DOI: 10.2352/ISSN.2470-1173.2016.14.IPMVA-384

Dense Correspondence using Multilevel Segmentation and Affine Transformation

*Sungil Choi, Kihong Park, Seungryong Kim, and *Kwanghoon Sohn*
School of Electrical and Electronic Engineering, Yonsei University, Seoul, Korea

Abstract

Establishing dense correspondence fields between images is an important issue with many computer vision and computational photography applications. Although there have been significant advances in estimating dense correspondence fields, it is still difficult to find reliable correspondence fields between a pair of images because of their geometric and photometric variations. In this paper, we propose an unified framework for establishing dense correspondences, consisting of sparse matching, multilevel segmentation, and derivation of affine transformations. Dense correspondence fields are estimated via winner-takes-all (WTA) optimization by utilizing affine transformations, derived from spare matching and multilevel segmentation. The proposed method reduces a size of label search space dramatically, and further extends the dimension of label search space, by leveraging affine transformation with the multilevel segmentation scheme. Our robust dense correspondence estimation is evaluated on extensive experiments, which show that our approach outperforms the state-of-the-art methods both qualitatively and quantitatively.

1. Introduction

Estimating dense correspondence fields is one of the most important tasks in computer vision applications. Conventionally, it has been studied by focusing on a stereo matching [4, 24, 25] and an optical flow estimation [26, 27]. Recently, dense correspondences for further general images, which have high variability such as different viewpoint, different scales, and different rotations, has been popularly studied. Although many approaches are proposed to estimate dense correspondence fields between two images [3, 6, 8, 10, 12, 16–19], dense correspondence fields estimation is still challenging because of two main reasons: (1) photometric variations due to illumination variations or camera setting, and (2) geometric variations due to viewpoint changes, object pose changes, and non-rigid deformation of objects [16].

To solve these problems, many researchers have attempted with various approaches [3, 6, 8, 10, 12, 16–19]. To provide a robustness to photometric variations, illumination invariant descriptors were employed such as SIFT [11], SURF [9], SID [13] and DAISY [7]. However, they have an inherent limitation on geometric variations. Some of previous efforts tried to solve these problems as global models [3, 6, 16–18], such as 2D pixel-level Markov random field (MRF) model and spatial pyramid model. Others tried to solve them by local methods [8, 10, 12, 19] such as nearest-neighbor search and PatchMatch schemes. SIFT flow [3], PatchMatch Filter (PMF) [10], and Deformable Spatial Pyramid (DSP) [6] set the matching cost in their models with fixed scale and rotation. Accordingly, they failed to find correspondence fields with scale and rotation variations. To settle the limitations caused by fixed scale and rotation, Recent works suggested DAISY Filter Flow (DFF) [12], variants of SIFT flow [17,18], and Generalized deformable spatial pyramid (GDSP) [16]. In these approaches, estimating dense correspondence fields can be casted by discrete labeling of displacement and even geometric variational fields (*e.g.* scale and rotation). They enhanced matching performances between images having geometric variations by extending the dimension of label search space, but still have limited performance because of quantized scale and rotation term. Furthermore, a large size of label search space might bring lots of complexity.

To alleviate these limitations, we propose an unified framework to deal with large photometric and geometric variations between images, by utilizing affine transformation and multilevel segmentation. To reduce the size of label search space, we employ a segmentation scheme based on the assumption that pixels on a segment should have similar correspondence fields. We then employ sparse feature matching to derive the candidates of labels. The label of each segment is built by deriving affine transformation matrix, which implies large geometric variations, with sparse matching points in the segment. To solve the problem due to lack of sparse matching points or wrong sparse matching points, we augment the candidates of labels by employing multilevel segmentation.

Our contributions are summarized as follows. First, we utilized the affine transformation matrix derived from sparse matching points in each segment, which broadens the dimension of label search space and enables itself to deal with large geometric variations while reducing computational time. Second, we employed multilevel segmentation to complement the inaccuracy of labels induced from sparse matching points in each segment and enrich the candidates of labels. Finally, we provided qualitative evaluations of recent studies including ours on various data sets.

The remainder of this paper is organized as follows. In section 2, we enumerated recent approaches estimating dense correspondence fields. We formulate the problem and propose the method in section 3. Experimental results are presented in section 4. Finally, we conclude the paper in section 5.

2. Related Work

Many approaches have been developed in the area of estimating dense correspondence fields. As a pioneering work, SIFT

DOI: 10.2352/ISSN.2470-1173.2016.14.IPMVA-384

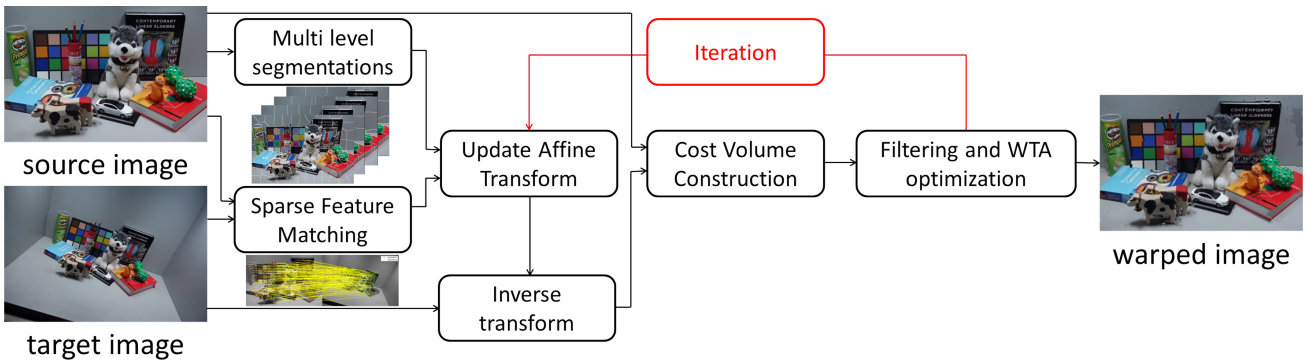


Figure 1. *Framework of the proposed method. Our approach first employ initial sparse feature matching and multi-level segmentation. Based on those fields, affine transformation fields are then estimated. With a cost volume construction and filtering, reliable correspondence fields can be estimated. By an iterative scheme, final dense correspondence fields can be estimated.*

flow [3] based on Markov random field (MRF) model gives inspiration to many further research directions. In SIFT flow [3], SIFT descriptor [11] of each pixel is extracted and the energy function, established including small displacement and spatial regularization, is minimized by a dual loopy belief propagation. However, its performance is limited because it does not deal with scale and rotation. To alleviate this limitation, based on SIFT flow, other approaches are proposed with extended label search spaces with considering scales or rotation. Scale-Less SIFT (SLS) [17] improved its performance with multiple scale SIFT descriptors, and scale-space SIFT flow (SSF) [18] increased the label search space by adding scale term in SIFT flow [3].

Unlike MRF based methods, other approaches based on PatchMatch [8] have been proposed. PatchMatch is much faster than MRF optimization due to its randomized scheme, but it does not guarantee smoothness of labels. In addition, PatchMatch is not robust to photometric variations because it does not utilize illumination invariant descriptors such as SIFT or DAISY. A generalized PatchMatch (GPM) [19], which generalized PatchMatch scheme, finds the fields by searching nearest neighbor fields with consideration of scale and rotation. However, it can not estimate accurate correspondence fields since it does not provide smoothness of the fields. PatchMatch Filter (PMF) [10] searches label by employing PatchMatch method and edge aware filtering with SIFT descriptor to solve the label's discontinuity and the photometric variation problem. Recently, DAISY Filter Flow (DFF) [12], which employs DAISY descriptor and PMF, extended label search space including scale and rotation. However, the computational complexity increases with extended label search space and even the extended label search space does not cover whole geometric variations.

Deformable spatial pyramid (DSP) [6] uses a pyramid graph model with SIFT descriptor. It can be also extended to multiple scale matching approach with a scale term and a scale smoothness term. More recently, generalized deformable spatial pyramid (GDSP) [16] which deals with scale and rotation variations is an extending version of DSP. It produces satisfactory results invariant to geometric deformations, but it causes seriously high complexity. Unlike those methods, our approach provides highly reliable matching results while providing very low computational time.

3. Proposed method

3.1. Motivation and overview

Given two images which have large photometric and geometric variations, our goal is to estimate dense correspondence fields $\mathbf{t}_p = (u_p, v_p)$ where p denotes a pixel in an image. To handle scale s_p and rotation θ_p for pixel p, the dimension of the search space is extended by 4-dimensional space $\mathbf{t}_p = (u_p, v_p, s_p, \theta_p)$. However, this 4-dimensional search space causes a lot of computational complexity. In addition, many discrete quantization scheme for geometric fields $(s_p,\ \theta_p)$ [12, 16, 18] might degrade the matching performance while providing artifacts.

To reduce the search space and guarantee robustness of the dense correspondence fields, we introduce an unified framework summarized in Figure 1. First of all, sparse feature matching is conducted in the source and target image, and multilevel segmentation is employed in the source image. The source image is segmented based on the assumptions that pixels in each segment are from the same object and their correspondence fields are similar. With sparse matching points in each segment, the affine transformation matrix is derived in each segment. Finally, to find the optimal dense correspondence fields, cost volume filtering and winner-takes-all (WTA) optimization are conducted, and this process is repeated until it converges. In the first iteration, the cost volume is constructed with the affine transformations from lower level linked-segments. After first iteration, cost volume is constructed with the affine matrices from adjacent segments. With some iterations which is from cost volume construction to optimization, more accurate dense correspondence fields are updated. In section 3.2, We describe the advantage of affine transformation approach and method. Also, we explain the multilevel segmentation method in section 3.3 and cost volume filtering in section 3.4.

3.2. Affine transformation between two images

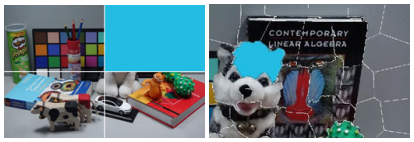
The affine transformation can be useful in estimating dense correspondence fields between images which have a variety of geometric deformations [23]. The affine transformation matrix $\mathbf{H}$ whose size is 3-by-3 can be decomposed as follows:

$$\mathbf{H} = \mathbf{T} \cdot \mathbf{S} \cdot \mathbf{R} \cdot \tilde{\mathbf{S}} \tag{1}$$

©2016 Society for Imaging Science and Technology
DOI: 10.2352/ISSN.2470-1173.2016.14.IPMVA-384

(a) (b) (c)

Figure 2. *Robustness of affine transformation. (a) and (b) are a pair of images which have geometric deformations, and (c) is a rotationally transformed image of (b). The yellow boxes shows that geometric deformation of (a) and (c) is not uniformly scaling. As shown in (c), affine transformation fields can deal with more challenging scenarios.*



(a) level-1 segmentation (b) level-2 segmentation

(c) level-3 segmentation (d) level-4 segmentation (e) level-5 segmentation

Figure 3. *Examples of multilevel segmentation. (e) shows the highest level segmentation. Pink segments are adjacent to the blue segment in (e). Sky-blue segments in (a) to (d) are lower level segments linked to the blue segment in (e).*

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{pmatrix}, \mathbf{S} = \begin{pmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tilde{\mathbf{S}} = \begin{pmatrix} 1 & sh_x & 0 \\ sh_y & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where $\mathbf{T}$ is the translation matrix, $\mathbf{S}$ is the scale matrix, $\mathbf{R}$ is the rotation matrix, and $\tilde{\mathbf{S}}$ is the shear matrix. Since the affine transformation matrix has various geometric variation elements, it can deal with more severe geometric deformations than conventional 4-dimensional label search space cases. As shown in Figure 2, unlike conventional methods that cannot deal with non-uniform scaling, the affine transformation approach could handle not only non-uniform scaling but also shear variation.

To estimate affine transformation between two images, there are many approaches to consider [23, 28, 29]. Recently, Park et al. [23] employed an affine transformation estimation with sparse matching points. Similarly, our approach employs SURF algorithm [9] to acquire spare matching points. To obtain the affine transform, at least 3 sparse matching points are needed since there exist 9 known parameters. RANSAC [5] is utilized for excluding outlier matching points in deriving the affine transformation matrix.

3.3. Multilevel segmentation

Our assumption is that each object in the image has its own affine transformation as it has its own dense correspondence field. With this assumption, we induce an affine transform matrix derived from sparse matching points in each segment by employing the segmentation method. Obtaining an affine transformation matrix from sparse matching points in each segment does not guarantee optimal dense correspondence field because some segmentations lack sparse matching points (*i.e.* lose information about geometric deformation.) or sparse matching points may have wrong matching (*i.e.* labels could be wrong.). Thus, we propose a method which updates dense correspondence fields exploiting multilevel segmentation and cost volume filtering.

Multilevel segmentation is performed by applying segmentation N times with different number of segments at each level. As shown in Figure 3, level-1 segmentation scheme divides a source image uniformly by quadrant, denoted by $\mathbf{S}_1$. Higher level segmentations are performed with more number of segment and denoted by $\mathbf{S}_2, \ldots, \mathbf{S}_N$. In this paper, we choose the segmentation method as SLIC [1] due to its efficiency and robustness. We denote $\mathbf{S}_k$ as a set $\mathbf{S}_k = \{S_{k1}, S_{k2}, \ldots, S_{kN}\}$ where S_{kn} is n-th segment in level-k segmentation.

3.4. Flow field inference with cost volume filtering

To get an optimal dense correspondence field at each pixel, we build feature descriptors on the source image and and inverse transformed image of the target image. Since it is impossible to extract feature descriptors from the target image considering nonuniform scales and shear, we build the descriptors on the inverse transformation mapping of the target image. With the affine transformation matrix, the matching cost can be calculated as follows :

$$C_{\mathbf{H}}(p) = \min\left(\left\|D(x_p, y_p) - D'(x'_p, y'_p)\right\|_1, \tau\right) \quad (2)$$

where

$$\begin{pmatrix} x'_p \\ y'_p \\ 1 \end{pmatrix} = \mathbf{H} \cdot \begin{pmatrix} x_p \\ y_p \\ 1 \end{pmatrix}$$

where $D(x_p, y_p)$ is a feature descriptor at (x_p, y_p) in the source image, and τ is a truncation value for robustness. $D'(x'_p, y'_p)$ is a feature descriptor at (x_p, y_p) in the inverse mapping of the target image. Because we produce the descriptor without scale and rotation term, only with inverse mapping, we can build a matching cost as shown in (2).

In this paper, we use the DAISY [7] descriptor as a feature descriptor. However, inferencing dense correspondence fields with only raw matching cost can cause the fields be not spatially smooth or not reliable because of its erroneous local pixels. To overcome these limitations, fast edge-aware smoothing filters [2, 14, 22], which is used to impose the spatial smoothness on the labels, are adopted. Given a raw matching cost $C_{\mathbf{H}}(p)$ computed at pixel p , the edge-aware filtered output $\overline{C}_{\mathbf{H}}(p)$ can be denoted as :

$$\overline{C}_{\mathbf{H}}(p) = \sum_{q \in W(p)} \lambda_{q,p}(I_s) C_{\mathbf{H}}(q) \quad (3)$$

where $W(p)$ is the local aggregation window centered at pixel p. The adaptive weight $\lambda_{q,p}(I_s)$ is defined based on how similar two pixels p and $q \in W(p)$ in the color space are with regard to the

DOI: 10.2352/ISSN.2470-1173.2016.14.IPMVA-384

Figure 4. *Results on DIML data set [21]. The first and second row show the warped images on images with scale and photometric variations, the third and fourth row are the results on images with rotation and photometric variations, and the fifth and last row are results about scale, rotation and photometric variations.*

source image I_s. Several edge-aware filtering [2, 14, 22] can be employed with different way in terms of defining $\lambda_{q,p}(I_s)$ such that $\sum_{q\in W(p)} \lambda_{q,p}(I_s) = 1$. In this paper, we use Guided Image Filtering [14] for computational efficiency.

The optimal dense correspondence fields can be updated with WTA scheme. In each segment in the highest level segmentation, the field in each pixel is optimized in filtered cost volume constructed with (1) set of lower level linked-segments $\mathscr{L}(S_{Nk})$ and (2) set of adjacent segments $\mathscr{N}(S_{Nk})$ as shown in Figure 3. In each highest level segment, lower level linked-segments $\mathscr{L}(S_{Nk})$ are segments which center point of the highest level segment belongs to in the lower level segmentation. For each pixel $p \in S_{Nk}$, the WTA optimization with these cost volume is defined as :

$$\mathbf{H}(p) = \text{argmin}_{\mathbf{H}\in\mathscr{F}(\mathscr{L}(S_{Nk}))} \overline{C}_{\mathbf{H}}(p) \tag{4}$$

$$\mathbf{H}(p) = \text{argmin}_{\mathbf{H}\in\mathscr{F}(\mathscr{N}(S_{Nk}))} \overline{C}_{\mathbf{H}}(p) \tag{5}$$

where N is the highest level and $\mathscr{F}(\mathscr{L}(S_{Nk}))$ denotes the affine transformation matrices and each of them is from a random sampled pixel from each element of $\mathscr{L}(S_{Nk})$. Also, $\mathscr{F}(\mathscr{N}(S_{Nk}))$ is defined in same way as $\mathscr{F}(\mathscr{L}(S_{Nk}))$. With the WTA optimization in (4), correspondence fields of pixels in each segment are refined by comparing the fields from lower level linked-segments. With refined labels obtained by WTA optimization in (4), some iterations of WTA optimization in (5) should be executed until the fields converge.

4. Experimental results

In order to verify the performance of the proposed method, we performed the experiment with DIML data set [21]. SURF matching [9], DAISY descriptor [7], SLIC [1] and guided filtering [14] are employed as we described above. We set the radius of DAISY descriptor 8. The window size of guided filtering was set to 21. We performed SLIC 5 times for conducting multilevel segmentation. In applying SLIC, the number of segments are proportional to the number of sparse matching points. We evaluated our method with DIML data set (sec 4.A) and compared the warped images on challenging image pairs (sec 4.B).

A. Results on DIML data set [21]. We compared our method with SID [13], segSID [15], SIFTflow [3], SSF [18], and segSIFT [15] on DIML data set [21]. To identify the dense correspondence fields of each algorithm, we warped each pixel of the target image using correspondence fields. Figure 4 ex-

DOI: 10.2352/ISSN.2470-1173.2016.14.IPMVA-384

Table 1. Evaluations on DIML data set [21]

	SID [13]	segSID [15]	SIFTflow [3]	SSF [18]	segSIFT [15]	Ours
Scale, Photometric	0.9612	0.9263	0.8745	0.7258	0.8223	**0.9626**
Rotation, Photometric	0.3792	0.5645	0.3512	0.4645	0.5339	**0.9628**
Scale, Rotation, Photometric	0.7969	0.7339	0.4548	0.2972	0.6447	**0.9605**
Average	0.7124	0.7416	0.5602	0.4958	0.6676	**0.9620**

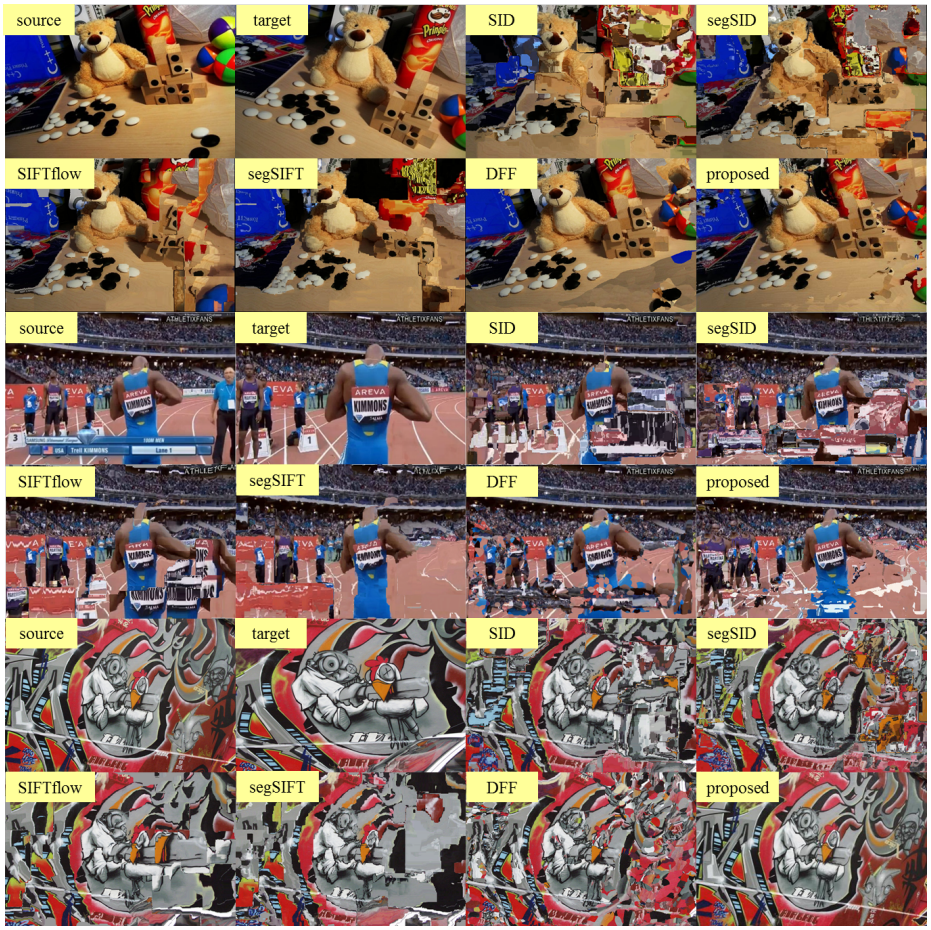


Figure 5. *Results on challenging image pairs including SID [13], segSID [15], SIFTflow [3], segSIFT [15], DFF [12] and proposed method.*

hibits the warped images produced by SID, segSID, SIFTflow, SSF, segSIFT and our method on DIML data set including scale, rotation and photometric variations. We adopted the evaluation method as label transfer accuracy (LT-ACC) between pixel correspondences [20]. Given the source and target images, the warped image can be produced with estimated correspondence field at each pixel. With a source and target image, we transfer the annotated class of the target pixels to the source ones using dense correspondence fields, and count how many pixels in the source image are correctly matched. Table 1 presents the LT-ACC evaluations on comparison methods and ours in various deformation scenarios. In Table 1 and Figure 4, it is shown that SID descriptors are more robust to geometric variation than SIFT. The segment based approach [15] shows more robust to geometric variation than SIFT flow [3]. Our approach estimated more reliable correspondence fields than any other methods.

B. Visual comparison on challenging image pairs. We also conducted an experiment on challenging image pairs, Figure 5 shows the warping results by estimating dense correspondence fields from the source image and the target image. SIFT flow [3] failed to handle the geometric variations since the warped image lost the details when rotation variations are prominent. SID [13] could not estimate properly in our experiment even though it is consisted that robustness to geometric variations. segSID and segSIFT [15],which use segmentation priors, improved their performance than SID [13] and SIFTflow [3] but have still limited performance. DFF [12] showed better results than other comparison methods but warping results failed to construct details in image pairs with complicated geometric variations due to its search space. Our method presented most plausible results with reliable estimated correspondence fields.

DOI: 10.2352/ISSN.2470-1173.2016.14.IPMVA-384

5. Conclusion

We have presented the robust unified framework for estimating dense correspondence field approach, using sparse matching, multilevel segmentation, and affine transformations. The proposed method was robust to geometric variations from an affine transformation matrix derived from sparse matching points in each segment. By leveraging multilevel segmentation, we enriched affine transformation candidates and solved the shortage of sparse matching points. In experiments, the proposed method demonstrated its outstanding performance by establishing dense correspondence fields between challenging image pairs which have photometric and geometric variations. In future work, our approach can be applied to image-level tasks such as object recognition or annotation.

References

[1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE TPAMI*, 34(11):2274-2282, 2012.

[2] E. Gastal, and M. Oliveira. Domain transform for edge-aware image and video processing. *ACM TOG*, 30(4):69, 2011.

[3] C. Liu, J. Yuen, and A. Torralba. SIFT flow: Dense correspondence across scenes and its applications. *IEEE TPAMI*, 33(5):978-994, 2011.

[4] K. Yoon, and I. Kweon. Adaptive support-weight approach for correspondence search. *IEEE TPAMI*, 28(4):650-656, 2006.

[5] Q. Tran, T. Chin, G. Carneiro, M. S. Brown, and D. Suter. In defence of ransac for outlier rejection in deformable registration. In *ECCV*, 2012.

[6] J. Kim, C. Liu, F. Sha, and K. Grauman. Deformable spatial pyramid matching for fast dense correspondences. In *CVPR*, 2013.

[7] E. Tola, V. Lenpetit, and P. Fua. Daisy: An efficient dense descriptor applied to wide-baseline stereo. *IEEE TPAMI*, 32(5):815-830, 2010.

[8] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman. PatchMatch: A randomized correspondence algorithm for structural image editing. In *ACM SIGGRAPH*, 2009.

[9] H. Bay, T. Tuytelaars, and L. V. Gool. Surf: Speeded up robust features. In *ECCV*, 2006.

[10] J. Lu, H. Yang, D. Min, and M. N. Do. Patchmatch filter: Efficient edge-aware filtering meets randomized search for fast correspondence field estimation. In *CVPR*, 2013.

[11] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(5):91-110, 2004.

[12] H. Yang, W. Lin, and J. Lu. Daisy filter flow: A generalized discrete approach to dense correspondences. In *CVPR*, 2014.

[13] I. Kokkinos, and A. Yuille. Scale invariance without scale selection. In *CVPR*, 2008.

[14] K. He, J. Sun, and X. Tang. Guided image filtering. *IEEE TPAMI*, 35(6):1397-1409, 2013.

[15] E. Trulls, I. Kokkinos, A. Sanfeliu, and F. M. Noguer. Dense segmentation-aware descriptors. In *CVPR*, 2013.

[16] J. Hur, H. Lim, C. Park, and S. C. Ahn. Generalized deformable spatial pyramid: Geometry-preserving dense correspondence estimation. In *CVPR*, 2015.

[17] T. Hassner, V. Mayzels, and L. Zelnik-Manor. On sifts and their scales. In *CVPR*, 2012.

[18] W. Qiu, X. Wang, X. Bai, A. Yuille, and Z. Tu. Scale-space sift flow. In *WACV*, 2014.

[19] C. Barnes, E. Shechtman, D. B. Goldman, and A. Finkelstein. The generalized patchmatch correspondence algorithm. In *ECCV*, 2010.

[20] C. Liu, J. Yuen, and A. Torralba. Nonparametric Scene Parsing via Label Transfer. *IEEE TPAMI*, 33(12):2368-2382, 2011.

[21] S. Kim, D. Min, B. Ham, S. Ryu, D. Minh, and K. Sohn. Dasc: Dense adaptive self-correlationdescriptor for multi-modal and multi-spectral correspondence. In *CVPR*, 2015.

[22] Q. Yang, K. Tan, and N. Ahuja. Real-time o(1) bilateral filtering. In *CVPR*, 2009.

[23] K. Park, S. Kim, S. Ryu and K. Sohn. Randomized global transformation approach for dense correspondence. In *BMVC*, 2015.

[24] M. Okutomi and T. Kanade. A multiple-baseline stereo. *IEEE TPAMI*, 15(4):353-363, 1993.

[25] C. Rhemann, A. Honsi, M. Bleyer, C. Rother, and M. Gelautz. Fast cost-volume filtering for visual corespondence and beyond. In *CVPR*, 2011.

[26] B. K. Horn and B. G. Schunck. Determining optical flow: a retrospective. *AI* 59(1):81-87, 1993.

[27] P. Weinzaepfel, J. Revaud, Z. Harchaoui and C. Schmid. Deepflow: Large displacement optical flow with deep matching. In *ICCV*, 2013.

[28] S. Jang, M. Pomplun, G. Kim and H. Choi. Adaptive robust estimation of affine parameters from block motion vectors. *IVC*, 23(14):1250-1263, 2005.

[29] S. H. Lai. Robust image matching under partial occlusion and spatially varying illumination change. *CVIU*, 78(1):84-98, 2000.

Author Biography

Sungil Choi received the B.S. degree from the School of Electronic Engineering , Yonsei University, Seoul, Korea, in 2015. He is currently pursuing the joint M.S. and Ph.D. degrees in the School of Electrical and Electronic Engineering, Yonsei University, Seoul, Korea. His research interests include computer vision, image visual correspondence, and machine learning.

Kihong Park received the B.S. degree from the School of Electronic Engineering , Sogang University, Seoul, Korea, in 2014. He is currently pursuing the joint M.S. and Ph.D. degrees in the School of Electrical and Electronic Engineering, Yonsei University, Seoul, Korea. His research interests include image visual correspondence, and machine learning.

Seungryong Kim received the B.S. degree in electrical and electronic engineering from Yonsei University, Seoul, Korea, in 2012, where he is currently working toward the joint M.S. and Ph.D. degrees. His research interests include 2-D/3-D computer vision, machine learning, and image visual correspondence.

Kwanghoon Sohn (M'92 SM'12) received the B.E. degree in electronic engineering from Yonsei University, Seoul, Korea, in 1983, the M.S.E.E. degree in electrical engineering from the University of Minnesota, Minneapolis, MN, USA, in 1985, and the Ph.D. degree in electrical and computer engineering from North Carolina State University, Raleigh, NC, USA, in 1992. He was a Post-Doctoral Fellow with the MRI Center, Medical School of Georgetown University, Washington, DC, USA, in 1994. He was a Visiting Professor with Nanyang Technological University, Singapore, from 2002 to 2003. He is currently a Professor in the School of Electrical and Electronic Engineering, Yonsei University. His research interests include 3D video processing, computer vision, and video communication. He is a a senior member of IEEE and a member of SPIE.