

視点合成映像を用いた Multiview DVC のための SI 生成に関する検討

志水信哉[†] 木全英明[†]

Distributed Video Coding(DVC)は近年注目を集めた映像符号化手法である。DVCの性能は復号側で作られる Side Information(SI)の品質に依存すると言われている。代表的な SI 生成法では、被写体の動きが等速直線運動であると仮定し、動き補償を行いながら、キーフレーム間を補間することで生成する。しかしながら、被写体の動きは等速直線運動であるとは限らないため、高品質な SI を生成することができなくなった。多視点 DVC では、時間の異なるフレームのほかに、視点の異なるフレームを使用することが可能となる。そこで、我々は視点の異なるフレーム間を補間して得られる視点合成映像を用いて、任意の動きを予測しながら動き補償予測を行うことで、高品質な SI を生成する手法を提案する。実験の結果、従来の最も高品質な SI と比べて、2dB 以上の品質改善を達成した。

Side Information Generation using View Synthesis for Multiview Distributed Video Coding

Shinya Shimizu[†] and Hideaki Kimata[†]

Distributed Video Coding (DVC) is an attractive video coding scheme. It is well-known that the quality of the side information (SI) strongly impacts the coding performance of DVC. One of the most popular SI generations is motion compensated temporal interpolation, where temporal interpolation is performed by assuming linear uniform motion. However, it is difficult to generate high quality SI because there are a lot of irregular motions. In the multiview DVC, it becomes possible to utilize inter-view correlations in addition to the temporal ones. Therefore, this report proposes a temporal frame interpolation method that yields the ability to compensate irregular motions by estimating motion on view synthesized pictures. Simulations show that the proposed method improves SI quality more than 2 dB.

1. はじめに

近年、複数のカメラを同期させて同じシーンを様々な位置や向きから撮影した映像を用いた映像アプリケーションが高い注目を集めている。特に奥行き感や立体感をユーザに提供することのできる立体映像は、その高いエンターテインメント性から、盛んに研究開発が進められている[1]。多視点映像を用いた立体映像以外の代表的な映像アプリケーションとしては、ユーザが自由にカメラを操作できる自由視点テレビや、様々な角度から撮影することでオクルージョン等によるセキュリティ低下を防ぐビデオ監視などがある。

多視点映像の撮像や表示技術の進歩は目覚しく、3D 映画は既に一般で楽しまれているほか、本年度中にも家庭用の立体ディスプレイが市場に現れると言われている。これらの立体映像はステレオ式であり、右目と左目の2視点の多視点映像であるため、そのデータ量は余り問題にならない。しかしながら、裸眼立体視可能な立体ディスプレイや自由視点テレビなどを実現するためには、非常に多くの視点数が必要となる。多視点映像のデータ量はその視点数に比例するため、高効率な圧縮符号化を実現することは、三次元映像アプリケーションを実現する上での大きな課題の1つである。

多視点映像の符号化方式としては、ISO/IEC JTC1/SC29/WG11 Moving Picture Experts Group (MPEG)と、ITU-T SG19 WP3 Q.6 Video Coding Experts Group (VCEG)の元で、MPEG-4 AVC/H.264の拡張規格として Annex. H Multiview Video Coding (MVC)が国際標準規格化されている[2]。MVC は視点間の相関を利用することで、それぞれを個別に符号化するよりも、効率的な圧縮符号化を実現している。しかしながら、複数視点の映像を同時に処理しなくてはならないため、その符号化処理の演算負荷は非常に高い。この問題に対処するために、視点ごとにプロセッサを割り当てて並列処理をすることも可能である。その場合は、視点間の冗長性を取り除くために、他の視点の符号化結果をプロセッサ間で交換する必要が生じるため、符号化遅延が増加してしまうという問題が新たに生じる。特に分散カメラ環境では、ネットワーク上を流れるデータの総量だけでなく、各カメラを物理的に繋ぐネットワークを構築しなくてはならず、大きな問題となる。

高効率な圧縮符号化、低演算符号化、視点ごとの独立した符号化を同時に実現する符号化手法として、Multiview Distributed Video Coding (MDVC)と呼ばれる手法がある[3, 4]。MDVCでは、Side information (SI)と呼ばれる予測画像を生成することで、復号側でのみ視点間の相関を利用する。MDVCの圧縮性能はこのSIの品質に大きく依存すると言われている。そのため、高品質なSIを生成するための方法が数多く研究されている。多視点映像の特徴を利用して、高品質なSIを生成する方法として、視点合成

[†] 日本電信電話株式会社 NTT サイバースペース研究所
NTT Cyber Space Laboratories, NTT Corporation

予測(VSP: View Synthesis Prediction)と呼ばれる方法が提案されている[5]. 視点合成予測では, シーンの奥行き情報とカメラパラメータの情報とを用いて, 別の視点の映像を正確にワーピングすることで SI を生成する. しかしながら, シーンの奥行き推定やカメラのキャリブレーションを正確に行うことは非常に困難であるため, 視点合成 SI の性能は非常に限られたものであると言われている[6].

MDVC では, 視点間相関を利用した SI だけでなく, 通常の DVC と同じように映像の時間相関を利用した SI を使用することも可能である. 時間相関を利用する場合, 被写体の動きの規則性を仮定するため, 不規則な動きを含む映像に対して高品質な SI を生成することは困難である. 一方, 視点間相関を利用する場合, 視差はカメラや被写体の位置関係によって決まるため, 幾何的な規則性を仮定する必要はない. しかしながら, 被写体が完全拡散反射すること, つまり, 同じ被写体から生じる光線は観測位置に依存しないと仮定しているため, カメラの個体差が大きい場合や鏡面反射など完全拡散反射と大きく異なる特性を持つ被写体が含まれている場合には, 高品質な SI を生成することは困難である. どちらの方式も完璧ではないため, 2 通りの方法で生成された SI を融合することで, より高品質な SI を生成する方法も数多く提案されている[5, 7, 8]. しかしながら, これらの方式では, 推定された動きの大きさなどの情報から, SI に重みをつけて合成するため, 動きが不規則な非完全拡散反射の被写体などに対しては, 高品質な SI を生成することができない.

本報告では提案する方法においても, 時間相関を利用した SI と視点間相関を利用した SI の両者を用いて高品質な SI を生成する. ただし, 従来の両者を融合する手法とは異なり, 時間相関を利用して SI を生成する際に, 視点間相関を利用した SI を補助情報として用いることで, SI の品質を改善する方法を提案する. 第 2 章で MDVC 及び既存の SI の生成方法について説明を行い, 第 3 章で提案する SI 生成手法について詳細を述べる. 提案手法の性能評価実験の結果を第 4 章で述べ, 第 5 章でまとめと今後の課題を述べて本報告を締めくくる.

2. Multiview Distributed Video Coding (MDVC)

Distributed Video Coding (DVC)は, Slepian と Wolf [9]や Wyner と Ziv [10]の研究を理論基盤とする MPEG や H.26x など従来のものとは異なる映像符号化方式であり, MDVC はその多視点映像拡張である. DVC では, 相関のある異なる信号を別々に符号化するが, 生成された符号データは相関性を利用しながら一緒に復号する. このとき, Wyner-Ziv 理論によると, ある特定の条件を満たす場合には, 相関を利用しながら一緒に符号化する従来方式と同等の符号化効率を実現できることが示されている. 一般的な DVC の実装では, 各フレームを Key frames (KFs)と Wyner-Ziv frames (WZFs)の 2 種類に分類し, それぞれを別々の方法で符号化する[11, 12]. KFs は従来方式と同

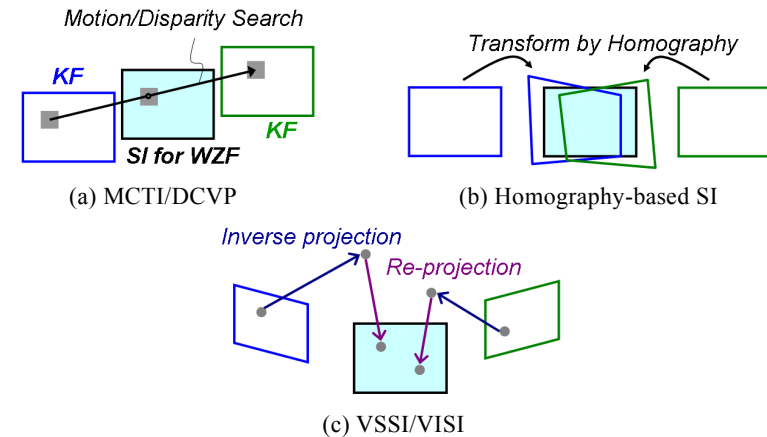


図 1 Various Side Information Generation Methods

様の方法で情報源符号化し, WZFs は通信路符号化した後, パリティビットのみを伝送する. 復号側では, KFs を従来方式で復号した後, それらを用いて SI と呼ばれる WZFs を予測したフレームを生成する. この SI を WZFs に通信ノイズが付加したものとみなし, パリティビットを用いて誤り訂正を行うことで WZFs が復号される. なお, 必要なパリティビットの量はノイズ量に依存するため, あらかじめ多く伝送するか, 誤り訂正が成功するまで符号化器に更なるビットを要求する必要がある. DVC についての詳細な説明は別の文献を参照されたい[13].

DVC で最もよく使われる SI 生成法は, 時間相関を利用した Motion Compensated Temporal Interpolation (MCTI)と呼ばれる方法である[14]. 図 1(a)で表されるとおり, MCTI では KFs 間で動き探索を行い, その動きが等速直線運動であると仮定して中間フレームを内挿補間する. フレームレートが高いなど, 動きが小さく少ない場合に高品質な SI を生成することが可能である.

MDVC では, 時間相関以外に視点間相関を利用できるため, それらを利用した様々な方法が提案されている. Disparity Compensation View Prediction (DCVP)は, 最も代表的な MDVC の SI 生成法であり, 図 1(a)に示す通り基本的なアルゴリズムは MCTI と同じである. つまり, MCTI を時刻の異なるフレーム間ではなく, 視点の異なるフレーム間に適用したものである[7]. ただし, MCTI と異なり, 内挿補間すべき位置が与えられないため, 先頭フレームなど全ての視点で復号画像が得られる時刻において, 最適な内分比を求め, 以後のフレームでは求められた内分比を用いて内挿補間を行う. DCVP で補償する視差は, カメラの位置だけでなく, カメラの向きや被写体の位置に



(a) 原画 (b) 合成画像 (c) デプスマップ
図 2 視点合成画像と合成に用いたデプスマップの品質

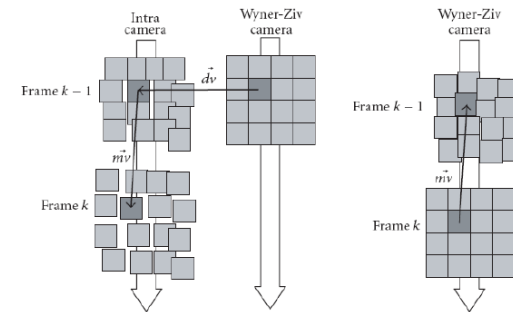


(a) 色の不一致 (b) フォーカスの不一致
図 3 視点間の映像信号の不一致例

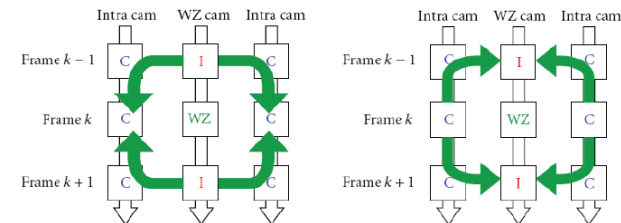
も依存したものであるため、視差を常に同じ割合で内分しても正しい視差補償を行うことができない。その結果、特殊なカメラ配置である場合を除いて、DCVP の性能は MCTI より低いと言われている。

より正確な視差補償を実現するために、Homography-based SI (HSI)が提案されている[6, 8]。ホモグラフィとはあるカメラ平面を別のカメラ平面へと変換する 3×3 行列のことである。HSI では、DCVP と同様に先頭フレームなど全ての視点で復号画像が得られる時刻において、KFs となる視点から WZFs とする視点へのホモグラフィを求め、そのホモグラフィを用いてワーピングされた KFs の平均値で SI を構成する(図 1(b))。ホモグラフィは全ての被写体が 1 つの平面上に存在している場合や、全てのカメラの光軸中心が同一の場合に正確な視差補償を実現することができる。

より一般的な状況において正確な視差補償を実現するために提案された方法が View Synthesis SI (VSSI)である。VSSI では、図 1(c)に示す通り、画素ごとに三次元情報を用いながら逆射影と再射影を繰り返すことで SI を生成する。逆射影とは、被写体のデプス情報を用いて、画素を画像平面から三次元空間へ投影することで三次元点を復元する処理であり、再射影とは、三次元点のあるカメラ平面へ投影するプロセスである。なお、各画素の被写体が属するデプス平面を同定するために、KFs を用いてステレオマッチングなどのデプス推定アルゴリズムを用いる必要がある。デプス平面ごとにその平面に属する画素をワーピングするのだが、その際にはどのデプス平面が補間視点において前面に来るのかを判定しながらワーピングを行う。各画素のデプス情



(a) 動き探索と動き補償予測スキーム



(b) 8通りの動き/視差推定パス

図 4 MVMEによるSI生成(文献[17]より)

報が正確に推定できた場合、VSSI は正確に視差を補償して予測画像を生成することが可能である。

しかしながら、VSSI の性能は非常に限られたものであることが報告されている[6]。これには様々な原因が考えられる。第一にデプス推定を正確に行うことがほぼ不可能ということである。多視点画像から三次元形状を復元する問題は、特にコンピュータビジョンの分野で数十年も取り組まれている課題であるが、未だに完全ではなく、更なる研究が行われている。しかしながら、図 2 を見る限り、別の視点を合成するという目的においては、得られたデプスは許容可能なレベルにあると考えられる。したがって、本報告においては、デプス推定における課題については取り扱わない。

別の原因としては、カメラの個体差による視点間の映像信号のミスマッチの影響がある。実環境において複数のカメラ間で全ての設定を同一にすることは不可能であるため、様々なミスマッチが生じる。例えば、もっとも単純な例ではカラーフィルターやゲイン値が異なれば、被写体が完全拡散反射としていたとしても異なる色や輝度値

が得られることになる (図 3(a)). 別の例としては, 露光, 焦点距離, シャッタースピードなどが異なれば, 焦点位置や被写界深度が一致せず, カメラごとに映像の周波数成分が異なってしまう (図 3(b)). 我々は, この問題に対処するために, カメラ間の mismatches を補正する適応フィルタを推定し, VSSI を補正してその品質を改善する手法を提案した[15]. しかしながら, 適応フィルタによって VSSI を補正するため, VSSI に大きなノイズが生じている場合に, そのノイズを他の画素へ拡散してしまうという問題があった. そこで, 本報告では, 別のアプローチでこの問題に対処する方法を提案する.

時間と視点間のどちらか一方の KFs を使うのではなく, 両者を同時に利用する方法として, Multiview Motion Estimation (MVME)がある[16]. MVME では, 別の視点で求めた動きベクトルを用いて, 動き補償を行うことで WZFs における SI を生成する. なお, 動き補償を行う前に, 動きベクトルはカメラの位置関係などによって変換する. 具体的には, 図 4 (a)に示す通り, まず, WZFs のカメラと別のカメラ間で視差ベクトルと, その視差ベクトルによって示されるブロックの動きベクトル推定する. そして, 得られた視差ベクトルを使って動きベクトルを WZFs のカメラへと投影することで, WZFs のブロックにおける動きベクトルを推定する. 図 4 (b)に示すように, MVME では 8 つの異なるパスで予測画像を異なる生成することができるため, それらを重み付け平均したものが最終的な SI として利用される.

本報告で提案するアプローチも時間と視点間の両方の KFs を同時に利用する点で, MVME に類似した方式である. しかしながら, MVME には被写体の動きが対象のカメラの光軸に対して, 垂直もしくは水平でなければ, 動きベクトルを正しく変換できないが, 提案方式ではそのような制限はなく高品質な SI を生成することが可能である. また, MVME は 16 回のブロックマッチングを行う必要があるが, 提案方式では, 最低 4 回のブロックマッチングで SI を生成することができる.

3. View Synthesis Motion Estimation

第 2 章で述べたように, 時間的にフレーム補間を行う MCTI の欠点は, 被写体の動きが等速直線運動であると仮定しなくてはならない点である. 一方, カメラ間でフレーム補間を行う VSSI の欠点は, 被写体が完全拡散反射であり, カメラが全く同一のものである仮定しなくてはならない点である. そこで, この 2 つの方法を組み合わせることで, 両者の欠点を補い高品質な SI を生成する方法を提案する. つまり, 被写体の運動モデルを仮定せずに, 時間方向に映像信号を補間する方式を提案する.

提案手法である View Synthesis Motion Estimation (VSME)の基本的なアイデアは図 5 に示す通りである. つまり, VSSI を WZFs と時間方向の KFs に対して生成し, VSSI 間の動きベクトルを用いて時間方向のフレーム補間を行う. VSSI は, 映像信号として

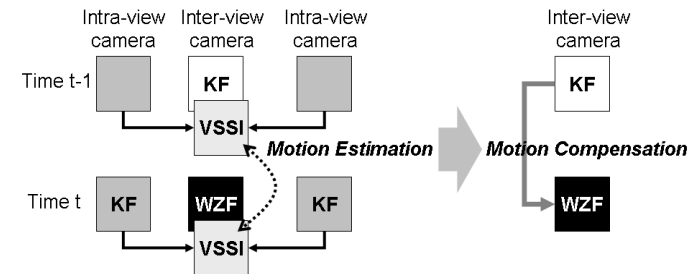


図 5 提案手法(VSME)の基本コンセプト

は誤りを含むが, 幾何的には高精度にフレーム補間を実現することができる. そのため WZFs に対して生成される VSSI は, WZFs の被写体やシーンの形状に対する高精度な予測とみなすことが可能である. そこで, この VSSI を用いて動きベクトルを求めることで, 被写体の運動モデルを仮定する必要がなくなる. なお, VSSI と時間方向の KFs 間で動き推定をすることも可能であるが, VSSI と時間方向の KFs とでは映像信号としての mismatches が存在しているため, その影響が大きい場合は高精度に動きベクトルを見つけることができず, 最終的な SI の品質が低下する. しかしながら, VSSI を生成する処理は非常に高負荷な演算であるため, 時間方向の KFs を直接使って高精度な動き推定を行う手法を見つけることは今後の課題である. 以下, 詳細に処理を説明する.

まず, WZF と時間方向の KFs に対して VSSI を生成する. VSSI の生成にはどのような手法を用いることが可能であるが, 幾何的に正しい VSSI を生成すると, 最終的な SI の品質が向上する. ここでは, WZFs の各画素に対して, デプスを推定し, そのデプスを用いて視点合成を行う手法を用いた. 具体的には, あるデプスを仮定した際の両隣のフレームにおける対応画素間の画素値の二乗誤差をマッチングコストとし, 隣接画素とのデプス値の差の絶対値に比例する平滑化コストとして設定したエネルギー関数を, グラフカットを用いて最小化してデプスを推定した. 図 2 (c)が推定されたデプスマップの例である. 隣接フレームに対してそれぞれデプスマップを求めても構わないが, その場合はデプス推定を 2 度行うことで演算量が増加する.

WZFs と時間方向の KFs に対する VSSI の生成が完了したら, VSSI にローパスフィルタを施す. 画素ベースのワーピング処理によって生じる VSSI における不自然な高周波や, もととの映像が持つノイズの影響をできるだけ少なくし, よりロバストに動きベクトルを探索するためである. 動き探索を行う際のパラメータとしては, ブロックサイズと探索窓, 探索領域がある. 探索窓とはブロックマッチングを行う際に, 対象ブロックのほかにも周囲何画素までを考慮して探索を行うかを示すものである. 実

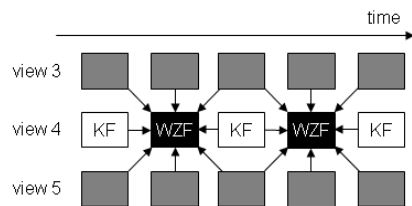


図 6 実験における KFs と WZFs の設定

験はブロックサイズを 1x1, 探索窓を 9x9, 探索領域を 32x32 で固定して行った。

動き探索においては隣接画素との関係を余り考慮しないため、その結果得られる動きベクトルの空間相関が低くなってしまっているときがある。実際の映像では、被写体の動きには空間的な相関が存在するため、Weighted Vector Median Filter を用いて、空間的に動きベクトルの外れ値を減少させる。得られたベクトルを用いて、KFs との間で動き補償予測を行うことで SI を生成する。

なお、複数の KFs を参照可能な場合は、それぞれ別々に動き補償予測を行い、得られた予測画像の平均を取って最終的な SI を生成する。単純な平均ではなく動きベクトルの信頼度に応じて重み付けを行うことで、より高品質な SI を生成可能であると考えられるため、最適な重みの決定方法は今後の課題である。

4. 実験

提案手法の評価実験は Microsoft Research が提供している多視点映像の breakdancers と ballet を用いて行った[17]。配布されているシーケンスの解像度は 1024x768 であるが、今回の評価実験では 256x192 にダウンサンプルしたものをを用いた。なお、どちらも 100 フレームからなる 15fps の比較的動きの激しい映像である。

KFs と WZFs の選択には様々な方法が考えられるが、今回は図 6 で示される最も単純な設定で行った。この設定では View3 と View5 は Intra-view として、カメラ間相関を使わずに処理されるものとし、View4 のみ Inter-view として、カメラ間相関を使った復号が行われるものとした。なお、今回は View4 のみを評価対象とし、View3 と View5 でどのような処理が行われるのかは関係ないものとした。View4 のフレーム構成としては、最も単純な設定ということで、1 フレームごとに KFs と WZFs とを入れ替えながら処理するものとした。

提案手法の SI の品質を評価することで行った。つまり、生成された SI とオリジナルの WZFs との PSNR で評価する。実験条件は文献[5]と同じ設定を採用し、全ての KFs と Intra-view のフレームはロスレス符号化されているものとした。なお、KFs や Intra-view のフレームに発生する符号化ノイズの影響も受けるため、代表的な DVC

Codec を用いて、実際に符号化を行いレート歪み性能を確認する必要があるため、今後の作業として評価を行う予定である。

従来手法として MCTI, DCVP, MVME, VSSI (VISI), AFVISI の実装を行い、同様の条件で SI を生成し比較を行った。複数の SI を融合してより高品質な SI を生成する方法も提案されているが、今回提案する VSME を融合する際の候補として採用することも可能であるため、それらの方式との比較は行わなかった。また、SI の品質を向上させる方法として Iterative SI generation という手法もあるが、これも VSME を初期 SI の一部として使用することで、両方の利点を生かすことができるため、今回は比較することをしなかった。

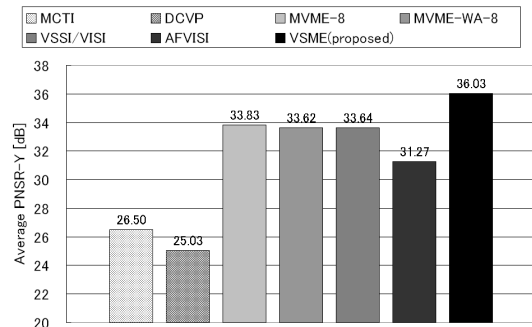
実験結果を図 7 に示す。どちらのシーケンスに対しても提案手法が最も高い品質を示しており、従来手法で最も高い品質を示した MVME よりも 2dB 以上もその品質を改善していることが分かる。また、視点合成の精度が比較的高い breakdancers のみでなく、視点合成の精度が低い ballet においても、その視点合成画像を用いて探索した動きベクトルを用いて動き補償予測を行うことで、非常に高い品質の SI の生成が可能になった。図 8 は生成された SI の例である。右側の背景のカーテン部分において、VSSI では視点合成がうまく行かず画像が崩れてしまっているのが、MVSE では時間方向に予測しなおすことで正しく推定が出来ていることがわかる。

5. おわりに

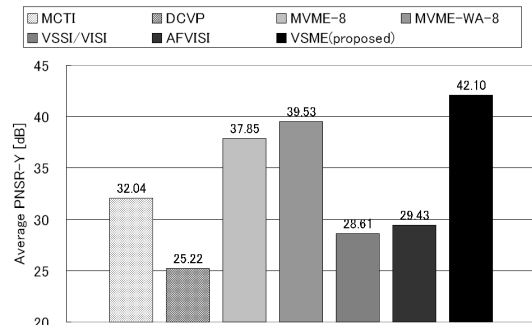
本報告では、効率的な MDVC の実現に向けて、高品質な SI の生成方法の提案を行った。提案方式である View Synthesis Motion Estimation では、WZF を撮影したカメラの映像を視点合成によって生成し、その視点合成映像上で動き探索を行い、その動きベクトルを用いて同じカメラの KF から動き補償予測を行うことでフレーム補間を行う。これによって、被写体が非線形な動きを有している場合にも、正しい動き補償を実現することが可能となり、従来に比べて高い品質の SI を生成可能となる。実験の結果、従来方式と比べて 2dB 以上の品質改善を実現できることが分かった。

今後は、提案手法を用いて生成した SI によって達成できる MDVC のレート歪み特性を明らかにする予定である。また、VSSI を生成する演算コストは非常に高いため、WZF 対してのみ VSSI を生成し、KFs に対しては生成せずに高精度な動きベクトル推定を達成するなど、演算コストを下げるための検討は今後の課題である。

謝辞 テストシーケンス”breakdancers”と”ballet”を提供して頂いた Microsoft Research Visual Media Group に深く感謝申し上げます。



(a) breakdancers



(b) ballet

図 7 各 SI の平均 PSNR



(a) VSSI

(b) VSME

図 8 生成された SI の例

参考文献

- 1) A. Smolic, K. Muller, P. Merkle, C. Fehn, P. Kauff, P. Eisert, and T. Wiegand, "3D video and free viewpoint video – technologies, applications and MPEG standards," in *Proc. ICME2006*, pp.2161-2164 (July 2006).
- 2) ISO/IEC 14496-10:2009 Coding of audio-visual objects – Part 10: Advanced Video Coding (edition 2009-05-15) Annex H. Multiview Video Coding.
- 3) X. Zhu, A. Aaron, and B. Girod, "Distributed Compression for Large Camera Arrays," in *Proc. IEEE Workshop on Statistical Signal Processing*, pp.30-33 (2003).
- 4) G. Toffetti, M. Tagliasacchi, M. Marcon, A. Sarti, S. Tubaro, and K. Ramchandran, "Image Compression in a Multi-camera System based on a Distributed Source Coding Approach," in *Proc. EUSIPCO2005* (September 2005).
- 5) X. Artigas, E. Angeli, and L. Torres, "Side Information Generation for Multiview Distributed Video Coding using a Fusion Approach," in *Proc. NORSIG2006*, pp.250-253 (June 2006).
- 6) M. Ouaret, F. Dufaux, and T. Ebrahimi, "Iterative Multiview Side Information for Enhanced Reconstruction in Distributed Video Coding," *EURASIP Journal on Image and Video Processing* (2009).
- 7) M. Ouaret, F. Dufaux, and T. Ebrahimi, "Multiview Distributed Video Coding with Encoder Driven Fusion," in *Proc. EUSIPCO2007* (2007).
- 8) M. Ouaret, F. Dufaux, and T. Ebrahimi, "Fusion-based Multiview Distributed Video Coding," in *Proc. 4th ACM international workshop on VSSN*, pp.139-144 (2006).
- 9) D. Slepian and J. K. Wolf, "Noiseless Coding of Correlated Information sources," *IEEE Transactions on Information Theory*, vol.19, no.4, pp.471-480 (July 1973).
- 10) A. D. Wyner and J. Ziv, "The Rate-distortion Function for Source Coding with Side Information at the Decoder," *IEEE Transaction on Information Theory*, vol.22, no.1, pp.1-10 (January 1976).
- 11) R. Puri and K. Ramchandran, "Prism: A Video Coding Architecture based on Distributed Compression Principles," EECE Department, University of California, Berkeley, Tech. Rep. UCB/ERL M03/6, 2003 [Online]. Available: <http://www.eecs.berkeley.edu/Pubs/TechRpts/2003/4055.html>
- 12) A. Aaron, R. Zhang, and B. Girod, "Wyner-ziv Coding of Motion Video," in *Proceedings of Asilomar Conference on Signals and Systems*, pp.240-244 (2002).
- 13) 高村誠之, "Distributed Video Coding の動向と今後," 情報処理学会研究報告 2006-AVM-54, pp.71-76 (2006).
- 14) J. Ascenso, C. Brites, and F. Pereira, "Improving Frame Interpolation with Spatial Motion Smoothing for Pixel Domain Distributed Video Coding," in *Proc. EC-SIPMCS2005* (July 2005).
- 15) S. Shimizu, Y. Tonomura, H. Kimata, and Y. Ohtani, "Improved View Interpolation for Side Information in Multiview Distributed Video Coding," in *Proc. ICDCS2009* (August 2009)
- 16) X. Artigas, F. Tarres, and L. Torres, "Comparison of Different Side Information Generation methods for Multiview Distributed Video Coding," in *Proc. SIGMAP2007* (July 2007).
- 17) C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High-quality video view interpolation using a layered representation," *ACM Trans. Graph.*, vol.23, no.3, pp.600-608 (2004).