# Simultaneous Surgical Visibility Assessment, Restoration, and Augmented Stereo Surface Reconstruction for Robotic Prostatectomy

Xiongbiao Luo[1(✉)], Ying Wan[5], Hui-Qing Zeng[2(✉)], Yingying Guo[1], Henry Chidozie Ewurum[1], Xiao-Bin Zhang[2], A. Jonathan McLeod[3], and Terry M. Peters[4]

[1] Department of Computer Science, Xiamen University, Xiamen, China
xbluo@xmu.edu.cn
[2] Zhongshan Hospital, Xiamen University, Xiamen, China
13606080893@139.com
[3] Intuitive Surgical Inc., Sunnyvale, CA, USA
[4] Robarts Research Institute, Western University, London, Canada
tpeters@robarts.ca
[5] School of Electrical and Data Engineering,
University of Technology Sydney, Sydney, China
joyee.wa@gmail.com

**Abstract.** Endoscopic vision plays a significant role in minimally invasive surgical procedures. The maintenance and augmentation of such direct in-situ vision is paramount not only for safety by preventing inadvertent injury, but also to improve precision and reduce operating time. This work aims to quantitatively and objectively evaluate endoscopic visualization on surgical videos without employing any reference images, and simultaneously to restore such degenerated visualization and improve the performance of surgical 3-D reconstruction. An objective no-reference color image quality measure is defined in terms of sharpness, naturalness, and contrast. A retinex-driven fusion framework was proposed not only to recover the deteriorated visibility but also to augment the surface reconstruction. The approaches of surgical visibility assessment, restoration, and reconstruction were validated on clinical data. The experimental results demonstrate that the average visibility was significantly enhanced from 0.66 to 1.27. Moreover, the average density ratio of surgical 3-D reconstruction was improved from 94.8% to 99.6%.

## 1 Endoscopic Vision

Noninvasive and minimally invasive surgical procedures often employ endoscopes inserted inside a body cavity. Equipped with video cameras and optical fiber light sources, an endoscope provides a direct in-situ visualization of the surgical field during interventions. The quality of in-situ endoscopic vision has a critical impact on the performance of these surgical procedures.


A. F. Frangi et al. (Eds.): MICCAI 2018, LNCS 11073, pp. 315–323, 2018.
https://doi.org/10.1007/978-3-030-00937-3_37

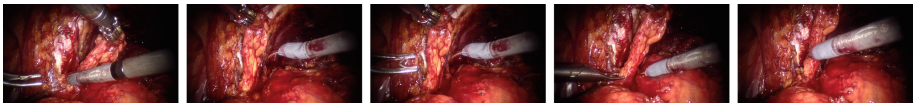


**Fig. 1.** Examples of degenerated surgical images due to small viewing field and very non-uniform and highly directional illumination in robotic prostatectomy

Unfortunately, the endoscope has two inherent drawbacks: (1) a relatively narrow field or small viewing angle and (2) very non-uniform and highly directional illumination of the surgical scene, due to limited optical fiber light sources (Fig. 1). These drawbacks unavoidably deteriorate the clear and high-quality visualization of both the organ being operated on and its anatomical surroundings. Furthermore, these disadvantages lead to difficultly in distinguishing many characteristics of the visualized scene (e.g., neurovascular bundle) and prevent the surgeon from clearly observing certain structures (e.g., subtle bleeding areas). Therefore, in addition to the avoidance of inadvertent injury, it is important to maintain and augment a clear field of endoscopic vision.

The objective of this work is to evaluate and augment on-site endoscopic vision or visibility of the surgical field and to simultaneously improve stereoscopic surface reconstruction. The main contributions of this work are as follows. To the best of our knowledge, this paper is the first to demonstrate objective quality assessment of surgical visualization or visibility in laparoscopic procedures, particular in robotic prostatectomy using the da Vinci surgical system. An no-reference color image quality assessment measure is defined by integrating three attributes of sharpness, naturalness, and contrast. Simultaneously, this work also presents the first study on surgical vision restoration to augment visualization in robotic prostatectomy. A retinex-driven fusion approach is proposed to restore the substantial visibility in the endoscopic imaging and on the basis of the surgical visibility restoration, this study further improves the performance of stereoscopic endoscopic field 3-D reconstruction.

## 2 Approaches

This section describes how to evaluate and restore the degraded endoscopic image. This restoration results not only in enhancing the visualization of the surgical scene but also improving the performance of 3-D reconstruction.

### 2.1 Visibility Assessment

Quantitative evaluation of surgical vision in an endoscopic video sequence is a challenging task because there are no “gold-standard” references for these surgical images in the operating room. Although no-reference color image quality evaluation methods are widely discussed in the computer vision community [1], it still remains challenging to precisely assess the visual quality of natural scene images.

(a) Original (b) Canny (c) LoG (d) Sobel (e) Prewitt

**Fig. 2.** Comparison of different edge detection algorithms

Meanwhile, in the computer-assisted interventions community, there are no publications reporting no-reference surgical vision assessment. This work defines an no-reference objective measure to quantitatively evaluate the surgical endoscopic visibility with three characteristics of image sharpness, naturalness, and contrast related to image illumination variation and textureless regions.

*Sharpness* describes the structural fidelity, i.e., fine detail and edge preservation. The human visual system (HVS) is sensitive to such structural information. The sharpness $\psi$ here is defined based on local edge gradient analysis [2]:

$$\psi = \frac{1}{M} \sum_{M} \sum_{\Omega} E(x,y),\ E(x,y) = \frac{D_{max}(x,y) + D_{min}(x,y)}{\cos \nabla(x,y)}, \tag{1}$$

where $E(x,y)$ is the computed edge width at pixel $(x,y)$ in patch $\Omega$ on image $\mathbf{I}(x,y)$ that is divided into $M$ patches, $D_m(x,y)$ and $D_{min}(x,y)$ indicate the distances between the edge pixel $(x,y)$ and the maximal and minimal intensity pixels $\mathbf{I}_{max}(x,y)$ and $\mathbf{I}_{min}(x,y)$ at patch $\Omega$, respectively, and $\nabla(x,y)$ denotes the angle between the edge gradient and the tracing direction. The computation of the edge width $E(x,y)$ should consider the impact of the edge slopes because humans perceive image contrast more than the strength of the local intensity [2]. In this work, a Canny edge detector was used as it provided a denser edge map compared to other edge detection methods (Fig. 2).

*Naturalness* depicts how natural surgical images appear. It is difficult to quantitatively define naturalness, which is a subjective judgment. However, by statistically analyzing thousands of images [1], the histogram shapes of natural images generally yield Gaussian and Beta probability distributions:

$$f_g(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-(z-\mu)/2\sigma^2,\ f_b(z;u,v) = u^{-1} z^u \mathcal{F}(u, 1-v; u+1; z), \tag{2}$$

where $\mathcal{F}(\cdot)$ is the hypergeometric function that is defined in the form of a hypergeometric series. As recent work has demonstrated, since image luminance and contrast are largely independent on each other in accordance with natural image statistics and biological computation [3], the naturalness $\chi$ should be defined as a joint probability density distribution of a Gaussian and a Beta functions:

$$\chi = \left(\max\left(f_g(z), f_b(z;u,v)\right)\right)^{-1} f_g(z) f_b(z;u,v). \tag{3}$$

*Contrast* is also an attribute important in human perception. In this work, contrast $\mathcal{C}$ is defined as the difference between a pixel $\mathbf{I}(x,y)$ and the average

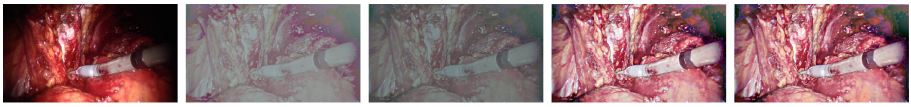
edge-weighted value $\omega(x, y)$ in a patch $\Omega$ (the detected edge $\phi(x, y)$) [4]:

$$\mathcal{C} = N^{-1} \sum_{N} \frac{|\mathbf{I}(x, y) - \omega(x, y)|}{|\mathbf{I}(x, y) + \omega(x, y)|}, \ \omega(x, y) = \sum_{\Omega(x,y)} \phi(x, y)\mathbf{I}(x, y). \tag{4}$$

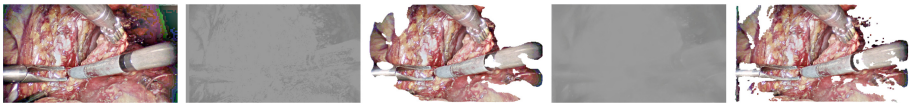
The proposed quality index, $\mathcal{Q}$, combines the metrics defined above into a single objective measure:

$$\mathcal{Q} = a\psi^{\alpha} + b\chi^{\beta} + (1 - a - b) \log \mathcal{C} \tag{5}$$

where $a$ and $b$ balance the three parts, and $\alpha$ and $\beta$ control their sensitivities.



(a) Outputs at four different steps in the surgical visibility restoration approach



(b) Disparity maps and reconstructed surfaces at cost construction and propagation

**Fig. 3.** Surgical visibility restoration and augmented stereo field reconstruction

### 2.2 Visualization Restoration

The retinex-driven fusion framework for surgical vision restoration contains four steps (1) multiscale retinex, (2) color recovery, (3) histogram equalization, and (4) guided fusion. Figure 3(a) illustrates the output of each step in this framework.

The output $\mathbf{R}^i(x, y)$ of the multiscale retinex processing is formulated as [5]:

$$\mathbf{R}^i(x, y) = \sum_{s=1}^{S} \gamma_s \left( \log \mathbf{I}^i(x, y) - \log \left( \mathbf{I}^i(x, y) \otimes G_s(x, y) \right) \right), \ i \in \{r, g, b\}, \tag{6}$$

where $s$ indicates the scale level, $S$ is set to 3, $\gamma_s$ is the weight of each scale, $\otimes$ denotes the convolution operator, and $G_s(\cdot)$ is the Gaussian kernel.

A color recovery step usually follows the multiscale retinex which deteriorates the color saturation and generates a grayish image, and its output $\tilde{\mathbf{R}}^i(x, y)$ is:

$$\tilde{\mathbf{R}}^i(x, y) = \eta \mathbf{R}^i(x, y)(\log \kappa \mathbf{I}^i(x, y) - \log \sum_{i} \mathbf{I}^i(x, y)), \tag{7}$$

where parameter $\kappa$ determines the nonlinearity and factor $\eta$ is a gain factor.

Unfortunately, the recovery step introduces the color inversion problem. To address this problem, histogram equalization processing was performed.

Although high-quality surgical visibility was achieved after the histogram equalization, it still suffers somewhat from being over-bright. A guided-filtering fusion is employed to tackle this problem. Let $\hat{\mathbf{R}}(x, y)$ be the output image after the histogram equalization processing. The guided-filtering fusion of the input image $\mathbf{I}(x, y)$ and $\tilde{\mathbf{R}}(x, y)$ first decomposes them into two-scale representation [7]: $\mathbf{M_I} = \mathcal{M}(\mathbf{I}(x, y))$, $\mathbf{M_R} = \mathcal{M}(\hat{\mathbf{R}}(x, y))$, $\mathbf{D_I} = \mathbf{I}(x, y) - \mathbf{M_I}$, $\mathbf{D_R} = \mathbf{R}(x, y) - \mathbf{M_R}$, where $\mathcal{M}$ denotes the mean-filtering operator, and then reconstructs the final output image $\breve{\mathbf{R}}(x, y) = (\mathbf{w_I^M M_I} + \mathbf{w_R^M M_R} + \mathbf{w_I^D D_I}, \mathbf{w_R^D D_R})$, where the weight maps $\mathbf{w_I^M}$,$\mathbf{w_R^M}$, $\mathbf{w_I^D}$,and $\mathbf{w_R^D}$ are calculated using the guided filtering method [8].

### 2.3 Augmented Disparity

The surgical visibility restoration significantly corrects the illumination non-uniformity problem (Fig. 3(a)) and should be beneficial to disparity estimation.

Based on recent work on the cost-volume filtering method [9], the surgical visibility-restoration stereo endoscopic images are used to estimate the disparity in dense stereo matching. The matching cost volume can be constructed by

$$\forall\ (x, y, d) \in \Re^3, \mathcal{V}((x, y), d) = \mathcal{F}_\delta((x, y), (x + d, y)), \tag{8}$$

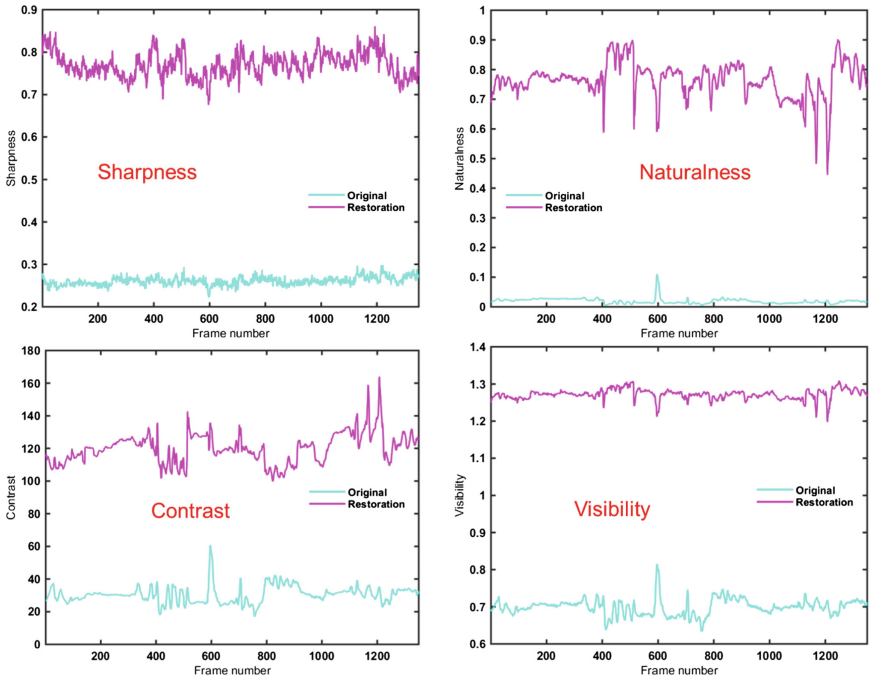



**Fig. 4.** Surgical vision quality assessment in terms of different measures

where $\Re^3 = \{x \in [1, W], y \in [1, H], d \in [d_{min}, d_{max}]\}$, $d_{min}$ and $d_{max}$ denote the disparity search range, and image size $W \times H$. The matching cost function $\mathcal{F}_\delta$ is

$$\mathcal{F}_\delta((x, y), (x + d, y)) = (1 - \delta)\mathcal{F}_c((x, y), (x + d, y)) + \delta\mathcal{F}_g((x, y), (x + d, y)), \quad (9)$$

where $\mathcal{F}_c$ and $\mathcal{F}_g$ characterize the color and gradient absolute difference between the elements of the image pair, and constant $\delta$ balances the color and gradient costs. After the cost construction, the rough disparity and the coarse reconstructed surface can be obtained (see the second and third images in Fig. 3(b)). The coarse disparity map contains image noise and artifacts. The cost propagation performs a filtering procedure that aims to remove these noise and artifacts:

$$\tilde{\mathcal{V}}((x, y), d) = \sum \Psi(\mathbf{I}(x, y))\mathcal{V}((x, y), d), \quad (10)$$

where weight $\Psi(\mathbf{I}(x, y))$ is exactly calculated by using guided filtering [8]. after the cost propagation, the disparity map and the reconstructed surface become better (see the forth and the fifth images in Fig. 3(b)). Finally, the optimal disparity can be achieved by an optimization procedure (e.g., winner-takes-all).

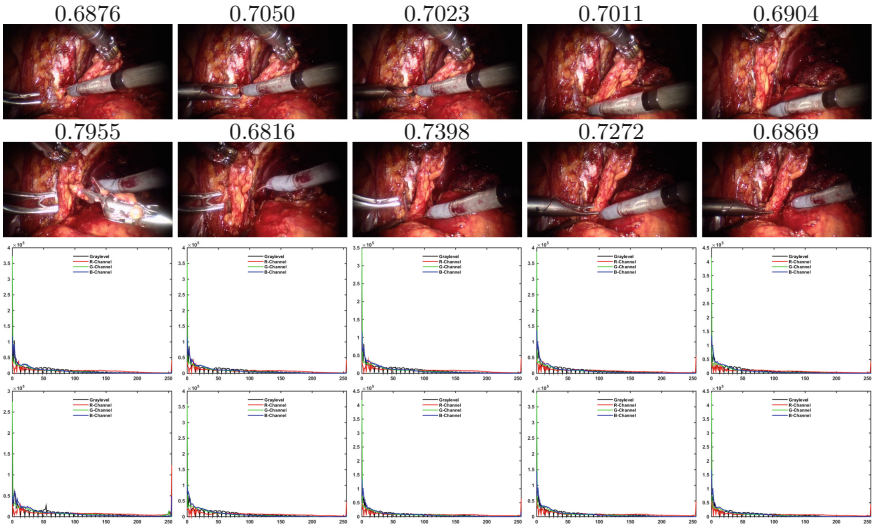
## 3 Validation

Surgical stereoscopic video sequences were collected during robotic-assisted laparoscopic radical prostatectomy using the da Vinci Si surgical system (Intuitive Surgical Inc., Sunnyvale, CA, USA). All images were acquired under a protocol approved by the Research Ethics board of Western University, London, Canada. Similar to the work [6], we set $a = 0.5$, $b = 0.4$, $\alpha = 0.3$, and $\beta = 0.7$. All the experiments were tested on a laptop installed with Windows 8.1 Professional 64-Bit Operating System, 16.0-GB Memory, and Processor Intel(R) Core(TM) i7 CPU×8 and were implemented with Microsoft Visual C++.

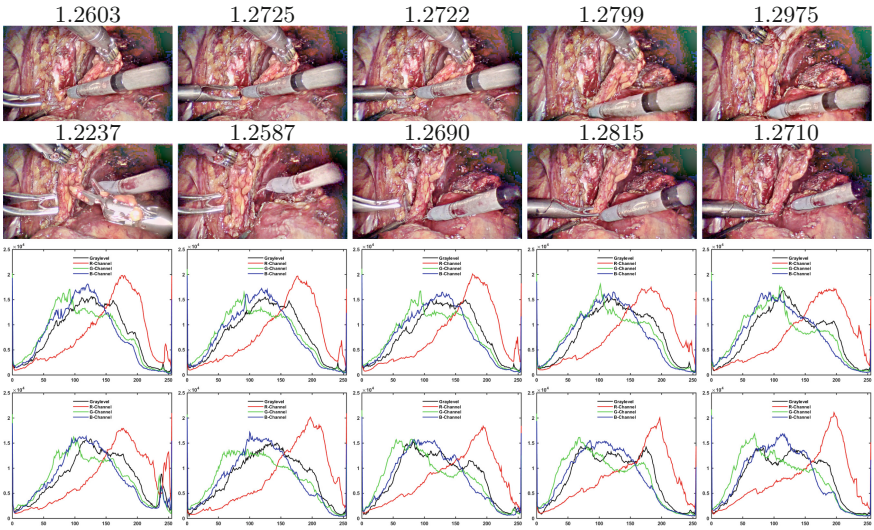
## 4 Results and Discussion

Figure 4 shows the surgical vision assessment in terms of the sharpness, naturalness, contrast, and the proposed joint quality measures. Figure 5 compares the visibility before and after surgical vision restoration by the proposed retinex-driven fusion. The visibility was generally enhanced from 0.66 to 1.27. Figure 6 illustrates the original stereo disparity and reconstruction augmented by visibility restoration. The numbers in Fig. 6 indicate the reconstructed density ratio (the number of reconstructed pixels by all the pixels on the stereo image). The average reconstructed density ratio was improved from 94.8% to 99.6%. Note that image illumination variations, low-contrast and textureless regions are several issues for accurate stereo reconstruction for which we used color and gradient (structural edges) aggregation for stereo matching. Our restoration method can improve color, contrast, and sharpness, resulting in improve the disparity.

The quality of endoscopic vision is of critical importance for minimally invasive surgical procedures. No-reference and objective assessment is essential for

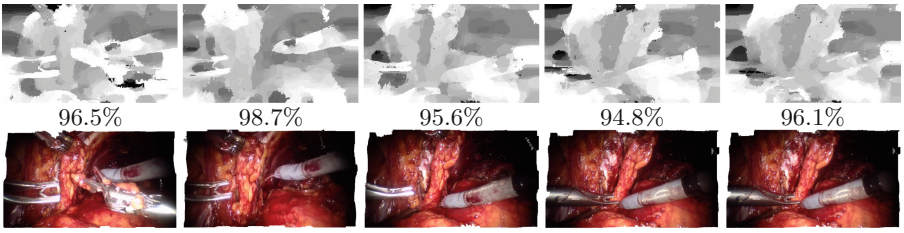



(a) Original surgical video images, their visibility and histograms: Original images in the first and second row correspond to the histograms in the third and fourth row

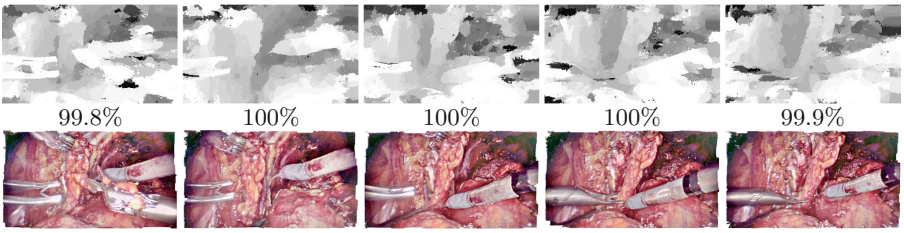



(b) Restored surgical video images, their visibility and histograms: Restored images in the first and second row correspond to the histograms in the third and fourth row

**Fig. 5.** Comparison of surgical endoscopic video images and their visibility and histograms before and after the restoration processing by using the proposed retinex-driven fusion framework. The numbers on the image top indicate the visibility values. The restoration greatly improved the surgical vision. In particular, the restored histograms generally yielded Gaussian and Beta functions.




(a) Disparity and reconstruction using original surgical stereo images




(b) Disparity and reconstruction using restored surgical stereo images

**Fig. 6.** Examples of disparity and reconstruction before and after the restoration

quantitatively evaluating the quality of surgical vision since references are usually unavailable in practice. Moreover, surgical vision and 3-D scene reconstruction are unavoidably deteriorated as a result of the inherent drawbacks of endoscopes. In these respects, the methods proposed above were able to quantitatively evaluate and simultaneously restore surgical vision and further improved endoscopic 3-D scene reconstruction, demonstrating that these approaches can significantly improve the visibility of visualized scenes, but also enhance 3-D scene reconstruction from stereo endoscopic images. On the other hand, the goal of surgical 3-D scene reconstruction is to fuse other imaging modalities such as ultrasound images. This fusion enables surgeons to simultaneously visualize anatomical structures on and under the organ surface. Hence, both visibility restoration and stereo reconstruction are developed to augment surgical vision.

## 5 Conclusions

This work developed an objective no-reference color image quality measure to quantitatively evaluate surgical vision and simultaneously explored a retinex-driven fusion method for surgical visibility restoration, which further augmented stereo reconstruction in robotic prostatectomy. The experimental results demonstrate that the average surgical visibility improved from 0.66 to 1.27 and the average reconstructed density ratio enhanced from 94.8% to 99.6%.

**Acknowledgment.** This work was partly supported by the Fundamental Research Funds for the Central Universities (No. 20720180062), the Canadian Institutes for Health Research, and the Canadian Foundation for Innovation.

## References

1. Wang, Z., et al.: Modern Image Quality Assessment. Morgan & Claypool, San Rafael (2006)
2. Feichtenhofer, C., et al.: A perceptual image sharpness metric based on local edge gradient analysis. IEEE Sig. Process. Lett. **20**(4), 379–382 (2013)
3. Mante, V., et al.: Independence of luminance and contrast in natural scenes and in the early visual system. Nat. Neurosci. **8**(12), 1690–1697 (2005)
4. Celik, T., et al.: Automatic image equalization and contrast enhancement using Gaussian mixture modeling. IEEE Trans. Image Process. **21**(1), 145–156 (2012)
5. Rahman, Z., et al.: Retinex processing for automatic image enhancement. J. Electron. Imaging **13**(1), 100–110 (2004)
6. Yeganeh, H., et al.: Objective quality assessment of tone-mapped images. IEEE Trans. Image Process. **22**(2), 657–667 (2013)
7. Li, S., et al.: Image fusion with guided filtering. IEEE TIP **22**(7), 2864–2875 (2013)
8. He, K., et al.: Guided image filtering. IEEE Trans. Pattern Anal. Mach. Intell. **35**(6), 1397–1409 (2013)
9. Hosni, A., et al.: Fast cost-volume filtering for visual correspondence and beyond. IEEE Trans. Pattern Anal. Mach. Intell. **35**(2), 504–511 (2013)