

背景マスクを用いた奥行き情報のフィルタリングによる自由視点 VoD 映像の高画質化

Improving Free Viewpoint VoD Images by filtering of Depth Information Using Background Mask

石川 彰夫 川田 亮一 小池 淳
Akio Ishikawa Ryoichi Kawada Atsushi Koike

1. まえがき

筆者らはこれまで、1 視点の映像とその奥行き情報から自由視点映像を合成する「自由視点 VoD システム」を検討してきた[1][2]。また、携帯端末などの低ビットレート伝送路を想定し、伝送量を低減するためのデータ圧縮手法も検証してきた。しかし、圧縮率を高くすると、主に奥行き情報の劣化に起因して、合成した自由視点映像の画質が劣化するという問題があった[3]。

そこで、本検討では、伝送速度が小さくても合成された自由視点映像の画質を保つことのできる手法を求める目的とする。

2. 従来手法の課題

自由視点 VoD システムのシステム構成を図 1 に示す。

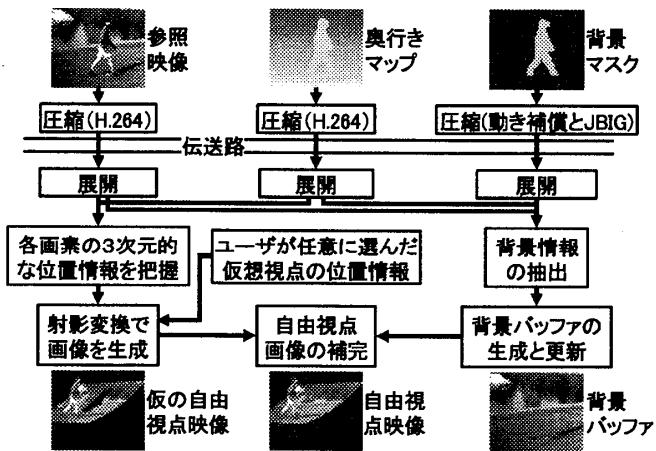


図1 システム構成

本手法は、以下の 3 種類の入力データから、自由視点映像を合成する。

- 参照映像：1 視点から撮影した映像
- 奥行きマップ：参照映像の各フレームの奥行き値
- 背景マスク：背景領域を指定する 2 値映像

本手法の処理手順は、以下の通りである。

1. 背景マスクを参照し、参照映像と奥行きマップの背景領域を抽出し、背景バッファに格納する
2. 奥行きマップに基づき参照映像の各画素の 3 次元的な位置を求め、ユーザが任意に選択した仮想視点位置の情報に基づき射影変換を施す
3. 手前の物体に隠蔽されて背景情報が無い領域を、背景バッファの背景情報を用いて補完する

いずれの入力データもそのままでは量が多く、携帯端末などの比較的狭帯域な通信では実時間の伝送が困難である。そこで、参照映像と奥行きマップを H.264 により圧縮し、背景マスクを一連の動画像とみなして 1 フレームを 1 個の動きベクトルで予測する簡単な動き補償と JBIG により可逆圧縮して伝送する[3]。

しかし、圧縮率を高めると、奥行きマップ内の輪郭部分（物体と背景の境界など）の周辺にモスキート雑音やブロック雑音などの誤差が多く生じるので、仮想視点が正面以外にある場合、針状の誤差が生じ、合成した自由視点映像の画質が著しく低下するという問題があった。

3. 提案手法

前述の問題を解決するため、奥行きマップの誤差が主に奥行き値の不連続に変化する部分の周辺に生じる点に着目し、展開後の奥行きマップのそれらの領域に対して平滑化フィルタを施す手法を提案する。これにより、針状の雑音を低減し、自由視点映像の画質を向上させる。

その際、物体と背景の境界を股いで平滑化フィルタを施すと、両者の奥行き値が連続に変化するようになって両領域が連結してしまい、かえって誤った映像が合成されてしまう可能性がある。そこで、図 2 に示すように、背景マスクに基づき境界の位置を把握して、境界に股がる平滑化フ

				物体と背景の境界	
				目標 画素	
				平滑化ウインドウ内の画素を 用いて目標画素の値を決定する	
				物体と背景の境界を跨ぐ 平滑化は行わない	
				目標 画素	
				境界を跨がない範囲の画素を 用いて目標画素の値を決定する	

ィルタは施さないようにする。

図2 提案手法の平滑化フィルタ

提案手法は、平滑化フィルタを、左右方向と上下方向のそれぞれの方向について施す。これにより、斜めや左右方向の境界に対しても、平滑化処理ができる。

4. シミュレーション実験

提案手法の性能を評価する計算機シミュレーション実験を行う。

H.264により圧縮した奥行きマップを展開した後、平滑化フィルタを施す場合と施さない場合とで、自由視点映像の画質を比較する。非圧縮の参照映像/奥行きマップから合成した映像を基準映像とし、それに対する評価対象映像のPSNRを評価値（以下、単に「PSNR」と呼ぶ）とする。平滑化フィルタとして、平均値フィルタと中央値フィルタの2種類のフィルタを使う。

シミュレーション実験は、2種類のコンテンツに対して行う。各コンテンツの参照映像/奥行きマップ/背景マスクのそれぞれ1フレームずつを図3に示す。



図3 評価対象映像

参照映像と奥行きマップに施すH.264のQP値を変えて、PSNRの変化を調べる。QP値は、6~40の内4種類の値を使用する。実験の結果を、以下に示す。

図4に、同じタップ数3で平滑化フィルタとしてそれぞれ中央値フィルタと平均値フィルタを用いる場合、および平滑化フィルタを用いない場合のPSNRを示す。いずれのタップ数でも、平均値フィルタよりも中央値フィルタの方が高いPSNRを示している。これは、奥行きの異なる複数の物体が重なって見えている画像（例えば、図3のFlamenco2では、右端の2人の人物が重なって見えている）などで平均値フィルタを用いると、境界部において本来存在しない奥行き値が得られてしまうためと考えられる。

図5に、中央値フィルタのタップ数を変えた場合のPSNRを示す。いずれの場合も、タップ数が短い方が高いPSNRを示していた。これは、QPが30以下では、殆どの雑音が1~2画素の針状であるため雑音除去の点ではタップ数の影響が少ない一方で、タップ数が小さい方がより近傍の比較的正しい画素値を求め易いためと考えられる。QPが40前後になると針状でない雑音も増えるため、近傍に正しい画素が少なくなり、タップ数に依らずPSNRが急激に低下している。

最もPSNRが高いタップ数3の中央値フィルタを用いた場合で比較すると、Golf2はQPが15以上、Flamenco2はQPが30以上で、フィルタを用いない場合に対してPSNRが約2~3dB上回った。以上により、低ビットレート伝送における本手法の有効性が示された。

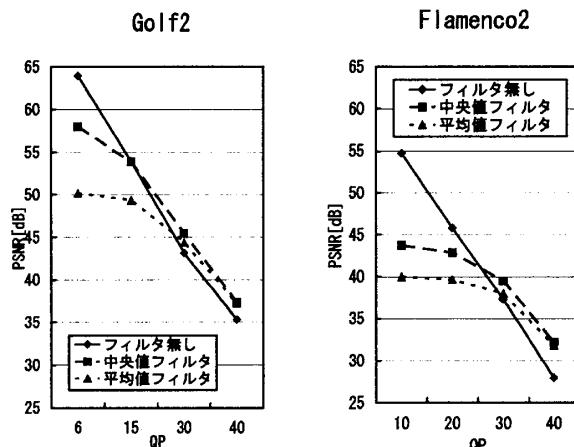


図4 中央値フィルタと平均値フィルタの比較

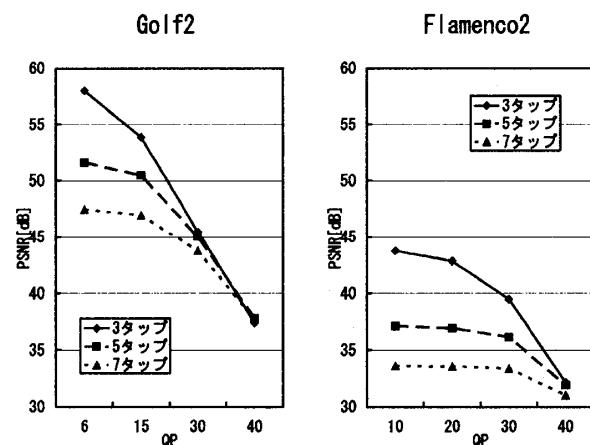


図5 タップ数による変化

5. おわりに

本稿では、圧縮された奥行きマップに平滑化フィルタを施すことにより、自由視点VoDシステムにおける映像の画質を改善する手法を提案した。提案手法は、圧縮により劣化した奥行きマップに対し、3タップの中央値フィルタを、物体と背景の境界を考慮しつつ、左右方向と上下方向にそれぞれ施す。これにより、奥行きマップの劣化を軽減する。この手法により、低ビットレートで伝送する場合、約2~3dB程度画質を改善することができる。

参考文献

- [1] 松村,内藤,川田,小池,松本,”動的に更新される背景バッファを用いた自由視点映像に対する補完方式”,映像情報メディア学会誌,Vol.59,No.6,pp.901-908,2005.
- [2] 石川,松村,川田,小池,”曲面背景バッファによる自由視点 VoD システムの高画質化,”IMPS2004-2.17,pp.47-48,2004.
- [3] 石川,川田,小池,”自由視点 VoD 映像の高画質化のための奥行き情報の伝送方式”,信学技報 IE2005-4,pp.19-24,2005.4.