A General Form of Illumination-Invariant Descriptors in Variational Optical Flow Estimation

Dinh-Hoan Trinh, Walter Blondel, and Christian Daul

Université de Lorraine and CNRS, CRAN, UMR 7039

2 avenue de la Forêt de Haye – TSA 60604, 54518 Vandœuvre-lès-Nancy, France

Scientific challenge

Determination of an accurate and dense optical flow field between two images in weakly textured scenes affected by strong illumination changes.

Contributions

1. Proposal of a criterion to check whether a patch-based descriptor is illumination invariant or not.
2. Theoretical proof of the illumination-invariance of some well-known descriptors (e.g. Census [1], Corr [2], MLDP [3], NND [4]) based on the proposed criterion.
3. Introduction of a generalized descriptor formulation facilitating the design of illumination invariant data-terms in variational optical flow.
4. Proposal of a new illumination-invariant descriptor (**NLDP** – Normalized Local Directional Pattern).
5. Comparison of the NLDP performance with that of the reference descriptors in the literature through experiments on both simulated data (with known ground truth flow) and real medical data (gastroscopic image sequences including complex illumination changes).

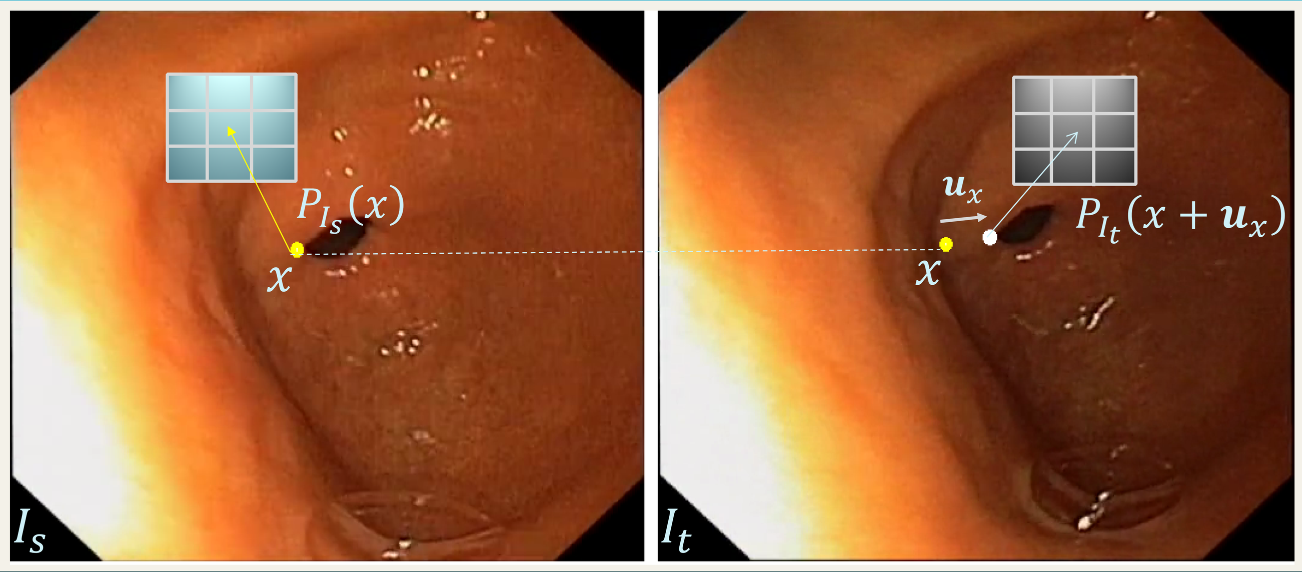
References

[1] R. Zabih, J. Woodll, "Non-parametric local transforms for computing visual correspondence," In: *ECCV*, Stockholm, Sweden, pp. 151-158, 1994.

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

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

[4] 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.



Descriptor-based variational OF Model

$$\min_{u}\left[E_{reg}(u) + \lambda E_{data}(I_s, I_t, u)\right] \quad (1)$$

where

$$E_{reg}(u) = \sum_{x\in\Omega}\sum_{x'\in N_x} w_x^{x'} \|u_x - u_{x'}\|_1 , \quad (2)$$

with

$$w_x^{x'} = exp\left(\frac{-\|x-x'\|_2^2}{2\sigma_1^2} + \frac{-\|L(x)-L(x')\|_2^2}{2\sigma_2^2}\right) \quad (3)$$

and

$$E_{data}(I_s, I_t, u) = \sum_{x\in\Omega}\left\| D\left(P_{I_s}(x)\right) - D\left(P_{I_t}(x+u_x)\right)\right\|_2^2 . \quad (4)$$

Illumination Invariance Condition

- Local/patch region illumination change model
$$P_{I_t}(x+u_x) = a_x P_{I_s}(x) + b_x \quad (5)$$
$$a_x, b_x \in \mathbb{R}, \text{ and } a_x > 0.$$
- Descriptor D is called illumination-invariant when
$$D(P_I(x)) = D(a_x P_I(x) + b_x). \quad (6)$$

General form of Illumination-Invariant Descriptors

- Suppose $D = (D_1, D_2, \dots, D_m)$, the aim is to construct descriptor D such that
$$D_i(P_I(x)) = D_i(a_x P_I(x) + b_x), \quad (7)$$
$$\forall a_x > 0, \forall b_x \in \mathbb{R}, \forall i = 1, \dots, m.$$
- To this end, D_i is defined as follows:
$$D_i(P_I(x)) = \begin{cases} \Psi\left(\dfrac{g_{1,i}(P_I(x))}{g_{2,i}(P_I(x))}\right) & \text{if } g_{2,i}(P_I(x)) \neq 0 \\ 0 & \text{otherwise.}\end{cases} \quad (8)$$

where Ψ is a non-constant function, $g_{1,i}$ and $g_{2,i}$ are functions such that

$$g_{1,i}(a_x P_I(x) + b_x) = h(a_x) g_{1,i}(P_I(x)) \quad (9)$$
$$g_{2,i}(a_x P_I(x) + b_x) = h(a_x) g_{2,i}(P_I(x)). \quad (10)$$

Then, one have $D_i(P_I(x)) = D_i(a_x P_I(x) + b_x)$.

In other words, D is invariant to local illumination changes defined in (5).

- Consider a patch $P_I(x_0)$ as an arranged set of pixel values $P_I(x_0) = \{I(x_0), I(x_1), \dots, I(x_n)\}$. The following general form can be used as functions $g_{1,i}$ and $g_{2,i}$:

$$g(P_I(x_0)) = \gamma\left(\sum_{i=0}^{N}\left(\sum_{j=0}^{n}\alpha_{i,j} I(x_j)\right)^{\tau}\right)^{\eta} . \quad (11)$$

where $\{\alpha_{i,j}\}_{j=0}^{n}$, $i = 0,1,\dots,N$ are non-zero and zero-sum sequences, and γ, τ and η are fixed positive numbers so that

$$g(a_{x_0} P_I(x_0) + b_{x_0}) = (a_{x_0})^{\tau\eta} g(P_I(x_0)).$$

Proposed Descriptor (NLDP)

$$M_4 = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -2 \end{bmatrix} \quad M_3 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix} = M_2$$

$$M_5 = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad \begin{bmatrix} I(x_4) & I(x_3) & I(x_2) \\ I(x_5) & I(x_0) & I(x_1) \\ I(x_6) & I(x_7) & I(x_8) \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} = M_1$$

$$M_6 = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix} \quad M_7 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} = M_8$$

$$M_i = \begin{array}{|c|c|c|}\hline \alpha_{i,4} & \alpha_{i,3} & \alpha_{i,2} \\ \hline \alpha_{i,5} & \alpha_{i,0} & \alpha_{i,1} \\ \hline \alpha_{i,6} & \alpha_{i,7} & \alpha_{i,8} \\ \hline \end{array} \qquad \sum_{j=0}^{8}\alpha_{i,j} = 0 \qquad M_i \otimes P_I(x_0) = \sum_{j=0}^{8}\alpha_{i,j} I(x_j)$$

$$D(P_I(x_0)) = \frac{V(P_I(x_0))}{\|V(P_I(x_0))\|_2} \quad (12)$$

with $V(P_I(x_0)) = (M_1 \otimes P_I(x_0), M_2 \otimes P_I(x_0), \cdots, M_8 \otimes P_I(x_0))$.

In this case,

$$g_{1,i} = M_i \otimes P_I(x_0) = \sum_{j=0}^{8}\alpha_{i,j} I(x_j) \quad (13)$$

$$g_{2,i} = \left(\sum_{i=1}^{8}\left(\sum_{j=0}^{8}\alpha_{i,j} I(x_j)\right)^2\right)^{1/2} \quad (14)$$

- $g_{1,i}(a_{x_0} P_I(x_0) + b_{x_0}) = M_i \otimes (a_{x_0} P_I(x_0) + b_{x_0})$ (15)
$= a_{x_0} M_i \otimes P_I(x_0) + b_{x_0} \sum_{j=0}^{8}\alpha_{i,j} = a_{x_0} g_{1,i}(P_I(x_0))$.

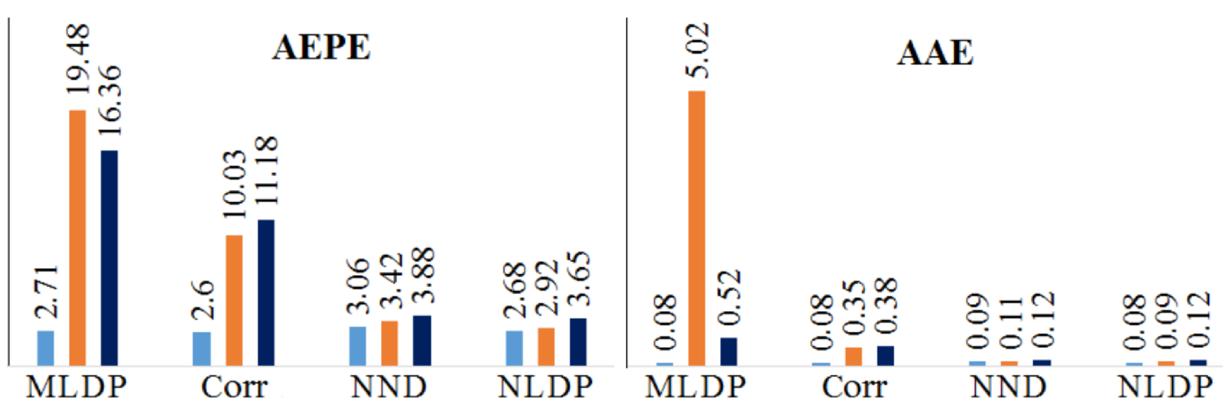
- $g_{2,i}(a_{x_0} P_I(x_0) + b_{x_0}) = \|V(a_{x_0} P_I(x_0) + b_{x_0})\|_2$
$= a_{x_0}\|V(P_I(x_0))\|_2 = a_{x_0} g_{2,i}(P_I(x_0))$. (16)

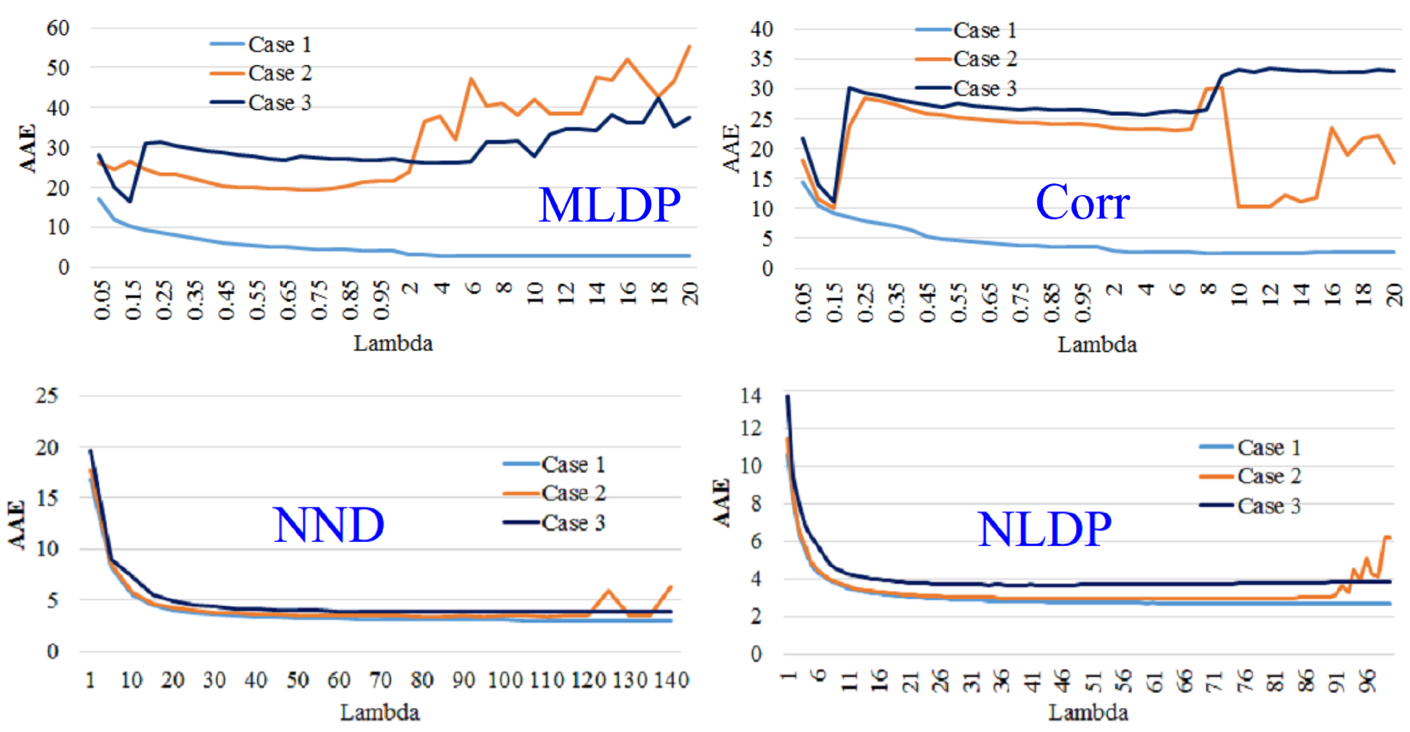
- Combining (12), (15) and (16),
$\Rightarrow D(a_{x_0} P_I(x_0) + b_{x_0}) = D(P_I(x_0))$

The proposed descriptor is illumination - invariant.

Results on synthetic data with ground truth flow (descriptor comparison)

Experiments on tree cases of illumination changes





AEPE and AAE comparison on training image pairs of the Middlebury data [1]

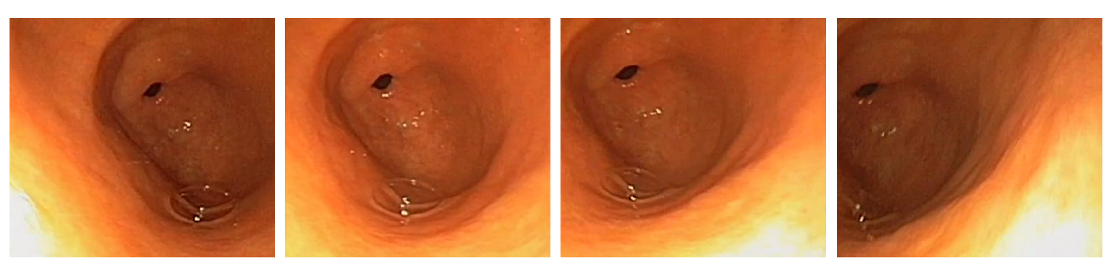
Descriptor	Dimetrodon		Grove2		Grove3		Hydrangea		RubberWhale		Urban2		Urban3		Venus	
	AEPE	AAE	AEPE	AAE	AEPE	AAE	AEPE	AAE	AEPE	AAE	AEPE	AAE	AEPE	AAE	AEPE	AAE
MLDP	0.13	2.35	**0.13**	1.84	0.48	5.02	**0.17**	2.06	**0.08**	2.71	**0.33**	**3.18**	0.57	4.18	**0.26**	**3.68**
Corr	0.23	4.86	0.20	2.53	0.49	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

$$\text{AEPE} = \frac{1}{N}\sum^{N}\sqrt{\left(u_{1,i} - u_{1,i}^{gt}\right)^2 + \left(u_{2,i} - u_{2,i}^{gt}\right)^2}$$

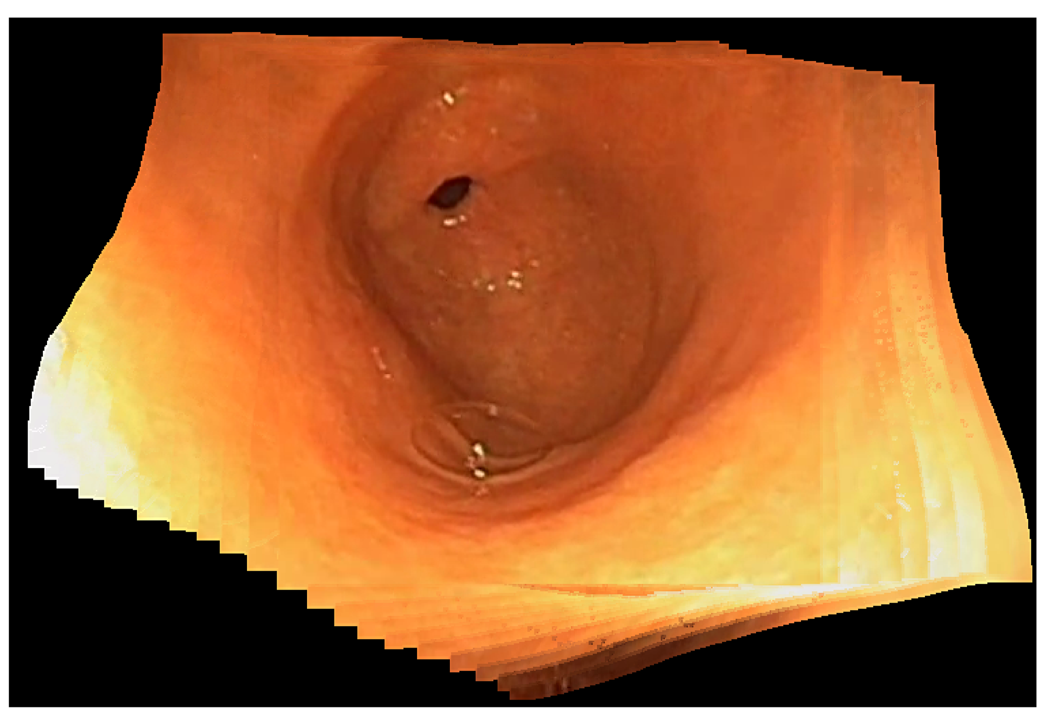
$$\text{AAE} = \frac{1}{N}\sum_{i=1}^{N}\arccos\left(\frac{u_{1,i}u_{1,i}^{gt} + u_{2,i}u_{2,i}^{gt}}{\sqrt{u_{1,i}^2 + u_{2,i}^2 + 1}\sqrt{u_{1,i}^{2,gt} + u_{2,i}^{2,gt} + 1}}\right)$$

[1] http://vision.middlebury.edu/flow/data/

Use of the NLDP descriptor for gastroscopic image mosaicing



Endoscopic images in the pyloric antrum region (epithelium of the internal stomach wall).



Mosaic built with 21 images extracted from a gastroscopic video-sequence.