JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 35, 23-39 (2019)
DOI: 10.6688/JISE.201901_35(1).0002

Artifact Reduction based Multi-view Generation Algorithm for Sparse Camera Configuration*

PEI-JUN LEE, TRONG-AN BUI AND CHIAO SU
Department of Electrical and Engineering
National Chi Nan University
Nantou, 545 Taiwan
E-mail: pjlee@ncnu.edu.tw; trongan93@gmail.com; aabbabcwinter@hotmail.com

A high quality multi-view generation algorithm is proposed to eliminate the camera parameter limits for a sparse camera configuration. The proposed algorithm includes multi-view generation and artifact reduction. The feature points of the foreground objects in the input images are used to find the disparity function for the central view and then we use this function to determine the disparity calibration function for all virtual views to replace the camera parameters. To improve the generated virtual image quality, artifact reduction based on a modified inpainting algorithm is proposed to reconstruct the holes and false contours.

Experimental results show that the quality of the virtual view images in multi virtual view generation and in artifact reduction can be raised by 2.67dB and 4.57dB, respectively in terms of PSNR. From the experimental results, it is found that the proposed algorithm can provide multi virtual views of high quality and can be adapted to different multi-view display systems.

***Keywords*:** multi-view video, virtual view synthesis, camera parameter estimation, image inpainting, artifact reduction

1. INTRODUCTION

In recent years, 3D technologies have become very popular. To improve the 3D viewing experience, multi-view video (MVV) has been developed for a new 3DTV application. Multi-view TV can provide multiple viewers who are in different positions relative to the display system a simultaneously real 3D perceptual experience without the need for further equipment such as 3D glasses. Since the cost of traditional video captured by multiple cameras is too high, especially in sports arenas, the cameras are set up in a sparse camera configuration as shown in Fig. 1 [1]. The depth image based rendering (DIBR) algorithm [2, 3] is a popular method to achieve virtual views for a multi-view display. However, there are two problems encountered in the process of virtual view generation for sparse camera configuration, one is the camera parameter estimation for virtual view generation, and the other is the existence of artifacts such as small cracks and holes in the generated views. The above two problems will cause the viewer to be dissatisfied.

The projection matrix of the camera parameter estimation is usually used to create the virtual view, but the difficulties of obtaining the projection matrix result in a more

Received July 28, 2017; revised October 12 & November 26, 2017; accepted December 8, 2017.
Communicated by Jing-Ming Guo.
* This work was supported in part by the Ministry of Science and Technology of Taiwan, and the Ministry of Economic Affairs of Taiwan, under Contract No. MOST 103-2221-E-260-033-MY3 and 101-EC-17-A-02-S1-201, respectively.

complex virtual view generation process. Since the misalignment is between color and depth, the false contour appears in the background. When generating virtual views, the previous works [4-6] generally used the camera parameters. In the paper [4], their algorithm uses the camera parameters to generate the projection matrix, and uses it to process 3D warping, and generates the virtual views. The papers [5, 6] use camera parameters for the horizontal warping and the linear interpolation function to find the relationship between the different views in order to generate the virtual views. The methods in the papers [4-6] all use the camera parameters to generate virtual views. However, the camera parameters are difficult to obtain, so that, the multi-view display becomes inconvenient in terms of use. To eliminate the camera parameter limits, the paper [7] uses the disparity estimation method based on region segmentation to replace the camera matrix calculation. To reduce the computation complexity of stereo matching, their algorithm uses arithmetic mean and standard deviation of depth value distribution to divide an image into three regions; which are the foreground region, medium region and background region. Then it uses feature points to perform stereo matching thus finding the disparity between the regions. However, their algorithm cannot accurately divide object regions using depth value distribution. The stereoscopic 3D effect is not sufficient when using only the disparity to create the virtual views.

Fig. 1. Sparse camera configuration.

Since the warping function of DIBR is a nonlinear function, some artifacts exist in the generated images. The false contour is a very common artifact. The paper [8] proposes an adaptive edge-oriented smooth preprocessing method to deal with geometric distortions with low computation complexity. The paper [9] estimates the maximum allowable depth distortion to synthesize a virtual view without introducing any geometry distortion so that the fast and efficient depth compressions are achieved. To enhance the quality in a region of the virtual view which is affected by an artifact, some studies [10, 11] focus on hole-filling for artifact reduction and use the inpainting algorithm to reconstruct the holed regions. In addition, false contour effects will appear in the background region, when a misalignment exists between the color image and the depth map. False contours also negatively impact 3D perception. The papers [12, 13] design a filter to smooth false contours. The hole dilation method is proposed in the paper [14], in which the method is applied before the image blending, then the hole is expanded to eliminate the false contour. These methods can reduce the false contours. However, if the hole

region is too large, the quality of the resulting virtual image will be decreased. In additional, objective crosstalk metric in stereoscopic displays by method applying a 2D structural similarity (SSIM) map result and depth map to consider perceptual 3D crosstalk attributes by Xing *et al*. [15]. To obtain the very high quality image, paper [16] adopts deep learning technology to remove false contour in HEVC compressed images. However, those algorithms are applied to the 2D images.

In this paper, a high-quality multi-view generation algorithm is proposed in which the virtual multi-views are generated for a multi-view display in a sparse camera configuration. This paper uses the feature points of the foreground objects to find the disparity of the input images. Then this value is used to determine the disparity calibration function for all virtual views to replace the camera parameters. Based on our previous work [7], this paper uses the disparity estimation method for finding the object regions to replace the camera matrix calculation, then the warping module and image blending [17] are used to create a high-quality virtual multi-view. To achieve a high-quality 3D viewing experience, a modified image inpainting algorithm is proposed for artifact reduction and then the bilateral filter [12] is applied to improve the quality of the hole and false contour regions.

This paper is organized as follows. In section 2, the multi-view synthesis algorithm using two pairs of images; one pair consisting of color images and another pair corresponding depth images, is proposed. The experiment result of the proposed algorithm is presented in section 3. In section 4, the conclusions are given.

2. THE PROPOSED ALGORITHM

To achieve virtual multi-views for a multi-view display in a sparse camera configuration, the proposed algorithm is applied to generate the virtual multi-views using a pair of color images and their corresponding depth images. The proposed algorithm is divided into two parts as shown in Fig. 2. One is virtual multi-views generation in which the input images and their corresponding disparity tables are used to generate the virtual multi-views for multi-view display. The other one is artifact reduction which is applied to improve the generated virtual image quality when the virtual image contains some artifacts such as holes or false contours.

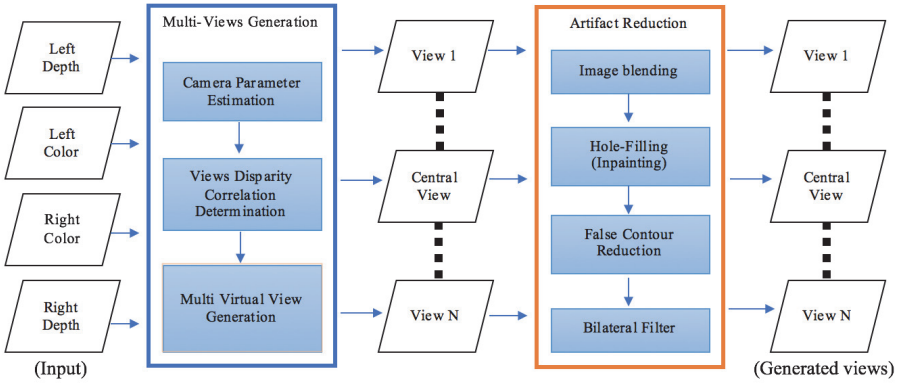


Fig. 2. Structure of the proposed algorithm.

2.1 Multi View Generation

The proposed virtual multi-views generation in sparse camera configuration includes camera parameter estimation, view disparity correlation and virtual multi-views generation.

(A) Camera Parameter Estimation

The distance difference between the two input images, which is called the disparity value, is an important parameter in the generation of virtual multi-views. We can use these disparity values and the warping module of DIBR to generate the virtual multi-views. To estimate the camera parameters, the proposed method uses the feature points of the left and right color input images and their depth map information to estimate the disparity of the input images. The flowchart of the feature based disparity matching is shown in Fig. 3.

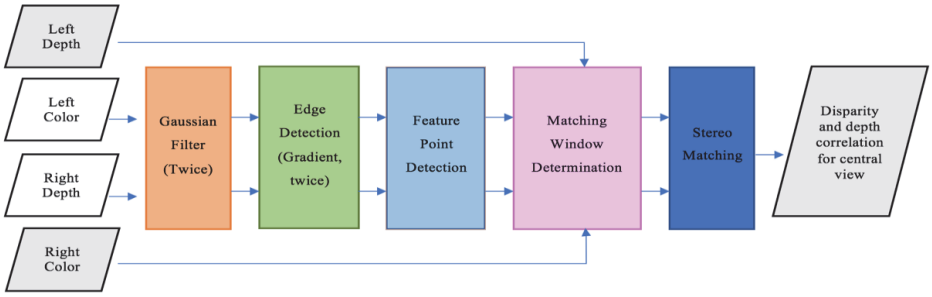


Fig. 3. Flowchart of camera parameter estimation.

To improve the accuracy of the disparity estimation and reduce the computation time of the stereo matching, the proposed method uses the feature points as the matching points for finding the disparity. For the feature point detection, the Gaussian filter is used to reduce the noise in the input images, and the gradient operator is used twice to detect the edge(s) as defined in Eq. (1).

$$s(x,y) = \sqrt{(s_h(x,y))^2 + (s_v(x,y))^2} \tag{1}$$

where $S_h = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * A;\ S_v = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A.$

A denotes the image which has been Gaussian filtered twice in order to reduce the noise. S_h and S_v denote the horizontal and vertical edge detection matrices, respectively, and $S_h(x, y)$ and $S_v(x, y)$ denote the elements of S_h and S_v at the location (x, y). * denotes the convolution operation.

$$(x^*, y^*) = \arg\max_{x,y}(|f_h(x,y)| + |f_v(x,y)|) \tag{2}$$

where $F_h = \begin{bmatrix} 0\,0\,0 \\ -1\,0\,1 \\ 0\,0\,0 \end{bmatrix} * S;\ F_v = \begin{bmatrix} 0\,1\,0 \\ 0\,0\,0 \\ 0\,-1\,0 \end{bmatrix} * S.$

Then feature point (x^*, y^*) defined in Eq. (2) is the point on the edge with the maximum gradient variation in each block. F_h and F_v denote the gradient matrices of the horizontal and the vertical, respectively; and $f_h(x, y)$ and $f_v(x, y)$ denote the elements of F_h and F_v at the location (x, y), respectively. The result of the feature point detection for the left and the right views are shown in Figs. 4 (a) and (b), respectively.

(a) Feature points of the left view. (b) Feature points of the right view.

Fig. 4. Feature point detection.

To estimate the disparity between the left and right input images, the adaptive matching window (*i.e.*, w_R) in the right color image is used to find a similar window (*i.e.*, w_L) in the left color image, and then we calculate the distance between the two windows. Therefore these feature points in the right view are the matching points and used to produce its matching windows in the stereo matching process. The size of w_R for the feature point is determined by the variance of the depth values Eqs. (3) and (4), where Eqs. (3) and (4) denote the window sizes on the horizontal axis and vertical axis, respectively.

$$|D_R(x, y) - D_R(x+1, y)| + |D_R(x, y) - D_R(x-1, y)| \le t, \tag{3}$$

$$|D_R(x, y) - D_R(x, y+1)| + |D_R(x, y) - D_R(x, y-1)| \le t, \tag{4}$$

where $D_R(x, y)$ ($D_L(x, y)$) denotes the depth value at the location of the pixel (x, y) in the right (left) view, and t is a threshold which is set as 5 here.

In order to find the disparity, the criterion of the disparity determination is defined in Eq. (5),

$$dis_c = \arg \left\{ \min_{d \in search\ range} \left[\sum_{y}^{y+w_{Ry}} \sum_{x}^{x+w_{Rx}} \left| C_R(x, y) - C_L(x+d, y) \right| \right] \right\} \tag{5}$$

where C_R and C_L denote the pixel values of the matching window in the right and left color images, respectively. d is the search region size. c is the central view. dis_c denotes the disparity value of the feature points. After the stereo matching process, the depth value is used to find the corresponding disparity in the depth map. If D_c denotes the depth value of the central view, then $D_c(dis_c)$ denotes the depth value of the corresponding disparity dis_c. Fig. 5 is an illustration of the disparity and depth correlation for the first frame in the central view of the Café sequence.

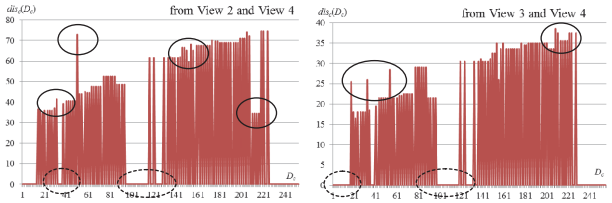


Fig. 5. Disparity and depth correlation for the central view.

(B) View Disparity Correlation Determination

For the 3D viewing experience, the correlation function of disparity and depth should be an increasing function. However, as can be seen in Fig. 5, there are two problems with the disparity and depth correlation function which arise from the camera parameter estimation. One is that some depth values do not have corresponding disparity values, as shown in the broken line circles of Fig. 5. In this case, we need to assign the disparity values using the disparity of the closest depth value. The other problem occurs when there is no correlation between the depth values and its corresponding disparity value as shown in the rigid circles of Fig. 5. In order to provide the viewer with a high-quality 3D viewing experience, this paper proposes a disparity refinement method to modify the correlation of the estimated disparity and the depth values as defined in Eq. (6). The results of the modified disparity and depth correlation for the central view as Fig. 5 are shown in Fig. 6.

$$D_c(dis_c) = \begin{cases} D_c(dis_c - 1), & \text{if } D_c(dis_c - 1) > D_c(dis_c) \text{ or} \\ & D_c(dis_c) > D_c(dis_c + 1) \\ D_c(dis_c), & \text{otherwise} \end{cases} \tag{6}$$

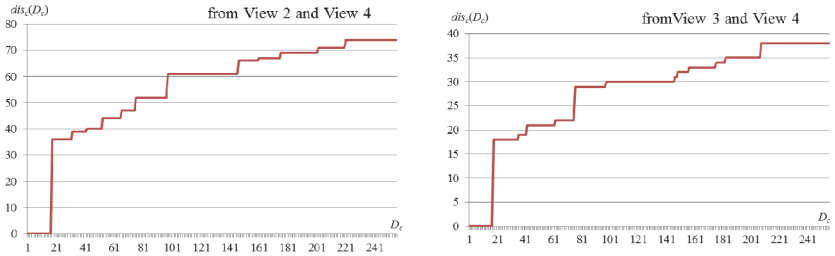


Fig. 6. Modified disparity and depth correlation for the central view.

When the correlation function of the disparity and the depth value for the central view is obtained from Eq. (6), the central disparity values (disc) and the three-dimensional space transform [18] are used to generate the correlation of the disparity and the depth value for the other views. Then the virtual multi-views can be generated using the corresponding disparity correlation. The correlations of the disparity and depth value for other views $D_c(dis_c)$ are defined in Eq. (7),

$$D_c(dis_c) = D_{c\pm i}(dis_c) \pm d_i,\ i = 1, 2, 3, 4 \tag{7}$$

where c is the central view number, $c\pm i$ denotes the view number. (*i.e.* $c+i$ views are the right side virtual views of the central view, $c-i$ views are the left side virtual views of the central view). d_i denotes the disparity calibration function as Eq. (8) with consideration of the 3D TV viewer's experience.

$$d_i = d_{i\_\max} \times \left(\frac{Z_{TV} - Z}{Z} \right),\ i = 1, 2, 3, 4 \tag{8}$$

Z_{TV} is the distance between the display and the viewer. $Z_{TV} = 500$ cm is set here. Z is a real-world distance between the object and the viewer. In order to find the disparity between two views, the depth value is transformed into the distance (Z) from the 3D display at which the viewer must sit to enjoy a satisfactory 3D viewing experience, and then we use an inverse transform to obtain the disparity between the different views as Eq. (9) [18].

$$Z = Z_{far} - \frac{D_c}{255}(Z_{far} - Z_{near}) \tag{9}$$

where Z_{far} and Z_{near} are the farthest distance in the positive parallax and nearest distance in the negative parallax, respectively. Here, we set $Z_{far} = 550$cm and $Z_{near} = 400$cm. $d_{i\_\max}$ is the maximum disparity between different views and is set $d_{i\_\max} = 3+2(c-i)$.

(C) Virtual Multi-views Generation based on Bi-directional Warping

For the virtual multi-views generation, the Z proposed algorithm uses the disparity function of each view to perform the bi-directional warping process of DIBR. The warping module uses forward warping and backward warping on the depth maps and the color images, respectively. The flowchart is shown in Fig. 7 [4]. There are two sets of generated virtual views at the same virtual view position.

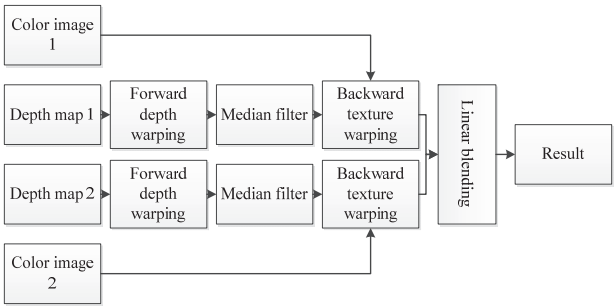
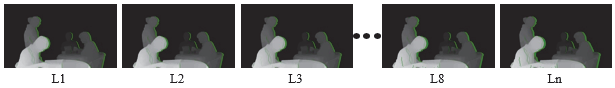
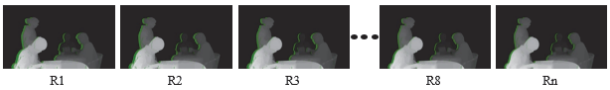


Fig. 7. Forward and backward warping based on DIBR [4].

The generated virtual depth maps for each view from the left and right depth images are shown in Figs. 8 (a) and (b), respectively. The generated virtual color images for each view from the left and right color images are shown in Figs. 8 (c) and (d), respectively.



(a) Depth maps from the left depth image.



(b) Depth maps from the right depth image.

(c) Virtual color images from the left color image.

(d) Virtual color images from the right color image.

Fig. 8. Forward and backward warping for virtual multi-views.

Since there are two sets of generated virtual views at the same virtual view position, the process of image blending for the pixels of the virtual view selection can be adapted to avoid mismatching the warped pixels and to improve the quality of the generated virtual view images. The method of image blending adopts the idea from the paper [18] which selects suitable pixels for the generated virtual view from two sets of virtual images. The idea of image blending for the virtual view is that the pixel value is selected from a side of the main virtual image, (*i.e. c+i* views use the right side, *c–i* views use the left side), if the pixel is not a hole (pixel value is available) in the main virtual image. There are three different types of image blending as shown in Table 1. In case 1, means the pixel values in the main virtual image are available and are used as the generated virtual view pixel values. In case 2, the hole is in the main virtual image and available pixels from another virtual image are chosen to be the generated virtual view pixel values. In case 3, the pixels in the two virtual images are holes, in this case, the proposed hole filling method is used to reconstruct the hole. The result of image blending for the pixels of the virtual view selection is shown in Fig. 9.

Table 1. Image blending cases.

	Pixel in main virtual image	Pixel in other virtual image	Pixel values used for virtual view
Case 1	Available	X (not required)	Side of the main virtual image
Case 2	Hole	Available	Side of the other virtual image
Case 3	Hole	Hole	Hole-Filling

Fig. 9. Result of image blending for the central view.

From the result of the image blending, we can see that two artifact effects (holes and false contours) have appeared in the image. The holes are enclosed by the red circles and the false contour is enclosed by the blue circle in Fig. 9. These artifacts will give the viewer an unsatisfactory experience. In this paper, an artifact reduction method is proposed to improve the quality of the generated virtual views.

2.2 The Proposed Artifact Reduction Algorithm Based on a Modified Image Inpainting Method

In this section, an artifact reduction algorithm based on a modified image inpainting method is proposed to reconstruct holes and eliminate false contours in the generated virtual views.

From the results of the forward warping and backward warping, the hole pixels can be flagged. Then the position of the hole pixels is known after the image blending process. The modified image inpainting algorithm is used to perform the hole filling.

False contour artifacts occur when a foreground edge (boy's head) appears in the background region as shown in the blue circle in Fig. 10. The false contour artifact will produce an unsatisfactory experience for 3D viewers. Therefore, false contour reduction is an important part of virtual view generation.

Fig. 10. False contour.

To detect the false contours, a gradient of pixel values is used to detect the edge as defined in Eq. (1) where A is now the image generated by the image blending process. However, this method detects both false and true edges. It is known that an edge which has a high depth value and is in the background region of the depth map is a false contour. If the depth values on both sides of the detected edge are deemed to belong to the background area, then that edge is considered to be a false contour.

To eliminate false contours and to reconstruct the artifact affected regions (*i.e.*, regions in which false contours and holes are found), this paper proposes a modified image inpainting algorithm which uses the concealment confidence (H_c) of the pixels surrounding the occluded pixels in the artifact affected regions. H_c is a measure of the amount of reliable information in the pixels surrounding the hole or false contour region within a patch τ_p. An illustration of the proposed algorithm is shown in Fig. 11 in which $a \in H_R$ is the center of the patch τ_p, H_R denotes the artifact affected regions, F_R denotes the foreground region, and B_R denotes the background region. To reduce the computation complexity of the confidence H_c and to avoid restarting the patching process, the proposed modified image inpainting selects the initial concealment patch from the first H_c

calculation for the artifact region reconstruction. The conceal patch candidate which has the highest confidence value is selected to be the concealment patch, and then the artifact pixels will be filled using data extracted from the source region from left to right. The concealment confidence H_c is defined in Eq. (10).

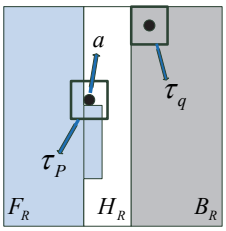


Fig. 11. Illustration of the modified image inpainting method.

$$H_c = \sum_{n=-3}^{n=3} \sum_{m=-3}^{m=3} B_H(x+m, y+n) \tag{10}$$

where B_H is a buffer to record the position of the hole or false contour region. If, $B_H(x, y) = 1$, then the location (x, y) is not a hole and is not in a false contour region, therefore, this pixel value is a suitable candidate for hole filling and false contour reduction. The candidate patch with the most similar τ_q, and the maximum H_c is determined from Eq. (11), where $SAD(\tau_p, \tau_q)$ denotes the sum of the absolute difference between the patches τ_p and τ_q. This patch is then used to fill the artifact $\hat{\tau}_p$

$$\hat{\tau}_p = \arg\min_{\tau_q \in B_R} SAD(\tau_p, \tau_q). \tag{11}$$

Finally, the bilateral filter [12] smooths the false contour and its surrounding region. The result of the artifact reduction is shown in Fig. 12. The artifact regions are already eliminated as shown in the blue circle in Fig. 12.

Fig. 12. The result for the central view.

The flowchart of the proposed false contour reduction algorithm based on a modified image inpainting method is shown in Fig. 13. An edge detection algorithm is applied to all depth maps and color images, then edge information in the color image and its corresponding depth edge is used to detect the false contours. The modified image inpainting and a bilateral filter are used to improve the quality of the generated virtual images.

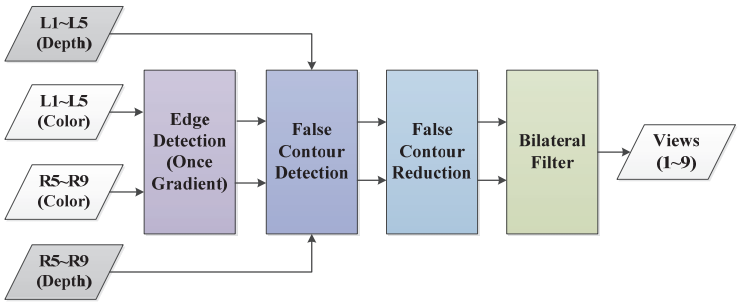


Fig. 13. Flowchart of artifact reduction.

3. EXPERIMENT RESULT

The comparison between the proposed algorithm and other existing algorithms can be divided into two parts. One tests the virtual multi-view generation. The comparison algorithm is proposed by the paper [7], which uses the disparity estimation method based on region segmentation to find the disparity of the two input views, then uses the estimated disparity to generate the virtual views. The other part of the comparison shows the results of the artifact reduction in the generated virtual views. The comparison algorithms for this part are proposed by the papers [12, 14], which use a bilateral filter and hole dilation to eliminate false contours.

The test sequences are Café, Lovebird, GT fly, Baby, Books and Bowling, and their frame sizes are 1920×1080, 1024×768, 1920×1088, 1240×1100, 1390×1100 and 1252×1100, respectively. Finally the nine virtual views are synthesized for multi view display, and the number of frames is 100. PSNR [19], SSIM [19] and CORR [20] indexes which are used to evaluate the proposed algorithm and the compared algorithms are as Eqs. (12)-(14), respectively.

$$PSNR(f, g) = 10\log_{10}(255^2/MSE(f, g)) \tag{12}$$

where $MSE(f,g)=\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(f_{ij}-g_{ij})^2$ is the mean square error between two images f and and g. f denotes the reference image, g is the test image, M and N denote the height and width of the image, respectively, f_{ij} and g_{ij} are pixel values. The *SSIM* index in Eq. (13) is a value belonging to the interval [0, 1], if the value is close to 1, the structural similarity is high.

$$SSIM(f, g) = |l(f, g)c(f, g)s(f, g)|, \tag{13}$$

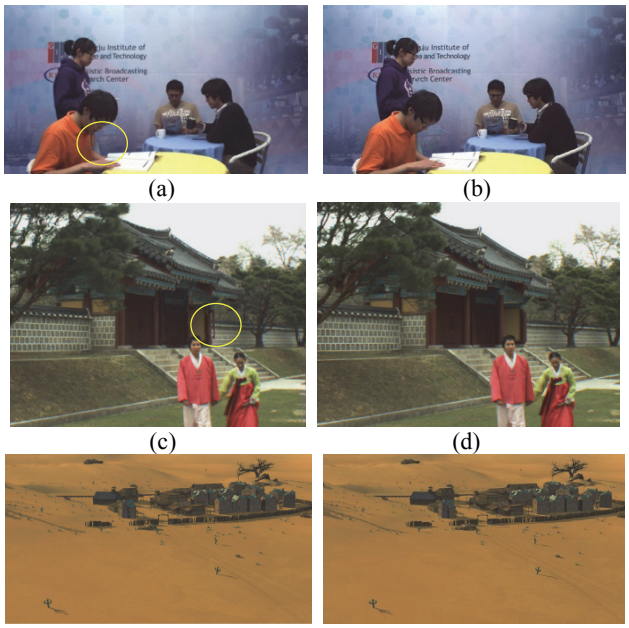
where $\begin{cases} l(f,g)=\frac{2\mu_f\mu_g+C_1}{\mu_f^2+\mu_g^2+C_1} \\ c(f,g)=\frac{2\sigma_f\sigma_g+C_2}{\sigma_f^2+\sigma_g^2+C_2} \\ s(f,g)=\frac{\sigma_{fg}+C_3}{\sigma_f\sigma_g+C_3} \end{cases}$, and $\begin{cases} \sigma_f=\frac{1}{MN}\sum_{i=0}^{M}\sum_{j=1}^{N}(f_{ij}-\mu_f)^2 \\ \sigma_{fg}=\frac{1}{MN}\sum_{i=0}^{M}\sum_{j=1}^{N}(f_{ij}-\mu_f)(g_{ij}-\mu_g) \end{cases}$,

μ_f and μ_g are the mean values of images f and g, respectively. σ_f and σ_g are the variances of images f and g respectively, σ_{fg} denotes the covariance of images f and g. The *CORR* index Eq. (14) is the correlation of the two images, if the value is close to 1, the correlation is higher.

$$CORR(f,g) = \left| \frac{\sigma_{fg}}{\sigma_f \times \sigma_g} \right| \tag{14}$$

3.1 Test for Virtual Multi-views Generation

The experiment results for the central view are shown in Fig. 14, and Table 2. Experimental results show that the proposed algorithm performs better than the other algorithms in the synthesized image quality. Due to the depth value discontinuity, the disparity estimation of the paper [7] will generate broken objects as shown in the yellow circle in Figs. 14 (a) and (c). On the Café sequence, the test results of [7] are better than those of the proposed algorithm. This is because the proposed algorithm estimates the disparity as zero in the background region, therefore, the background regions of the virtual views are filled using the background information of the input images. However, the disparity of the background region in the ground truth images may not be equal to zero, thus producing lower test results than that of [7]. Yet even in this special case, a good 3D experience is maintained in the generated virtual views.



(a) (b) (c) (d) (e) (f)

Fig. 14. Result of the central virtual view; (a) Result of literature [7] for Café; (b) Result of the proposed algorithm for Café; (c) Result of literature [7] for Lovebird; (d) Result of proposed algorithm for Lovebird; (e) Result of literature [7] for GT_fly; (f) Result of proposed algorithm for GT_fly.

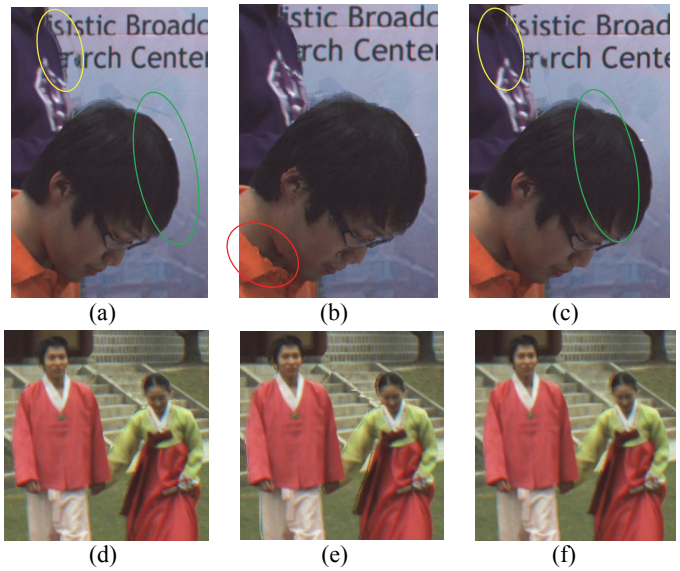
Table 2. Test score of multi virtual view generation.

Café	PSNR(dB)	SSIM	CORR
Literature [7]	33.0024	0.9956	0.9956
Proposed Algorithm	31.5693	0.9872	0.9871
Lovebird	PSNR(dB)	SSIM	CORR
Literature [7]	22.6656	0.9516	0.9517
Proposed Algorithm	24.9571	0.9716	0.9717
GT_fly	PSNR(dB)	SSIM	CORR
Literature [7]	26.8965	0.8285	0.8287
Proposed Algorithm	27.2756	0.8892	0.8892
Books	PSNR(dB)	SSIM	CORR
Literature [7]	18.0647	0.760072	0.760077
Proposed Algorithm	19.9389	0.8403	0.8405
Baby	PSNR(dB)	SSIM	CORR
Literature [7]	24.4138	0.815305	0.815331
Proposed Algorithm	24.5105	0.8278	0.8282
Bowling	PSNR(dB)	SSIM	CORR
Literature [7]	25.006569	0.933182	0.933253
Proposed Algorithm	29.2366	0.9754	0.9755

3.2 Test for Artifact Reduction

The result of a comparison of the artifact reduction performed by different algorithms on generated virtual views is shown in Fig. 15, and the test scores are shown in Table 3. In this experiment, the virtual views generated by the proposed virtual view generation algorithm are used for all comparison artifact reduction algorithms. Observing the results of different algorithms, all false contours are reduced by different methods, but there are still some phenomena as follows. The filter [12] has caused some blurring in the background, as shown in the yellow circle in Figs. 15 (a) and (c). In Fig. 15 (b), all the false contours are reduced using the method proposed in [14], but due to the different color information in the input images, the color discontinuity appears as in the red circle. In Fig. 15 (c), the results of the proposed algorithm reduce the false contours better than those of the other algorithms. In the Lovebird sequence, the depth value is not uniformly distributed; therefore, some false contours do not appear in the background region. This phenomenon will influence the image quality after false contour reduction. But the test score of the proposed algorithm is still better than those of the other algorithms.

Experimental results show that the quality of the virtual view image in the virtual multi-view generation and artifact reduction can be increased by 2.67dB and 4.57dB in PSNR, 0.08 and 0.09 in SSIM, and 0.081 and 0.091 in CORR, respectively. From the experimental results, the proposed algorithm can provide virtual multi-views of high quality, and can be adapted to different multi-view display systems.



(a) (b) (c)

(d) (e) (f)

Fig. 15. (a) Result of literature [12] for Café; (b) Result of literature [14] for Café; (c) Result of proposed algorithm for Café; (d) Result of literature [12] for Lovebird; (e) Result of literature [14] for Lovebird; (f) Result of proposed algorithm for Lovebird.

Table 3. Test score of artifact reduction.

Café	PSNR(dB)	SSIM	CORR
Bilateral Filter [12]	18.9289	0.883736	0.883790
Dilation [14]	19.8204	0.906193	0.906195
Proposed Algorithm	24.8693	0.9716	0.9714
Lovebird	PSNR(dB)	SSIM	CORR
Bilateral Filter [12]	21.369795	0.933599	0.933612
Dilation [14]	21.390028	0.934086	0.934094
Proposed Algorithm	24.9571	0.9716	0.9717
GT_fly	PSNR(dB)	SSIM	CORR
Bilateral Filter [12]	27.28	0.89	0.89
Dilation [14]	25.769239	0.886740	0.886749
Proposed Algorithm	27.3756	0.8992	0.8992
Books	PSNR(dB)	SSIM	CORR
Bilateral Filter [12]	19.94	0.84	0.84
Dilation [14]	19.555493	0.824367	0.824760
Proposed Algorithm	19.9689	0.8503	0.8505
Baby	PSNR(dB)	SSIM	CORR
Bilateral Filter [12]	24.315	0.82	0.82
Dilation [14]	24.352859	0.810036	0.810209
Proposed Algorithm	24.5105	0.8378	0.8382
Bowling	PSNR(dB)	SSIM	CORR
Bilateral Filter [12]	29.24	0.98	0.98
Dilation [14]	28.852294	0.973215	0.973275
Proposed Algorithm	29.4366	0.9854	0.9855

4. CONCLUSION

In this paper, an effective algorithm without camera parameters for virtual view generation in sparse camera configuration has been presented. The algorithm uses the feature points of the foreground edges to find the disparity function of the input images and then to find the disparity calibration function for all virtual views. To improve the quality of the generated virtual views, an artifact reduction algorithm based on the modified image inpainting algorithm is proposed. The proposed algorithm can not only generate virtual views in sparse camera configuration without camera parameters, but also effectively reduce the artifacts. Experimental results have shown that the proposed method can improve the quality by 0.029-5.940 PSNR, 0.005-0.088 SSIM and 0.006-0.088 CORR index depending on the sequences. In addition, the proposed method can be easily applied to different multi-view displays.

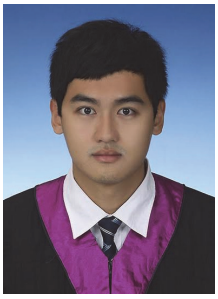
REFERENCES

1. http://www.hksilicon.com/kb/articles/504318.
2. C. Fehn, "Depth-Image-Based-Rendering (DIBR), compression and transmission for a new approach on 3D-TV," in *Proceedings of International Society for Optical Engineering Conference on Stereoscopic Displays and Virtual Reality Systems*, Vol. 5291, 2004, pp. 93-104.
3. P. Ndjiki-Nya, M. Koppel, D. Doshkov, H. Lakshman, P. Merkle, K. Muller, and T. Wiegand, "Depth image-based rendering with advanced texture synthesis for 3-D video," *IEEE Transactions on Multimedia*, Vol. 13, 2011, pp. 453-465.
4. ISO/IEC JTC1/SC29/WG11 MPEG/M15672, "View synthesis software and assessment of its performance," Hannover, Germany, 2008.
5. C. J. Kuo, C. Liao, and C. C. Jin, "Stereoscopic image generation based on depth images," in *Proceedings of International Conference on Image Processing*, 2004, pp. 2993-2996.
6. C. Lü, H. Wang, H. Ren, and Y. Shen, "Virtual view synthesis for multi-view 3D display," in *Proceedings of the 3rd International Joint Conference on Computational Science and Optimization*, 2010, pp. 444-446.
7. P. J. Lee, Y. C. Wang, C. Su and H. L. Lin, "Artifact reduction warping algorithm for virtual view generation," https://sites.google.com/view/viplab-publication/home/artifact-reduction-warping-algorithm-for-virtual-view-generation/, 2017.
8. P. J. Lee and Effendi, "Non-geometric distortion smoothing approach for depth map pre-processing," *IEEE Transactions on Multimedia*, Vol. 13, 2011, pp. 246-255.
9. S. Ma, S. Wang, and W. Gao, "Low complexity adaptive view synthesis optimization in HEVC based 3D video coding," *IEEE Transactions on Multimedia*, Vol. 16, 2014, pp. 266-271.
10. A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," in *Proceedings of IEEE International Conference on Image Processing*, 2004, pp. 1200-1212.

11. L. Wang and Y. Wan, "Multiresolution image inpainting," in *Proceedings of IEEE International Conference on Computer Science and Automation Engineering*, 2012, pp. 22-26.
12. H. R. Choi, J. W. Lee, R. H. Park, and J. S. Kin, "False contour reduction using directional dilation and edge-preserving filtering," *IEEE Transactions on Consumer Electronics*, Vol. 52, 2006, pp. 1099-1106.
13. J. W. Lee, B. R. Lim, and R. H. Park, "Two-stage false contour detection algorithm using re-quantization and directional contrast features and its application to adaptive false contour reduction," in *Proceedings of International Conference on Consumer Electronics*, 2006, pp. 377-378.
14. M. Li, H. Chen, R. Li, and X. Chang, "An improved virtual view rendering method based on depth image," in *Proceedings of the 13th IEEE International Conference on Communication Technology*, 2011, pp. 381-384.
15. L. Xing, J. You, T. Ebrahimi, and A. Perkis, "Assessment of stereoscopic crosstalk perception," *IEEE Transactions on Multimedia*, Vol. 14, 2012, pp. 326-337.
16. Q. Huang, H. Y. Kim, W. J. Tsai, S. Y. Jeong, J. S. Choi, and C.-C. J. Kuo, "Understanding and removal of false contour in HEVC compressed images," *IEEE Transactions on Circuits and Systems for Video Technology*, 2016, p. 99.
17. P. K. Tsung, P. C. Lin, L. F. Ding, S. Y. Chien, and L. G. Chen, "Single iteration view interpolation for multiview video applications," in *Proceedings of 3DTV Conference: The True Vision – Capture*, *Transmission and Display of 3D Video*, 2009, pp. 1-4.
18. L. Yu, S. Xiang, H. Deng, and P. Zhou, "Depth based view synthesis with artifacts removal for FTV," in *Proceedings of the 6th International Conference on Image and Graphics*, 2011, pp. 506-510.
19. A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *Proceedings of International Conference on Pattern Recognition*, 2010, pp. 2366-2369.
20. P. Singh and D. M. Chandler, "F-MAD: a feature-based extension of the most apparent distortion algorithm for image quality assessment," in *Proceedings of SPIE 8653 on Image Quality and System Performance X*, Vol. 8653, 2013.

Pei-Jun Lee received the Ph.D. degree from the Department of Electrical Engineering, National Taiwan University in 2004. Since August 2015, she has been Distinguished Professor of the Department of Electrical Engineering, National Chi Nan University, Taiwan. She became a Senior Member of IEEE in 2014. In 2008, she was the recipient of FIRST RUNNER-UP and outstanding advisor of the Macronix Golden Silicon Awards Semiconductor Design and Application Competition from MXIC Co., Ltd. Her research interests include 2D/3D image conversion, 3D/multi-view video coding, video communication, video compression, image processing, robot fish and FPGA system application.

Trong-An Bui received the B.Sc degree from the Faculty of Information Technology, Ho Chi Minh University of Education, Vietnam. He is currently pursuing the M.S. degree in Department of Electrical Engineering, National Chi Nan University, Taiwan. His research interests include image processing, video coding, and computer vision.



Chiao Su received the M.S. degree from Department of Electronic Engineering National Chi Nan University in 2014. His research interests multiview video generation.