

Foreground and Shadow Segmentation Based on a Homography-Correspondence Pair

Haruyuki Iwama, Yasushi Makihara, and Yasushi Yagi

Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka, 567-0047, Japan

Abstract. A static binocular camera system is widely used in many computer vision applications; and being able to segment foreground, shadow, and background is an important problem for them. In this paper, we propose a homography-correspondence pair-based segmentation framework. Existing segmentation approaches, based on homography constraints, often suffer from occlusion problems. In our approach, we treat a homography-correspondence pair symmetrically, to explicitly take the occlusion relationship into account, and we regard the segmentation problem as a multi-labeling problem for the homography-correspondence pair. We then formulate an energy function for this problem and get the pair-wise segmentation results by minimizing them via an α - β swap algorithm. Experimental results show that accurate segmentation is obtained in the presence of the occlusion region in each side image.

1 Introduction

In many computer vision applications, foreground segmentation is important preprocessing for subsequent processing such as object detection, localization, identification and tracking. For this purpose, background subtraction has been widely used for scenes with a static camera [1]. The methods, however, often extract not only the objects but also their shadows, which can be problematic. Consequently, many shadow segmentation, detection, or removal techniques have been proposed [2] [3] [4] [5] [6] [7] [8] [9], based mainly on the following two properties of shadow color: (a) The shadow region is darker than the original background region, (b) The color vector direction of the shadow region is similar to that of the original background region. The property (b) is not reliable in the case of a strong shadow because the color vector direction becomes unstable, and furthermore, it is not true under multiple-color illumination conditions, such as an outdoor scene with daylight. Eventually, any state of the art method using a color based approach fails if the foreground regions contain exactly the same color as the shadows, such as a black-haired human head. This is the essential limitation of color-based segmentation.

On the other hand, a static binocular camera system is widely used in many computer vision applications, such as surveillance systems, traffic monitoring systems, and total coverage camera systems in soccer stadiums. Taking advantage of the multi-view framework, several geometric approaches have been applied to the foreground/shadow segmentation. One well known approach is foreground separation from shadow based on disparities [10] [11]. However, it often suffers from mis-correspondence problems and cannot be applied to scenes with no texture.

Alternatively, a homography constraint is also popular as a geometric constraint between multiple viewpoints [12] [13] [14] [15]. Approaches based on homography aim mainly to distinguish standing objects from ground plane objects including shadow. Their mechanism is described as follows. Suppose that one side image is a base image to be processed. If the color of a pixel in the base side image is similar to that of the homography-correspondence pixel in the other side image, the pixel in the base side image is regarded as part of the ground plane objects, and vice versa. Although the homography approach is quite effective despite the low computational cost, most of the existing methods, however, have a serious disadvantage. Let us consider the following case: A pixel belongs to the ground plane objects in the base side image and the homography-correspondence pixel is occluded by a foreground object in the other side image. In such a case, because the colors of the pair of pixels are different, the pixel in the base image is incorrectly regarded as a foreground object.

In the field of stereo correspondence problems, symmetric correspondence based approaches have been proposed to handle such occlusions appropriately [16] [17]. These approaches explicitly take the occlusion relationship into account by treating a stereo correspondence pair in a symmetric way.

Inspired by the symmetric approaches, we propose a symmetric segmentation framework based on a homography constraint with occlusion handling. Our goal is “*how to segment foreground, shadow, and background*”, and we regard this segmentation problem as a homography-correspondence pair labeling problem. Then, we solve this in an energy minimization framework together with a graph-cut algorithm [18]. Considering the homography-correspondence symmetrically, we cannot only segment the occluded region correctly, but also acquire additional information about the occluded region, such as, what label is assigned to the occluded region, *shadow* or *background*. This kind of information is valuable for many multi-view applications.

The remainder of this paper is organized as follows. Section 2 introduces our segmentation framework. Section 3 describes the detailed implementation of the proposed method. Section 4 demonstrates the effectiveness of the proposed method using experiments. Finally, Section 5 concludes our work.

2 Segmentation Based on Homography-Correspondence Pairs

2.1 Problem setting

In this paper, the following conditions are assumed in our segmentation problem.

- A scene is captured by a static calibrated binocular camera system.
- The background of the scene is modeled as a pixel-wise Gaussian distribution.
- An object in the foreground stands on the ground plane and its shadow appears on the ground plane.

Our goal is to segment the target region as *foreground* (“*F*”) or *shadow* (“*S*”) or *background* (“*B*”), that is to say, to assign one of the three labels “*F*”, “*S*”, or “*B*” to each pixel in both side images in a conformal manner. Note that “*S*” and “*B*” lie on the ground plane while “*F*” stands on the ground plane.

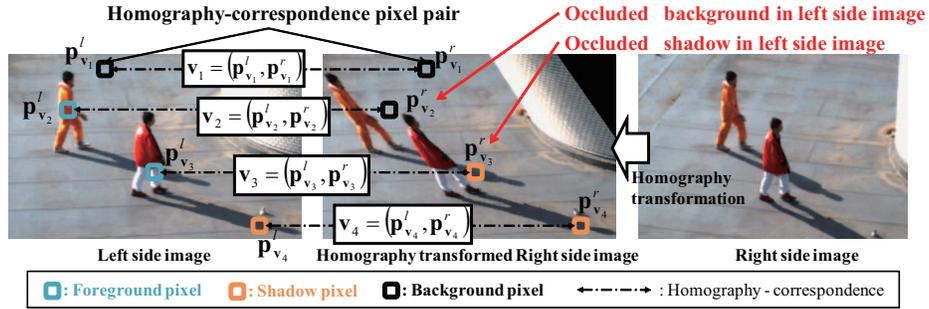


Fig. 1: Homography-correspondence pair

2.2 Asymmetric treatment of homography constraint

Let us consider the homography-correspondence pair on the ground plane in the binocular camera. According to the homography constraint, if a pixel belongs to the ground plane on one side image, the color of the pixel is strictly consistent with that of the homography-correspondence pixel in the other side image under the condition that ideally any standing object does not exist on the ground plane. This is a very useful property to distinguish the standing objects on the ground plane from the ground plane objects. Many object detection techniques, for example, obstacle detection [12], shadow detection [13] [14] [15], have been proposed based on this property.

In segmentation problems, this property is also useful when assigning a label to each pixel. Some examples of the homography-correspondence pairs are shown in Fig. 1. First, suppose that the left side image is a base image to be segmented. Because v_1 and v_4 have similar colors between each correspondence pixel, $p_{v_1}^l$ and $p_{v_4}^l$ are labeled as “S” or “B” in the left side image. On the other hand, because the pixel pairs v_2 and v_3 have different colors between each pair of correspondence pixels, $p_{v_2}^l$ and $p_{v_3}^l$ are labeled as “F” in the left side image. Next, suppose that the right side image is a base image to be segmented in turn, pixel pairs $p_{v_1}^r$ and $p_{v_4}^r$ are labeled as “S” or “B”, and the pixel pairs $p_{v_2}^r$ and $p_{v_3}^r$ are labeled as “F” in the right side image in the same way. The true labels of $p_{v_2}^r$ and $p_{v_3}^r$ are, however, not “F” but “B” and “S”.

This mislabeling often arises in cases where a pixel belongs to the ground plane in one side image and where the corresponding pixel’s ground plane point in the other side image is occluded by a foreground object as shown in this example. Therefore, the existing asymmetric homography-based approaches suffer from the mislabeling due to occlusion.

2.3 Symmetric approach based homography-correspondence pair

In our framework, the homography-correspondence is treated symmetrically to cope with the occluded regions and to segment them correctly.

Taking the occlusion relationship into consideration, the labeling strategy is as follows. If the pixels are labeled “S” or “B” in one side image, their homography-correspondence pixels in the other side image are given either the same label (not the occluded case) or “F” (the occluded case). If the pixels in one side image are labeled

Table 1: The pair-wise label sets for a homography-correspondence pair

Left-side label	Right-side label		
	F	S	B
F	FF	FS	FB
S	SF	SS	-(prohibited)
B	BF	-(prohibited)	BB

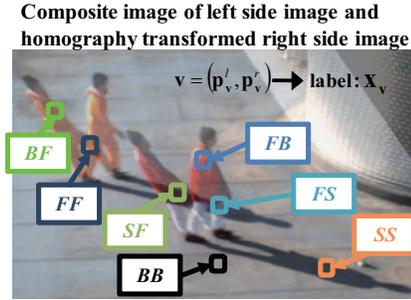


Fig. 2: Labeling examples

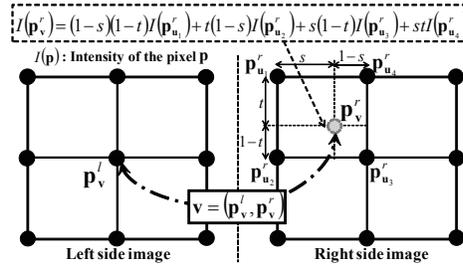


Fig. 3: Homography-correspondence detail

“ F ”, their homography-correspondence pixels in the other side image are possibly labeled “ F ”, “ S ”, or “ B ”, because the standing object is not constrained by homography. From this observation, the possible pair-wise label for the homography-correspondence pair are defined in Tab. 1. Note that the pair-wise label “ SB ” and “ BS ” are prohibited because it is not possible for a ground plane object to occlude another ground plane object. Thus our segmentation problem is regarded as a *multi-labeling problem for homography-correspondence pair pixels*, and the labeling results provide all the relationships between homography-correspondence pair of pixels. For example, the label “ FS ” means the foreground occludes the shadow in the left side image, and also means the shadow in the right side image is occluded by the foreground in the left side image. Example of pair-wise labeling are shown in Fig. 2.

2.4 Problem formulation

We formulate the pair-wise multi-labeling problem in a framework that minimizes energy. Let us define the site $\mathbf{v} = (\mathbf{p}_v^l, \mathbf{p}_v^r)$ which represents a homography-correspondence pair as described in the previous subsection. Then, the label set is defined as,

$$\mathbf{L} = \{FF, FS, FB, SF, SS, BF, BB\}, \quad (1)$$

and the label assigned to a site \mathbf{v} as $\mathbf{x}_v \in \mathbf{L}$. Then our goal is to assign each site \mathbf{v} a label \mathbf{x}_v from the set \mathbf{L} . Generally, this problem is formulated in an energy minimization framework as follows,

$$E(\mathbf{x}) = w_g \sum_{\mathbf{v} \in \mathbf{V}} \mathbf{g}(\mathbf{x}_v) + w_h \sum_{(\mathbf{u}, \mathbf{v}) \in \mathbf{E}} \mathbf{h}(\mathbf{x}_u, \mathbf{x}_v) \quad (2)$$

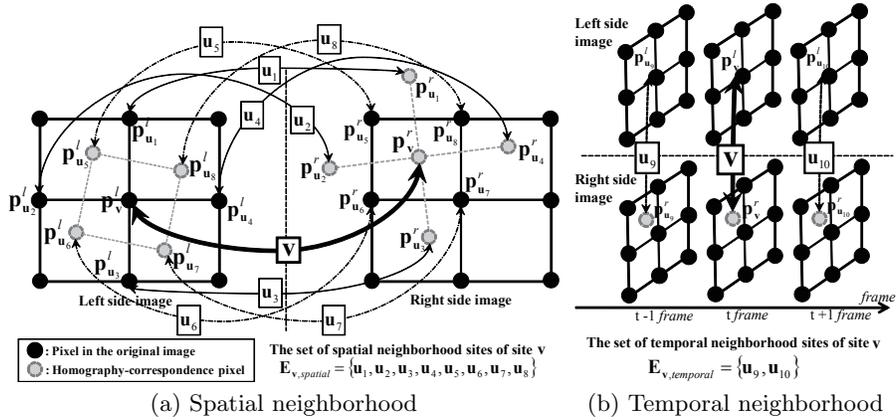


Fig. 4: Spatio-temporal neighborhood sites for smoothness term

where the first and the second terms are data and smoothness terms, w_g and w_h are the weights of each term, \mathbf{x} is a configuration (label combination), \mathbf{V} is a set of all sites, and \mathbf{E} is all the combinations of the neighborhood sites. This energy function is minimized via graph-cut algorithms such as the α -expansion or α - β swap algorithms [18].

Note that, the homography-correspondence positions are calculated using sub-pixel order and the color of the sub-pixel position is spatially interpolated by their 4-neighborhood pixels as shown in Fig. 3. In addition, as shown in Fig. 4, we consider 10-neighborhood sites in a spatio-temporal 3D domain composed of spatial 8-neighborhood, and temporal 2-neighborhood sites.

3 Implementation

3.1 Seed generation

Given background subtraction regions as potential regions of shadow and foreground, the foreground seed is provided as the union of the following two regions; one is the intersection of the potential region and background region projected by homography from the other image, and the other is the region which has a largely different color direction from the background one. Then, the shadow seed is decided based on homography consistency and color-based shadow likelihood (see Chapter 3.2 for detail).

3.2 Data term

The data term is defined by the log of the likelihood as,

$$\mathbf{g}(\mathbf{x}_{\mathbf{v}}) = -\log\left(P(\mathbf{x}_{\mathbf{v}}|\mathbf{c}(\mathbf{v}))\right) = -\log\left(\frac{P(\mathbf{c}(\mathbf{v})|\mathbf{x}_{\mathbf{v}})P(\mathbf{x}_{\mathbf{v}})}{\sum_{l_i \in \mathbf{L}} P(\mathbf{c}(\mathbf{v})|\mathbf{x}_{\mathbf{v}}=l_i)P(\mathbf{x}_{\mathbf{v}}=l_i)}\right), \quad (3)$$

where $P()$ is probability and $\mathbf{c}(\mathbf{v})$ is a six dimensional color vector at site \mathbf{v} composed of a pair of RGB vectors in each image as where $P()$ is probability and $\mathbf{c}(\mathbf{v})$ is a six dimensional color vector at site \mathbf{v} composed of a pair of RGB vectors in each image as

$\mathbf{c}(\mathbf{v}) = [\mathbf{c}(\mathbf{p}_v^l), \mathbf{c}(\mathbf{p}_v^r)]^T$, and $\mathbf{c}(\mathbf{p})$ is color vector at pixel \mathbf{p} . Then the pair-wise color observation model $P(\mathbf{c}(\mathbf{v}) | \mathbf{x}_v)$ is decomposed into $\prod_i P(\mathbf{c}(\mathbf{p}_v^i) | \mathbf{x}_v^i)$, where \mathbf{x}_v^i is the one side label and i ($i = l, r$) is the camera identifier.

Foreground model The foreground color is approximated by a pixel-wise GMM which is trained by k-means clustering from *foreground seed* pixels, and the foreground observation model is expressed as,

$$P(\mathbf{c}(\mathbf{p}_v^i) | \mathbf{x}_v^i = F) = \mathcal{N}(\mathbf{c}_f^{k^*}, \Sigma_f^{k^*}) \tag{4}$$

$$k^* = \arg \min_k \left(\left(\mathbf{c}(\mathbf{p}_v^i) - \mathbf{c}_f^k \right)^T \Sigma_f^{k-1} \left(\mathbf{c}(\mathbf{p}_v^i) - \mathbf{c}_f^k \right) \right), \tag{5}$$

where \mathbf{c}_f^k and Σ_f^k are a mean vector and a covariance matrix of the k th cluster, and \mathcal{N} is the Gaussian distribution.

Shadow-Background model First, a linear color transformation matrix from the background color to the shadow color is estimated from the *shadow seed* colors and their modeled background colors. This matrix is modeled as following a finite-dimensional linear model [19],

$$\mathbf{c}_s(\mathbf{p}) = \mathbf{A} \tilde{\mathbf{c}}_{bg}(\mathbf{p}), \tag{6}$$

where \mathbf{c}_s is a color vector of a shadow seed, $\tilde{\mathbf{c}}_{bg}$ is an extended color vector of a modeled background, $\tilde{\mathbf{c}}_{bg} = [\mathbf{c}_{bg}^T, 1]$, and \mathbf{A} is a 3 by 4 shadow transformation matrix. Then, the color transformation matrix \mathbf{A} is obtained by minimizing the following objective function S ,

$$\mathbf{e}(\mathbf{p}) = \mathbf{A} \tilde{\mathbf{c}}_{bg}(\mathbf{p}) - \mathbf{c}_s(\mathbf{p}) \tag{7}$$

$$S = \sum_{\mathbf{p} \in \mathbf{P}_s} \mathbf{e}(\mathbf{p})^T \left(\sum_{bg}(\mathbf{p}) \right)^{-1} \mathbf{e}(\mathbf{p}), \tag{8}$$

where \mathbf{e} and \sum_{bg} are the color transformation error vector and covariance matrix of the modeled background color, and \mathbf{P}_s is a set of shadow seed pixels.

Next we define the vector \mathbf{c}_r which is the nearest color to an input color \mathbf{c} on the line segment between the modeled background color \mathbf{c}_{bg} and the estimated shadow color $\hat{\mathbf{c}}_s = \mathbf{A} \tilde{\mathbf{c}}_{bg}$ in RGB color space as shown in Fig. 5. Then, the vector \mathbf{c}_r is expressed as

$$\mathbf{c}_r(\mathbf{p}_v^i) = \hat{t} \hat{\mathbf{c}}_s(\mathbf{p}_v^i) + (1 - \hat{t}) \mathbf{c}_{bg}(\mathbf{p}_v^i) \tag{9}$$

$$t = \frac{(\hat{\mathbf{c}}_s(\mathbf{p}_v^i) - \mathbf{c}_{bg}(\mathbf{p}_v^i))^T (\mathbf{c}(\mathbf{p}_v^i) - \mathbf{c}_{bg}(\mathbf{p}_v^i))}{\| \hat{\mathbf{c}}_s(\mathbf{p}_v^i) - \mathbf{c}_{bg}(\mathbf{p}_v^i) \|} \tag{10}$$

$$\hat{t} = \min\{1, \max\{t, 0\}\}, \tag{11}$$

Finally the background and shadow observation models are introduced based on the interpolation on the line segment as,

$$P(\mathbf{c}(\mathbf{p}_v^i) | \mathbf{x}_v^i = S) = \hat{t} \mathcal{N}(\mathbf{c}_r(\mathbf{p}_v^i), \Sigma_r') \tag{12}$$

$$P(\mathbf{c}(\mathbf{p}_v^i) | \mathbf{x}_v^i = B) = (1 - \hat{t}) \mathcal{N}(\mathbf{c}_r(\mathbf{p}_v^i), \Sigma_r') \tag{13}$$

$$\Sigma_r'(\mathbf{p}_v^i) = \Sigma_r(\mathbf{p}_v^i) + \Sigma_e(\mathbf{p}_v^i), \tag{14}$$

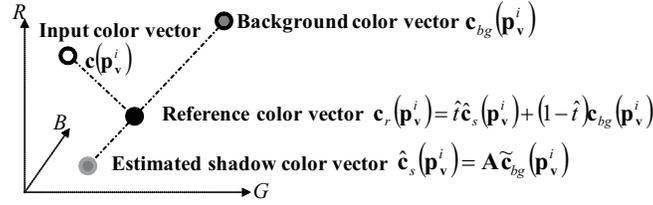


Fig. 5: Shadow-background model

where \sum_r and \sum_e are covariance matrices of the reference color \mathbf{c}_r and color transformation error \mathbf{e} .

3.3 Smoothness term

The smoothness term considering intensity value normalization is defined as,

$$\mathbf{h}(\mathbf{x}_v, \mathbf{x}_u) = \begin{cases} \exp(-\kappa d_e^l(\mathbf{x}_v, \mathbf{x}_u)) & \text{Left side label is different} \\ \exp(-\kappa d_e^r(\mathbf{x}_v, \mathbf{x}_u)) & \text{Right side label is different} \\ \exp(-\kappa\sqrt{d_e^l(\mathbf{x}_v, \mathbf{x}_u)d_e^r(\mathbf{x}_v, \mathbf{x}_u)}) & \text{Both side labels are different} \\ 0 & \text{Otherwise} \end{cases}, \quad (15)$$

where d_e^i is an edge intensity criteria given by,

$$d_e^i(\mathbf{x}_v, \mathbf{x}_u) = \frac{\|\mathbf{c}(\mathbf{p}_v^i) - \mathbf{c}(\mathbf{p}_u^i)\|^2}{\|\mathbf{c}(\mathbf{p}_v^i) + \mathbf{c}(\mathbf{p}_u^i)\|^2 + \epsilon}, \quad (16)$$

where κ and ϵ are coefficients for this term.

4 Experiments

4.1 Data set and parameters

We carried out experiments using sequences of people walking outdoors. Tab. 2 shows the details of the data set. Every sequence contains some men or women with strong shadows. A total of 3 images were provided for graph-cut segmentation in a block. Note that in some figures in this section, the results of the experimental images are trimmed around the segmentation target region because page space is limited.

In these experiments, the data terms were spatially smoothed in response to the magnitude of the edge pixels. Because the pixel color is quite variable, and it is unstable near the edge, the reliability of the data terms is very low for such pixels. The segmentation process was done iteratively, and there were 2 iterations. The parameters of the proposed method were experimentally set at $w_g = 3.0$, $w_h = 0.3$, $\kappa = 4.0$, and $\epsilon = 10^{-7}$. Initially the prior of each label is set as follows: $P(FB) = P(BF) = P(SS) = 0.16$, $P(FS) = P(SF) = P(FF) = 0.14$, $P(BB) = 0.1$. In addition, the distribution number of GMM was set at 6 for *SeqA* and at 10 for *SeqB* and *SeqC*. We adopted the α - β swap algorithm [18] to minimize our energy function Eq. (??).

Table 2: Data set for experiments

Sequence set	Image size	Image number	Frame rate
<i>SeqA</i>	640×480	32	30 fps
<i>SeqB</i>	620×280	12	9 fps
<i>SeqC</i>	620×280	24	9 fps

4.2 Benchmark

We compared the segmentation performance of the following three approaches,

- *Color*: the color-based method and its implementation is as follows. First we generate the *foreground seed* and *shadow seed* based on the shadow color properties [5] [7]. Then we label each pixel in one side image as “*F*”, “*S*”, or “*B*”.
- *Color + Homography (asymmetric)*: the method integrating color and homography, and its implementation, follows. The seed generation process is the same as the proposed method, and the energy of the color similarities between the homography-correspondence pixels are integrated into a data term as *homography data term*. Then we label each pixel in the same way as the color-based method.
- *Color + Homography (symmetric)*: the proposed method.

Note that while the *Color + Homography (symmetric)* is a binocular-based symmetric framework, *Color* and *Color + Homography (asymmetric)* are not, so they are implemented as a “multi-labeling problem for a pixel” in each side of the image.

4.3 Results

First, the multi-labeling results of the proposed method for each data set are shown in Fig. 6, Fig. 7 and Fig. 8. In each result, the labeling results are good even for the occlusions.

Second, the performance comparison results for the ground truth which is created manually are shown in Fig. 9. In this figure, we can see that *Color* tends to fail at the pixels whose colors are quite similar to the shadow color. In situations where there is a black school bag (Fig. 9(a)) or a black-haired human head (Fig. 9(b)), this region is initially mislabeled *shadow seed*. Furthermore, it is a problem of the color-based approach that the pixels whose colors are similar to the shadow cannot be identified in the foreground GMM model, because such pixels are poorly labeled as a *foreground seed*, so the foreground data terms of such pixels are very low, and as a result, these are mislabeled as shadows. This is inevitable for the color-based approach.

On the other hand, for the results of the *Color + Homography (asymmetric)* and *Color + Homography (symmetric)* approaches, such pixels are correctly labeled as foreground. This is because such pixels are initially labeled as foreground seeds by the homography constraints, and the foreground GMM model includes such color information, and foreground data terms are high. As for the results of *Color + Homography (asymmetric)*, however, we can see the occlusion problem as described in Section 2. In contrast, *Color + Homography (symmetric)* segments them correctly.

The quantitative performance comparisons are shown in Tab. 3. The performance of each method is evaluated by *F-measure*, which is defined as,

$$F = \frac{2PR}{P + R}, \quad (17)$$

Table 3: Quantitative evaluation results

Method	SeqA		SeqB		SeqC	
	<i>f</i>	<i>s</i>	<i>f</i>	<i>s</i>	<i>f</i>	<i>s</i>
<i>Color</i>	0.890	0.863	0.825	0.747	0.817	0.776
<i>Color + Homography (asymmetric)</i>	0.938	0.923	0.904	0.824	0.874	0.824
<i>Color + Homography (symmetric)</i>	0.940	0.900	0.919	0.858	0.899	0.864

f: foreground, *s*: shadow

where F is F -measure, and P and R are precision and recall.

In the tables, we see that the *Color + Homography (symmetric)* approach totally outperforms the other methods. For the *SeqA*, there is little difference between the *Color + Homography (asymmetric)* and the *Color + Homography (symmetric)* approaches. This is because the color data and homography data terms are well balanced in this sequence.

4.4 Discussions

Effective use of extracted shadow By making effective use of extracted shadow, our approach can obtain consistent labeling as well as information as to whether the occluded region belongs to the *shadow* or *background*. This means that we can get additional scene information. For example, because a whole shadow silhouette including the occluded shadow, can be seen as another projection from the viewpoint of a light source, we can say that one more different-view of the whole silhouette of the target foreground objects is extracted as shown in Fig. 10. This is quite valuable for many computer vision applications, especially silhouette based applications, like gait recognition, gesture recognition, 3D reconstruction by shape from silhouettes and so on. As for gait recognition, it is reported in [20] that the different views of silhouettes improve recognition, and more, shadow-based gait recognition scheme is proposed in [21].

In addition, homography-based object localization techniques have been proposed [15], where the position of the object is localized by estimating the intersection point of the object region and the shadow region. Hence, if the occluded shadow region is also extracted by the proposed method, the object localization accuracy is improved.

Extension to more complex scene or moving platform Although the assumption that the shadow appears on the ground plane may seem to be a heavy constraint, our method can be extended to more complex scenes by modeling scenes as piecewise facets and by calibrating the homography for each facet.

Furthermore, our method can be applied to a mobile platform such as a vehicle binocular video system, and an intelligent robot with a combination of state of the art dynamic background modeling, ego-motion, and image stabilizing techniques. For example, we can acquire a background model for each frame of the image sequence by using dynamic background modeling, and we can calibrate the geometric relationship between the binocular camera system and the target plane by using ego-motion and image stabilizing techniques.

5 Conclusions

In this paper, we propose a homography-correspondence pair based segmentation framework. We treat homography-correspondence pairs symmetrically, and formulate the segmentation problem as a multi-labeling problem for a homography-correspondence pair to explicitly take the occlusion relationship into account. Then we obtain the segmentation result by minimizing the energy function via the α - β swap algorithm. In our experiments, it turns out that the segmentation results of the proposed method outperform the existing color-based and asymmetric homography-based methods.

Acknowledgement. This work was supported by Grant-in-Aid for Scientific Research(S) 21220003.

References

1. Stauffer, C., Grimson, W.: Adaptive background mixture models for real-time tracking. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. (1999) 246–252
2. Choi, J., Jun, Y., Choi, J.Y.: Adaptive shadow estimator for removing shadow of moving object. *Computer Vision and Image Understanding* **114** (2010) 1017–1029
3. Cucchiara, R., Grana, C., Piccardi, M., Prati, A.: Detection moving objects, ghosts, and shadows in video streams. *IEEE Trans. Pattern Analysis and Machine Intelligence* **25** (2003) 1337–1342
4. Horprasert, T., Harwood, D., Davis, L.S.: A robust background subtraction and shadow detection. In: Proc. of the 4th Asian Conference on Computer Vision. (2000) 983–988
5. Huang, J.B., Chen, C.S.: Moving cast shadows detection using physics-based features. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. (2009) 2310–2317
6. Kakuta, T., Vinh, L.B., Kawakami, R., Oishi, T., Ikeuchi, K.: Detection of moving objects and cast shadows using a spherical vision camera for outdoor mixed reality. In: Proc. on 15th ACM Symposium on Virtual Reality Software and Technology. (2008) 219–222
7. Porikli, F., Thornton, J.: Shadow flow: A recursive method to learn moving cast shadows. In: Proc. IEEE Int. Conf. on Computer Vision. Volume 1. (2005) 891–898
8. Prati, A., Mikic, I., Trivedi, M.M., Cucchiara, R.: Detecting moving shadows: Algorithms and evaluation. *IEEE Trans. Pattern Analysis and Machine Intelligence* **25** (2003) 918–923
9. Tanaka, T., Shimada, A., Arita, D., Taniguchi, R.: Non-parametric background and shadow modeling for object detection. In: Proc. on 8th Asian Conference on Computer Vision. (2007) 159–168
10. Gordon, G., Darrell, T., Harville, M., Woodfill, J.: Background estimation and removal based on range and color. In: Proc. of the 18th Int. Conf. on Pattern Recognition. (1999) 459–464
11. Madsen, C.B., Moeslund, T.B., Pal, A., Balasubramanian, S.: Shadow detection in dynamic scenes using dense stereo information and an outdoor illumination model. In: Proc. on the DAGM 2009 Workshop on Dynamic 3D Imaging. (2009) 110–125
12. Batavia, P.H., Singh, S.: Obstacle detection using adaptive color segmentation and color stereo homography. In: Proc. of IEEE Int. Conf. on Robotics and Automation. Volume 1. (2001) 705–710

13. Hamid, R., KrishanKumar, R., Grundmann, M., Kim, K., Essa, I., Hodgins, J.: Player localization using multiple static cameras for sports visualization. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. (2010) 731–738
14. Jeong, K., Jaynes, C.: Moving shadow detection using a combined geometric and color classification approach. In: Proc IEEE Workshop on Motion and Video Computing 2005. Volume 2. (2005) 36–43
15. Kasuya, N., Kitahara, I., Kameda, Y., Ohta, Y.: Robust trajectory estimation of soccer players by using two cameras. In: Proc. of the 19th Int. Conf. on Pattern Recognition. (2008) 1–4
16. Kolmogorov, V., Zabih, R.: Computing visual correspondence with occlusions via graph cuts. In: Proc. IEEE Int. Conf. on Computer Vision. (2001) 508–515
17. Kolmogorov, V., Zabih, R.: Graph cut algorithms for binocular stereo with occlusions. *Mathematical Models in Computer Vision: The Handbook* (2005) 423–438
18. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Analysis and Machine Intelligence* **23** (2001) 1222–1239
19. Marimont, D.H., Wandell, B.A.: Linear models for surface and illumination spectra. *Journal of the Optical Society of America* (1992) 1905–1913
20. Sugiura, K., Makihara, Y., Yagi, Y.: Gait identification based on multi-view observations using omnidirectional camera. In: Proc. on 8th Asian Conference on Computer Vision. (2007) 452–461
21. Iwashita, Y., Stoica, A.: Gait recognition using shadow analysis. In: 2009 Bio-inspired Learning and Intelligent Systems for Security. (2009) 26–31

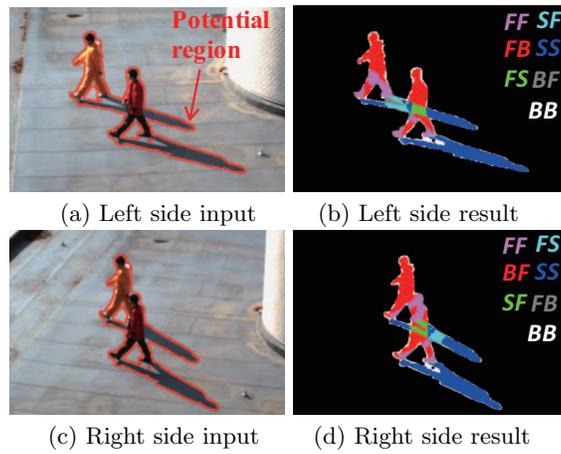


Fig. 6: Input and segmentation results of *SeqA*

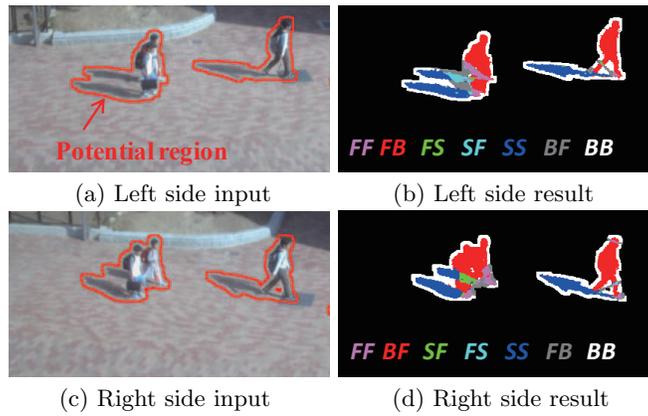


Fig. 7: Input and segmentation results of *SeqB*

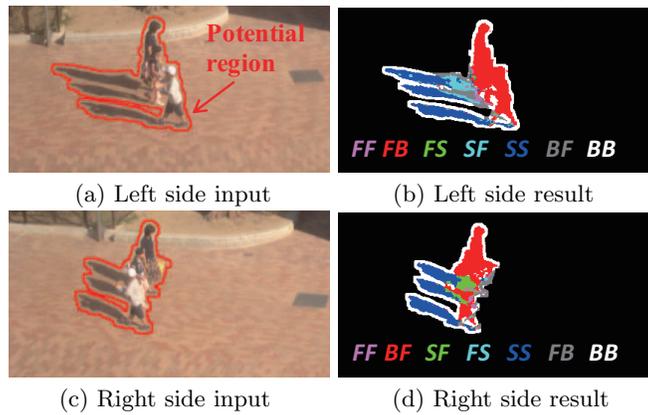


Fig. 8: Input and segmentation results of *SeqC*

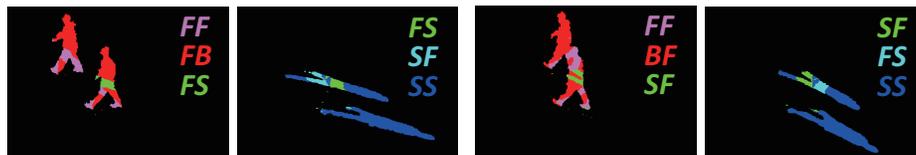
	<i>Seed</i>	Segmentation result (left side image)	
		foreground	shadow
<i>Color</i>			
<i>Color + Homography (asymmetric)</i>			
<i>Color + Homography (symmetric)</i>			

(a) *Seq B*

	<i>Seed</i>	Segmentation result (left side image)	
		foreground	shadow
<i>Color</i>			
<i>Color + Homography (asymmetric)</i>			
<i>Color + Homography (symmetric)</i>			

(b) *Seq C*

Fig. 9: Comparison Results



(a) Left side result

(b) Right side result

Fig. 10: Extracted foreground and a whole shadow including occluded shadow for *SeqA*