

デプスベース三次元映像符号化方式の国際標準化動向 International Standardization of Depth-based 3D Video Coding

志水 信哉[†]
Shinya Shimizu

1 まえがき

本稿では、ISO/IEC JTC1 SC29/WG11(MPEG)とITU-T SG16(VCEG)の合同で結成されたJoint Collaborative Team on 3D Video Coding Extension Development(JCT-3V)の下で進められている三次元映像符号化方式について概説する。まず、想定するシナリオと共に、符号化対象となる映像データについて説明する。その後、HEVC[1]の拡張規格として規格化が進められている3D-HEVCの代表的な符号化技術を概観する。なお、符号化技術の詳細については、3D-HEVCテストモデル文書を参照されたい[2]。

JCT-3Vでは、自由視点映像や立体映像など三次元映像の効率的な圧縮符号化の実現を目的とする。これらの三次元映像で必要となる非常に多くの視点に対する映像を限られた符号量で表現するために、三次元映像を多視点映像とデプスマップを用いて表現するフレームワークを想定している。多視点映像とデプスマップに対して、Depth-Image Based Renderingと呼ばれる映像合成の手法を適用することで、各視点間の高品質な映像を軽量の演算で生成できる。これによって、限られた視点数の情報のみで一定範囲内の任意の視点の映像を表現し、効率的な三次元映像の蓄積・伝送を実現する。なお、多視点映像とデプスマップで表現されたMVD(Multiview Video plus Depth map)と呼ばれるデータに対する圧縮符号化方式のみが標準化対象であり、映像合成に用いる手法は標準化対象外である。

2 3D-HEVC テストモデル

3D-HEVCは、MVDデータを効率的に圧縮符号化するために、HEVCをベースとして検討されている標準規格の略称である。2012年6月時点での規格案には、H.264/AVCの多視点映像符号化拡張規格MVCに採用されている視差補償予測に加えて、様々な符号化技術が採用されている。ここでは、代表的な符号化技術について紹介する。なお、3D-HEVCでは、各視点において映像、デプスマップの順に符号化を行う前提である点に留意されたい。

2.1 視差ベクトル推定

多視点映像を効率的に符号化するためには、視点間の対応関係を示す視差ベクトルを効率よく獲得することが重要である。MVCでは使用する視差ベクトルは動きベクトルと同様に符号化する必要があったのに対して、3D-HEVCでは時空間で隣接するブロックやデプスマップの情報を用いて推定する技術を採用している。時空間で隣接するブロックから推定された視差ベクトルは

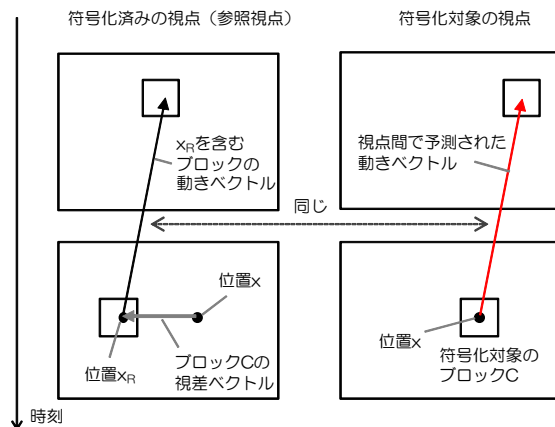


図1: 視点間動き予測

NBDV(Neighboring Block Disparity Vector)と呼ばれており、後述する様々な符号化ツールで使用されている。

NBDVの導出は、既に符号化済みのブロックを予め定められた順にチェックし、そのブロックが視差補償予測で符号化されている場合などに蓄積されている視差ベクトルを同定することで行われる。符号化済みのブロックのチェックは、時間隣接ブロックから開始し、空間隣接ブロックを左、上の順に行う。

NBDVをデプスマップの情報を用いて更新する技術も採用されている。更新された視差ベクトルはDoNBDV(Depth-oriented NBDV)と呼ばれる。DoNBDVは、NBDVによって対応付けられた領域における符号化済みデプスマップの値を視差ベクトルへ変換することで行われる。

2.2 視点間動き予測

視点間動き予測(IVMP: Inter-View Motion Prediction)は、その名の通り、別の視点に対する動き情報を用いて、現在のブロックの動き情報を生成する方法である。図1に示すように、現在のブロックに対応する別の視点のブロックをNBDVまたはDoNBDVを用いて同定し、そのブロックにおいて符号化された動き情報が予測された動き情報となる。なお、動き情報は現在のブロックを分割したサブブロック毎に求める。

IVMPは動き情報の符号化を省略することで、全体の符号量を削減する。ただし、別の視点を符号化した際の動き情報を蓄積し、参照する必要があるため、メモリ量とメモリアクセス量は増加する。なお、IVMPは、3D-HEVCに採用されている符号化技術の中で、最も符号化効率への影響が大きい技術である。

2.3 視点合成予測

視点合成予測(VSP: View Synthesis Prediction)は、符号化済みのデプスマップを用いて、高精度な視差補償

[†]日本電信電話株式会社 メディアインテリジェンス研究所
NTT Media Intelligence Laboratories, NTT Corporation

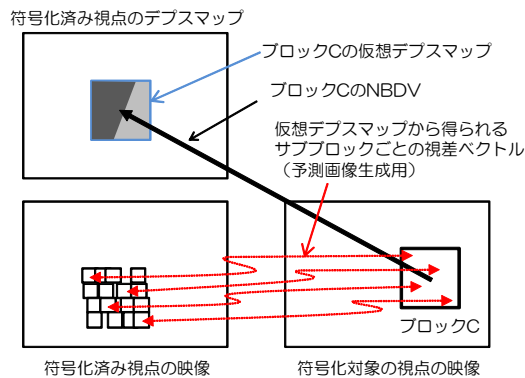


図 2: 視点合成予測

予測を実現する手法である。通常の視差補償予測と異なり、視差ベクトルを符号化する必要がなくなるため、小さなブロック毎に異なる視差を補償した高精度な予測を少ない付加情報量で実現することができる。

3D-HEVCでは、図2に示すように2段階の視差補償予測によってVSPを実現する。1段階目は、現在のブロックの仮想デプスマップを生成するための処理であり、NBDVによるデプスマップ上の視差補償予測である。2段階目は予測信号を生成するための視差補償予測である。具体的には、現在のブロックを4x8画素または8x4画素に分割したサブブロック毎に代表となるデプス値を求め、そこから得られた視差ベクトルを用いて視差補償予測を行う。

2.4 残差予測

残差予測は、動き補償予測や視差補償予測の結果として得られる予測残差に対する予測である。動き補償予測の予測残差を視点間で予測するモードと、視差補償予測の予測残差を時間的に予測するモードがある。図3に動き補償予測残差を視点間で予測する場合を示す。

NBDVによって得られる別の視点上の対応ブロックにおいて、現在のブロックと同じ動き情報を用いた動き補償予測を行うことで、残差予測に用いる参照予測残差を生成する。なお、図3の場合の現在のブロックに対する最終的な予測信号は $B_t + w * (B_{iv} - B_{iv+t})$ で表される。 w は残差予測の重み係数 (1 又は 0.5) である。

残差予測は従来の予測よりも多くのブロックをアクセスしなくてはならず、デコーダ負荷が高い。そのため、補間にバイリニアフィルタを用いるなど、デコード負荷低減のために様々な工夫がされている。

2.5 デプスペース領域分割

一般に、動きや視差の情報は写っている被写体に依存するため、被写体毎に異なるベクトル情報を用いることで、予測残差の量を削減できると言われている。しかしながら、被写体形状に応じた任意の形状を表現するには、非常に多くの付加情報が必要となり、トータルとしては効率的な符号化を実現することは困難である。しかしながら、デプスマップが被写体を表現していると考えられるため、領域分割の情報を別途符号化せずに、被写体形状に応じた領域分割を実現することが可能である。

デプスペース領域分割 (DBBP: Depth-Based Block

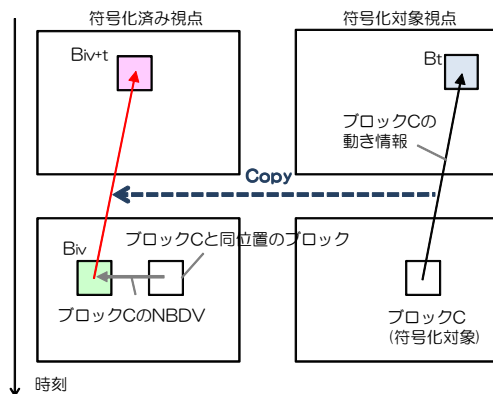


図 3: 残差予測

Partitioning) は、そのような被写体形状に応じた領域分割を用いたインター予測を実現する方法である。DBBPでは、DoNBDVによって現在のブロックに対応付けられた仮想デプスマップを2値化することで、現在のブロックを2つの領域に分割する。その後、分割された領域ごとに最適なベクトルを推定し予測画像を生成する。

2.6 領域ベースデプス符号化

デプスマップは、映像と異なり、被写体内では値が一定または滑らかに変化し、被写体間では値が大きく変化する特徴を持つ。領域ベースデプス符号化 (SDC: Segment-based Depth Coding) は、そのようなデプスマップを効率よく符号化するための技術である。

SDCでは、ブロック内の画素と最大2つの領域に分割し、それぞれの領域のデプスが定数であるとして符号化する。領域分割の方法には2つの方法が採用されている。1つ目は、直線を用いてブロック内を2つの領域に分割する方法である。分割に使用する直線はWedgeletによってパラメータ化して符号化する。2つ目は、同じ位置の映像を2値化することで、より柔軟な形状の領域へと分割する方法である。

領域内のデプス値が定数で近似できない場合は、予測残差を符号化することで、領域内のデプス値の変化を符号化することも可能となっている。なお、そのような場合の予測をSDCと区別して、DMM (Depth Modeling Mode) と呼んでいる。

3 まとめ

本稿では、JCT-3Vで行われているデプスペース三次元映像符号化方式3D-HEVCについて紹介した。3D-HEVCでは、視点間の相関やデプスマップの特性に基づく新しい符号化技術を採用することで、通常の高視点映像符号化と比較して20%以上の符号量を削減する。なお、本稿で紹介した技術以外に、映像とデプスマップ間の相関を利用する方法や、合成映像の品質を考慮したデプスマップ符号化制御なども検討されている。

参考文献

- [1] Recommendation ITU-T H.265: High efficiency video coding, 2013.
- [2] Y. Chen, *et al.*: "Test Model 8 of 3D-HEVC and MV-HEVC", JCT3V-H1003, 8th JCT-3V meeting (April 2014)