

**INTERNATIONAL ORGANISATION FOR STANDARDISATION  
ORGANISATION INTERNATIONALE DE NORMALISATION  
ISO/IEC JTC1/SC29/WG11  
CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11  
MPEG2011/M21282  
July 2011, Torino, IT**

**Source: GIST (Gwangju Institute of Science and Technology)**  
**Status: Proposal**  
**Title: Temporally Consistent Hole Filling for View Extrapolation**  
**Author: Cheon Lee and Yo-Sung Ho**

## 1. Introduction

Dealing with holes in view synthesis is very important issue for 3D video systems. Particularly, visual artifacts due to hole in view extrapolation is severe. In the last Geneva meeting, we have implemented a hole filling method for view extrapolation in VSRS3.5.alpha [1]. However, the generated virtual viewpoint videos still have some flickering artifacts. In this document, we propose a modified hole filling method using gradient map and temporal enhancement using three consecutive frames for view extrapolation. The experimental results show that the flickering artifacts are significantly reduced and holes are filled more accurately.

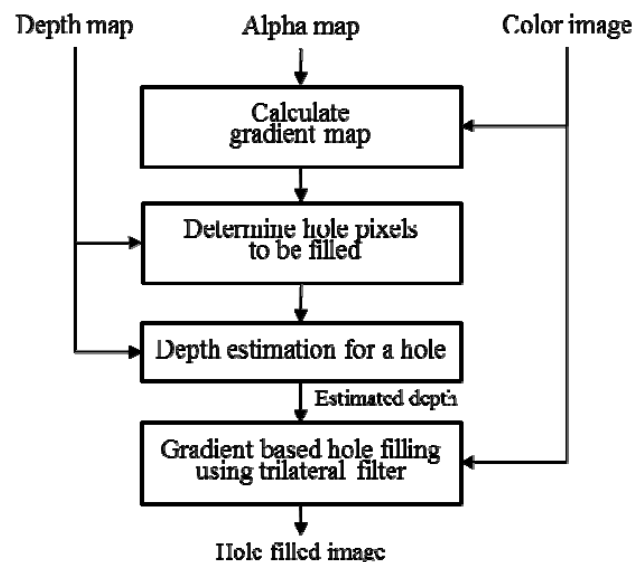


Fig. 1. View extrapolation method using gradient map

## 2. Gradient based hole filling method for view extrapolation

The former hole filling method for view extrapolation in VSRS3.5.alpha uses the bilateral filter considering the depth similarity and distance between the target pixel and reference pixels. Hence, the determined textured are blurred; the directions of textures are not survived. In this sense, we modified and expanded the bilateral filter to a trilateral filter. By taking it into account, directions of clear edges can be survived. Figure 1 describes the procedure of the proposed hole filling method for view extrapolation.

### 2.1 Gradient map of synthesized image

The gradient value is a vector representing magnitude and direction of pixel variations, hence we can utilize it to the hole filling to keep the directions of reference textures. We used a typical 3x3 masks to generate a gradient map except for the hole regions. Figure 1 shows both the synthesized with holes and gradient map. Gradient values of the gradient map are quantized into 256 levels to show the gradient values as an image. The direction of hole filling is the same with the former version of hole filling in [1]; from background to foreground.

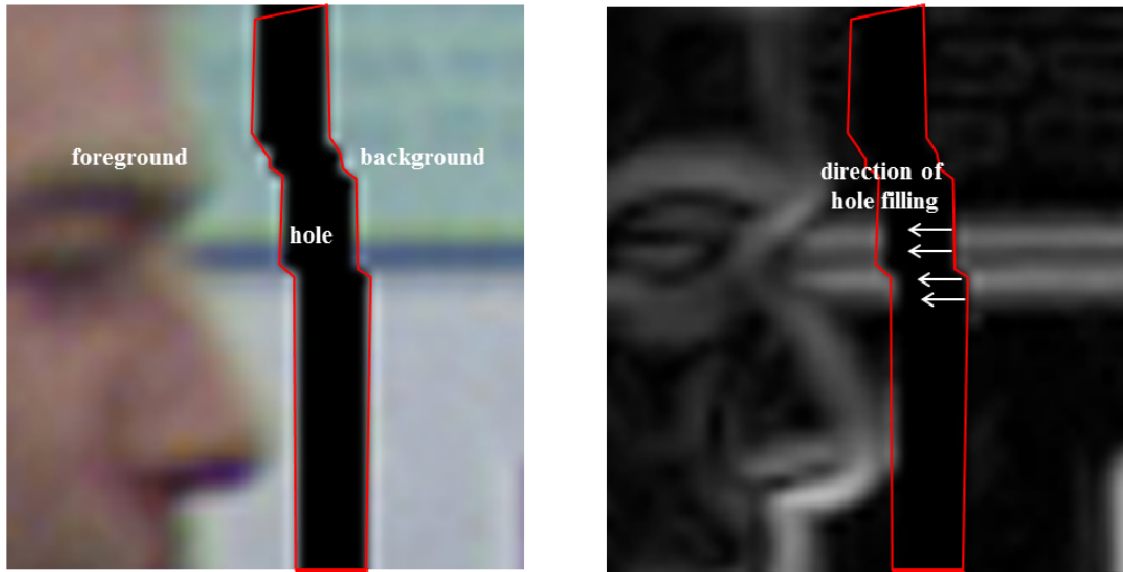


Fig. 2. Synthesized image with hole (left) and gradient map (right)

### 2.2 Estimating a virtual gradient value of a hole pixel

Similarly with the hole filling method proposed in VSRS3.5.alpha, we use a virtual gradient value to determine a pixel value for a hole. In order to estimate the most probable gradient value for a hole pixel, we first calculate an average gradient value using Eq. (1). In a certain window, we add all available gradient value, then we divide it by the number of available gradient values.

$$\bar{g} = \frac{1}{\sum_{i=0}^{N-1} \alpha(i)} \sum_{i=0}^{N-1} \alpha(i) \cdot g(i) \quad (1)$$

where  $\alpha(i)$  is an alpha value indicating hole. Then, we compare the directions of reference gradient vectors with the average gradient value. The virtual gradient value is determined by taking the most similar gradient value among the reference gradient values as shown in Eq. (2).

$$\hat{g} = \arg \min_{g(i)} |g(i) - \bar{g}| \quad (2)$$

### 2.3 Hole filling with trilateral filter

After estimating a virtual gradient value for a hole pixel, we determine a virtual depth value, which is the same with the previous method in [1]. Using these virtual values, we designed a trilateral filter as shown in Eq. (3), which is an expanded version of the former bilateral filter.

$$C(x, y) = \frac{\sum_{u \in \tilde{u}_p} \sum_{v \in \tilde{v}_p} W(u, v, \hat{d}, \hat{g}) \cdot C(u, v)}{\sum_{u \in \tilde{u}_p} \sum_{v \in \tilde{v}_p} W(u, v, \hat{d}, \hat{g})} \quad (3)$$

where

$$W(u, v, \hat{d}, \hat{g}) = \exp\left(-\frac{\|\hat{d}, D(u, v)\|^2}{2\sigma_D^2}\right) \exp\left(-\frac{(x-u)^2 + (y-v)^2}{2\sigma_r^2}\right) \exp\left(-\frac{\|\hat{g}, G(u, v)\|^2}{2\sigma_G^2}\right) \quad (4)$$

$$C(u, v) = \alpha(u, v) \cdot I(u, v) \quad (5)$$

## 3. Temporal enhancement for hole region

In the former hole filling method for view extrapolation, we employed a temporal refinement method for reducing a flickering effects on the hole regions by comparing two consecutive frames; a current frame and a previous frame. However, those artifacts were still generated due to the uncovered hole regions. To solve this problem, we modified the temporal refinement method using three consecutive frames, and we named this method temporal enhancement. This method is effective for a static background scene.

### 3.1 Temporal enhancement using three consecutive frames

If the background is static and the foreground objects have motions, we can find the corresponding textures of holes from the other frames. Figure 3 describes the concept of temporal refinement. The typical hole filling process is conducted for every frame, then the temporal enhancement uses three frames; one past frame ( $F_{N-1}$ ), one current frame ( $F_N$ ) and one future frame ( $F_{N+1}$ ). The input data for temporal enhancement are a

synthesized virtual image, a depth map, and an alpha map for every frame. The depth map is hole filled one with estimated virtual depth values at the hole filling process. The alpha map indicates all hole regions.

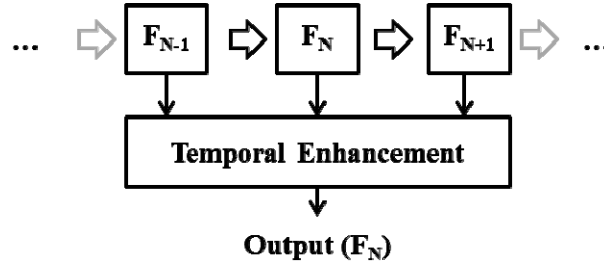


Fig. 3. Temporal enhancement using three consecutive frames

For a hole pixel of the current frame, we check the existence of texture at both past and future frames. Table 1 contains four possible cases of holes and reference frame to be used for each case. The Case 1 uses both past and future frames, but those textures are synthesized. Hence, we just blend by averaging. In case of Case 2 and Case 3, we select only one reference frame. When we refer to the reference frame, we check the depth differences between the current frame and the reference frame. If the difference is higher than a certain threshold, we skip the temporal refinement. In the experiments, we set the threshold value to 5.0

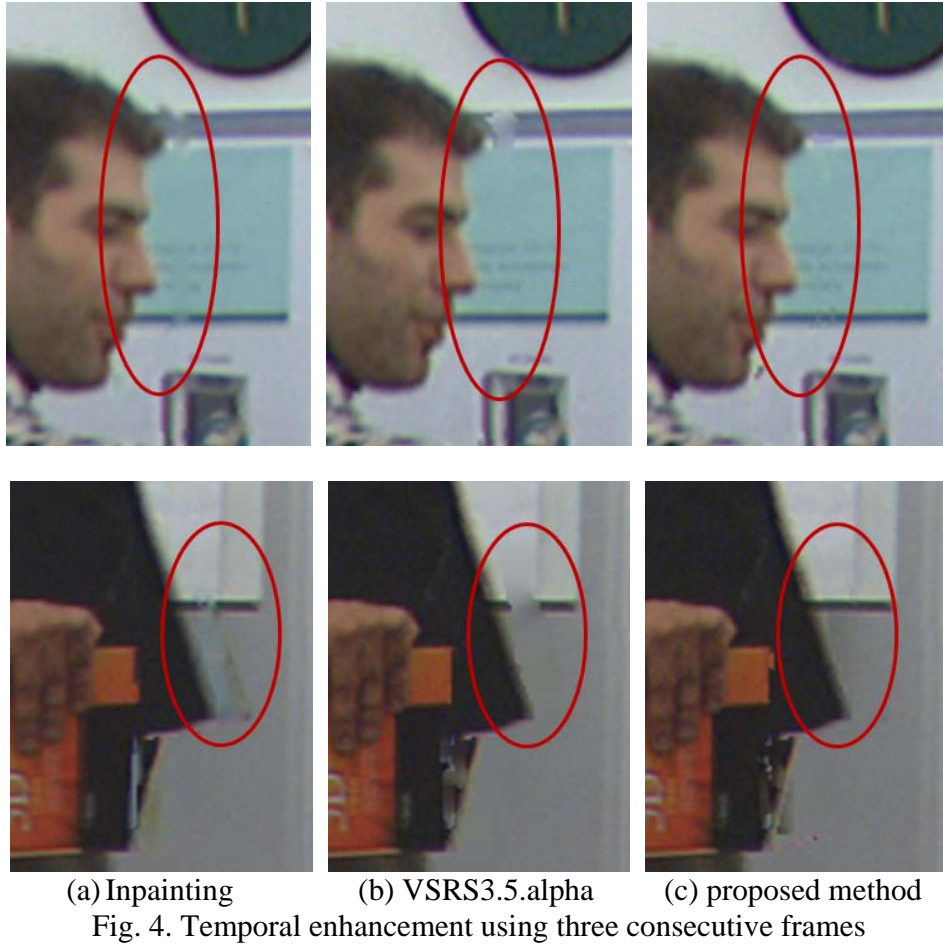
Table 1. Selection of reference frame for the current frame

Case	Past frame ( $F_{N-1}$ )	Future frame ( $F_{N+1}$ )	Reference Frame to be used
1	Hole	Hole	Both: $F_{N-1}$ , $F_{N+1}$
2	Hole	Non-hole	$F_{N+1}$
3	Non-hole	Hole	$F_{N-1}$
4	Non-hole	Non-hole	(no reference)

## 4. Experimental results

### 4.1. Experimental results on gradient based hole filling method

The proposed hole filling method using a gradient map is efficient for an image that has lots of directional textures. Figure 4 shows the results of hole filling on ‘Book\_arrival’ sequence. The first column shows the results of the inpainting method implemented in VSRS3.5. The second column shows the results of the hole filling method that was proposed in the last meeting in VSRS3.5.alpha. The last columns show the results of the updated hole filling method using a gradient map. Obviously, the proposed method filled the holes more efficiently.



#### 4.2. Experimental results on temporal enhancement

For evaluation of the proposed temporal enhancement method, we conducted view extrapolation according to the description of EE document [2]. All viewpoints were tested in full length of frames for 3D viewing. Figure 5 and 6 show the visual comparisons. Three consecutive frames are shown for each method. As we shown in the results, visual artifacts around the hole regions are significantly reduced.





(b) Hole filling method (with temporal enhancement in VSRS3.5.alpha)



(c) Proposed method

Fig. 5. Results of temporal enhancement on 'Mobile' sequence: #167, #168, #169



(a) Inpainting method (without temporal enhancement)



(b) Hole filling method (with temporal enhancement in VSRS3.5.alpha)



(c) Proposed method

Fig. 6. Results of temporal enhancement on 'Poznan\_street' sequence: #7, #8, #9

## 5. Conclusion

In this document, we proposed an efficient hole filling method for view extrapolation. By taking into account the gradient map during the hole filling method, we keep the directional textures using the proposed trilateral filter. After conducting hole filling process for every frame, we added the temporal enhancement part. It uses three consecutive frames to reduce flickering effects. By the experiments, we generated improved virtual views for extra views. To show the experimental results, we prepared three test sequences for 3D viewing.

## 6. Acknowledgements

This research was supported by a grant "Establishing the Foundation of International Standardization for Mobile Devices and Multimedia Convergence Technology" from Improvement of Standards Technology Program funded by Korean Agency for Technology and Standards, Ministry of Knowledge Economy of Korea.

## 7. References

- [1] ISO/IEC JTC1/SC29/WG11, "Implementation of Hole Filling Methods for VSRS 3.5.alpha," in *MPEG input document M20005*, March 2011.
- [2] ISO/IEC JTC1/SC29/WG11, "Description of Exploration Experiments in 3D Video Coding," in *MPEG output document N12037*, March 2011.