Stereoscopic image inpainting considering the consistency of texture similarity

Ayako ABE and Ikuko SHIMIZU

Tokyo University of Agriculture and Technology,
2-24-16 Nakacho, Koganei-City, Tokyo, Japan
50010834301@st.tuat.ac.jp, ikuko@cc.tuat.ac.jp

Abstract. Recently, as 3D display devices become popular, automatic editing methods for stereoscopic image become important. In this paper, we propose an inpainting method for a stereoscopic image considering the texture similarity. By extending the conventional image inpainting method for an image considering the texture similarity to for a stereoscopic image by considering the consistency of two images, natural results are obtained by our method. As in the conventional methods for a stereoscopic image inpainting, first the depth maps are recovered and then the color images are inpainted based on the depth maps in our method. Consistency of the depth maps is guaranteed in our method which are important for the inpainting of color image. For color image inpainting, similar textures are searched considering not only the depth layer but also the depth values of respective pixels. Experimental results demonstrates the efficiency of our method.

Keywords: stereoscopic image, inpainting, texture consistency, depth, disparity, 3D warping

1 Introduction

In recent years, techniques for automatic generation and efficiency editing of stereoscopic images are required to be developed, because many 3D displays become very popular which enables users to feel depth by showing different images to the left eye and right eye respectively. A pair of the images for showing to left eye and right eye is called a 'stereoscopic image'.

Among many editing techniques, image inpainting is one of the most important ones. It reconstructs the input image which contains "hole(s)", so that it looks natural. In this paper, we denote an inpainting method for an (ordinary) image as '2D inpainting method' and an inpainting method for stereoscopic image as '3D inpainting method'.

There are a lot of 2D inpainting methods proposed in the literature. The 2D inpainting method can be classified into two categories: one is to inpaint from around the hole(s) considering the color continuity [1], and another is to inpaint using the texture of other part in the same image [2, 3]. Recently the latter is consider to be better because it can be apply to an image with a big hole.

For 3D inpainting, even if we apply the best 2D inpainting methods to respective images, the result will be unnatural because it leads to the inconsistent intensity between the two images consisting a stereoscopic image and the unnatural depth in the hole. Therefore, 3D inpainting methods has been proposed [4–6]. In these methods, not only color images but also depth maps of two images are inpainted. In Wang's method [4], both color images and depth maps are inpainted iteratively. However, the unnatural artifacts occurs because this method does not ensure the convergence of the iteration. Furthermore it allows that the corresponding pixels have slightly different the intensity values each other. A method by Hervieu [5] ensures the convergence and the color consistency by two step algorithm: in the first step, the depth maps are inpainted, and in the second step, the color images are inpainted under constraints that the corresponding pixels have the same intensity value using the depth maps. However, this method can be applied only to the images whose depth map is relatively smooth. Then, Hervieu [6] proposed another method which is applicable to the image whose depth maps have various values by introducing the assumption that the depth maps consist of some depth planes. However, depth map inpainting in this method does not guarantee the consistency of corresponding pixels. It leads to an unnatural color image inpainting because the depth maps are important to ensure the color image consistency. And, this method extended the 2D inpainting method [2] in color image inpainting, but it has a drawback in that it tends to induce discontinuous textures when the image has complex textures. In addition, the similar pixels are searched without considering the depth. Then, if two or more objects are contained in one depth layer, the inpainted pixels will be unnatural. To prevent this, user may tune the parameter, but it is difficult for users.

In this paper, a 3D inpainting method is proposed by considering the texture similarity to achieve consistent inpainting between two images and natural inpainting when the input image has complex texture. Our method guarantees consistency of color images and depth maps between two images. To obtain natural results for image with complex texture, 2D inpainting method [3] is extended to 3D which is applicable to the image with complex texture and does not induce the discontinuous texture by considering the similarity of texture. In this step, similar pixels are searched in the same depth layer considering the depth of the pixel. By our method, the similar textures are obtained which are supposed to belongs to the same object without tuning parameters.

2 Overview of proposed method

The overview our 3D inpainting method is shown in Fig.1. In our method, two step algorithm is employed as in the conventional methods [5,6]: in the first step depth maps are inpainted, and in the second step color images are inpainted to ensures the convergence. Our method guarantees the consistency between corresponding pixels both of the color images and the depth maps of two images which consists a stereoscopic image.

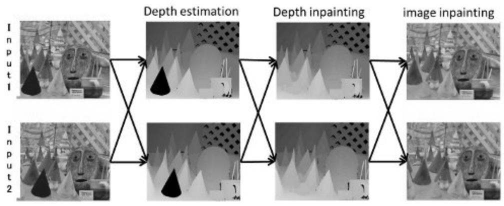


Fig. 1. The overview of the proposed method.

3 Estimation of depth map

First, we estimate the depth map of each image consisting a stereoscopic image. There are many methods for estimation of the depth maps. For example, graphcut algorithm [7] can be applicable.

Two depth maps for respective images are estimated by establishing correspondence between two images for each pixel of each image. Note that there are some holes in the depth map which correspond to the holes in the image in this step.

4 Inpainting of depth map

Hole of the depth maps are inpainted by following two steps: initialization by 3D warping described in Sec.4.1 and inpainting by extension of depth layer described in Sec.4.2.

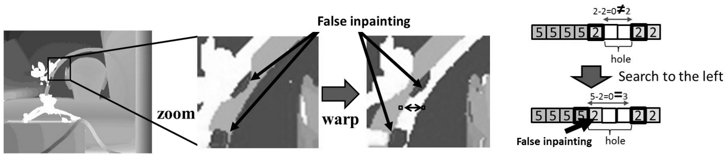
4.1 Initialization of depth map

We denote the pixels in the left image and right image by (x^l, y) and (x^r, y) and the depth maps of the respective images by D_l and D_r. The relation between the corresponding pixels of the depth maps are as follows:

$$D_l(x^l, y) = D_r(x^l - D_l(x^l, y), y), \quad D_r(x^r, y) = D_l(x^r + D_r(x^r, y), y). \quad (1)$$

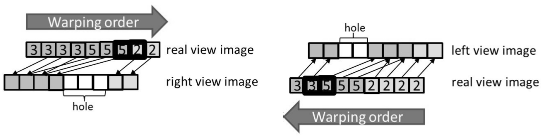
If a pixel (x^l, y) in the left image is in the hole and the corresponding pixel in the right image is not in the hole, $D_l(x^l, y)$ is obtained by $D_r(x^l - D_l(x^l, y), y)$. This transformation is called 3D warping. If $D_r(x^r, y)$ is in the hole and the corresponding pixel is not in the hole, $D_r(x^r, y)$ is obtained by 3D warping in the same way. When multiple pixels are warped to one pixel, the largest depth is selected since it corresponds to the closest object to the camera.

However, sometimes initialization by 3D warping is failed because the closer object to camera is occluded by the object in the hole and the smaller depth than the true depth is recovered (Fig.2). To eliminate such a false depth value, 3D warping property [8], that the width of the hole is the same as the difference of the depths between the pixels in both side of the hole, is used. The pixel which does not satisfy this property is supposed to have false depth value and the depth value is eliminated. For depth map of the left image, the false pixel is searched from the left side of the hole(Fig.3(a)). On the other hand, for depth map of the right image, the right side of the hole is searched(Fig.3(b)).



(a) (b)

Fig. 2. Error of the depth map.



(a) (b)

Fig. 3. (a)3D warping of the right image. (b) 3D warping of the left image. The number represents the depth for each pixel.

4.2 Inpainted by extension of depth layer

Next, the rest of the hole of the depth map is inpainted. We assume that the depth map consists of multiple depth planes, called depth layers [6]. The hole of the depth map is inpainted by extending its neighboring depth layer as shown in [6].

However, the recovered depth of each pixel does not guarantee to satisfy Eqn.(1). Therefore, our method iteratively estimates the depth of the hole until all pixels satisfy Eqn.(1). Concretely, the depth which does not satisfy Eqn.(1 is eliminated and the neighboring depth layer is extended to the hole again. These processes are repeated. The hole of the depth will be smaller when these processes are repeated and the depth of the hole will be converged to satisfy Eqn.(1).

5 Inpainting of color image

After completing to inpaint depth maps, color images are inpainted by two steps: initialization by 3D warping described in Sec.5.1 and inpainting by minimizing energy function considering texture similarity describe in Sec.5.2.

5.1 Initialization of color image

We denote the intensities of the respective images by I_l and I_r. The relation between the corresponding pixels of the color images are as follows:

$$I_l(x^l, y) = I_r(x^l - D_l(x^l, y), y), \quad I_r(x^r, y) = I_l(x^r + D_r(x^r, y), y). \tag{2}$$

As for the depth map, the hole of the color image is initialize by 3D warping using Eqn.(2). In this step, pixels in the hole which are observed in one of the stereoscopic image are recovered.

5.2 Inpainting of color image considering of consistency of the corresponding pixels' intensity and texture similarity

Next, the rest of the hole is inpainted by minimizing energy function considering texture similarity. Among many 2D inpainting method, a method by Kawai [3] is known to have flexibility to texture pattern and give natural results by considering the similarity of texture even when there is no exactly the same texture in non-missing region as in the hole. Therefore, we extend this method to 3D inpainting.

Extension to the 3D inpainting is done by adding constraints that the corresponding pixels between the left image and the right image has the same intensities. The energy function for 3D inpainting is as follows:

$$E = E_l + E_r, \quad \text{s.t.} \;\; g = I_l(x_i^l, y_i) - I_r(x_i^r, y_i) = 0, \tag{3}$$

where E_l and E_r is the energy in left and right image respectively, and they are same as the energy functions defined in 2D inpainting [3], and (x_i^l, y_i) and $(x_i^r, y_i) = (x_i^l + D_l(x_i^l, y_i), y_i)$ is the i-th corresponding pixel pair.

The minimization of E is done by greedy algorithm as follows: (i) update the similar texture patterns for respective pixels $(x_i^r, y_i), (x_i^l, y_i)$, (ii) update the intensity values in the hole. These two steps are repeated until convergence. In addition, we create an image pyramid and minimize the energy from coarse to fine.

In step (i), similar texture patterns are searched in only the depth layer which belongs to each pixel in the hole. In this step, our method restricts the search region of the similar pixel to the same depth layer considering the depth value of the pixel. By considering the depth value of each pixel, our method prevents to search the different object in the same layer. By searching the similar pixel for only the pixels belonging to the same depth layer, unnatural inpainting is supposed to prevent. The computational time will be reduced by restricting the searching region.

In step (ii), energy E which is a sum of both images' energy is minimized. The relation between E can be rewritten using energy elements $E_l(\boldsymbol{x}_i^l), E_r(\boldsymbol{x}_i^r)$ of respective images' respective pixels is as follows:

$$E = \sum_{i=1}^{N_\Omega} (E_l(\boldsymbol{x}_i^l) + E_r(\boldsymbol{x}_i^r)) + C, \tag{4}$$

where $\boldsymbol{x}_i^l = (x_i^l, y_i)$, $\boldsymbol{x}_i^r = (x_i^r, y_i) = (x_i^l - d, y_i)$, N_Ω is the number of pixels included in the hole, Ω' is the region including hole Ω and the region adjacent to Ω , and C is the energy in $\overline{\Omega} \cap \Omega'$. Note that C is a constant value in this step, because the similar pattern is fixed. To minimize E in Eqn.(4), the Lagrange multiplier are utilized. We define $J = E + \lambda g$ with constraint $g = 0$ in Eqn.(3) and the variables are $I_l(x_i^l, y_i), I_r(x_i^r, y_i)$. Then, the minimization of E is equivalent to these partial differential equations as follows:

$$\frac{\partial J}{\partial I_l(x_i^l, y_i)} = \frac{\partial E}{\partial I_l(x_i^l, y_i)} + \lambda = 0, \tag{5}$$

$$\frac{\partial J}{\partial I_r(x_i^r, y_i)} = \frac{\partial E}{\partial I_r(x_i^r, y_i)} - \lambda = 0, \tag{6}$$

$$I_l(x_i^l, y_i) - I_r(x_i^r, y_i) = 0. \tag{7}$$

$I_l(x_i^l, y_i), I_r(x_i^r, y_i)$ is obtained analytically from Eqn.(5), (6), (7) by eliminating λ from Eqn.(5) and Eqn.(6):

$$I_l(x_i^l, y_i) = I_r(x_i^r, y_i) =$$
$$\frac{\sum_{\boldsymbol{p}\in W}(w_l(\boldsymbol{x}_i^l+\boldsymbol{p})\alpha(\boldsymbol{x}_i^l+\boldsymbol{p})f(\boldsymbol{x}_i^l+\boldsymbol{p})+w_r(\boldsymbol{x}_i^r+\boldsymbol{p})f(\boldsymbol{x}_i^r+\boldsymbol{p}))I_l(f(\boldsymbol{x}_i^l+\boldsymbol{p})-\boldsymbol{p})}{\sum_{\boldsymbol{p}\in W} w_l(\boldsymbol{x}_i^l+\boldsymbol{p})+w_r(\boldsymbol{x}_i^r+\boldsymbol{p})}, \tag{8}$$

where w_r and w_l are the weights for each pixel of respective images, $f(\boldsymbol{x})$ is a similarity pattern of $\boldsymbol{x}$, and α is a correction parameter which allows the variation of the intensity as defined in [3].

6 Experiment results

In experiments in this section, we used the images and depth maps from middleburry stereo datasets(`http://vision.middlebury.edu/stereo/data/`).

6.1 Initialization of depth map

As mentioned in 4.1, this initialization process consists of two steps;(i) 3D warping, (ii) elimination of the false inpainting. In this section, the depth map by our method comparing with depth maps by method without these steps to show the effectiveness of our initialization of the depth map.

The depth map without initialization of the depth map Fig.4(a) and its inpainted depth map is shown in Fig.4(d). The depth map by only 3d warping without elimination of the false inpainting Fig.4(b) and its inpainted depth map is shown in Fig.4(e). The depth map initialized by our method Fig.4(c) and its inpainted depth map is shown in Fig.4(f). From these results, effectiveness of our depth initialization is shown, especially the region around round-arched object.

6.2 The consistency of the corresponding pixels' depth values

In our method, the depth inpainting is repeated until the the corresponding pixels have the same depth values, as described in Sec.4.2. The depth map without consistency check of the corresponding pixels is shown in Fig.5(a). In contrast, Fig.5(b) shows the result by our method. Better result is obtained by our method.

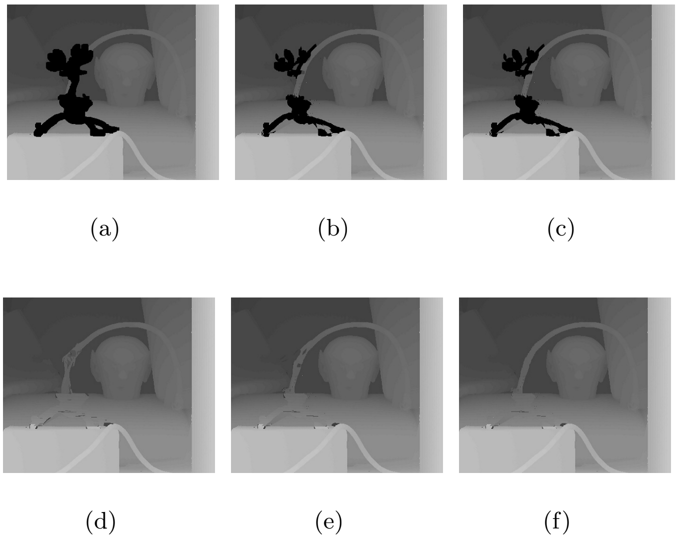


Fig. 4. The results of initialization of depth map;(a) without initialization, (b) only by the 3D warping (c) by out with method, (d),(e),(f) are the respective results after expanding of the depth layer from (a),(b),(c)

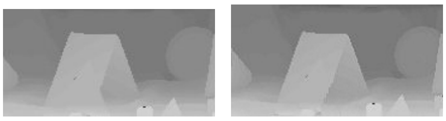
6.3 The comparison with our method and the conventional method

The results with our method and the conventional method are shown in Fig.6. In Fig.6(a), bottom of a red cone is not natural. In contrast, results by our method shown in Fig.6(b), looks natural. One of the reason is that our method initializes the color images by 3D warping based on the consistent depth maps.

7 Conclusion

We proposed an inpainting method for a stereoscopic image considering the similarity of texture based on the consistency of corresponding pixels between two images consisting a stereoscopic image. As conventional methods for stereoscopic image inpainting, our method consists of two steps to guarantee the convergence: the first step is inpainting of depth maps and the second step is inpainting of color images based on the depth maps.

In the depth map inpainting, our method guarantees the consistency of the depth maps which play an important role in the second step. In the color image inpainting, the conventional inpainting method [3], which does cause the discontinuous and unclear texture even for the input image with complex texture, were extended to 3D inpainting considering consistency of corresponding pixels between two images. In addition, similar textures are searched considering not only the depth layer but also the depth values of respective pixels.



(a) (b)

Fig. 5. The result considering the consistency of the corresponding pixels' depth

(a) (b)

Fig. 6. The comparison of 3D inpainting;(a)conventional method [6],(b)proposed method

We confirmed that more natural results of stereoscopic image were obtained by our method than by the conventional method. To extend our method for the stereoscopic video will be one of the future works by considering that the depth varies smoothly in time-line.

References

1. T.F.Chan ,J.Shen,"Non-Texture Inpainting by Curvature-Driven Diffusions (CDD)",Journal of Visual Communication and Image Representation, Vol.12, pp.436–449, 2001
2. A. Criminisi ,P.Perez et al, "Region filling and object removal by exemplar-based inpainting",IEEE Trans. on IP, Vol.13, No.9, pp.1200–1212, 2004
3. N.Kawai, T.Sato et al,"Image Inpainting Considering Brightness Change and Spatial Locality of Textures", Proc. of VISAPP, pp. 66–73, vol.1, 2008
4. L.Wang, H.Jin et al,"Stereoscopic inpainting: Joint color and depth completion from stereo images", Proc. of CVPR, pp.1–8, 2008
5. A.Hervieu ,N.Papadakis et al,"Stereoscopic Image Inpainting: Distinct Depth Maps and Images Inpainting", Proc. of ICPR, pp.4101–4104, 2010
6. A.Hervieu ,N.Papadakis et al,"Stereoscopic image inpainting using scene geometry", Proc. of ICME, pp.1–6, 2011
7. Y.Boykov , V.Kolmogorov, "An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision", IEEE Trans. on PAMI, Vol.26, No.9, pp.1124–1137, 2004
8. Y.R.Horng, Y.C.Tseng et al,"Stereoscopic images generation with directional Gaussian filter",Proc. of ISCAS, pp.2650–2653, 2010