A GENERAL FORM OF ILLUMINATION-INVARIANT DESCRIPTORS IN VARIATIONAL OPTICAL FLOW ESTIMATION

Dinh-Hoan Trinh, Walter Blondel, and Christian Daul

Université de Lorraine, CRAN, UMR 7039
2 avenue de la Forêt de Haye, 54518 Vandœuvre-lès-Nancy, France
CNRS, CRAN, UMR 7039, 54518 Vandœuvre-lès-Nancy cedex, France

ABSTRACT

This paper introduces a generalized descriptor formulation facilitating the design of illumination invariant data-terms in variational optical flow. This contribution also proposes a criterion to check whether a patch-based descriptor is illumination invariant or not. To do so, a local model is used to simulate complex illumination changes between images. As an example, it is shown how a novel patch-based descriptor can be derived from the generalized form. The performances of this descriptor are compared to those of reference descriptors in the literature using data sets with and without strong illumination changes. The accuracy and robustness of the proposed descriptor is also demonstrated through tests on gastroscopic image sequences including complex illumination changes.

***Index Terms*—** Variational optical flow, illumination-invariant descriptor, endoscopic image registration.

1. INTRODUCTION

Although variational Optical Flow (OF) has been studied for decades in numerous applications, estimating accurately OF under complex illumination changes is still a challenge. Generally, variational OF methods compute the dense flow field $\mathbf{u} = (u_x, u_y)$ between source image I_s and target image I_t by optimizing

$$E(\mathbf{u}) = E_{reg}(\mathbf{u}) + \lambda E_{data}(I_s, I_t, \mathbf{u}), \tag{1}$$

where E_{reg} is a regularization term that assumes smoothness of solution $\mathbf{u}$, E_{data} is a data-term that measures the similarity of pixels in I_s and I_t, and $\lambda > 0$ is a parameter controlling the relative importance of the data- and regularization terms.

Among the numerous formulations proposed in the literature for E_{reg}, regularization terms using the total variation [1], the non-local total variation [2–4], or the non-local total generalized variation [5] of the flow are known as the most effective ones. They notably allow for discontinuities in the flow field.

The data-term, E_{data}, is related to illumination changes but the well known brightness constancy assumption (BCA) [6] is no longer valid for complex ones. For this reason, several other data-terms were proposed combining BCA and the gradient constancy assumption and/or higher order constancy assumptions such as the Laplacian and the Hessian [7–9]. In other approaches, an image preprocessing step was used prior to OF estimation [10–13]. In real scenes (e.g. outdoor or medical) the illumination changes between consecutive images of a sequence are very complex, but the nature of the illumination variation is barely constant for all homologous pixels of two images. Thus, it cannot be represented effectively by a global constancy assumption or treated effectively by a preprocessing step.

More recently, descriptor-based OF methods have proved their effectiveness in handling the illumination change issue. The underlying idea of these methods is to construct a robust data-term based on descriptors mathematically represented by vectors whose components give an illumination invariant representation of a small image region. Some works [14–16] are based on binary-descriptors, other ones used real value vectors to define descriptors [17, 18]. However, although these descriptors were mathematically defined in most contributions, their performances under changing illumination was only validated by experiments. In the existing literature, no theoretical framework is provided neither to understand why such descriptors are independent to a given illumination change model nor to facilitate the design of new descriptors.

This paper proposes a local model representing effectively complex illumination changes (Section 2). This local model can be used as a required condition to check whether a descriptor is illumination invariant or not. Moreover, in Section 3, a generalized formulation for a class of illumination-invariant descriptors is introduced and a novel illumination-invariant descriptor is derived from it. In Section 4, the results obtained with the new descriptor are compared to those of well-known descriptors to illustrate the effectiveness of the proposed approach in terms of OF accuracy.

This work was funded by the Agence Nationale de la Recherche in the frame of the EMMIE project (ANR-15-CE17-0015).

2. ILLUMINATION INVARIANCE CONDITION

Let us first consider the problem of modelling illumination changes between consecutive images of a sequence. An appropriate illumination change model has to reflect as closely as possible complex variations as occurring in outdoor or medical scenes, for instance. Negahdaripour [19] proposed the Generalized Dynamic Image Model (GDIM) to simulate the illumination changes. The GDIM is a pixel-wise model given by

$$I_t(\mathbf{x} + \mathbf{u}_\mathbf{x}) = a_\mathbf{x} I_s(\mathbf{x}) + b_\mathbf{x}, \tag{2}$$

where $a_\mathbf{x}$ and $b_\mathbf{x}$ are two real parameters of the model. This model is able to capture complex variations. However, as shown in [20], integrating this model into a variational calculus algorithm leads to a complex optimization problem and high computation times since, besides a flow vector, parameters $a_\mathbf{x}$ and $b_\mathbf{x}$ have to be estimated for each pixel.

Usually, two consecutive images in a sequence have the property of local stationarity such that all pixels in a small image region may share the same parameters in the illumination variation model. Therefore, the assumption is made that the illumination changes of all corresponding pixels in small homologous neighborhoods in images I_s and I_t can be accurately represented by a model with constant parameters. Thus, instead of using the model in (2) at pixel-level, a patch-based model is used to describe locally the illumination changes between homologous neighborhoods:

$$P_{I_t}(\mathbf{x} + \mathbf{u}_\mathbf{x}) = a_\mathbf{x} P_{I_s}(\mathbf{x}) + b_\mathbf{x} \tag{3}$$

with $a_\mathbf{x} \in \mathbb{R}_{>0}, b \in \mathbb{R}$, $P_I(\mathbf{x})$ denotes the patch centered at pixel $\mathbf{x}$ in image I, and $\mathbf{u}_\mathbf{x}$ is the displacement vector from pixel $\mathbf{x}$ in source image I_s. Unlike in the GDIM model, parameter $a_\mathbf{x}$ is assumed to be greater than 0, since the intensity values of pixels are non-negative. Rather than computing the parameter values $a_\mathbf{x}$ and $b_\mathbf{x}$, the model (3) is only used to design illumination independent descriptors.

From (3) one can see that a descriptor $\mathbf{D}$ is invariant to illumination changes if it satisfies the condition

$$\mathbf{D}(a_\mathbf{x} P_I(\mathbf{x}) + b_\mathbf{x}) = \mathbf{D}(P_I(\mathbf{x})), \tag{4}$$

for all pixel $\mathbf{x}$ in image I, and for all $a_\mathbf{x} \in \mathbb{R}_{>0}$, $b_\mathbf{x} \in \mathbb{R}$. Here, $\mathbf{D}(P_I(\mathbf{x})) \in \mathbb{R}^m$ stands for the descriptor vector computed for patch $P_I(\mathbf{x})$ centered on pixel $\mathbf{x}$. Equation (4) could be considered as the necessary condition for illumination-invariant descriptors.

3. A GENERAL FORM OF ILLUMINATION-INVARIANT DESCRIPTORS

This section presents a generalized form of illumination-invariant descriptors based on the elimination of the model parameters $a_\mathbf{x}$ and $b_\mathbf{x}$ in (4). The aim is to construct descriptor

$$\mathrm{M}_4 = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -2 \end{bmatrix} \quad \mathrm{M}_3 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix} = \mathrm{M}_2$$

$$\mathrm{M}_5 = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad \begin{bmatrix} I(\mathrm{x}_4) & I(\mathrm{x}_3) & I(\mathrm{x}_2) \\ I(\mathrm{x}_5) & I(\mathrm{x}_0) & I(\mathrm{x}_1) \\ I(\mathrm{x}_6) & I(\mathrm{x}_7) & I(\mathrm{x}_8) \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} = \mathrm{M}_1$$

$$\mathrm{M}_6 = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix} \quad \mathrm{M}_7 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \mathrm{M}_8$$

Fig. 1. Robinson kernels used to define the NLDP descriptor. The patch corresponds to a 3×3 neighborhood centered on pixel $\mathbf{x}_0$.

$\mathbf{D}$ such that:

$$\mathbf{D}^i(a_\mathbf{x} P_I(\mathbf{x}) + b_\mathbf{x}) = \mathbf{D}^i(P_I(\mathbf{x})), \forall i = 0, 1, \ldots, m, \tag{5}$$

where $\mathbf{D}^i$ is the i-th component of descriptor vector $\mathbf{D}$.

To this end, $\mathbf{D}^i$ is defined as follows:

$$\mathbf{D}^i(P_I(\mathbf{x})) = \Psi\left(\frac{g_{1,i}(P_I(\mathbf{x}))}{g_{2,i}(P_I(\mathbf{x}))}\right), \tag{6}$$

where $\Psi(\cdot)$ is a non-constant function and $g_{1,i}, g_{2,i} : \mathbb{R}^{n+1} \to \mathbb{R}$ are functions such that

$$g_{1,i}(a_\mathbf{x} P_I(\mathbf{x}) + b_\mathbf{x}) = h(a_\mathbf{x}) g_{1,i}(P_I(\mathbf{x})), \tag{7}$$

$$g_{2,i}(a_\mathbf{x} P_I(\mathbf{x}) + b_\mathbf{x}) = h(a_\mathbf{x}) g_{2,i}(P_I(\mathbf{x})). \tag{8}$$

Note that the fraction in (6) is set to 0 when denominator $g_{2,i}(P_I(\mathbf{x}))$ equals 0. Then, for each component i of a descriptor, we have

$$\mathbf{D}^i(a_\mathbf{x} P_I(\mathbf{x}) + b_\mathbf{x}) = \Psi\left(\frac{g_{1,i}(P_I(\mathbf{x}))}{g_{2,i}(P_I(\mathbf{x}))}\right) = \mathbf{D}^i(P_I(\mathbf{x})) \tag{9}$$

and consequently, $\mathbf{D}(a_\mathbf{x} P_I(\mathbf{x}) + b_\mathbf{x}) = \mathbf{D}(P_I(\mathbf{x}))$. In other words, the descriptor defined by (6)-(8) is invariant to local illumination changes given in (3).

Many functions can be used as $g_{1,i}$ and $g_{2,i}$ in (7) and (8). Let us consider patch $P_I(\mathbf{x}_0) = \{I(\mathbf{x}_i)\}_{i=0}^n$ centered on pixel $\mathbf{x}_0$. A first function type is given by:

$$g(P_I(\mathbf{x}_0)) = \sum_{i=0}^{n} \alpha_i I(\mathbf{x}_i), \tag{10}$$

where n is the number of pixels in $P_I(\mathbf{x}_0)$, $\{\alpha_i\}_{i=0}^n$ is a non-zero sequence and $\sum_{i=0}^n \alpha_i = 0$. In this case, $g(a_{\mathbf{x}_0} P_I(\mathbf{x}_0) + b_{\mathbf{x}_0}) = a_{\mathbf{x}_0} g(P_I(\mathbf{x}_0))$. The second function type is given by:

$$g(P_I(\mathbf{x}_0)) = \gamma \left(\sum_{i=0}^{N} \left(\sum_{j=0}^{n} \alpha_{i,j} I(\mathbf{x}_j) \right)^\tau \right)^\eta, \tag{11}$$

where $\{\alpha_{i,j}\}_{j=0}^n$, $i = 0, 1, \ldots, N$ are non-zero and zero-sum sequences, and γ, τ and η are fixed positive numbers. In this case, $g(a_{\mathbf{x}_0} P_I(\mathbf{x}_0) + b_{\mathbf{x}_0}) = (a_{\mathbf{x}_0})^{\tau\eta} g(P_I(\mathbf{x}_0))$.

Table 1. AEPE and AAE values obtained with different descriptors under weak illumination changes.

Descriptor	Dimetrodon		Grove2		Grove3		Hydrangea		RubberWhale		Urban2		Urban3		Venus	
	AEPE	AAE	AEPE	AAE	AEPE	AAE	AEPE	AAE	AEPE	AAE	AEPE	AAE	AEPE	AAE	AEPE	AAE
MLDP	0.13	2.35	**0.13**	1.84	0.48	5.02	**0.17**	2.06	**0.08**	2.71	**0.33**	**3.18**	0.57	4.18	**0.26**	**3.68**
Corr	0.23	4.86	0.20	2.53	0.49	5.21	**0.17**	**2.01**	**0.08**	**2.60**	0.34	3.65	0.75	5.67	0.28	4.26
NND	0.22	4.62	0.19	2.77	0.63	6.50	0.18	2.29	0.09	3.06	0.52	3.93	0.59	4.28	0.34	5.30
NLDP	**0.11**	**2.09**	**0.13**	**1.80**	**0.46**	**4.79**	**0.17**	2.07	**0.08**	2.68	0.35	3.30	**0.48**	**3.55**	**0.26**	3.88

Using the general form, several illumination-invariant descriptors can be created. As an example, a novel descriptor named NLDP (Normalized Local Direction Pattern) is proposed. This descriptor is based on the Robinson edge kernels $M_1, M_2, \ldots, M_8$ given in Fig. 1. Robinson kernels can be considered as non-zero and zero-sum sequences of coefficients. At each pixel $\mathbf{x}_0$ in image I, descriptor NLDP is defined based on patch $P_I(\mathbf{x}_0)$ of size 3×3,

$$\mathbf{D}_{NLDP}(P_I(\mathbf{x}_0)) = \frac{\mathcal{A}_{\mathbf{x}_0}}{\|\mathcal{A}_{\mathbf{x}_0}\|_2}, \tag{12}$$

where $\mathcal{A}_{\mathbf{x}_0} = [M_1 \otimes P_I(\mathbf{x}_0), \ldots, M_8 \otimes P_I(\mathbf{x}_0)]^T \in \mathbb{R}^8$ and operator $\otimes$ gives the sum of element-wise products of two matrices. Referring to (6), (7) and (8), the i-th component of $\mathbf{D}_{NLDP}$ is defined by $\Psi(x) = x$, $g_{1,i}(P_I(\mathbf{x}_0)) = M_i \otimes P_I(\mathbf{x}_0)$ and $g_{2,i}(P_I(\mathbf{x}_0)) = \|\mathcal{A}_{\mathbf{x}_0}\|_2$.

4. PERFORMANCE EVALUATION

Experiments are performed on three data types: 1) images of the Middlebury data set [21] with weak illumination changes, 2) images with strong and simulated illumination changes, and 3) gastroscopic images with strong and real changes. The proposed NLDP descriptor is compared to descriptors MLDP [16], Corr [17], and NND [18]. For a fair comparison, all descriptors are placed in the same variational model for OF estimation.

4.1. OF test model

In the experiments, the OF field is estimated by

$$\mathbf{u} = \arg\min_{\mathbf{u}} \left(E_{reg}(\mathbf{u}) + \lambda E_{data}(I_s, I_t, \mathbf{u})\right) \tag{13}$$

where the data and regularization terms are defined as:

$$E_{data} = \sum_{\mathbf{x} \in \Omega} \|\mathbf{D}(P_{I_s}(\mathbf{x})) - \mathbf{D}(P_{I_t}(\mathbf{x} + \mathbf{u}_{\mathbf{x}}))\|_2^2, \tag{14}$$

$$E_{reg} = \sum_{\mathbf{x} \in \Omega} \sum_{\mathbf{x}' \in \mathcal{N}_{\mathbf{x}}} w_{\mathbf{x}}^{\mathbf{x}'} \|\mathbf{u}_{\mathbf{x}} - \mathbf{u}_{\mathbf{x}'}\|_1. \tag{15}$$

In (14) and (15), Ω stands for the image domain, $\mathcal{N}_{\mathbf{x}}$ is the set of neighbor pixels centered on pixel $\mathbf{x}$, and $w_{\mathbf{x}}^{\mathbf{x}'}$ is the weight depending on the similarity of pixels $\mathbf{x}$ and $\mathbf{x}'$. Similarly to [17], the weights $w_{\mathbf{x}}^{\mathbf{x}'}$ are defined as follows:

$$w_{\mathbf{x}}^{\mathbf{x}'} = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|_2^2}{2\sigma_1^2} - \frac{\|L(\mathbf{x}) - L(\mathbf{x}')\|_2^2}{2\sigma_2^2}\right), \tag{16}$$

where σ_1 and σ_2 are parameters controlling the similarity measure, and $L(\mathbf{x})$ is the color vector in the CIE Lab space.

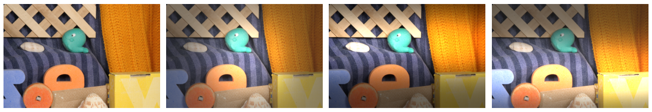
The minimization in (13) is performed with the projected-proximal-point algorithm [22] and the well-known coarse-to-fine warping strategy is used to cope with large displacements. The number of pyramid levels depends on a scale parameter Py_s. At each pyramid level, we use 5 warps and 40 iterations per warp for optimizing the energy. In (14), the default size of patch $P_I(\mathbf{x})$ is 3×3 pixels, except for the patch of descriptor NND where the minimal size is 5×5 (see [18]). The size of $\mathcal{N}_{\mathbf{x}}$ in (15) is systematically set to 5×5, the values of parameters σ_1 and σ_2 in (16) are constantly set to 7.

4.2. Middlebury data set

To compare the performance of the descriptors, the λ-parameter in (13) and the scale parameter (Py_s) of the coarse-to-fine OF approach have to be optimally adjusted. To do so, the OF was estimated on 8 training image pairs (with known ground truth) for different parameter combinations ($\mathrm{Py}_s \in \{0.5, 0.6, \ldots, 0.9\}$ and $\lambda \in \{1, 2, 3, \ldots, 120\}$). The average end-point error (AEPE) and the average angular error (AAE) of the estimated OF were determined for each image pair and (Py_s, λ) combination. Experimental results showed that the optimal values for the pair (Py_s, λ) of descriptors MLDP, Corr, NND and NLDP are (0.5, 10), (0.5, 12), (0.7, 100) and (0.8, 70), respectively. Table 1 gives the AEPE and AAE values obtained with the optimal values of Py_s and λ. As it can be noticed, most of the best results (numbers in bold) were obtained with the proposed NLDP descriptor.

4.3. Synthetic illumination changes

In order to illustrate the effectiveness of the descriptors, the RubberWhale images in the Middlebury data-set were used to simulate strong illumination changes in two image pairs using both multiplicative and additive factors. In Fig. 2, the center of image "Target 1" in (b) is much brighter than the center in image (a), while the image pair (c, d) includes a strong vertical intensity gradient. This test used all parameter



(a) Source 1 (b) Target 1 (c) Source 2 (d) Target 2

Fig. 2. The image pairs with strong illumination changes.

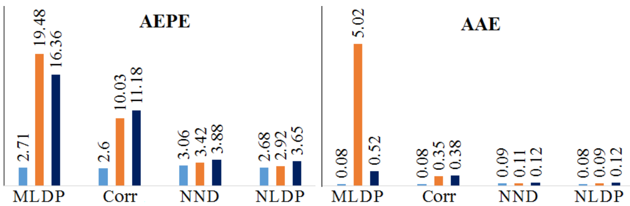
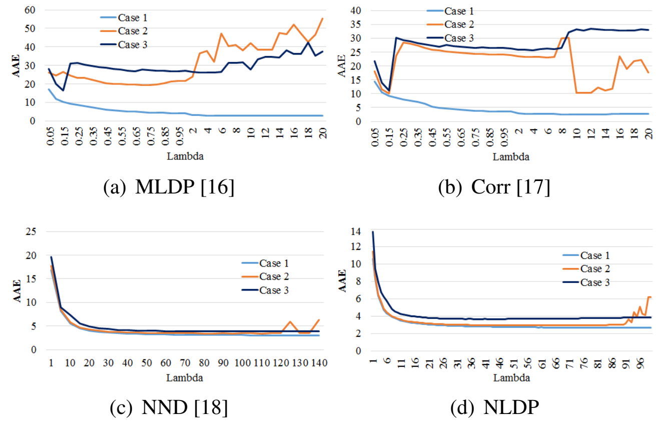
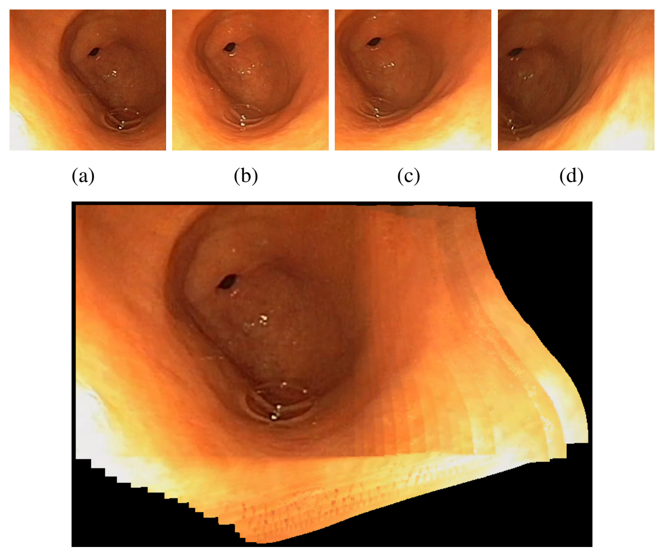


Fig. 3. OF accuracy under three illumination changes with light blue, orange and dark blue bars indicating weak changes in the original RubberWhale images, strong vignetting in Figs. 2(a)-2(b) and strong vertical gradient in Figs. 2(c)-2(d), respectively.



(a) MLDP [16] (b) Corr [17] (c) NND [18] (d) NLDP

Fig. 4. AAE values with respect to the λ-parameter values (the values of Py_s are those adjusted in Section 4.2). The curve colors have the same meaning as in Fig. 3.



(a) (b) (c) (d)

(e) Mosaiced image

Fig. 5. Mosaic built with 17 images of a gastroscopic sequence (4 images of the sequence are also given).

values of Sections 4.1 and 4.2, except the λ value which was adjusted for the strong illumination changes.

Fig. 3 shows the best OF results of the descriptors on the original image pair (weak illumination changes) and the two strong illumination changes given in Fig. 2. It is noticeable that the errors remained small when passing from weak to strong illumination changes only for descriptors NND and NLDP, while for descriptors MLDP and Corr the AEPE and AAE values varied significantly. The AEPE and AAE values of descriptor NLDP were globally the lowest for all illumination conditions. It means that the OF estimated with descriptor NLDP leads to the most accurate and robust results.

The three image pairs were also used to evaluate the impact of the value of parameter λ on the OF accuracy. As demonstrated in Fig. 4, the optimal values of λ corresponding to descriptors NND and NLDP are nearly invariant to illumination changes (i.e. the errors are small and constant for a large range of λ values), while for descriptors MLDP and Corr it is difficult to determine the optimal value of λ due to the large and changing AAE values.

4.4. Descriptor DNLP in gastroscopic image mosaicing

A major challenge when mosaicing endoscopic images [23, 24] is to register them robustly under changing illumination, the images being poorly textured and also affected by specular reflections (see Figs. 5(a)-5(d)). The image registration uses the flow field obtained with descriptor NLDP using (Py_s, λ)=(0.8 and 70) adjusted with Fig. 4. Fig. 5 shows a panoramic image computed with 17 images of the pyloric antrum region. The precise mosaic of Fig. 5(e) (without structure discontinuities) confirms the accuracy and robustness of the proposed descriptor. The illumination discontinues at image borders were intentionally not corrected as in [25] to visualize the data superimposition.

5. CONCLUSION

This paper introduced a general form of illumination-invariant descriptors for robust OF estimation. This general form is based on a novel illumination change model used as a necessary condition during the design of descriptors. A new descriptor (NLDP) was derived from the general form. OF estimation with descriptor DNLDP led to high accuracy, invariance to illumination changes, and easy to adjust the parameters. The effectiveness of the descriptor NLDP for estimating the OF on real and complex data confirmed the relevance of the proposed approach.

6. REFERENCES

[1] C. Zach, T. Pock, and H. Bischof, "A duality based approach for realtime TV-L^1 optical flow," in *Pattern Recognition, 29th DAGM Symposium*, Heidelberg, Germany, 2007, pp. 214–223.

[2] M. Werlberger, T. Pock, and H. Bischof, "Motion estimation with non-local total variation regularization," in *IEEE CVPR*, San Francisco, CA, USA, 2010, pp. 2464–2471.

[3] D. Sun, S. Roth, and M. J. Black, "Secrets of optical flow estimation and their principles," in *IEEE Conf. on Computer Vision and Pattern Recognition*, San Francisco, CA, USA, 2010, pp. 2432–2439.

[4] P. Krähenbühl and V. Koltun, "Efficient nonlocal regularization for optical flow," in *ECCV*, Florence, Italy, 2012, pp. 356–369.

[5] R. Ranftl, K. Bredies, and T. Pock, "Non-local total generalized variation for optical flow estimation," in *ECCV*, Switzerland, 2014, pp. 439–454.

[6] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, no. 1-3, pp. 185–203, 1981.

[7] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, "High accuracy optical flow estimation based on a theory for warping," in *ECCV*, Prague, Czech Republic, 2004, pp. 25–36.

[8] L. Xu, J. Jia, and Y. Matsushita, "Motion detail preserving optical flow estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 9, pp. 1744–1757, 2012.

[9] N. Papenberg, A. Bruhn, T. Brox, S. Didas, and J. Weickert, "Highly accurate optic flow computation with theoretically justified warping," *International Journal of Computer Vision*, vol. 67, no. 2, pp. 141–158, 2006.

[10] Y. Mileva, A. Bruhn, and J. Weickert, "Illumination-robust variational optical flow with photometric invariants," in *Pattern Recognition, 29th DAGM Symposium*, Heidelberg, Germany, 2007, pp. 152–162.

[11] A. Wedel, T. Brox, C. Zach, H. Bischof, and D. Cremers, "An improved algorithm for TV-L^1 optical flow," in *Dagstuhl Motion Workshop*, 2008.

[12] H. Zimmer, A. Bruhn, and J. Weickert, "Optical flow in harmony," *International Journal of Computer Vision*, vol. 93, pp. 368–388, 2011.

[13] A. Kumar, F. Tung, A. Wong, and D. A. Clausi, "A decoupled approach to illumination-robust optical flow estimation," *IEEE Trans. Image Processing*, vol. 22, no. 10, pp. 4136–4147, 2013.

[14] F. Stein, "Efficient computation of optical flow using the census transform," in *26th DAGM Symposium*, Tübingen, Germany, 2004, pp. 79–86.

[15] O. Demetz, D. Hafner, and J. Weickert, "The complete rank transform: A tool for accurate and morphologically invariant matching of structures," in *BMVC 2013*, Bristol, UK, 2013.

[16] M. A. Mohamed, H. A. Rashwan, B. Mertsching, M. A. García, and D. Puig, "Illumination-robust optical flow using a local directional pattern," *IEEE Trans. Circuits Syst. Video Techn.*, vol. 24, no. 9, pp. 1499–1508, 2014.

[17] M. Drulea and S. Nedevschi, "Motion estimation using the correlation transform," *IEEE Trans. Image Processing*, vol. 22, no. 8, pp. 3260–3270, 2013.

[18] S. Ali, C. Daul, E. Galbrun, and W. Blondel, "Illumination invariant optical flow using neighborhood descriptors," *Computer Vision and Image Understanding*, vol. 145, pp. 95–110, 2016.

[19] S. Negahdaripour, "Revised definition of optical flow: Integration of radiometric and geometric cues for dynamic scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 9, pp. 961–979, 1998.

[20] Y. Kim, A. M. Martínez, and A. C. Kak, "Robust motion estimation under varying illumination," *Image Vision and Computing*, vol. 23, no. 4, pp. 365–375, 2005.

[21] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," *International Journal of Computer Vision*, vol. 92, no. 1, pp. 1–31, 2011.

[22] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[23] T. Weibel, C. Daul, R. Rösch, and D. Wolf, "Endoscopic bladder image registration using sparse graph cuts," in *IEEE ICIP*, Hong-Kong, China, 2010, pp. 157–160.

[24] S. Ali, C. Daul, T. Weibel, and W. Blondel, "Fast mosaicing of cystoscopic images from dense correspondence: Combined SURF and TV-L1 optical flow method," in *IEEE ICIP*, Melbourne, Australia, 2013, pp. 1291–1295.

[25] T. Weibel, C. Daul, D. Wolf, and R. Rösch, "Contrast-enhancing seam detection and blending using graph cuts," in *ICPR*, Japan, 2012, pp. 2732–2735.