Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11
3rd Meeting: Geneva, CH, 17–23 Jan. 2013

Document: JCT3V-C0070

Title:	**3DV-ATM: View synthesis quality improvement for in-loop view-synthesis based inter-view prediction**
Status:	Input Document to JCT-3V
Purpose:	Proposal

Author(s) or Contact(s):	Chun-Fu Chen, Gwo Giun (Chris) Lee, I-Shuo Chen, Ciao-Siang Siao, Jhen-Yue Hu, Bo-Syun Li No. 1, University Road, Tainan, Taiwan	Tel: Email:	+886-6-2360663 n28991146@mail.ncku.edu.tw
Source:	National Cheng Kung University (NCKU)		

Abstract

This proposal is the work merging the JCT3V-B0129 [1] which utilizing the adaptive directional depth filtering and the structure-based hole filling. The added functionality of this proposal is motion-compensated hole filling which utilized the motion vectors generated in the coding process to select the suitable temporal reference frame and furthermore, the suitable temporal reference pixels are selected by motion information and parallax. The result is summarized as follows. Since the proposed technique would not modify the depth data, the coding result of texture data and synthesized data are concerned. For all videos, there is 0.02 % DB-rate increasing on texture data and 0.21% BD-rate increasing on synthesized data and 34% decoding time increasing in average; however, for the full-HD videos, there is 0.25% BD-rate decreasing on texture data, 0.03% BD-rate increasing on synthesized data and 28% decoding time increasing in average. Although the coding gain is not improved overall, the subjective viewing result is better than the anchor for some specific scenes.

1 Introduction and Problem Identification

During view synthesis, the hole occurs at a discontinuance of the depth map due to the parallax difference. In order to fill the hole, the usual process is to use spatial information to fill the hole. Because the geometric distortion of the foreground is more visible to human eyes, and the geometric distortion of the background is not easily visible. Therefore, the hole can be filled by utilizing the background information of temporal or spatial reference.

During the hole filling processing, the way to fetch the reference from temporal or spatial is critical since the dis-occlusion region may occur at previous frame or future frame. By analyzing the motion information and spatial structure of hole region, this proposal presented a motion-compensated hole filling algorithm and combined the previous work, JCT3V-B0129 [1], which utilizing the structure-based hole filling to provide better visual quality of synthesis view.

This is an example showing that the information of hole region can be obtained from different time instant according to the motion of the foreground. Fig. 1 (a) is texture data at frame #25 of the "Dancer" of view 5 and Fig. 1 (b) is the synthesized view prior to fill the hole. By observing the sequence, the hole region should be the window appeared at different time instants. Thus, the blue box in Fig. 1 (a) is enlarged at different time instants, including frame #9, frame #24, and frame #25 and these enlarged figures are shown in Fig. 1 (d), Fig. 1 (e), and Fig. 1 (f). At Fig. 1 (d), the window is appeared without any occluded area and Fig. 1 (e) contains some part of window. By analyzing the motion of the pillar, the window should be appeared at previous frame; therefore, the temporal reference can be obtained from previous frame (frame #24 in this case). Hence, with the help of temporal reference, the hole in Fig. 1 (b) can be filled and provide better visual quality as shown in Fig. 1 (c).

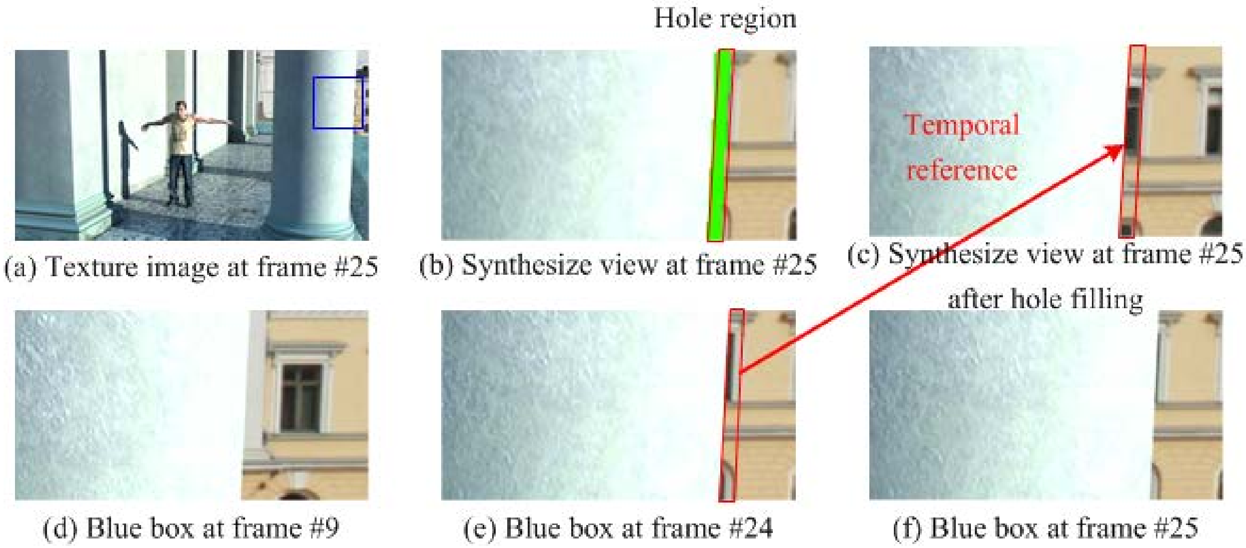


Fig. 1 Synthesis left view with hole at frame #25

2 Proposed algorithm

In this proposal, a motion-compensated hole filing algorithm that uses motion information is presented. The algorithm uses the uncovered region in the temporal reference frame to fill the hole. The benefit of this method is that more information in temporal reference frame fills the hole which does not exist in the current frame.

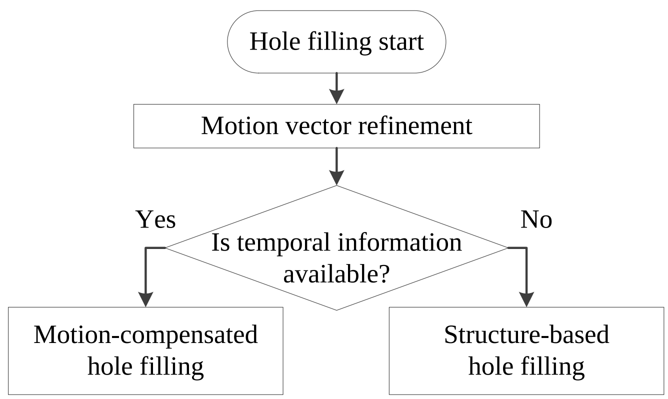
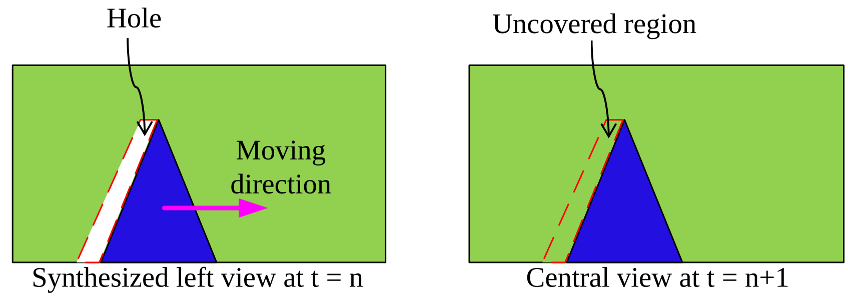


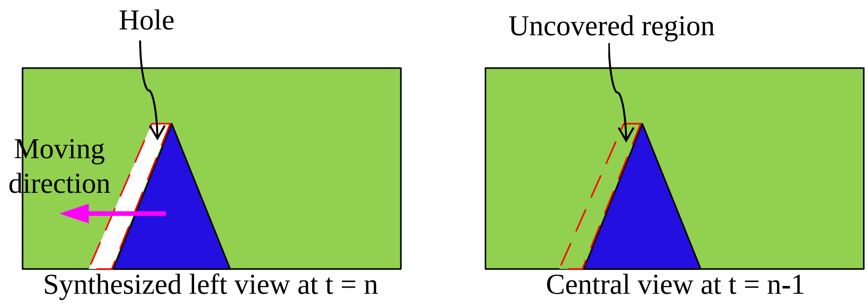
Fig. 2 Flowchart of proposed hole filling algorithm

The flowchart of the proposed hole filling algorithm is shown in Fig. 2. At first, the motion vector (MV) is refined to enhance the accuracy. Subsequently, the refined MV is utilized to determine the temporal reference frame, such as backward or forward reference and the pixels would be referenced. Finally, the hole-filling method is chosen and hole is filled.

Since motion estimation is block-based, when the background and foreground located in the same block, then the MV of one of the background and foreground would be wrong. Therefore, prior to motion-compensated hole filling, the MV should be refined to pixel accuracy. The MV refinement processing of the block located at the boundary of background and fore ground referenced the neighboring left and right blocks and the criterion of MV refinement is sum-of-absolute-differences (SAD). In this manner, the MV of background and foreground would be accurate.



(a) Synthesized left view, foreground moves from left to right



(b) Synthesized left view, foreground moves from right to left

Fig. 3 Temporal reference determination

Having the refined MV, the direction of MV is capable of determining that the available uncovered information is located at future frame or previous frame and range of reference pixels. Due to the difference in the warping direction (left view or right view) and the difference in the moving direction of the foreground, the uncover region in the current frame can be available in a different temporal reference frames. Fig. 3 illustrates an example when warping the image to right. In this case, when the foreground is moving to right, the available background information is located at future frame as shown in Fig. 3 (a); on the other hand, the opposite case is the foreground moving to left, the temporal reference frame should be the previous frame as shown in Fig. 3 (b). It is an inverse process when the image is warped to left.

After the temporal reference frame is determined, motion-compensated hole filling method is used to fill the hole with the uncovered region in the temporal reference frame. Fig. 4 illustrates in detail the process used to fill holes with temporal information. The red region in the synthesized view is the hole region. In this example, the foreground moves from left to right during t=n to t=n+1; the yellow rectangular moves from right to left during t=n to t= n+1, the hole in the synthesized view at t=n is an uncovered region in the texture data at t=n+1. Therefore, the hole can be filled directly from the temporal reference frame by the MV of the background. In Fig. 5, it depicted the pixels on the black dash line in Fig. 4. The hole size of the synthesized view is 4 pixel and there are 3 pixels can be fetched from the temporal reference view by the MVs. Because the movement distance might be smaller than the parallax, the uncovered region in the temporal reference frame might be smaller than the hole; therefore, the entire hole cannot be filled with temporal information. The remaining holes are filled with the previous proposal, JCT3V-B0129, structure-based hole-filling method.

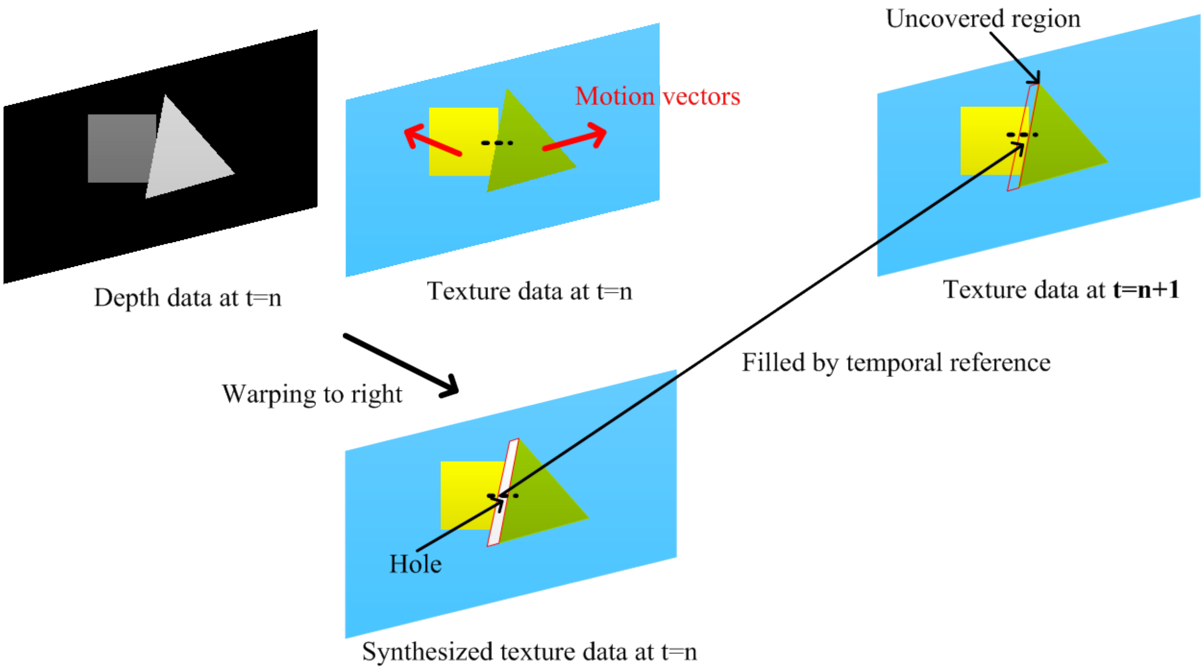


Fig. 4 Hole filling method determination

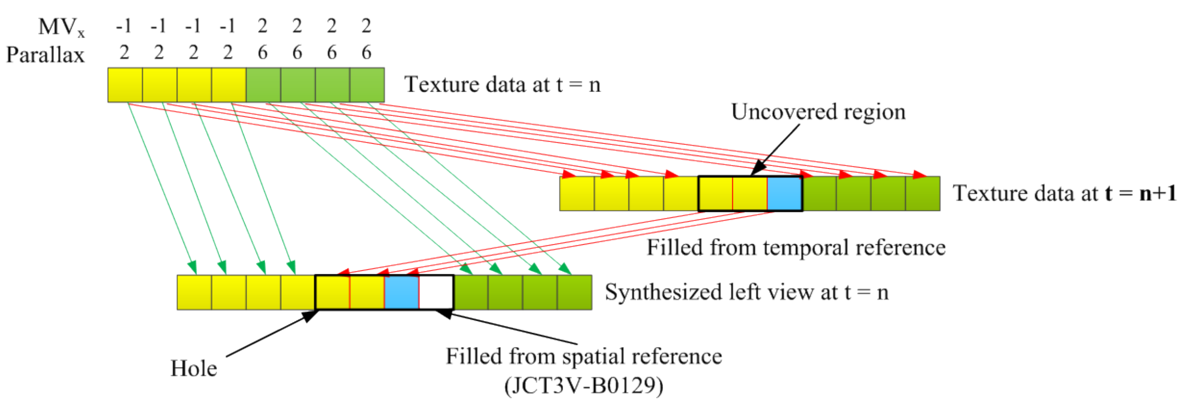


Fig. 5 The black dash line in Fig. 4

3 Experimental results

The presented techniques are built based on the reference software 3DV-ATMv6.0r1. We turn off the backward projection proposed by Nokia and USTC (NOKIA_USTC_B_VSP_M26069 = 0). The experiments are tested according to Common Test Condition. Since the proposed techniques are applied to generate the view for in-loop VSP and the depth data would not be modified, and hence the results listed in Table 1 would not include the coding result of depth data. The overall experimental results can be found at the attached spreadsheet. Moreover, the improvement on the subjective viewing results is illustrated at the appendix section to justify the proposed algorithm combining JCT3V-B0129 can provide better visual quality.

Table 1 Experimental results based on Common Test Condition

	Texture Coding		Total (Coded PSNR)		Total (Synthesed PSNR)		Complexity estimate	Complexity estimate
	dBR(%)	dPSNR(dB)	dBR(%)	dPSNR(dB)	dBR(%)	dPSNR(dB)	Encoding time(%)	Decoding time(%)
S01	-0.19	0.01	-0.17	0.01	-0.04	0.00	102.57%	133.66%
S02	-0.20	0.01	-0.19	0.01	0.03	0.00	103.50%	135.71%
S03	-0.49	0.02	-0.46	0.02	-0.08	0.00	103.16%	129.15%
S04	-0.06	0.00	-0.06	0.00	0.16	-0.01	103.07%	120.43%
S05	0.52	-0.02	0.40	-0.02	0.57	-0.02	103.17%	135.57%
S06	0.48	-0.02	0.38	-0.02	0.70	-0.03	102.39%	132.41%
S08	0.06	0.00	0.03	0.00	0.15	-0.01	102.73%	153.56%
Ave.	0.02	0.00	-0.01	0.00	0.21	-0.01	102.94%	134.36%
Ave. (fullHD)	-0.25	0.01	-0.24	0.01	0.03	0.00	103.24%	128.43%

4 Conclusion

A new technique, motion-compensated hole filling, is integrated to the previous proposal, JCT3V-B0129, utilizing adaptive directional depth filtering and structure based hole filling is presented in this proposal. The proposed technique consider that the enhancement on the synthesized view for in-loop VSP would have coding gain and better visual quality. The results contain both Common Test Condition and subjective testing. According to the common test condition, there is 0.02 % DB-rate increasing on texture data, 0.21% BD-rate increasing on synthesized data and 34% decoding time increasing in average; however, for the full-HD video, there is 0.25% BD-rate decreasing on texture data, 0.03% BD-rate increasing on synthesized data and 28% decoding time increasing in average. Moreover, for the subjective viewing testing, proposed algorithm is confirmed with edge information maintenance and the correctness of hole by the adaptive directional depth filtering, structure based hole filling, and motion-compensated hole filling algorithm. Although the coding gain is not improved overall, the subjective viewing result is better than the anchor for some specific scenes.

5 Acknowledge

Thanks for the cross-check by NCTU (JCT3V-C0072) and ITRI (JCT3V-C0125)

6 Reference

[1] C.-F. Chen, G. G. (Chris) Lee, I-S. Chen, C. Cui, Y. Huo, J.-Y. Hu, “3DV-ATM: Adaptive directional depth map smoothing and structure based hole filling for in-loop view-synthesis based inter-view prediction,” ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V-B0129, Shanghai, CN, 13-19 October 2012

[2] JCT3V-B1100, “Common Test Conditions of 3DV Core Experiments,” Shanghai, CN, 13-19 October 2012

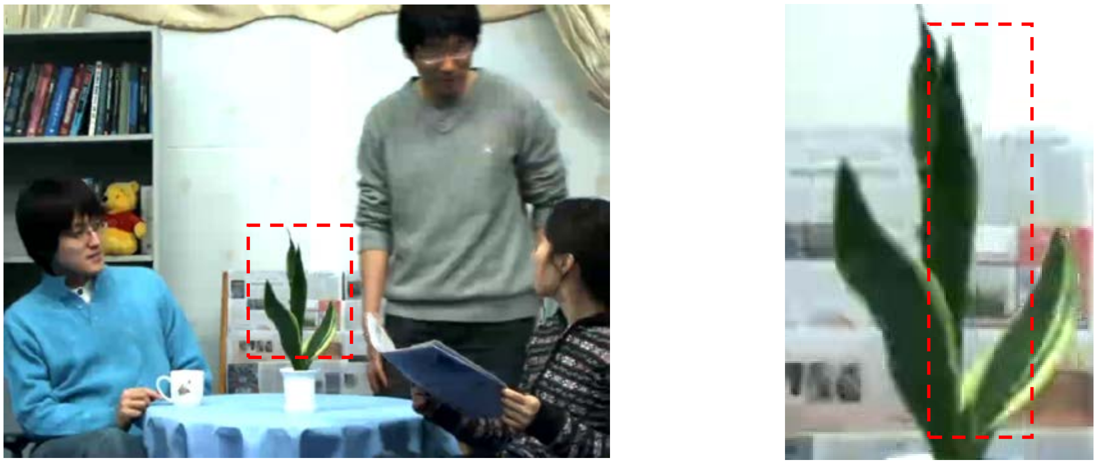
[3] JCT3V-B1003, “3D-AVC Test Model 4,” Shanghai, CN, 13-19 October 2012.

7 Appendix

(a) Anchor result

(b) Proposed result

Fig. 6 Newspaper, QP41, synthesized view 04, frame 101

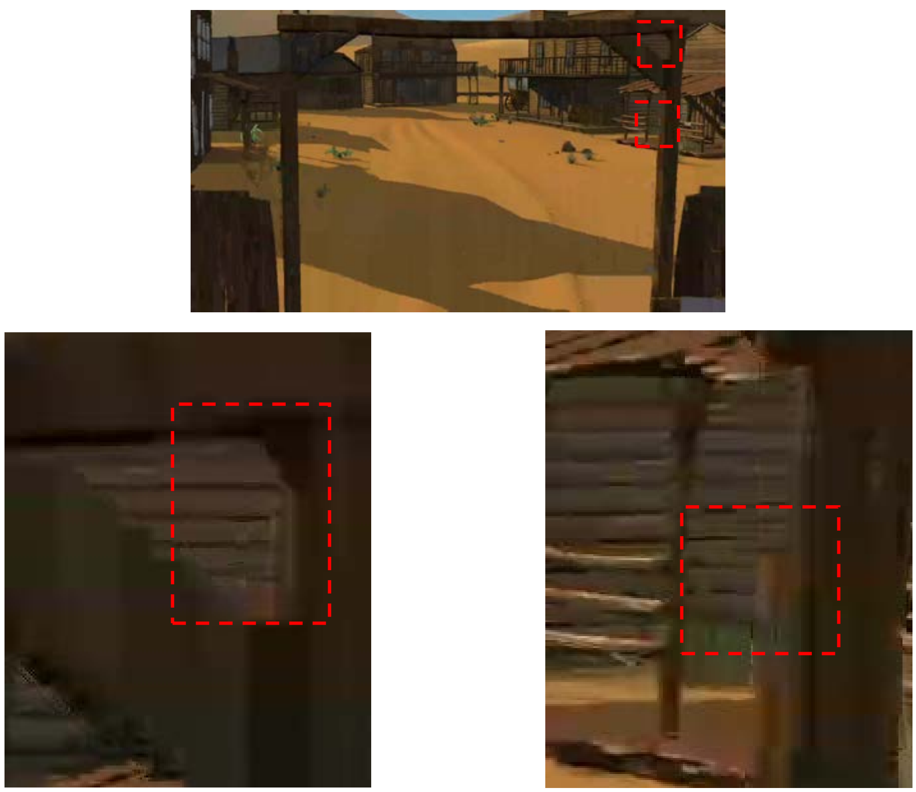


(a) Anchor result

(b) Proposed result

Fig. 7 Newspaper, QP41, synthesized view 06, frame 274

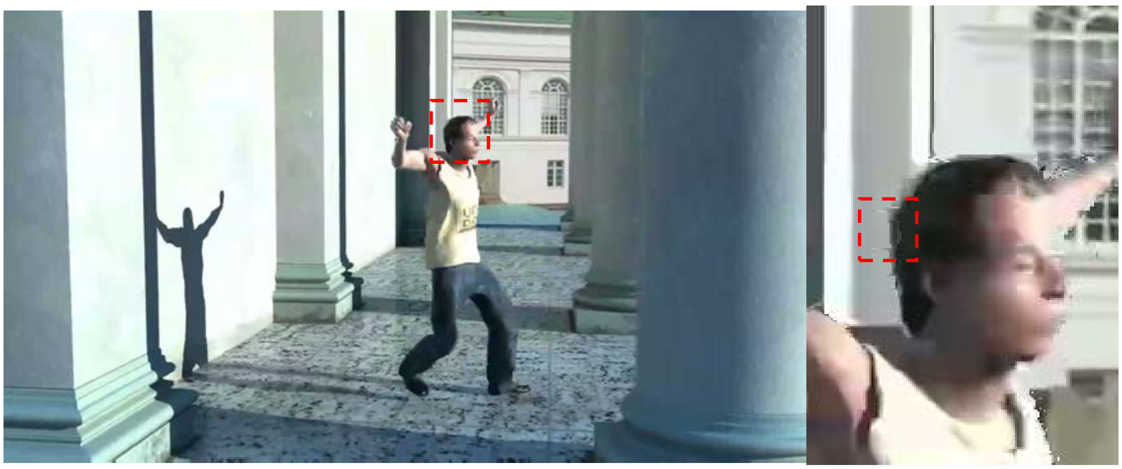
(a) Anchor result



(b) Proposed result

Fig. 8 GT_Fly, QP41, synthesized view 04, frame 157

(a) Anchor result



(b) Proposed result

Fig. 9 Dancer, QP41, synthesized view 04, frame 139

8 Patent rights declaration(s)

(NOTE – Activities in the JCT and contributions to the JCT are subject to the common patent policy for ITU-T/ITU-R/ISO/IEC. A statement of that policy can be found at http://www.itu.int/ITU-T/dbase/patent/patent-policy.html, with further information available at http://www.itu.int/ITU-T/ipr/index.html and in the ISO/IEC Directives. The form to be used for the formal reporting of patent rights to ITU-T/ITU-R/ISO/IEC can be found at http://www.itu.int/ITU-T/ipr/index.html. Contributions to the JCT proposing normative technical content shall contain a non-binding informal notice of whether the submitter may have patent rights that would be necessary for implementation of the resulting standard. The provided informal notice shall indicate the category of anticipated licensing terms according to the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form. This obligation to provide an informal notice is supplemental to, and does not replace, any existing obligations of parties with technology included in a final or draft standard to submit formal IPR declarations to ITU-T/ITU-R/ISO/IEC. Two examples of such an informal IPR notification statement for a contribution are provided below.)

NCKU may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).