

I-068

## 仮想視点合成に適した多視点映像及び奥行き情報の圧縮符号化

## Multiview Video and Depth Compression for Virtual View Synthesis

樋口雄太<sup>†</sup> Mehrdad Panahpour teherani<sup>†</sup> 圓道 知博<sup>‡</sup>藤井 俊彰<sup>†</sup> 谷本 正幸<sup>†</sup>Yuta Higuchi<sup>†</sup> Mehrdad Panahpour tehrani<sup>†</sup> Tomohiro Yendo<sup>‡</sup>Toshiaki Fujii<sup>†</sup> Masayuki Tanimoto<sup>†</sup>

## 1 はじめに

本稿では、複数の視点映像及び奥行き情報を入出力とする N-View N-Depth 符号化方式について扱う。また復号側では、FTV[1] 等の多視点映像システムにより、復号情報を用いて仮想視点合成を行うことを考える。多視点映像符号化方式として、MPEG/JVT より MVC(Multiview Video Coding) と呼ばれる方式が標準化されている [2]。MVC は視点映像のみを入出力とする方式であり、奥行き情報を符号化することは想定されていない。その為、復号側に対し低負荷で多視点映像システムを提供する為には、仮想視点合成に適した入力となる様、奥行き情報を別途伝送する必要がある。加えて、一般に伝送できる総ビットレートには制限があり、多視点映像及び奥行き情報各々のビットレートを、復号後に行われる仮想視点合成の精度が最大となる様に調整しなくてはならない。そこで本報告では、復号後の仮想視点映像生成に適した符号化方式を検討する初期段階として、視点映像及び奥行き情報の各伝送レートの違いが、仮想視点合成の精度にどう影響するか分析した。視点映像及び奥行き情報の符号化には MVC を用い、各々独立に符号化を行った。

## 2 Multiview Video Coding

多視点映像や 3D ビデオ等のシーケンスは複数の視点映像から構成されており、一般的な動画シーケンスと比べ膨大なデータ量を持つ。その為、FTV や 3DTV 等のアプリケーションをユーザーに提供するには、より効率の良い圧縮方式が必要不可欠である。MVC( Multiview Video Coding ) はマルチカメラで撮影された複数の映像を入出力とし、フレーム内及びフレーム間だけでなく、視点方向への相関を利用して圧縮する符号化方式である [2]。MVC では従来の圧縮符号化方式である H.264/AVC において、視点方向の予測に対し視差補償予測を採用した形で表される。ここで視差補償予測とは、既に符号化済みかつ別視点の映像と符号化対象としている映像とを比較し、より相関度の高い映像と符号化対象映像との差分を符

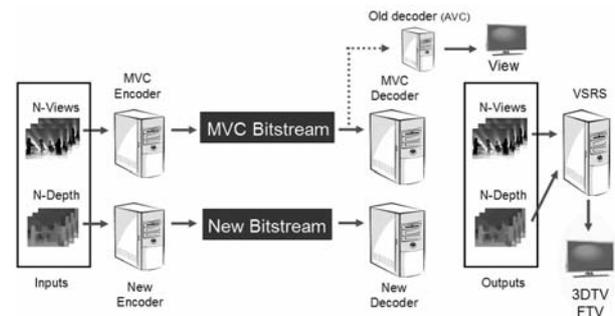


Fig 1: Diagram of N-View N-Depth coding system which we consider

号化するものである。これらの技術により、H.264/AVC をそれぞれの視点映像に対し独立に符号化する simulcast と比べ、約 20%~30% の圧縮率向上を実現している。

## 3 仮想視点合成に適した多視点映像符号化方式

MVC は多視点映像のみを符号化する方式である為、MVC を用いてユーザーに対し多視点映像を用いるアプリケーションを提供するには、復号側で奥行き情報を別途取得する必要がある。本報告では奥行き情報を圧縮伝送する事を考え、復号側で高精度な仮想視点合成を行える様、適した多視点映像及び奥行き情報の圧縮符号化方式について検討する。

想定する符号化システム図を Fig. 1 に示す。復号側で 3DTV や FTV 等の多視点映像を用いたアプリケーションを用いることを想定した、N-view N-Depth 符号化方式を考える。入力は、カメラキャリブレーションやレクティフィケーション済であり、かつカラーバランスの影響等が取り除かれた多視点映像及び奥行き情報とする。多視点映像及び奥行き情報は同数とし、入力と同数の多視点映像及び奥行き情報を出力する。復号側では、復号情報より仮想視点合成を行う様な、FTV 等の多視点映像を用いるアプリケーションを提供する事を考える。多視点映像の符号化には、既に標準化されている規格に対し下位互換性を持つよう、MVC を用いて符号化する。奥行き情報の符号化には、本報告では新たな符号化方式を

<sup>†</sup> 名古屋大学 工学研究科

<sup>‡</sup> 長岡技術科学大学 工学部 電気系

用いることを検討する。また符号化側では、多視点映像及び奥行き情報の総伝送量が定められた環境で、各々のビットレートを仮想視点映像の精度が最大となる様調整し、適したシーケンスの組み合わせを選択する必要がある。

#### 4 実験

前章で述べた符号化方式において、復号後に行う仮想視点合成に適した符号化方式を検討する初期段階として、視点映像及び奥行き情報各々の伝送レートの違いが、仮想視点合成の精度にどう影響するか分析した。視点映像及び奥行き情報の符号化にはMVCを用い、それぞれ独立に符号化を行った。得られた複数ビットレートの復号情報を基に仮想視点映像を生成し、同位置で撮影した映像と比較を行った。用いるシーケンスには、名古屋大学より提供されている”kendo” (1024×768) から10cm間隔で配置された視点1, 3, 5の0~31フレームを使用し、3視点の映像及び奥行き情報を伝送する事を想定する。評価に使用するMVCプログラムとして、JVTより提供されているJMVC8.3を用い、量子化パラメータを30~51の間で変動させ、異なるビットレートにて復号した。またGoPsizeは15とし、視点1, 5, 3の順で符号化した。復元後の仮想視点合成には、MPEGから提供されているVSRS3.5を用いた。仮想視点合成では、復号された多視点映像及び奥行き情報を基に、視点1と3から中間に位置する視点2を、視点3と5から中間に位置する視点4を生成した。視点2において多視点映像及び奥行き情報の量子化パラメータを変動させた場合に得られる、生成映像と元映像とのビットレート対PSNR特性のグラフをFig. 2に示す。ここで、グラフにおける視点映像及び奥行き情報のビットレートは、ある量子化パラメータで符号化した際における、3視点それぞれのビットレートの平均を指す。伝送できる総ビットレート量が制限される場合を考える際は、任意の制限値 $const.$ とし、 $Bitrate(texture) + Bitrate(depth) = const.$ となる平面で切った断面において、一番精度の高い点を選択すればよい。

Fig. 2より、仮想視点合成の精度が、殆ど視点映像の符号化精度に依存している事が見て取れる。奥行き情報のビットレート変動に着目すると、奥行き情報の平均ビットレートが0.2 Mbpsを下回る付近より合成精度が下がっている。これらの事より、MVCを各々独立に用いた方式では、視点映像に比べ奥行き情報の伝送レートが低くても、仮想視点映像生成の精度に影響を与えにくいと言える。視点映像及び奥行き情報の間には強い相関関係が存在する為、視点映像の情報を付加情報とし、奥行き情報が再現出来れば、より符号化効率の向上を見込める可能性がある。よって、復号後の仮想視点映像合成の

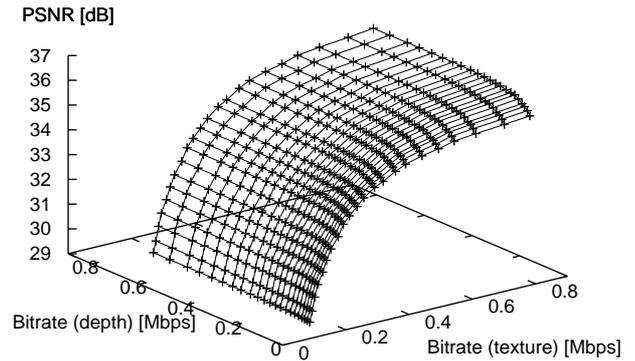


Fig 2: Bitrate (texture/depth) vs PSNR in view 2

精度を意識した、多視点映像及び奥行き情報の圧縮特性の向上を図る為には、多視点映像及び奥行き情報それぞれ独立で着目せず、双方の強い相関性を活用し、両方の符号化器で一部情報を共有化する様な形が望ましいと考えられる。

#### 5 まとめ

本報告では、複数の視点映像及び奥行き情報を入力とする $N$ -view  $N$ -Depth符号化方式において、復号後に行う仮想視点合成に適した符号化方式を検討する初期段階として、各々の伝送レートの違いが、仮想視点合成の精度にどう影響するか分析した。結果として、MVCを独立に用いた符号化方式では、視点映像に比べ奥行き情報は低レートでも仮想視点合成の精度に影響を与えにくい事が判明した。また、視点映像と奥行き情報の間には強い相関性が存在する為、双方の持つ情報を活用するよう符号化器を設計すべきであるという結論が得られた。今後の方針として、双方の情報を一部共有化する事で、それらを付加情報としてそれぞれの映像を再現する様な符号化器の設計及び提案が挙げられる。

#### 参考文献

- [1] T.Fujii and M.tanimoto, "Free-viewpoint Television based on the Ray-Space representation", SPIE ITCOM'02, pp. 175-189, Aug.2002.
- [2] Y.Chen, M.M.Hannuksela, L.Zhu, A.Hallapuro, M.Gabbouj and H.Li, "Coding Techniques in Multiview Video Coding and Joint Multiview Video Model", PCS2009, May.2009.