

## 自由視点映像通信のための多視点映像符号化の一検討

## A Study on Multiple View Video Coding for Free Viewpoint Video Communication

木全 英明 北原 正樹 志水 信哉 上倉 一人 八島 由幸

Hideaki Kimata Masaki Kitahara Shinya Shimizu Kazuto Kamikura Yoshiyuki Yashima

## 1. まえがき

自由視点映像はユーザが自由に視点位置や視線方向を変えて映像を見ることが出来る新しい映像メディアである[1]. 自由視点映像は複数のカメラ入力映像(多視点映像)から作成される. 従来は主に, IBR (Image Based Rendering)に関連して, 多視点映像から撮影されていない視点位置や視線方向の画像情報を生成する技術を中心に研究が行われていた. 我々は自由視点映像を通信するために適した表現フォーマットや多視点映像符号化方式や通信方式を研究している[2][3]. 本稿では多視点映像符号化について, 復号に要する遅延を抑えつつ符号化効率を上げる手法を提案する.

## 2. 自由視点映像通信の概要

我々は自由視点映像通信の映像データを多視点映像(multi-view video)と補間により構成する手法を提案している[2]. 補間は Ray Space[4] (光線空間)で行うことを想定している. 多視点映像は複数カメラで撮影された view (real view)であり, 撮影されていない視点位置や視線方向のview (virtual view)を補間により生成する. ユーザが所望の視点位置の view を見る場合には, 多視点映像データのうち必要な複数の real view を選択して, それら virtual view を生成する.

必要な real view の画像情報のみ得られれば良いので, 多視点映像デコーダはその view のみを復号すれば良い. またデコーダは予め設定した遅延時間内で所望の画像情報を出力する必要があることから, 復号する real view を選択する必要がある. 我々は復号する real view の数を固定して, かつ real view のランダムアクセスを実現するために GoGOP 構造を提案している[2]. GoGOP 構造では, 必ず復号する Base GOP と, 適宜復号する Inter GOP を設定する. Inter GOP に含まれるフレームは Base GOP または Inter GOP の復号画像を参照する. 本稿では, Inter GOP の符号化効率を向上するための手法を提案する.

## 3. 提案 Inter GOP 符号化方式

最も低遅延な多視点符号化は, 全ての view (GOP)を独立に符号化する, すなわち Base GOP として符号化する方式 (AB)である. 提案手法では, この AB 方式に比べて数フレームの復号遅延を許容することで, 符号化効率を向上する. まず 1 フレームだけ遅延を許す Single Reference Inter GOP 方式 (SR)を提案する. SR では, Inter GOP は Base GOP または同一 Inter GOP に属するフレームのみから参照画像を選択する(図 1). 図 2 に SR 方式でのデコーダ構成を示す. 特に最初のフレームは必ず Base GOP のフレームを参照する. Base GOP を 1 つ以上参照することも出来, 参照する GOP 番号を GOP ヘッダに含めて符号化する. 複数カメラのうち, どれを Base GOP とするのかは, カメラ配置に依存す

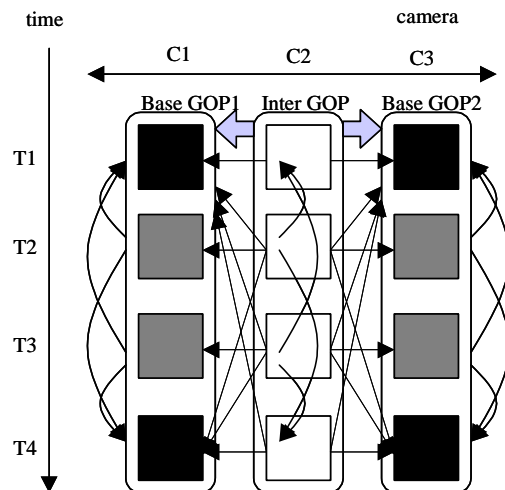


Figure 1: Example of GoGOP

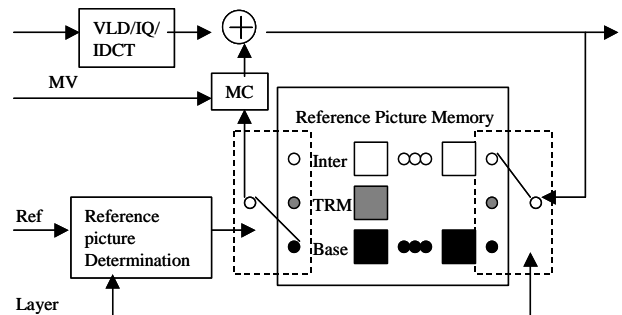


Figure 2: Decoder configuration of SR method

る. SR 方式では Base GOP が復号されていれば, そこから 1 フレーム後に所望の Inter GOP 中のフレームを復号できる. すなわち view のランダムアクセスが頻繁に発生する場合に適している.

時間方向について MPEG-2 B-picture のように参照画像メモリに蓄積しないフレーム (NS, Non Stored)を設定する. 参照画像メモリに B-picture を蓄積しないために符号化効率が低下するものの, 時間方向のランダムアクセスを容易にしている. Base GOP での NS フレームを復号した後, 提案方式では, この復号画像を Inter GOP 用の参照画像として, TRM (temporarily reference picture memory)に蓄積する. これにより, Inter GOP で同時刻のフレームを符号化する際に, NS フレームを参照画像に使用することが出来る. TRM に蓄積した NS フレームの復号画像は対応するフレームを符号化した後で, TRM から削除する.

更に提案方式では Base GOP と Inter GOP 間で算出される視差ベクトル(動きベクトル)を蓄積するメモリを持つ. 参照 GOP 数の動きベクトルメモリを持っており, ブロック単位に, 参照するフレームが Base GOP である時に, この

Table 1: Simulation conditions

# of Reference pictures	Total 5 for forward prediction or backward prediction
# of NS pictures	2
Coding method	H.264
QP	24, 28, 32, 36
Motion and disparity estimation	Accuracy: 1/4 pel Range: -32 to +32 pixels
GOP time interval	1 sec
Mode decision	Rate distortion optimization with Lagrange manipulator

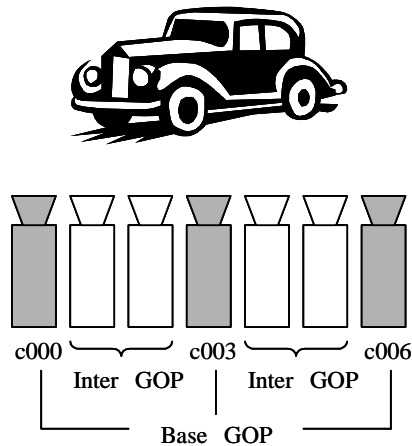


Figure 3: Camera arrangement for Race2

動きベクトルを使用する。また使用した動きベクトルは動きベクトルメモリに蓄積する。この動きベクトルを使用するかどうかをフレーム毎に切り替える。

#### 4. 実験結果と考察

表 1 に実験条件を示す。実験に使用した画像 Race2 は、8 眼のカメラを水平に並べた撮像系(図 3)を使って撮影されたものであり、各カメラの解像度は 320x240 である[5]。両端と中央のカメラ位置を Base GOP で符号化し、その他を Inter GOP で符号化した。図 4 に 8 眼のカメラ映像のうち、Inter GOP で符号化した c001, c002, c004, c005 について、AB 方式と SR 方式で符号化した場合の平均 PSNR を示す。この結果より、SR 方式は 300kbps 以下で約 1dB の符号化利得が得られることが分かる。

また図 5 に、Base GOP を参照するブロックについて、動きベクトルメモリに蓄積した動きベクトルを使用するように選択したフレームの割合を示す。QP が大きいほど選択率が増えることが分かる。これは QP が大きいほど、動きベクトルの符号量を削減するように、動きベクトルメモリの動きベクトルを選択することに起因すると考えられる。

#### 5. まとめ

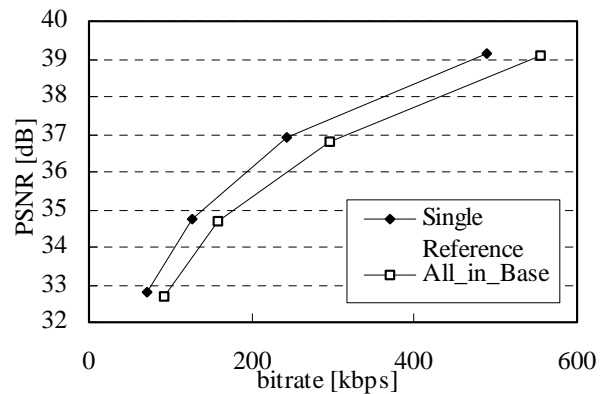


Figure 4: Average PSNR for 8 camera views

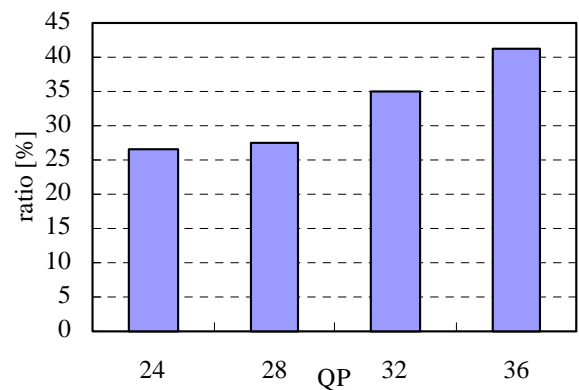


Figure 5: Selection ratio of motion vector memory

自由視点映像通信を想定した多視点映像符号化において、GoGOP を使った参照画像メモリ管理方式と動きベクトル符号化方式を提案した。実験により提案方式は、各カメラ映像を独立に符号化する手法よりも高い符号化効率を実現できることを示した。

今後は更に遅延を許す場合の符号化方法を検討する。

#### 参考文献

- [1] "Applications and Requirements for 3DAV," document N5877 MPEG Trondheim Meeting, 2003.
- [2] H. Kimata, M. Kitahara, K. Kamikura, Y. Yashima, T. Fujii, and M. Tanimoto, "System Design of Free-viewpoint Video Communication," CIT2004, Sep., 2004.
- [3] H. Kimata, "Preliminary study on multiple view coding for the Ray Space representation (3DAV EE2)," document M10054 MPEG Brisbane Meeting, 2003.
- [4] T. Kobayashi, T. Fujii, T. Kimoto, M. Tanimoto, "Interpolation of Ray-Space Data by Adaptive Filtering," IS&T/SPIE Electronic Imaging 2000, 2000.
- [5] R. Kawada, "KDDI multiview video sequences for MPEG 3DAV use," M10533, MPEG Munich document, Mar., 2004.