Telecommunications and Information Exchange Between Systems

ISO/IEC JTC 1/SC 6

Document Number:	N14176
Date:	2010-01-07
Replaces:	
Document Type:	Liaison organization contribution
Document Title:	Liaison Statement from JTC 1/SC 29 to ISO/IEC JTC 1/SC 6 on the applications and requirements on 3D video coding
Document Source:	JTC 1/SC 29 Secretariat
Project Number:	
Document Status:	For consideration at the SC 6 HoDC and WG 7 meeting in Barcelona.
Action ID:	FYI
Due Date:	
No. of Pages:	6

ISO/IEC JTC1/SC6 Secretariat Ms. Jooran Lee, KSA (on behalf of KATS)

Korea Technology Center #701-7 Yeoksam-dong, Gangnam-gu, Seoul, 135-513, Republic of Korea ;

Telephone: +82 2 6009 4808 ; Facsimile: +82 2 6009 4819 ; Email : jooran@kisi.or.kr

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

ISO/IEC JTC1/SC29/WG11
MPEG2009/N11061
Xian, CN,
October 2009

Source: **Video and Requirements**
Title: **Applications and Requirements on 3D Video Coding**
Status: **Approved**

---

1 Introduction

This document describes the applications and requirements of a new 3D Video (3DV) initiative underway in MPEG. Our vision is a new 3DV format that goes beyond the capabilities of existing standards to enable both advanced stereoscopic display processing and improved support for auto-stereoscopic N-view displays [1].

3DV is the second phase of FTV (free-viewpoint TV), which is a new framework that started back in 2001. FTV is a framework that allows viewing of a 3D world by freely changing the viewpoint. Multi-view Video Coding (MVC) was the first phase of FTV, which enabled the efficient coding of multiple camera views.

2 Application Scenarios

The 3DV format targets two specific application scenarios:

- Enabling stereo devices to cope with varying display types and sizes, and different viewing preferences. This includes the ability to vary the baseline distance for stereo video to adjust the depth perception, which could help to avoid fatigue and other viewing discomforts.
- Support for high-quality auto-stereoscopic displays, such that the new format enables the generation of many high-quality views from a limited amount of input data, e.g. stereo and depth.

The targets of the 3DV format are illustrated in Fig 1.

Due to limitations in the production environment, the 3DV data format is assumed to be based on limited camera inputs; stereo content is most likely, but more views might also be available.

The rate required for transmitting the 3DV format should be fixed to the distribution constraints. For support of a large range of high-quality displays, there should not be an increase in the rate simply because the display requires a higher number of views to cover a larger viewing angle. In this way, the transmission rate and the number of output views are decoupled. For support of displays providing a significantly limited number of output views, it should be possible to decrease the bit rate by excluding information that is not required for rendering from the transmission.

In order to support a wide range of auto-stereoscopic displays, it should be possible for a large number of views to be generated from the reconstructed data format. In addition, advanced stereoscopic processing that requires view generation at the display would also be supported by this format.

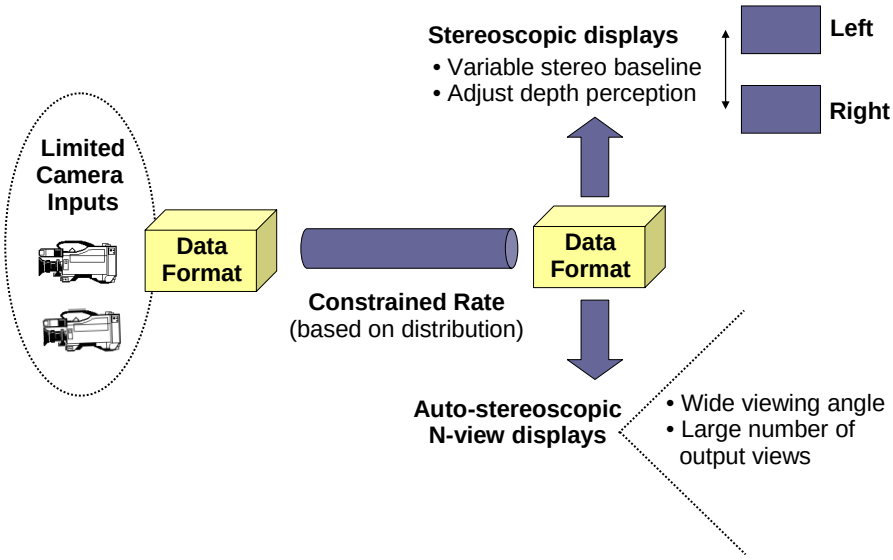

Fig. 1. Target of 3D Video format illustrating limited camera inputs and constrained rate transmission according to a distribution environment. The 3DV data format aims to be capable of rendering a large number of output views for auto-stereoscopic N-view displays and support advanced stereoscopic processing.

3 Requirements of 3DV

3.1 Requirements for Data Format

3.1.1 Video data

The uncompressed data format shall support stereo video, including samples from left and right views as input and output. The source video data should be rectified to avoid misalignment of camera geometry and colors. Other input and output configurations beyond stereo should also be supported.

3.1.2 Supplementary Data

Supplementary data shall be supported in the data format to facilitate high-quality intermediate view generation. Examples of supplementary data include depth maps, segmentation information, transparency or specular reflection, occlusion data, etc. Supplementary data can be obtained by any means from a predetermined set of input videos.

3.1.3 Metadata

Metadata shall be supported in the data format. Examples of metadata include extrinsic and intrinsic camera parameters, scene data, such as near and far plane, and others.

3.1.4 Low complexity for editing

The data format should allow for editing with low complexity.

3.1.5 Applicability

The data format shall be applicable for both natural and synthetic scenes.

3.2 Requirements for Compression of 3DV Data Format

3.2.1 Compression efficiency

Video and supplementary data should not exceed twice the bit rate of state-of-the-art compressed single video. It should also be more efficient than state-of-the-art coding of multiple views with comparable level of rendering capability and quality.

3.2.2 Synthesis accuracy

The impact of compressing the data format should introduce minimal visual distortion on the visual quality of synthesized views. The compression shall support mechanisms to control overall bitrate with proportional changes in synthesis accuracy.

3.2.3 Backward compatibility

The compressed data format shall include a mode which is backwards compatible with existing MPEG coding standards that support stereo and mono video. In particular, it should be backwards compatible with MVC.

3.2.4 Stereo/Mono compatibility

The compressed data format shall enable the simple extraction of bitstreams for stereo and mono output, and support high-fidelity reconstruction of samples from the left and right views of the stereo video. Solutions that do not enable simple extraction of bitstreams for stereo and mono output must demonstrate a significantly higher performance in terms of both: coding efficiency and rendering capability.

3.2.5 Forward compatibility

The compressed data format should be designed to enable the use in combination with future mono and stereo coding standards.

3.2.6 View scalability

The data format shall have configurations that allow for extraction of bitstreams that represent sub-sets of the views available in the compressed data format. This includes extraction of backward compatible bitstreams according to section 3.2.3 as well as more general extraction of bitstreams that allows e.g. for rendering of a limited viewing angle.

3.3 Requirements for Rendering of 3DV Data Format

3.3.1 Rendering Capability

The data format should support improved rendering capability and quality compared to existing state-of-the-art representations. The rendering range should be adjustable.

3.3.2 Low complexity

The data format shall allow real-time synthesis of views.

3.3.3 Display types

The data format shall be display-independent. Various types and sizes of displays, e.g. stereo and auto-stereoscopic N-view displays of different sizes with different number of views shall be supported. The data format shall be adaptable to the associated display interfaces.

3.3.4 Variable baseline

The data format shall support rendering of stereo views with a variable baseline.

3.3.5 Depth range

The data format should support an appropriate depth range, which is [-W/2;+W] or less, where W is the width of the display. Very large ranges are not needed, and even undesirable in order to maintain a good visual comfort and prevent viewers from feeling uncomfortable. It is important to avoid fatigue (e.g., due to framing effect, eye convergence and accommodation mismatch).

3.3.6 Adjustable depth location

The data format should support display-specific shift of depth location, i.e., whether the perceived 3D scene (or parts of it) are behind or in front of the screen.

3.3.7 Scalability

The data format should support scalable rendering to support real-time processing.

4 References

[1] Video and Requirements Group, “Vision on 3D Video,” ISO/IEC JTC1/SC29/WG11 N10357, Lausanne, CH, February 2008. Available online: http://www.chiariglione.org/mpeg/visions/3dv/index.htm