, and we do not aim to reconstruct the HDR image data. The flow estimation is formulated as minimizing a variational cost functional which is a sum of a data term and a set of terms that enforce regularity conditions on the flow solution [1, 14]. Different data term formula-
35 tions have been evaluated in our previous work [15], and the best among those candidates is adopted in this paper. OF methods, including ours, typically use pointwise data cost term correspondences and thus rely on a spatial regularity condition in order to obtain a minimization problem that has a unique solu-
40 tion. The spatial regularization term is formulated to penalize deviations from a piecewise smooth flow solution, based on the statistical observation that scenes are well represented by a set of objects whose respective points move in a similar manner [16]. Conventional OF methods generally estimate motion between pairs of consecutive frames using only those two frames as input data. Some methods, however, include additional frames and enforce temporal coherence of
45 the motion vectors in the estimation process [17, 18]. Whether such an approach improves the flow estimation performance depends on to what degree the image data fulfills the assumption of temporally smooth motion. Our method relies on temporal regularization since more than two input images are used.

In the context of OF estimation, there are a number of aspects related to
50 whether or not an OF method is robust. Firstly, it has to do with using robust cost expressions for the data term and the regularization terms, as opposed to the traditional L^2 -norm which deals poorly with data outliers [19, 20]. Secondly, it has to do with the ability to handle challenging scenes that lead to natural illumination changes and large flow magnitudes when captured in an image se-
55 quence [21]. Finally, since the typical objective of OF methods is to provide information of the real motion within the scene, it is required that the input image sequence contains the relevant information, particularly in the case of HDR scenes. This aspect is not necessarily about robustness of the OF method itself, but definitely about the robustness of using OF estimates in an application.

60 1.1. Contribution

The performance of resulting flow estimates depends directly on the quality of the input image data [9, 15]. This is particularly the case for HDR scenarios, for which any given image contains saturated image regions due to over- or underexposure. To address the issue of saturation, the optical flow estimation
65 problem is extended to use image sequences that contain differently exposed frames. The OF data term formulation for such a scenario is based on our previous work [15]. In this paper, we propose a complete OF method that includes handling of natural illumination changes, that occur even between images that are captured with the same exposure setting, as well as flow estimation of ob-
70 jects with large displacements. Deviations of the flow estimate from flow data obtained from pre-matched, sparse image features are penalized in a feature matching cost term, which mitigates the effects of local minima in the iterative minimization of the total OF cost functional. Existing OF methods for HDR scenarios do not address natural illumination changes or large flow mag-
75 nitudes [12, 13]. Furthermore, our method is formulated for a more general camera model that allows any changes to the camera exposure settings to be made. The algorithm for computing the minima is based on an efficient primal-dual method [22] (see also [23, 24]), that handles non-differentiable expressions of robust penalty functions [25] without the need for approximation.

80 1.2. Outline of the paper

The paper is structured as follows. Our camera model is described in Section 2. The proposed method is presented in Section 3 and the primal-dual optimization used to obtain flow estimates is presented in Section 4. Experimental results are provided in Section 5 and the paper is concluded in Section 6.

85 2. Camera model

We assume that an image is generated by a camera according to the model

$$\tilde{I}_f(\mathbf{x}) = \text{CRF}(\Phi_f(R(\mathbf{x}) + N_f(\mathbf{x}))), \quad (1)$$

where $\mathbf{x} = (x, y) \in \Omega \subset \mathbb{R}^2$ are continuous image coordinates, $R(\mathbf{x})$ is the (filtered) illuminance incident on the sensor for the specific lightning condition of the imaged scene at the time instance of the image \tilde{I}_f and $N_f(\mathbf{x})$ models sensor noise. The camera response function (CRF) clips the sensor exposure
90 $\Phi_f(R(\mathbf{x}) + N_f(\mathbf{x}))$ outside of the operating interval of the sensor, limited by its dynamic range. The function Φ_f models the specific exposure setting used for image f .

If the brightness (illuminance) of any given point is assumed to be constant along its motion trajectory, the standard assumption on which OF methods are based, another image \tilde{I}_{f+1} of the same scene can be related to the non-occluded regions of \tilde{I}_f through

$$\tilde{I}_{f+1}(\mathbf{x} + \mathbf{u}_f(\mathbf{x})) = \text{CRF}(\Phi_{f+1}(R(\mathbf{x}) + N_{f+1}(\mathbf{x}))), \quad (2)$$

where \mathbf{u}_f denotes the optical flow of point \mathbf{x} in \tilde{I}_f . The functions Φ_f, Φ_{f+1} can refer to arbitrary exposure settings, such as using flash illumination every other frame [15]. The local illumination effects that are caused by e.g. flash illumination, however, makes it difficult to mathematically formulate expressions for Φ_f, Φ_{f+1} that can relate the image intensities of points in an image taken with flash illumination to an image one taken without. This is discussed further at the end of the section. For the special case where the two images are taken with different exposure durations (the specific scenario that is treated e.g. in [12, 13]), represented by the positive scalars $\Delta t_1, \Delta t_2$, they are given by

$$\begin{aligned} \tilde{I}_f(\mathbf{x}) &= \text{CRF}(\Delta t_1(R(\mathbf{x}) + N_f(\mathbf{x}))), \\ \tilde{I}_{f+1}(\mathbf{x} + \mathbf{u}_f(\mathbf{x})) &= \text{CRF}(\Delta t_2(R(\mathbf{x}) + N_{f+1}(\mathbf{x}))). \end{aligned} \quad (3)$$

These images can be aligned photometrically by inverting the effect of the CRF in their non-saturated regions, followed by scaling with the inverse of the respective exposure durations. Then, the optical flow \mathbf{u}_f can be estimated between
95 I_f, I_{f+1} for points \mathbf{x} that are non-saturated in both images. Hence forth, we use I_f to denote images that, if possible, have been aligned photometrically. Furthermore, even such cases where the differently exposed frames cannot be

aligned photometrically can be treated by the proposed optical flow method
 100 (see [15] for experimental results on such image sequences). This is a novelty of
 the proposed method, which is achieved by taking four frames as input to the
 flow estimation and mathematically relating the image intensities of two pairs of
 similarly exposed frames as well as using shared flow variables for the respective
 pairs.

105 3. Optical flow estimation for differently exposed input images

The OF method proposed in this section utilizes an image sequence of four
 consecutive grayscale input frames I_f , $f = 1, \dots, 4$ for the objective of esti-
 mating the optical flow field at the reference frame, fixed here as I_2 . Exposure
 setting I is used for I_1, I_3 , whereas exposure setting II is used for I_2, I_4 . Different
 110 exposure settings are used for the two respective image pairs in order to capture
 high dynamic range information. For example, they could be set to use long
 and short exposure durations, respectively. The non-saturated regions of each
 respective image pair are denoted by $\Omega^{\text{Exp.I}} \subset \Omega$ and $\Omega^{\text{Exp.II}} \subset \Omega$. Data cost
 terms are formulated between the respective pairs of similarly exposed frames.
 115 The data costs are parameterized by flow terms such that the flow at the ref-
 erence frame coordinates is included as a variable [18]. If $\Omega^{\text{Exp.I}} \cup \Omega^{\text{Exp.II}} = \Omega$,
 the two data cost terms between the image pairs I_1, I_3 and I_2, I_4 successfully
 cover the whole image region. In other words, there are no points that are sat-
 urated for both exposure settings. Furthermore, an additional data cost term
 120 is included to relate any mutually non-saturated points $\Omega^{\text{Int}} = \Omega^{\text{Exp.I}} \cap \Omega^{\text{Exp.II}}$
 in the pair I_2, I_3 . This term is particularly beneficial due to its higher frame
 rate [15]. This data cost term requires that the non-saturated regions of I_2, I_3
 are photometrically aligned (recall the discussion in Section 2).

The cost functional that should be minimized for flow estimation, aside from
 the mentioned data term, consists of a spatial regularization term and a tem-
 poral regularization term, denoted E_S and E_T respectively, as well as a feature
 matching term E_M . Furthermore, to handle natural illumination changes, two

approaches are considered. The first approach is to include illumination offset variables in the data term, along with a corresponding spatial regularization term E_L [26, 22]. The second is to use a data term expression that is robust to natural illumination changes without explicitly modeling them [20]. We evaluate the second approach experimentally and provide the (minor) necessary modifications, but stick to the first approach throughout the current presentation. Thus, the task is to find the minimizer of

$$\begin{aligned} \min_{\{\mathbf{u}_f\}, \{l_f\}} E := & E_D(\{\mathbf{u}_f\}, \{l_f\}) + E_S(\{\nabla \mathbf{u}_f\}) + E_L(\{\nabla l_f\}) + \\ & + E_T(\{\mathbf{u}_f\}) + E_M(\{\mathbf{u}_f\}), \end{aligned} \quad (4)$$

where, for $f = 1, 2, 3$, $\mathbf{u}_f = (u_f(\mathbf{x}), v_f(\mathbf{x})) : \Omega \rightarrow \mathbb{R}^2$ are the flow variables and $l_f(\mathbf{x}) : \Omega \rightarrow \mathbb{R}$ are illumination offset variables. The gradient operator applied to a vector has the meaning $\nabla \mathbf{u}_f = ((\nabla u_f)^T, (\nabla v_f)^T)^T$. The primary interest is the estimation performance of the flow variable \mathbf{u}_2 , which is the (time-discrete) flow at the reference frame I_2 . The illumination offsets are included to account for natural illumination changes between image pairs, thereby modeling deviations from the assumption that the brightness intensity of any point \mathbf{x} is constant along its motion trajectory [26, 22, 27]. Introducing the concatenated variables $\mathbf{w}_f = (\mathbf{u}_f, l_f)$, the respective terms of the cost functional (4) are

$$\begin{aligned} E_D(\{\mathbf{w}_f\}) = & E_{D13}(\mathbf{w}_1, \mathbf{w}_2) + E_{D24}(\mathbf{w}_2, \mathbf{w}_3) + E_{D23}(\mathbf{w}_2) = \\ = & \alpha_D \int_{\Omega} \theta_{13} |I_3(\mathbf{x} + \mathbf{u}_2) - I_1(\mathbf{x} - \mathbf{u}_1) + \beta(l_1 + l_2)| + \\ & + \theta_{24} |I_4(\mathbf{x} + \mathbf{u}_2 + \mathbf{u}_3) - I_2(\mathbf{x}) + \beta(l_2 + l_3)| + \\ & + \theta_{23} |I_3(\mathbf{x} + \mathbf{u}_2) - I_2(\mathbf{x}) + \beta l_2| \, d\mathbf{x}, \end{aligned} \quad (5a)$$

$$\begin{aligned} E_S(\{\nabla \mathbf{u}_f\}) = & E_{S1}(\nabla \mathbf{u}_1) + E_{S2}(\nabla \mathbf{u}_2) + E_{S3}(\nabla \mathbf{u}_3) = \\ = & \alpha_S \sum_{f=1}^3 \int_{\Omega} \|\nabla \mathbf{u}_f\| \, d\mathbf{x}, \end{aligned} \quad (5b)$$

$$E_L(\{\nabla l_f\}) = E_{L1}(\nabla l_1) + E_{L2}(\nabla l_2) + E_{L3}(\nabla l_3) =$$

$$= \alpha_L \sum_{f=1}^3 \int_{\Omega} \|\nabla l_f\| d\mathbf{x}, \quad (5c)$$

$$\begin{aligned} E_T(\{\mathbf{u}_f\}) &= E_{T1}(\mathbf{u}_1, \mathbf{u}_2) + E_{T2}(\mathbf{u}_2, \mathbf{u}_3) = \\ &= \alpha_T \int_{\Omega} \|\mathbf{u}_2 - \mathbf{u}_1\| + \|\mathbf{u}_3 - \mathbf{u}_2\| d\mathbf{x}, \end{aligned} \quad (5d)$$

$$\begin{aligned} E_M(\{\mathbf{u}_f\}) &= E_{M1}(\mathbf{u}_1) + E_{M2}(\mathbf{u}_2) + E_{M3}(\mathbf{u}_3) = \\ &= \alpha_M \sum_{f=1}^3 \int_{\Omega} m_f \|\mathbf{u}_f - \mathbf{u}_f^{\text{Match}}\| d\mathbf{x}, \end{aligned} \quad (5e)$$

where $\|\cdot\|$ is the L^2 -norm, each term has a constant weight $\alpha_D, \alpha_S, \alpha_L, \alpha_T, \alpha_M \geq 0$, $\theta_{13}(\mathbf{x}), \theta_{24}(\mathbf{x}), \theta_{23}(\mathbf{x}) \geq 0$ are spatially dependent data term weights and β is a constant model parameter. Importantly, $\theta_{23}(\mathbf{x})$ is set to zero for image regions where either of the differently exposed I_2, I_3 is saturated to avoid false data correspondences that lead to poor flow estimates. If it is not possible to photometrically align images taken with exposure setting I and II respectively (e.g. if one image is taken with flash illumination and the other is not), $\theta_{23}(\mathbf{x})$ is simply set to zero $\forall \mathbf{x}$ and the illumination offset l_2 is removed completely from E so that each image pair has a separate illumination-offset variable in its data term. The overall data term E_D thus only contains correspondences between pairs of similarly exposed frames, but the method can still provide good estimates of the shared flow field \mathbf{u}_2 [15]. Each flow component, \mathbf{u}_f , has a corresponding separate spatial regularization term, E_{Sf} , that enforces the flow solution to be piecewise smooth by penalizing $\|\nabla \mathbf{u}_f\|_1 = \int_{\Omega} \|(u_{fx}, u_{fy}, v_{fx}, v_{fy})\| d\mathbf{x}$, the robust L^1 -norm of the magnitude of the first-order derivatives of the flow vector. This smoothness term is called the Total Variation (TV) semi-norm and is designed to preserve discontinuities, flow edges in this case. In certain scenarios, e.g. the second order Total Generalized Variation (TGV2) expression is more suitable [25]. For example, it provides a more appropriate description of the flow solution for scenes that consist largely of flat surfaces. It is straightforward to replace the TV expression by the TGV2 expression in (5b) as well as in the minimization scheme. The illumination terms are penalized similarly by

their respective TV expressions in E_L , in order to encourage spatially piece-wise smooth illumination offset estimates. The temporal regularization term E_T enforces temporal coherence of the flow along the motion trajectories of the points in the reference frame, using a suitable parametrization of the flow fields $\{\mathbf{u}_f(\mathbf{x})\}$ as increments relative to the reference frame positions [18, 15]. The feature matching term E_M is discussed in the next section.

3.1. Flow information from sparse feature matches

A feature matching term (5e) is included in the total OF cost functional E in (4) to penalize deviations of the estimated flow from the flow data in $\mathbf{u}_f^{\text{Match}}$, $f = 1, 2, 3$. Its inclusion helps to mitigate the difficulty of finding the global minima of the OF cost functional in the iterative solution strategy [21, 28]. Thus, a sparse set of points in each image, represented by some feature descriptors, are matched by a (approximate) nearest neighbor (NN) search as a pre-processing step to the OF estimation and are used to derive the flow information in $\mathbf{u}_f^{\text{Match}}$. The aim is that these matches will provide information about large displacements or complex motion patterns, particularly of small-scale objects, that are otherwise prone to be poorly estimated. Specifically, we extract HoG (histogram of oriented gradients) feature descriptors on a uniformly spaced grid, consisting of every 4th pixel location (per dimension), excluding image regions that are over- or underexposed. Then, using a fast approximate NN method [29], the descriptors of the pairs of similarly exposed images I_1, I_3 and I_2, I_4 are matched separately. Forward-backward consistency checks are performed to discard inconsistent matches. Furthermore, matches at feature locations \mathbf{x} where there is minor image structure are discarded. To achieve this, the 2×2 structure tensor is computed for each feature location based on the respective 7×7 neighborhoods. The locations where the smaller eigenvalue $eig_2(\mathbf{x})$ of the structure tensor satisfies $eig_2(\mathbf{x}) > (eig_1(\mathbf{x}) + eig_2(\mathbf{x}))/10$ are kept and the others are discarded [30, 31]. The two best NN matches are retrieved for each kept feature location, and used to set the weights $m_f(\mathbf{x})$ in (5e) according to the confidence level of each respective NN match, as the

relative feature distance of the second best NN match to the best NN match of each feature descriptor, and $m_f(\mathbf{x}) = 0$ holds for non-matched points [21]. The retained matches between I_1, I_3 , as well as those between I_2, I_4 , are stored in \mathbf{c}_{13} and \mathbf{c}_{24} respectively, in terms of the locations of the matched points in the reference image I_2 , their matching confidence and their displacements. Note that the second best NN matches are not used to determine the displacements themselves, only to determine their respective weights. Further implementation details are given in Section 4.3.

4. Flow estimation by primal-dual optimization

This section describes how (4) is minimized in order to obtain flow estimates. Successive local linear approximations of the data terms $E_{D13}, E_{D24}, E_{D23}$ about the current estimates (also called warping points in the OF literature), $\mathbf{u}_f^0(\mathbf{x})$, are made to find the minimizer of (4) iteratively [14]. The original data term expressions of (5a) are clearly non-convex in the flow variables, which in practice means that the search space contains undesired local minima. The standard approach to mitigate the effects of local minima in OF estimation methods, with or without the inclusion of pre-matched image features, is to use a coarse-to-fine minimization strategy as outlined in Section 4.3. Downsampled images are used to find rough estimates that are then refined for finer image resolution levels.

4.1. Linearized data terms

The linearized data terms (reusing the same names) are

$$\begin{aligned} E_{D13} &= \alpha_D \int_{\Omega} \theta_{13} |I_{t13} + \nabla I_3(\mathbf{u}_2 - \mathbf{u}_2^0) + \nabla I_1(\mathbf{u}_1 - \mathbf{u}_1^0) + \beta(l_1 + l_2)| \, d\mathbf{x}, \\ E_{D24} &= \alpha_D \int_{\Omega} \theta_{24} |I_{t24} + \nabla I_4(\mathbf{u}_2 + \mathbf{u}_3 - \mathbf{u}_2^0 - \mathbf{u}_3^0) + \beta(l_2 + l_3)| \, d\mathbf{x}, \\ E_{D23} &= \alpha_D \int_{\Omega} \theta_{23} |I_{t23} + \nabla I_3(\mathbf{u}_2 - \mathbf{u}_2^0) + \beta l_2| \, d\mathbf{x}, \end{aligned} \quad (6)$$

where

$$\begin{aligned}
I_{t13} &= I_3(\mathbf{x} + \mathbf{u}_2^0) - I_1(\mathbf{x} - \mathbf{u}_1^0), \\
I_{t24} &= I_4(\mathbf{x} + \mathbf{u}_2^0 + \mathbf{u}_3^0) - I_2(\mathbf{x}), \\
I_{t23} &= I_3(\mathbf{x} + \mathbf{u}_2^0) - I_2(\mathbf{x}), \\
\nabla I_1 &= \nabla(I_1(\mathbf{x} - \mathbf{u}_1^0)), \\
\nabla I_3 &= \nabla(I_3(\mathbf{x} + \mathbf{u}_2^0)), \\
\nabla I_4 &= \nabla(I_4(\mathbf{x} + \mathbf{u}_2^0 + \mathbf{u}_3^0)).
\end{aligned} \tag{7}$$

4.2. Sequential minimization

The minimization problem (4), which is convexified due to replacing E_D in (5a) by the linearized data terms (6), is approximated by a sequence of simpler minimization problems, following the approach by Estellers et al. [32], by updating the estimates of \mathbf{w}_f sequentially for each f . That is, fixed estimates \mathbf{w}_f^k are taken as inputs to a given iteration k . The sequential update is performed by solving

$$\mathbf{w}_f^{k+1} \leftarrow \min_{\mathbf{w}_f} P(\mathbf{w}_f) := E(\mathbf{w}_f) + G(\mathbf{w}_f), \tag{8}$$

for each f separately, where the term

$$G(\mathbf{w}_f) = \frac{\alpha_G}{2} \int_{\Omega} \|\mathbf{w}_f - \mathbf{w}_f^k\|^2 d\mathbf{x}, \tag{9}$$

is added to constrain the step length of each separate update. To encourage convergence to the minima of the joint problem, (4), the sequential minimization
200 of \mathbf{w}_f , $f = 1, 2, 3$, is iterated over $k = 0, \dots, K - 1$.

4.3. Pseudo-algorithm

In summary, the pseudo-algorithm for minimizing (4) is given in Table 1. The sequential minimization scheme described in Section 4.2 is incorporated into a coarse-to-fine multi-resolution strategy. The flow- and illumination variables are
205 initiated to zero-vectors at the initial pixel resolution ratio^($S-1$) of the coarse-to-fine image pyramid, where ratio < 1 is the re-sampling factor and S is the number of resolution levels. Thus, at the initial stage, significantly downsampled

input images are used to obtain flow estimates at a coarse resolution level. These are re-scaled and used as inputs as the flow estimation progresses at the
210 next, finer resolution level up until the original pixel resolution of the images I_f . As flow estimates are refined, the images are warped according to the current estimates, as expressed in (7). In the numerical implementation, non-integer arguments of I_1, I_3, I_4 that result from subpixel precision optical flow estimates are evaluated using bicubic interpolation. At a given warping point
215 \mathbf{u}_f^0 , estimates are updated $\mathbf{w}_f^k \rightarrow \mathbf{w}_f^{k+1}$ according to the iterative primal-dual solver presented in the next section.

Pseudo-algorithm of proposed OF method

Initialization: $\mathbf{w}_f^0 = \mathbf{0}, \forall f, \mathbf{x}$
for scale = ratio^(S-1,...,0)
 for warp = 0, ..., W - 1
 compute $I_{t13}, I_{t24}, I_{t23}, \nabla I_1, \nabla I_3, \nabla I_4$ as in (7).
 set $\mathbf{w}_f^{\{k=0\}} = \mathbf{w}_f^0, \forall f$
 for $k = 0, \dots, K - 1$
 for $f = 1, 2, 3$
 $\mathbf{w}_f^k \rightarrow \mathbf{w}_f^{k+1}$, according to Section 4.4.
 end
 end
 $\mathbf{u}_f^K \rightarrow \mathbf{u}_f^0, \forall f$, new warping point.
 end
end

Table 1: Sequential minimization method for optical flow estimation.

The sparse pre-computed feature matches are originally expressed in \mathbf{c}_{13} and \mathbf{c}_{24} (see Section 3.1) with respect to pixel coordinates at the full pixel resolution. To be integrated into the coarse-to-fine method, their associated locations in I_2
220 and their flow magnitudes are re-sampled at the current resolution level, and

the locations are rounded to the nearest integer pixel coordinate. At a coarse resolution, the locations of multiple feature matches may be rounded to the same integer pixel coordinate. Because the spacing of the feature grid is 4 points, any level finer than a fourth of the original resolution has at most one
225 candidate. Thus, the relative influence of the feature matches is high at coarse resolution levels and then naturally decreases as the minimization progresses to finer resolution levels. At any given resolution level, each integer location \mathbf{x} in $\mathbf{u}_2^{\text{Match}}$ is set according to the match in \mathbf{c}_{13} and \mathbf{c}_{24} that has the highest confidence among those that are associated with the particular location \mathbf{x} . The
230 flow data in $\mathbf{u}_1^{\text{Match}}$ as well as $\mathbf{u}_3^{\text{Match}}$ is constructed similarly, but only selecting among the feature matches in \mathbf{c}_{13} or \mathbf{c}_{24} respectively.

4.4. Primal-Dual update for given flow component

The update step $\mathbf{w}_f^k \rightarrow \mathbf{w}_f^{k+1}$ in Table 1 is performed using an efficient first-order primal-dual method by Chambolle and Pock [22], that they demonstrate on a set of image processing tasks, including optical flow estimation. We describe the update for $f = 2$. The updates for $f = 1, 3$ are analogous but include fewer cost terms. The minimization (8) for $f = 2$ is

$$\begin{aligned} \mathbf{w}_2^{k+1} \leftarrow \min_{\mathbf{w}_2} P(\mathbf{w}_2) := & E_{D13}(\mathbf{w}_1^{k+1}, \mathbf{w}_2) + E_{D24}(\mathbf{w}_2, \mathbf{w}_3^k) + E_{D23}(\mathbf{w}_2) + \\ & + E_{S2}(\nabla \mathbf{u}_2) + E_{L2}(\nabla l_2) + E_{T1}(\mathbf{u}_1^{k+1}, \mathbf{u}_2) + \\ & + E_{T2}(\mathbf{u}_2, \mathbf{u}_3^k) + E_{M2}(\mathbf{u}_2) + G(\mathbf{w}_2). \end{aligned} \quad (10)$$

The minimization problem (10) is equivalent to the constrained problem

$$\begin{aligned} \min_{\mathbf{w}_2, \mathbf{w}} P(\mathbf{w}_2, \mathbf{w}) := & E_{D13}(\mathbf{w}_1^{k+1}, \mathbf{w}_2^{(D13)}) + E_{D24}(\mathbf{w}_2^{(D24)}, \mathbf{w}_3^k) + E_{D23}(\mathbf{w}_2^{(D23)}) + \\ & + E_{S2}(\mathbf{u}_2^{(S2)}) + E_{L2}(l_2^{(L2)}) + E_{T1}(\mathbf{u}_1^{k+1}, \mathbf{u}_2^{(T1)}) + \\ & + E_{T2}(\mathbf{u}_2^{(T2)}, \mathbf{u}_3^k) + E_{M2}(\mathbf{u}_2^{(M2)}) + G(\mathbf{w}_2), \end{aligned} \quad (11)$$

where the argument includes a set of auxiliary variables

$$\mathbf{w} \triangleq (\mathbf{w}_2^{(D13)}, \mathbf{w}_2^{(D24)}, \mathbf{w}_2^{(D23)}, \mathbf{u}_2^{(S2)}, l_2^{(L2)}, \mathbf{u}_2^{(T1)}, \mathbf{u}_2^{(T2)}, \mathbf{u}_2^{(M2)})$$

with equality constraints

$$\begin{aligned} \mathbf{w}_2^{(D13)} &= \mathbf{w}_2^{(D24)} = \mathbf{w}_2^{(D23)} = \mathbf{w}_2, \quad \mathbf{u}_2^{(S2)} = \nabla \mathbf{u}_2, \\ l_2^{(L2)} &= \nabla l_2, \quad \mathbf{u}_2^{(T1)} = \mathbf{u}_2^{(T2)} = \mathbf{u}_2^{(M2)} = \mathbf{u}_2. \end{aligned} \quad (12)$$

The constrained problem (11) can be written on the form

$$\min_{\mathbf{w}_2} P(\mathbf{w}_2) := E(K\mathbf{w}_2) + G(\mathbf{w}_2), \quad (13)$$

where $\mathbf{w} = K\mathbf{w}_2$ and K is a linear operator that relates each auxiliary variable to \mathbf{w}_2 according to (12). This is the class of convex problems studied by Chambolle and Pock [22]. The term $E(K\mathbf{w}_2)$ is the linearized OF cost functional with $\mathbf{w}_1, \mathbf{w}_3$ kept fixed. The problem (13) is in primal form and has a corresponding primal-dual formulation

$$\max_{\boldsymbol{\lambda}} \min_{\mathbf{w}_2} \text{PD}(\mathbf{w}_2, \boldsymbol{\lambda}) := -E^*(\boldsymbol{\lambda}) + G(\mathbf{w}_2) + \int_{\Omega} \langle \boldsymbol{\lambda}, K\mathbf{w}_2 \rangle d\mathbf{x}, \quad (14)$$

where E^* is the convex conjugate of E , such that \mathbf{w}_2^* in the saddle point $(\mathbf{w}_2^*, \boldsymbol{\lambda}^*)$ for PD is the minimum of P. The dual variables

$$\boldsymbol{\lambda} \triangleq (\boldsymbol{\lambda}^{(D13)}, \boldsymbol{\lambda}^{(D24)}, \boldsymbol{\lambda}^{(D23)}, \boldsymbol{\lambda}^{(S2)}, \boldsymbol{\lambda}^{(L2)}, \boldsymbol{\lambda}^{(T1)}, \boldsymbol{\lambda}^{(T2)}, \boldsymbol{\lambda}^{(M2)})$$

are the Lagrange multipliers to each of the equality constraints (12) of the auxiliary primal variables [32]. The iterative algorithm to find \mathbf{w}_2^* , that is the new \mathbf{w}_2^{k+1} in the sequential minimization summarized in Table 1, is:

$$\begin{aligned} \boldsymbol{\lambda}^{n+1} &\leftarrow \min_{\boldsymbol{\lambda}} \left(\sigma E^*(\boldsymbol{\lambda}) + \frac{1}{2} \|\boldsymbol{\lambda} - (\boldsymbol{\lambda}^n + \sigma K\mathbf{z}^n)\|^2 \right), \\ \mathbf{w}_2^{n+1} &\leftarrow \min_{\mathbf{w}_2} \left(\tau G(\mathbf{w}_2) + \frac{1}{2} \|\mathbf{w}_2 - (\mathbf{w}_2^n - \tau K^* \boldsymbol{\lambda}^{n+1})\|^2 \right), \\ \mathbf{z}^{n+1} &= \mathbf{w}_2^{n+1} + \theta(\mathbf{w}_2^{n+1} - \mathbf{w}_2^n), \end{aligned} \quad (15)$$

where $\tau, \sigma > 0$ are primal and dual step lengths, $\theta \in [0, 1]$ controls the amount of over-relaxation in $\mathbf{z} = (z_u, z_v, z_l)$, and K^* is the adjoint operator of K [22]. Maximization of PD in (14) w.r.t. $\boldsymbol{\lambda}$ is substituted by minimization of $-\text{PD}$. Note that subscripts of $\boldsymbol{\lambda}$ and \mathbf{z} are omitted to lighten up notation, although these variables are specific to each f . The reformulation of the original problem

(10) leads to that E^* is separable in each of its variables, that is

$$\begin{aligned} E^*(\boldsymbol{\lambda}) = & E_{D13}^*(\boldsymbol{\lambda}^{(D13)}) + E_{D24}^*(\boldsymbol{\lambda}^{(D24)}) + E_{D23}^*(\boldsymbol{\lambda}^{(D23)}) + E_{S2}^*(\boldsymbol{\lambda}^{(S2)}) + \\ & + E_{L2}^*(\boldsymbol{\lambda}^{(L2)}) + E_{T1}^*(\boldsymbol{\lambda}^{(T1)}) + E_{T2}^*(\boldsymbol{\lambda}^{(T2)}) + E_{M2}^*(\boldsymbol{\lambda}^{(M2)}). \end{aligned} \quad (16)$$

Thus, the dual update equation in (15) is solved separately for each dual variable in $\boldsymbol{\lambda}$. Efficient closed-form solutions exist for each separate minimization problem. The update steps on the top two lines of (15) can equivalently be formulated as

$$\begin{aligned} \boldsymbol{\lambda}^{n+1} &= \mathbf{prox}_{\sigma E^*}(\tilde{\boldsymbol{\lambda}}), \quad \tilde{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^n + \sigma K \mathbf{z}^n, \\ \mathbf{w}_2^{n+1} &= \mathbf{prox}_{\tau G}(\tilde{\mathbf{w}}_2), \quad \tilde{\mathbf{w}}_2 = \mathbf{w}_2^n - \tau K^* \boldsymbol{\lambda}^{n+1}, \end{aligned} \quad (17)$$

where the minimizations correspond to evaluating the proximal operators of $\sigma E^*(\boldsymbol{\lambda})$ and $\tau G(\mathbf{w}_2)$ about the proximal points $\tilde{\boldsymbol{\lambda}}$ and $\tilde{\mathbf{w}}_2$. Such proximal operator notation is used in Appendix A where the update equations corresponding to each separate dual term in (16) are derived [33]. The resulting expressions are given in Table 2.

5. Experimental Results

Whether or not the proposed approach, to use differently exposed input frames, has merit depends upon if the imaged scene contains significant HDR content, as discussed at further length in [15]. To exemplify, consider the image sequence in Figure 1 (a). The four input images are frames from the *Tennis* sequence, of pixel resolution 380×530 , courtesy of the computer vision group at Freiburg University [34]. The frames have been altered to simulate a HDR scenario by clipping the pixel values above $a_1 = 0.6$ in I_1, I_3 , setting them equal to a_1 , which corresponds to exposure setting I (Exp.I). Similarly, pixel values below $a_2 = 0.3$ are clipped for I_2, I_4 , which is taken as exposure setting II (Exp.II). This achieves the same effect as using two different exposure durations every other frame and photometrically aligning them as discussed in Section 2. The estimated flow field based on the described sequence is given in Figure 1 (b). The flow is color encoded according to the inlaid circular chart, whose flow

Update equations for (15)

Initialization: $\mathbf{w}_2^{\{n=0\}} = \mathbf{w}_2^k$

for $n = 0, \dots, N - 1$

$\boldsymbol{\lambda}^{(D13),n+1}, \boldsymbol{\lambda}^{(D24),n+1}, \boldsymbol{\lambda}^{(D23),n+1}$ updated with (25)

$\boldsymbol{\lambda}^{(S2),n+1} = \tilde{\boldsymbol{\lambda}}^{(S2)} / \max(1, \|\tilde{\boldsymbol{\lambda}}^{(S2)}\|/\alpha_S), \quad \tilde{\boldsymbol{\lambda}}^{(S2)} = \boldsymbol{\lambda}^{(S2),n} + \sigma \nabla(z_u^n, z_v^n)$

$\boldsymbol{\lambda}^{(L2),n+1} = \tilde{\boldsymbol{\lambda}}^{(L2)} / \max(1, \|\tilde{\boldsymbol{\lambda}}^{(L2)}\|/\alpha_L), \quad \tilde{\boldsymbol{\lambda}}^{(L2)} = \boldsymbol{\lambda}^{(L2),n} + \sigma \nabla z_l^n$

$\boldsymbol{\lambda}^{(T1),n+1} = \mathbf{d} / \|\mathbf{d}\| \min(\alpha_T, \|\mathbf{d}\|),$

$\mathbf{d} = \tilde{\boldsymbol{\lambda}}^{(T1)} - \sigma \mathbf{u}_1^{k+1}, \quad \tilde{\boldsymbol{\lambda}}^{(T1)} = \boldsymbol{\lambda}^{(T1),n} + \sigma(z_u^n, z_v^n)$

$\boldsymbol{\lambda}^{(T2),n+1} = \mathbf{d} / \|\mathbf{d}\| \min(\alpha_T, \|\mathbf{d}\|),$

$\mathbf{d} = \tilde{\boldsymbol{\lambda}}^{(T2)} - \sigma \mathbf{u}_3^k, \quad \tilde{\boldsymbol{\lambda}}^{(T2)} = \boldsymbol{\lambda}^{(T2),n} + \sigma(z_u^n, z_v^n)$

$\boldsymbol{\lambda}^{(M2),n+1} = \mathbf{d} / \|\mathbf{d}\| \min(\alpha_M m_2, \|\mathbf{d}\|),$

$\mathbf{d} = \tilde{\boldsymbol{\lambda}}^{(M2)} - \sigma \mathbf{u}_2^{\text{Match}}, \quad \tilde{\boldsymbol{\lambda}}^{(M2)} = \boldsymbol{\lambda}^{(M2),n} + \sigma(z_u^n, z_v^n)$

$\mathbf{u}_2^{n+1} = \frac{1}{\tau \alpha_G + 1} \left(\tau \alpha_G \mathbf{u}_2^k + \mathbf{u}_2^n - \tau (\boldsymbol{\lambda}_{u,v}^{(D13),n+1} + \boldsymbol{\lambda}_{u,v}^{(D24),n+1} + \boldsymbol{\lambda}_{u,v}^{(D23),n+1} + \right.$
 $\left. + \boldsymbol{\lambda}^{(T1),n+1} + \boldsymbol{\lambda}^{(T2),n+1} + \boldsymbol{\lambda}^{(M2),n+1} - \text{div} \boldsymbol{\lambda}^{(S2),n+1}) \right)$

$l_2^{n+1} = \frac{1}{\tau \alpha_G + 1} \left(\tau \alpha_G l_2^k + l_2^n - \tau (\boldsymbol{\lambda}_l^{(D13),n+1} + \boldsymbol{\lambda}_l^{(D24),n+1} + \boldsymbol{\lambda}_l^{(D23),n+1} - \right.$
 $\left. - \text{div} \boldsymbol{\lambda}^{(L2),n+1}) \right)$

$\mathbf{z}^{n+1} = \mathbf{w}_2^{n+1} + \theta(\mathbf{w}_2^{n+1} - \mathbf{w}_2^n)$

end

$\mathbf{w}_2^N \rightarrow \mathbf{w}_2^{k+1}$

Table 2: The steps to solve (8) exemplified for $f = 2$.

magnitude at its outer radius corresponds to 24 pixels. The use of alternating exposure settings is analyzed by comparing the resulting flow estimate to the ones obtained from image sequences whose frames are all taken with a fixed exposure setting, either Exp.I or Exp.II. Certain objects are hardly visible in any of the frames for a given exposure setting, such as the tennis ball in the frames taken with Exp.I, thus estimating its motion may not be feasible. Figure 1 (c), (d) show the flow estimates that result from using the input sequences with similarly exposed, with data clipped above a_1 or below a_2 respectively for all 4 frames. The data term weights used are $\alpha_D = 3$, $\beta = 0.01$, $\theta_{13} = 1$, $\mathbf{x} \in \Omega^{\text{Exp.I}}$, $\theta_{24} = 1$, $\mathbf{x} \in \Omega^{\text{Exp.II}}$, $\theta_{23} = 1$, $\mathbf{x} \in \Omega^{\text{Int}}$ and $\theta_{13}, \theta_{24}, \theta_{23}$ are zero for unspecified points. The other weights are $\alpha_S = 0.2$, $\alpha_L = 0.2$, $\alpha_T = 0.02$, $\alpha_M = 1$ and $\alpha_G = 1$. Furthermore, Figure 1 (e)-(g) show the results obtained from performing the same experiments as in (b)-(d) but with increased amount of clipped data, where $a_1 = 0.5$ and $a_2 = 0.5$. The motion of the tennis ball is actually captured in the flow estimate based on the Exp.I sequence for $a_1 = 0.6$, but not anymore when $a_1 = 0.5$. The flow estimate for Exp.II is good for $a_2 = 0.3$ but degrades severely for $a_2 = 0.5$. The performance of the flow estimate that is based on the image sequence with differently exposed frames degrades more gracefully as the amount of saturation increases, due to the lost information being available in the other image pair.

5.1. Image sequence with large displacements

The use of the feature matching term in order to estimate flow of objects with large displacements is discussed here. First, note that if $\alpha_M = 0$ and $\beta = 0$, the current method becomes similar to our previous method, although here we use sequential minimization (Section 4.2) with the flow updates obtained from a primal-dual solver (Section 4.4) that does not require differentiable cost expressions, and thus the exact TV expression can be used. We validate the sequential minimization approach by comparing to the quantitative experiments in [15]. In terms average endpoint error (AEPE), we obtain equally good results on the Sintel [35] sequences for which the regularization weight in [15] was

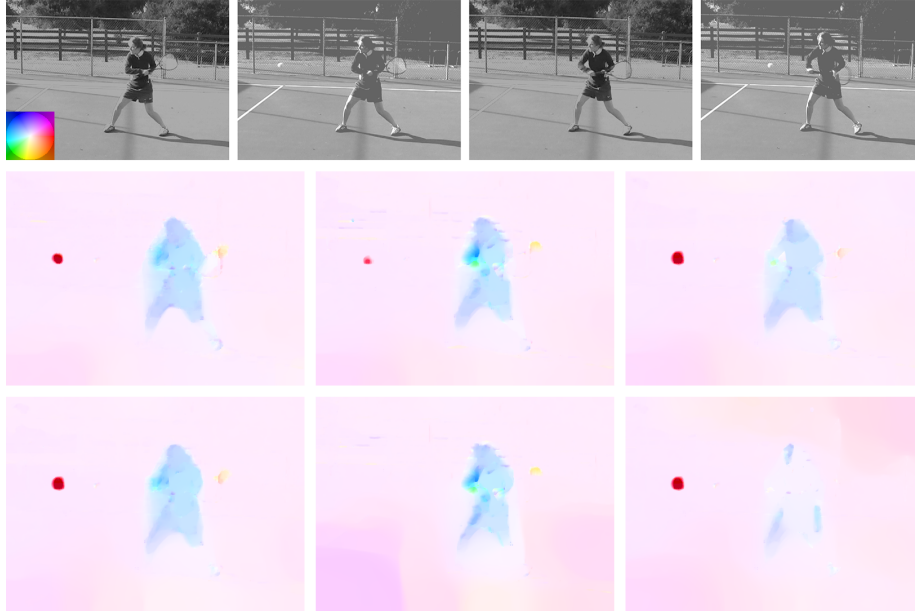


Figure 1: Top row: (a) The image sequence $\{I_f\}$ with differently exposed frames, and $a_1 = 0.6$, $a_2 = 0.3$. Middle row, from left to right: (b) Flow estimate based on the images in (a). (c) Flow estimate when all I_f are taken with Exp.I. (d) Flow estimate when all I_f are taken with Exp.II. Bottom row: (e)-(g) Similar to the middle row, but with clipping thresholds $a_1 = a_2 = 0.5$.

optimized and slightly better results on the Middlebury [36] sequences. Thus, sequential minimization is applicable.

The inclusion of the feature matching term in the coarse-to-fine minimiza-
 285 tion of (4) provides the ability to estimate large displacements. However, wise
 selection is needed in the feature extraction and matching steps to mitigate the
 occurrence of false matches with high confidence weights. The flow informa-
 tion $\mathbf{w}_2^{\text{Match}}$ from the feature matches is visualized in Figure 2 at the resolution
 level 88×123 . Figure 2 (a), (b) show $\mathbf{w}_2^{\text{Match}}$ before discarding feature locations
 290 with minor image structure, as well as the corresponding confidence weight map
 m_2 . The confidences are plotted for the range $[0, 1]$, although the maximum is
 2.42, as normalizing by the maximum leads to poorer visibility of other weights.
 Figure 2 (c), (d) show $\mathbf{w}_2^{\text{Match}}$ and m_2 after discarding feature locations with

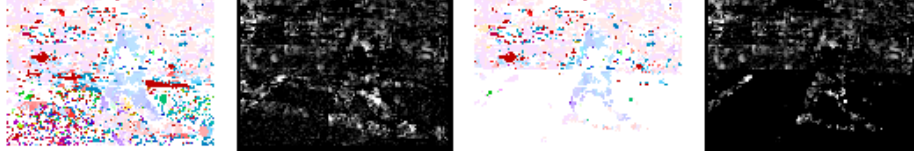


Figure 2: From left to right: (a) The flow $\mathbf{w}_2^{\text{Match}}$ at the resolution level 88×123 , without the selection strategy of discarding locations with low amount of structure. (b) The matching confidences m_2 corresponding to (a). (c) The flow $\mathbf{w}_2^{\text{Match}}$ without including feature matches in image regions that lack structure. (d) The confidences corresponding to (c).

low amount of structure. In Figure 3, the benefit of the feature matches and
 295 a wise selection strategy is demonstrated in terms of the flow estimates on the sequence in Figure 1 (a). First, Figure 3 (a) shows the flow estimate when

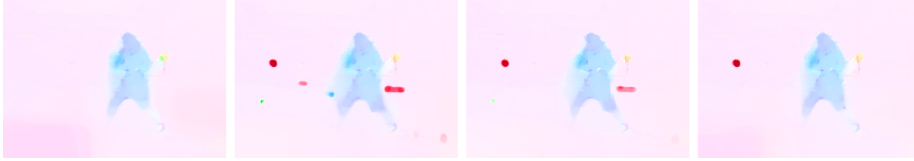


Figure 3: From left to right: (a) Flow estimate for $\alpha_M = 0$. (b) Flow estimate with $\mathbf{w}_2^{\text{Match}}$ as shown in Figure 2 (a). (c) Flow estimate for image-driven spatial regularization. (d) Flow estimate with $\mathbf{w}_2^{\text{Match}}$ as shown in Figure 2 (c).

$\alpha_M = 0$ and (b) shows the flow estimate when including all feature matches in
 Figure 2 (a), that successfully captures the motion of the ball but at the cost
 of poor flow estimates in several regions, that result from false feature matches.
 300 In Figure 3 (c), the influence of the false matches is reduced, for this partic-
 ular example, by using a spatially varying, image-driven regularization term
 $\alpha_S(\mathbf{x}) = 0.4 \exp(-3\|\nabla I_2(\mathbf{x})\|^{0.5})$, which smooths regions that do not coincide
 with image edges to a higher degree. Finally, Figure 3 (d) shows the result
 of the proposed selection strategy, that only uses the feature matches in Fig-
 305 ure 2 (c). Although the proposed feature extraction and matching strategy gives
 good performance in this example, the DeepFlow method by Weinzaepfel et al.
 is particularly designed to be robust in challenging scenarios, such as repetitive

image structures [37].

5.2. Robustness to natural illumination changes

310 In this section, the benefit of accounting for natural illumination changes in our method is demonstrated on two image sequences from the KITTI raw data sequences (that unlike the flow annotated KITTI sequences contain more than 2 frames per sequence) [38]. Aside from using the data term presented in Section 3, we run a separate experiment using the CSAD data term introduced in [20],
 315 which is adapted for inclusion in our proposed method. The CSAD expression is an approximation of the popular Census Transform that is adjusted to fit the variational minimization of the OF cost functional, and performs equally well as the original Census expression. A HDR scenario is simulated similarly as before, with $a_1 = 0.6$ and $a_2 = 0.3$ as clipping thresholds. The image
 320 resolution used is 621×188 . The two sequences are shown at the top row of the respective columns in Figure 4. Image sequences from the KITTI benchmark suite [39] are known to typically include natural illumination changes between the captured frames. As a result, the estimation performance is improved when accounting for illumination changes in the OF cost functional modeling [27, 40,
 325 20]. Due to the nature of the KITTI data, consisting largely of flat surfaces, we exchange the spatial TV regularization in E_S by the second order Total Generalized Variation [25] that penalizes deviations from (piecewise) affine flow solutions. The mathematical TGV2 expression and the modifications that follow to the update equations in Table 2 are given in Appendix B, along with the
 330 details of implementing the CSAD data term. In the experiments of this section, $\alpha_M = 0$ is used except where otherwise stated. Figure 4 (b) shows initial flow estimates for the case when using the data cost in (5a) with $\beta = 0$. The flow is color coded similarly as in Figure 1. The maximum flow magnitudes used in the color encoding is 20 pixels for the left column and 30 pixels for the
 335 right column. In Figure 4 (c) and (d), the flow estimates obtained when using $\beta = 0.01$, i.e. including the illumination offset terms, are shown along with plots of l_2 estimates, for which 0 offset is represented by 0.5 on the normalized

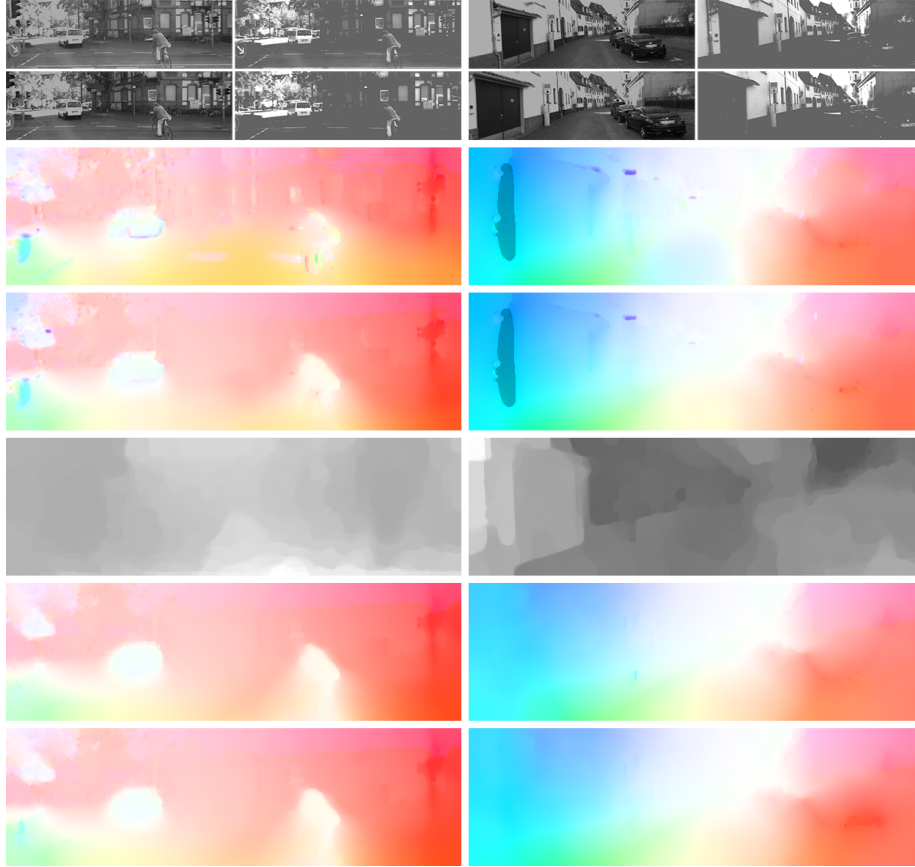


Figure 4: For each column, from top to bottom: (a) Input images I_1, \dots, I_4 . (b) Flow estimates when using the data term in (5a) and $\beta = 0$, and (c) with $\beta = 0.01$. (d) Plots of l_2 , where 0 is coded as the middle grayscale intensity). (e) Flow estimates when using the CSAD data term [20]. (f) Same as (e) but with $\alpha_M = 1$.

interval $[0, 1]$. Figure 4 (e) shows the flow estimates when using the CSAD data term, and finally (f) shows the result with CSAD and $\alpha_M = 1$. Numerous
 340 flow artifacts are seen at flow- and image edges, and the road surfaces are poorly estimated for both image sequences in Figure 4 (b) where no handling of illumination changes is included. The flow estimates are improved significantly when illumination variables are included in (c). The results in Figure 4 (e) show similar performance when using the CSAD term. Less flow artifacts are

345 present compared to (c), but the motion of the road sign in the bottom left
 corner of the left column is not well captured in the flow estimates. Because
 no large displacements of small scale objects are present, the aim of the feature
 matching term is primarily to avoid negative impacts in such a scenario. As can
 be seen in Figure 4 (f), no false feature matches impact the final flow estimates.
 350 Finally, flow estimates obtained for the same image sequences as in Figure 4
 (a) but with all 4 frames taken either with exposure setting I or II are shown
 in Figure 5 (a) and (b) respectively. These estimates are produced using the
 CSAD data term. The flow estimates on the image sequence in the left column
 are particularly poor, regardless of whether Exp.I or Exp.II was used. As seen
 355 in Figure 4, using the sequence of alternating exposure settings clearly leads to
 better results in that case.

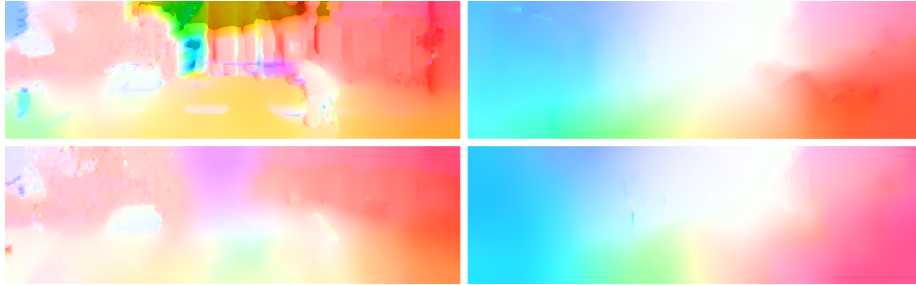


Figure 5: Flow estimates obtained using Exp.I for all 4 frames (top row) and Exp.II (bottom
 row) respectively.

6. Conclusions

In this paper, a framework has been proposed for OF estimation on im-
 age sequences with differently exposed frames, intended for motion analysis in
 360 HDR scenarios. Various aspects of robust OF estimation have been discussed
 and demonstrated. Experimental results show the benefit, with regard to flow
 estimation performance, of using the proposed setup to mitigate the negative
 impact of saturated image data. Our approach allows to re-use techniques from

the rich literature on conventional OF estimation, as they are easily integrated
 365 to tailor the proposed OF method to a given application. For instance, total gen-
 eralized variation (TGV) regularization, which has shown superior performance
 particularly on the KITTI image sequences, can straightforwardly replace the
 first order TV penalty. A spatial regularizer that takes the underlying HDR
 image data into account, as suggested by Hafner et al., also deserves considera-
 370 tion [13], as well as the feature matching approach DeepFlow [37]. A database
 of sequences with alternating exposure settings (every other frame), based on a
 measurement campaign of real-world HDR scenes, is highly desired.

A. Proximal operators

In this appendix, we demonstrate how to evaluate the proximal operators of
 375 σE^* and τG , presented in (17). The proximal operator of the function $\sigma E^*(\boldsymbol{\lambda})$,
 whose full expression is given in (16), is separable and can thus be evaluated
 sequentially for each variable contained in $\boldsymbol{\lambda}$ as shown in (16). Derivations
 are provided here for the proximal operators of $\sigma E_{D24}^*(\boldsymbol{\lambda}^{(D24)})$, $\sigma E_{S2}^*(\boldsymbol{\lambda}^{(S2)})$,
 $\sigma E_{T1}^*(\boldsymbol{\lambda}^{(T1)})$, a subset of the terms in E^* , as well as for that of $\tau G(\mathbf{w}_2)$. The
 380 proximal operators of the remaining terms in E^* are analogue to either of the
 presented ones and their expressions are included in Table 2.

A.1. The proximal operator of a general function

The proximal operator of a function σF maps a proximal point, \tilde{p} , to another
 point p^{n+1} according to [41]

$$\tilde{p} \rightarrow p^{n+1} = \mathbf{prox}_{\sigma F}(\tilde{p}) = \arg \min_p \left(\sigma F(p) + \frac{1}{2} \|p - \tilde{p}\|^2 \right). \quad (18)$$

By the Moreau decomposition, the mapping is equivalent to

$$\tilde{p} \rightarrow \tilde{p} - \sigma \mathbf{prox}_{\frac{1}{\sigma} F^*} \left(\frac{\tilde{p}}{\sigma} \right), \quad (19)$$

where $\mathbf{prox}_{\frac{1}{\sigma} F^*}$ is the proximal operator of $\frac{1}{\sigma} F^*$ [41, 33].

A.2. The proximal operator of σE_{D24}^*

The proximal operator of σE_{D24}^* for the proximal point $\tilde{\boldsymbol{\lambda}}^{(D24)} = \boldsymbol{\lambda}^{(D24),n} + \sigma K_{D24} \mathbf{z}^n = \boldsymbol{\lambda}^{(D24),n} + \sigma \mathbf{z}^n$, using the Moreau decomposition as in (19), gives

$$\begin{aligned} \boldsymbol{\lambda}^{(D24),n+1} &= \mathbf{prox}_{\sigma E_{D24}^*}(\tilde{\boldsymbol{\lambda}}^{(D24)}) = \tilde{\boldsymbol{\lambda}}^{(D24)} - \sigma \mathbf{prox}_{\frac{1}{\sigma} E_{D24}}\left(\frac{\tilde{\boldsymbol{\lambda}}^{(D24)}}{\sigma}\right) = \\ &= \tilde{\boldsymbol{\lambda}}^{(D24)} - \sigma \mathbf{s}^*, \end{aligned} \quad (20)$$

where \mathbf{s}^* is obtained by pointwise (w.r.t. \mathbf{x}) evaluation of the proximal operator of E_{D24}/σ . To obtain the final expression for $\boldsymbol{\lambda}^{(D24),n+1}$ we evaluate \mathbf{s}^* further. Taking E_{D24} from (6) and introducing a_0, \mathbf{a}, ρ for shorter notation,

$$\begin{aligned} \frac{E_{D24}(\mathbf{s})}{\sigma} &= \frac{\alpha_D}{\sigma} \int_{\Omega} \theta_{24} |a_0 + \mathbf{a}^T \mathbf{s}| d\mathbf{x}, \\ a_0 &= I_{t24} + (\nabla I_4)^T (\mathbf{u}_3^k - \mathbf{u}_2^0 - \mathbf{u}_3^0) + \beta l_3^k, \\ \mathbf{a}^T &= ((\nabla I_4)^T, \beta), \\ \rho(\mathbf{s}) &\triangleq a_0 + \mathbf{a}^T \mathbf{s}, \end{aligned} \quad (21)$$

we get (using (18)), for each point \mathbf{x} ,

$$\mathbf{s}^* \leftarrow \min_{\mathbf{s}} \left(\frac{\alpha_D \theta_{24}}{\sigma} |\rho(\mathbf{s})| + \frac{1}{2} \|\mathbf{s} - \tilde{\mathbf{s}}\|^2 \right), \quad (22)$$

where $\tilde{\mathbf{s}} = \tilde{\boldsymbol{\lambda}}^{(D24)}/\sigma$. The solution is given by a simple and effective thresholding scheme on the values of $\rho(\tilde{\mathbf{s}})$, which is derived by replacing $|\rho(\mathbf{s})|$ by a dual variable ρ' and solving the inequality constrained problem

$$\begin{aligned} \min_{\mathbf{s}} \quad & \frac{1}{2} \|\mathbf{s} - \tilde{\mathbf{s}}\|^2 + \eta \rho', \\ \text{s.t.} \quad & \rho(\mathbf{s}) - \rho' \leq 0, \\ & -\rho(\mathbf{s}) - \rho' \leq 0, \end{aligned} \quad (23)$$

where $\eta \triangleq \alpha_D \theta_{24}/\sigma$. Its solution is obtained by checking when the KKT conditions [42] of the dual formulation (23) hold for the three possible cases of $\rho(\mathbf{s}^*)$ in the primal formulation; $\rho(\mathbf{s}^*) \leq 0$, $\rho(\mathbf{s}^*) = 0$ and $\rho(\mathbf{s}^*) \geq 0$. These steps are shown in detail in [3], and yield

$$\mathbf{s}^* = \tilde{\mathbf{s}} + \begin{cases} \eta \mathbf{a}, & \rho(\tilde{\mathbf{s}}) < -\eta \mathbf{a}^T \mathbf{a} \\ -\eta \mathbf{a}, & \rho(\tilde{\mathbf{s}}) > \eta \mathbf{a}^T \mathbf{a} \\ -\mathbf{a} \rho(\tilde{\mathbf{s}})/(\mathbf{a}^T \mathbf{a}), & |\rho(\tilde{\mathbf{s}})| \leq \eta \mathbf{a}^T \mathbf{a} \end{cases} \quad (24)$$

Using the abbreviations in (21), this expression for \mathbf{s}^* inserted in (20) to get the corresponding thresholding steps for the update

$$\boldsymbol{\lambda}^{(D24),n+1} = \begin{cases} -\alpha_D \theta_{24} \mathbf{a}, & \rho(\tilde{\boldsymbol{\lambda}}^{(D24)}/\sigma) < -\eta \mathbf{a}^T \mathbf{a} \\ \alpha_D \theta_{24} \mathbf{a}, & \rho(\tilde{\boldsymbol{\lambda}}^{(D24)}/\sigma) > \eta \mathbf{a}^T \mathbf{a} \\ \sigma \frac{\rho(\tilde{\boldsymbol{\lambda}}^{(D24)}/\sigma)}{\mathbf{a}^T \mathbf{a}} \mathbf{a}, & |\rho(\tilde{\boldsymbol{\lambda}}^{(D24)}/\sigma)| \leq \eta \mathbf{a}^T \mathbf{a} \end{cases} \quad (25)$$

385 Another enlightening interpretation of the solution to (22) is given in [20].

A.3. The proximal operator of σE_{T1}^*

The Moreau decomposition (19) is used again as in (20) to derive the proximal operator of σE_{T1}^* , which gives $\boldsymbol{\lambda}^{(T1),n+1} = \tilde{\boldsymbol{\lambda}}^{(T1)} - \sigma \mathbf{s}^*$, where in this case

$$\mathbf{s}^* = \mathbf{prox}_{\frac{1}{\sigma} E_{T1}} \left(\frac{\tilde{\boldsymbol{\lambda}}^{(T1)}}{\sigma} \right) \leftarrow \min_{\mathbf{s}} \left(\frac{\alpha_T}{\sigma} \|\mathbf{s} - \mathbf{u}_1^{k+1}\| + \frac{1}{2} \left\| \mathbf{s} - \frac{\tilde{\boldsymbol{\lambda}}^{(T1)}}{\sigma} \right\|^2 \right), \quad (26)$$

and $\tilde{\boldsymbol{\lambda}}^{(T1)} = \boldsymbol{\lambda}^{(T1),n} + \sigma K_{T1} \mathbf{z}^n = \boldsymbol{\lambda}^{(T1),n} + \sigma(z_u^n, z_v^n)$. Depending on the magnitude of α_T/σ relative to the distance between \mathbf{u}_1^{k+1} and $\tilde{\boldsymbol{\lambda}}^{(T1)}/\sigma$, either the L^2 expression or the L^2 -squared will dominate in the minimization of (26), according to

$$\mathbf{s}^* = \begin{cases} \mathbf{u}_1^{k+1}, & \alpha_T \geq \|\mathbf{d}\| \\ \mathbf{u}_1^{k+1} + (\mathbf{d}/\sigma)(1 - \alpha_T/\|\mathbf{d}\|), & \text{otherwise} \end{cases} \quad (27)$$

introducing $\mathbf{d} \triangleq \tilde{\boldsymbol{\lambda}}^{(T1)} - \sigma \mathbf{u}_1^{k+1}$, and thus

$$\boldsymbol{\lambda}^{(T1),n+1} = \begin{cases} \mathbf{d}, & \alpha_T \geq \|\mathbf{d}\| \\ \alpha_T \mathbf{d} / \|\mathbf{d}\|, & \text{otherwise} \end{cases} \quad (28)$$

A.4. The proximal operator of σE_{S2}^*

The proximal operator of σE_{S2}^* is given (without using the Moreau decomposition) by

$$\boldsymbol{\lambda}^{(S2),n+1} \leftarrow \min_{\boldsymbol{\lambda}^{(S2)}} \left(\sigma E_{S2}^*(\boldsymbol{\lambda}^{(S2)}) + \frac{1}{2} \|\boldsymbol{\lambda}^{(S2)} - \tilde{\boldsymbol{\lambda}}^{(S2)}\|^2 \right), \quad (29)$$

where $\tilde{\boldsymbol{\lambda}}^{(S2)} = \boldsymbol{\lambda}^{(S2),n} + \sigma K_{S2} \mathbf{z}^n = \boldsymbol{\lambda}^{(S2),n} + \sigma \nabla(z_u^n, z_v^n)$ and

$$\sigma E_{S2}^*(\boldsymbol{\lambda}^{(S2)}) = \delta_P(\boldsymbol{\lambda}^{(S2)}) = \begin{cases} 0, & \boldsymbol{\lambda}^{(S2)} \in P \\ \infty, & \text{otherwise} \end{cases} \quad (30)$$

where $P = \{\boldsymbol{\lambda}^{(S2)} \in \mathbb{R}^4 : \|\boldsymbol{\lambda}^{(S2)}\|/\alpha_S \leq 1\}$ [42, 22]. Finally,

$$\boldsymbol{\lambda}^{(S2),n+1} = \Pi_P(\tilde{\boldsymbol{\lambda}}^{(S2)}) = \frac{\tilde{\boldsymbol{\lambda}}^{(S2)}}{\max(1, \|\tilde{\boldsymbol{\lambda}}^{(S2)}\|/\alpha_S)}, \quad (31)$$

where Π_P denotes the projection onto the set P .

A.5. The proximal operator of τG

It remains to show the proximal operator for τG (the expression for G is given in (9),

$$\mathbf{w}_2^{n+1} \leftarrow \min_{\mathbf{w}_2} \left(\frac{\tau \alpha_G}{2} \|\mathbf{w}_2 - \mathbf{w}_2^k\|^2 + \frac{1}{2} \|\mathbf{w}_2 - \tilde{\mathbf{w}}_2\|^2 \right), \quad (32)$$

where $\tilde{\mathbf{w}}_2 = \mathbf{w}_2^n - \tau K^* \boldsymbol{\lambda}^{n+1}$. In this case, the minimizer is given directly by differentiation, as

$$\mathbf{w}_2^{n+1} = (1 + \tau \alpha_G)^{-1} (\tau \alpha_G \mathbf{w}_2^k + \tilde{\mathbf{w}}_2). \quad (33)$$

The expression for the proximal point $\tilde{\mathbf{w}}_2 = (\tilde{\mathbf{u}}_2, \tilde{l}_2)$ is given by evaluating $K^* \boldsymbol{\lambda}^{n+1}$ in (17). Thus,

$$\begin{aligned} \tilde{\mathbf{u}}_2 &= \mathbf{u}_2^n - \tau (\boldsymbol{\lambda}_{u,v}^{(D13),n+1} + \boldsymbol{\lambda}_{u,v}^{(D24),n+1} + \boldsymbol{\lambda}_{u,v}^{(D23),n+1} + \\ &\quad + \boldsymbol{\lambda}^{(T1),n+1} + \boldsymbol{\lambda}^{(T2),n+1} + \boldsymbol{\lambda}^{(M2),n+1} - \operatorname{div} \boldsymbol{\lambda}^{(S2),n+1}) \\ \tilde{l}_2 &= l_2^n - \tau (\boldsymbol{\lambda}_l^{(D13),n+1} + \boldsymbol{\lambda}_l^{(D24),n+1} + \boldsymbol{\lambda}_l^{(D23),n+1} - \operatorname{div} \boldsymbol{\lambda}^{(L2),n+1}), \end{aligned} \quad (34)$$

390 where subscript u, v for $\boldsymbol{\lambda}^{(D13)} \in \mathbb{R}^3$, etc., denote the two elements of the dual variable that correspond to (u_2, v_2) , and $\boldsymbol{\lambda}_l^{(D13)}$ denotes the third element, i.e. the dual of l_2 in E_{D13} .

B. TGV2 and CSAD update equations

If the TV expression (also known as the first order TGV) in the spatial regularization term (5b) is replaced by the second order TGV2, we get

$$E_S = \min_{\mathbf{q}_f} \sum_{f=1}^3 \int_{\Omega} \alpha_S \|\nabla \mathbf{u}_f - \mathbf{q}_f\| + \alpha_{S2} \|\nabla \mathbf{q}_f\| \, d\mathbf{x}. \quad (35)$$

In the resulting update equations, $\tilde{\boldsymbol{\lambda}}^{(S2)}$ in (29) is changed to $\tilde{\boldsymbol{\lambda}}^{(S2)} = \boldsymbol{\lambda}^{(S2),n} + \sigma(\nabla(z_u^n, z_v^n) - \bar{\mathbf{q}})$, omitting the subscript for \mathbf{q} . Equation (31) for $\boldsymbol{\lambda}^{(S2),n+1}$ does not change. The following set of update equations are added for \mathbf{q} , its over-relaxation expression $\bar{\mathbf{q}}$ and its dual $\boldsymbol{\lambda}_{\mathbf{q}}$:

$$\begin{aligned}\tilde{\boldsymbol{\lambda}}_{\mathbf{q}} &= \boldsymbol{\lambda}_{\mathbf{q}}^n + \sigma \nabla \bar{\mathbf{q}}^n, \\ \boldsymbol{\lambda}_{\mathbf{q}}^{n+1} &= \frac{\tilde{\boldsymbol{\lambda}}_{\mathbf{q}}}{\max(1, \|\tilde{\boldsymbol{\lambda}}_{\mathbf{q}}\|/\alpha_{S2})}, \\ \mathbf{q}^{n+1} &= \mathbf{q}^n + \tau(\boldsymbol{\lambda}^{(S2),n+1} + \operatorname{div} \boldsymbol{\lambda}_{\mathbf{q}}^{n+1}), \\ \bar{\mathbf{q}}^{n+1} &= 2\mathbf{q}^{n+1} - \mathbf{q}^n.\end{aligned}\tag{36}$$

Details of how to implement e.g. the derivative operators numerically are specified in [25].

Next, the update equations are derived for the CSAD data term introduced in [20]. For each of the data terms in (5a), as exemplified for E_{D24} in (22), the optimum \mathbf{s}^* is given by

$$\mathbf{s}^* \leftarrow \min_{\mathbf{s}} \left(\eta |a_0 + \mathbf{a}^T \mathbf{s}| + \frac{1}{2} \|\mathbf{s} - \tilde{\mathbf{s}}\|^2 \right),\tag{37}$$

This can be reduced to a one dimensional problem in the direction of \mathbf{a} [43, 20], such that $\mathbf{s}^* = \tilde{\mathbf{s}} + \delta^* \mathbf{a} / \|\mathbf{a}\|$, where

$$\delta^* = \arg \min_{\delta} \left(\eta \|\mathbf{a}\| \left| \frac{a_0 + \mathbf{a}^T \tilde{\mathbf{s}}}{\|\mathbf{a}\|} + \delta \right| + \frac{1}{2} \delta^2 \right).\tag{38}$$

This is a problem of the form

$$\arg \min_{\delta} \left(\sum_{i=1}^1 \nu_i |\delta - b_i| + \xi (\delta - f)^2 \right),\tag{39}$$

in [43]. Identifying $\xi = 1/2$, $f = 0$, its solution is given by

$$\delta^* = \operatorname{median}\{b_1, c_0, c_1\},\tag{40}$$

where $b_1 = -(a_0 + \mathbf{a}^T \tilde{\mathbf{s}} / \|\mathbf{a}\|)$, $c_0 = \nu_1 = \eta \|\mathbf{a}\|$, $c_1 = -\nu_1 = -\eta \|\mathbf{a}\|$. Inserting into $\mathbf{s}^* = \tilde{\mathbf{s}} + \delta^* \mathbf{a} / \|\mathbf{a}\|$ gives the same expression as derived in (24). Exemplifying

for the CSAD data term corresponding to E_{D24} , its linearized expression is

$$\begin{aligned} \frac{\text{CSAD}_{24}(\mathbf{s}, \mathbf{x})}{\sigma} &= \sum_{\mathbf{x}' \in \mathcal{N}(\mathbf{x})} \eta |a_0(\mathbf{x}') + \mathbf{a}^T \mathbf{s}|, \\ a_0(\mathbf{x}') &= I_4(\mathbf{x} + \mathbf{u}_2^0(\mathbf{x}) + \mathbf{u}_3^0(\mathbf{x})) - I_2(\mathbf{x}) - \\ &\quad - I_4(\mathbf{x}' + \mathbf{u}_2^0(\mathbf{x}) + \mathbf{u}_3^0(\mathbf{x})) - I_2(\mathbf{x}') + \\ &\quad + \mathbf{a}^T (\mathbf{u}_3^k(\mathbf{x}) - \mathbf{u}_2^0(\mathbf{x}) - \mathbf{u}_3^0(\mathbf{x})), \end{aligned} \quad (41)$$

where $\mathcal{N}(\mathbf{x})$ is a 3×3 neighborhood of \mathbf{x} . Notice that the illumination terms are removed, and thus $\mathbf{a} = \nabla I_4$. Similar to (37), we get

$$\mathbf{s}^* \leftarrow \min_{\mathbf{s}} \left(\sum_{\mathbf{x}' \in \mathcal{N}(\mathbf{x})} \eta |a_0(\mathbf{x}') + \mathbf{a}^T \mathbf{s}| + \frac{1}{2} \|\mathbf{s} - \tilde{\mathbf{s}}\|^2 \right), \quad (42)$$

which using the solution formula to (39) given in [43] gives

$$\begin{aligned} \delta^* &= \arg \min_{\delta} \left(\sum_{i=1}^1 \eta \|\mathbf{a}\| \left| \frac{a_0(\mathbf{x}') + \mathbf{a}^T \tilde{\mathbf{s}}}{\|\mathbf{a}\|} + \delta \right| + \frac{1}{2} \delta^2 \right) = \\ &= \text{median}\{b_1, \dots, b_8, c_0, \dots, c_8\}, \end{aligned} \quad (43)$$

where each $b_i = -(a_0(\mathbf{x}') + \mathbf{a}^T \tilde{\mathbf{s}} / \|\mathbf{a}\|)$, and the indexes $i = 1, \dots, 8$ correspond to the 8 neighboring points \mathbf{x}' of a given point \mathbf{x} . The c_i are all multiples of $\eta \|\mathbf{a}\|$ (independent of i) by the factors $\{8, 6, 4, 2, 0, -2, -4, -6, -8\}$. The CSAD variants of E_{D13}, E_{D23} are similar to (41) and the solution follows the same
400 derivation, but with different a_0 and \mathbf{a} .

Acknowledgements

This work is funded by the Swedish research agency VINNOVA under project 2013-04702, and by Volvo Cars.

References

- 405 [1] B. Horn, B. Schunck, Determining optical flow, Artificial intelligence 17 (1) (1981) 185–203.

- [2] J. Barron, D. Fleet, S. Beauchemin, Performance of optical flow techniques, *International Journal of Computer Vision* 12 (1) (1994) 43–77. doi:10.1007/BF01420984.
 410 URL <http://dx.doi.org/10.1007/BF01420984>
- [3] A. Wedel, D. Cremers, *Stereo scene flow for 3D motion analysis*, Springer, 2011.
- [4] D. Cremers, S. Soatto, Motion competition: A variational approach to piecewise parametric motion segmentation, *International Journal of Computer Vision* 62 (3) (2005) 249–265.
 415
- [5] A. Giachetti, M. Campani, V. Torre, The use of optical flow for road navigation, *IEEE Transactions on Robotics and Automation* 14 (1) (1998) 34–48.
- [6] W. Crum, T. Hartkens, D. Hill, Non-rigid image registration: theory and practice, *The British Journal of Radiology* 77 (2004) 140–153.
 420
- [7] E. Allen, S. Triantaphillidou, *The Manual of Photography and Digital Imaging*, Taylor & Francis, 2012.
- [8] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, K. Myszkowski, *High dynamic range imaging: acquisition, display, and image-based lighting*, Morgan Kaufmann, 2010.
 425
- [9] D. Geronimo, A. Lopez, A. D. Sappa, T. Graf, Survey of pedestrian detection for advanced driver assistance systems, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32 (7) (2010) 1239–1258.
- [10] O. Tursun, A. Akyüz, A. Erdem, E. Erdem, The state of the art in hdr deghosting: A survey and evaluation, in: *Computer Graphics Forum*, Vol. 32, 2015, pp. 348–362.
 430
- [11] K. Hadziabdic, J. Telalovic, R. Mantiuk, Comparison of deghosting algorithms for multi-exposure high dynamic range imaging, in: *Proceedings of the 29th Spring Conference on Computer Graphics*, ACM, 2013, pp. 21–28.

- 435 [12] H. Zimmer, A. Bruhn, J. Weickert, Freehand HDR imaging of moving scenes with simultaneous resolution enhancement, *Computer Graphics Forum* 30 (2) (2011) 405–414.
- [13] D. Hafner, O. Demetz, J. Weickert, Simultaneous hdr and optic flow computation, in: *International Conference on Pattern Recognition (ICPR)*, 440 IEEE, 2014, pp. 2065–2070.
- [14] T. Brox, A. Bruhn, N. Papenberg, J. Weickert, High accuracy optical flow estimation based on a theory for warping, in: *European Conference on Computer Vision (ECCV)*, Springer Berlin Heidelberg, 2004, pp. 25–36.
- [15] T. Bengtsson, T. McKelvey, K. Lindström, Optical flow estimation on im- 445 age sequences with differently exposed frames, *Optical Engineering* 54 (9).
- [16] J. Weickert, C. Schnörr, A theoretical framework for convex regularizers in pde-based computation of image motion, *International Journal of Computer Vision* 45 (3) (2001) 245–264.
- [17] A. Salgado, J. Sánchez, Temporal constraints in large optical flow estima- 450 tion, in: *Computer Aided Systems Theory–EUROCAST 2007*, Springer, 2007, pp. 709–716.
- [18] S. Volz, A. Bruhn, L. Valgaerts, H. Zimmer, Modeling temporal coherence for optical flow, in: *IEEE International Conference on Computer Vision (ICCV)*, 2011, pp. 1116–1123.
- 455 [19] M. Black, P. Anandan, The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields, *Computer vision and image understanding* 63 (1) (1996) 75–104.
- [20] C. Vogel, S. Roth, K. Schindler, An evaluation of data costs for optical flow, in: *German Conference on Pattern Recognition (GPCR)*, Springer, 460 2013, pp. 343–353.

- [21] T. Brox, J. Malik, Large displacement optical flow: descriptor matching in variational motion estimation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33 (3) (2011) 500–513.
- [22] A. Chambolle, T. Pock, A first-order primal-dual algorithm for convex
465 problems with applications to imaging, *Journal of Mathematical Imaging and Vision* 40 (1) (2011) 120–145.
- [23] M. Zhu, T. Chan, An efficient primal-dual hybrid gradient algorithm for total variation image restoration, *UCLA CAM Report* (2008) 08–34.
- [24] E. Esser, X. Zhang, T. Chan, A general framework for a class of first order
470 primal-dual algorithms for convex optimization in imaging science, *SIAM Journal on Imaging Sciences* 3 (4) (2010) 1015–1046.
- [25] K. Bredies, Recovering piecewise smooth multichannel images by minimization of convex functionals with total generalized variation penalty, in: *Efficient Algorithms for Global Optimization Methods in Computer Vision*,
475 Springer, 2014, pp. 44–77.
- [26] D. Shulman, J.-Y. Hervé, Regularization of discontinuous flow fields, in: *Workshop on Visual Motion*, 1989, pp. 81–86.
- [27] O. Demetz, M. Stoll, S. Volz, J. Weickert, A. Bruhn, Learning brightness transfer functions for the joint recovery of illumination changes and optical
480 flow, in: *Computer Vision–ECCV 2014*, Springer, 2014, pp. 455–471.
- [28] J. Revaud, P. Weinzaepfel, Z. Harchaoui, C. Schmid, Epicflow: Edge-preserving interpolation of correspondences for optical flow, *arXiv preprint arXiv:1501.02565*.
- [29] M. Muja, D. Lowe, Fast approximate nearest neighbors with automatic
485 algorithm configuration, in: *International Conference on Computer Vision Theory and Applications (VISAPP)*, 2009, pp. 331–340.

- [30] C. Harris, M. Stephens, A combined corner and edge detector., in: Alvey vision conference, Vol. 15, Citeseer, 1988, p. 50.
- [31] M. Stoll, S. Volz, A. Bruhn, Adaptive integration of feature matches into
 490 variational optical flow methods, in: Asian Conference on Computer Vision (ACCV), Springer, 2013, pp. 1–14.
- [32] V. Estellers, S. Soatto, Detecting occlusions as an inverse problem, *Journal of Mathematical Imaging and Vision* (2015) 1–18.
- [33] N. Parikh, S. Boyd, Proximal algorithms, *Foundations and Trends in Op-
 495 timization* 1 (3) (2013) 123–231.
- [34] Computer vision group, University of Freiburg, Image sequences.
 URL <http://lmb.informatik.uni-freiburg.de/resources/datasets/sequences.en.html>
- [35] D. Butler, J. Wulff, G. Stanley, M. Black, A naturalistic open source movie
 500 for optical flow evaluation, in: European Conference on Computer Vision (ECCV), Springer, 2012, pp. 611–625.
- [36] S. Baker, D. Scharstein, J. Lewis, S. Roth, M. Black, R. Szeliski, A database and evaluation methodology for optical flow, *International Journal of Computer Vision* 92 (1) (2011) 1–31.
- [37] P. Weinzaepfel, J. Revaud, Z. Harchaoui, C. Schmid, et al., Deepflow: Large
 505 displacement optical flow with deep matching, in: IEEE International Conference on Computer Vision (ICCV), 2013.
- [38] A. Geiger, P. Lenz, C. Stiller, R. Urtasun, Kitti vision benchmark suite.
 URL <http://www.cvlibs.net/datasets/kitti/>
- [39] A. Geiger, P. Lenz, R. Urtasun, Are we ready for autonomous driving? the
 510 kitti vision benchmark suite, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012, pp. 3354–3361.

- [40] T. Müller, C. Rabe, J. Rannacher, U. Franke, R. Mester, Illumination-robust dense optical flow using census signatures, in: Pattern Recognition, Springer, 2011, pp. 236–245.
- [41] T. Rockafellar, Convex analysis, Princeton university press, 1997.
- [42] S. Boyd, L. Vandenberghe, Convex optimization, Cambridge university press, 2009.
- [43] Y. Li, S. Osher, A new median formula with applications to pde based denoising, Commun. Math. Sci 7 (3) (2009) 741–753.