

知っておきたいキーワード

多視点映像符号化

(正会員) 山口 宗明[†]

[†] 株式会社日立製作所 中央研究所 組込みシステム基盤研究所

"Multi-view Video Coding" by Muneaki Yamaguchi (Embedded System Platform Research Laboratory, Central Research Laboratory, Hitachi, Ltd., Yokohama)

キーワード：多視点映像，映像符号化，標準規格，MPEG

多視点映像符号化とは

多視点映像符号化 (MVC: Multi-view Video Coding) は、複数の映像を統合して符号化する技術です。この技術は、自由視点映像や3Dテレビ等の映像アプリケーションに利用でき、複数の視点からの映像を効率よく符号化することなどを目的として、標準化が現在も進められています¹⁾。

多視点映像符号化では、複数のカメラ等の視点からの入力映像を統合して符号化します。例えば、図1 (a) に示すように、三つの視点(視点1, 視点2, 視点3)がある場合に、図1 (b), ~図1 (d) に示す視点1, 視点2, 視点3の各視点の映像を符号化することにより、復号後に視点を選択して表示することが可能になります。

さらに、多視点映像符号化の標準化における要求条件¹⁾には、例えば図1 (a) の「生成したView」のように、符号時に設定された視点の制限を越えて、自由な視点からの映像表示を実現するための機能が含まれています。また、

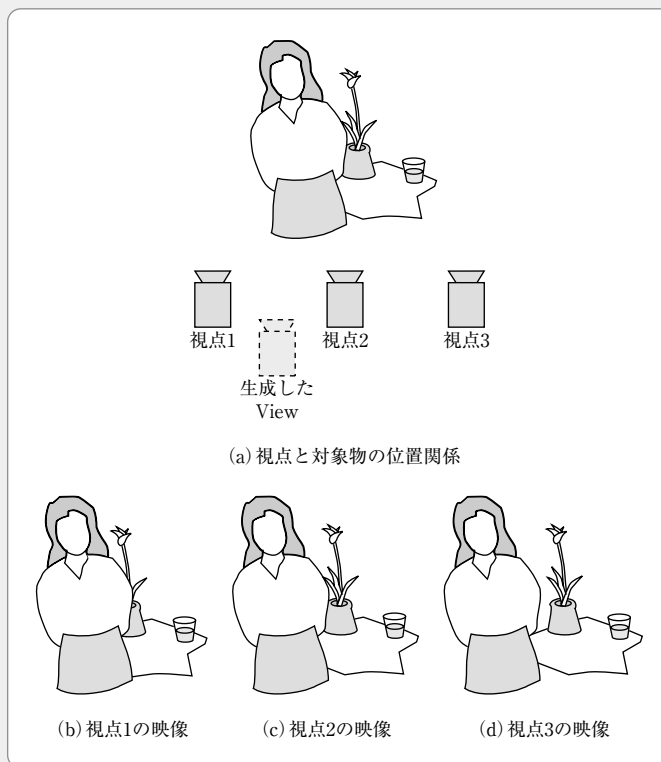


図1 三つの視点による多視点映像の例

選択した視点の映像抽出を容易にする技術や、異なる視点の映像の並列処理

を実現する技術などの実現についても、要求条件に挙げられています¹⁾²⁾。

多視点映像符号化規格策定の現状

多視点映像符号化の標準化は、MPEG-4 AVC/H.264³⁾の拡張として進められています。2007年12月における草案⁴⁾の規格案は以下の通りです。

まず、ベースとなるMPEG-4 AVC/H.264を多視点に拡張するために、視点ごとに識別番号を割り振り、データのそれぞれがどの視点の映像のデータであるかを明確にします。この識別番号として“view_id”というパラメータを割当てます。view_idは、ビットストリームの中のNAL (Network Abstraction Layer) ヘッダに格納されます。

NALはMPEG-4 AVC/H.264にて採用された仕様で、さまざまなフォーマットで利用しやすいように、圧縮データを分割しパケット化するものです。NALには、各種ヘッダ、映像の圧縮データが処理単位毎に格納されています。

多視点映像符号化では、NALヘッダ中のview_idによって、NALに格納された圧縮データがどの視点に属するものであるかを示します。言い換えると、同一のview_idであるNALに格納されている圧縮データを復号したものが、同一視点の映像となります。

次に、多視点映像符号化における圧縮率向上方法について、一般的なフレーム間予測と比較しながら説明します。

MPEG-2などの一般的な映像符号化では、フレーム間予測を用いて圧縮率を向上しています^{5) 6)}。これは、時間的に前後するフレームでは、絵柄が似ているという性質を利用したものです。図2に示すように、表示順が前(過去)のフレームから前方予測を行うPフレームと、表示順で前後のフレームから双方向予測を行うBフレームがあります。これにフレーム内で予測を行うIフレームを併せた3種類が、予測構造の基本的な構成です。

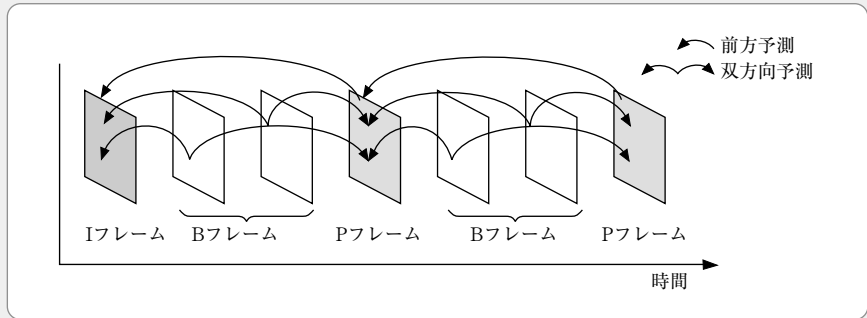


図2 MPEG-2などにおける予測構造の例

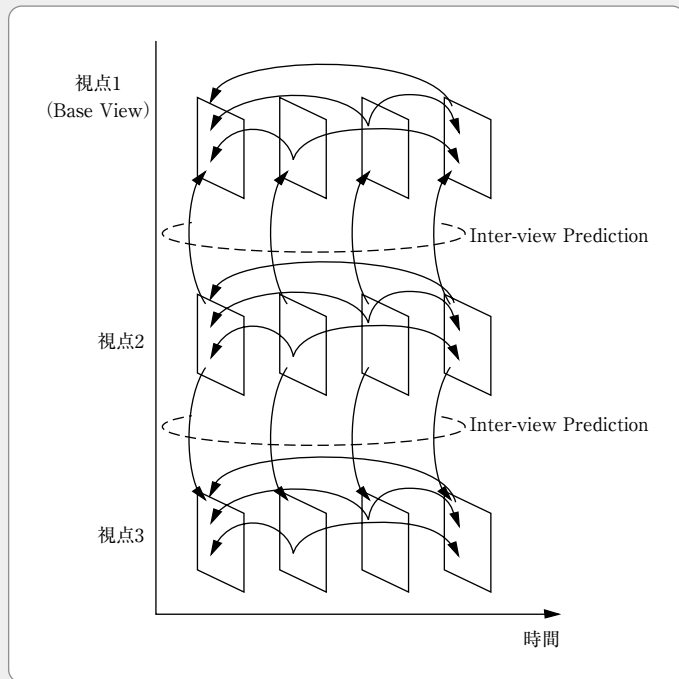


図3 多視点映像符号化における予測構造の例

一方、多視点映像符号化では、複数の視点があることを利用して、上記仕組みを視点 (View) 間の映像に取り入れることが提案されています。図1に示した視点と出力映像の関係をもう一度見てみましょう。図1 (b) ~ 図1 (d) の映像は、非常に似た絵柄となっています。このように、時間的に前後するフレームと同様に、視点間においても隣り合う視点は非常に似た絵柄となる可能性が高くなります。したがって、フレーム間予測に加え、視点間の予測を

行うことにより圧縮率の向上が期待できます。この視点間での参照を“Inter-view Prediction”と呼んでいます。

フレーム間の予測と視点間の予測を組み合わせると、一例として図3に示すような参照関係となります。なお、図3の視点1のように、他の視点からの予測を行わず、同一の視点のフレーム間予測のみで構成される視点を“Base View”と呼びます。Base Viewは、映像の復号の際に他の視点の復号の必要がない点の特徴です。

標準化スケジュール

多視点映像符号化の標準化は、2005年にISO/IEC MPEGにて開始され、ISO/IECとITU-Tの共同プロジェクトであるJVT (Joint Video Team)

にて行われています。現在は、草案の段階で最終決定ではなく、最終規格案が2008年7月に発行される予定となっています。

また、多視点映像符号化に関連する標準化として、FTV (Free viewpoint

TV)の検討がISO/IEC MPEGで開始されています⁷⁾。

今後、上記二つの標準化活動を中心として、自由視点映像が実現していくと思われます。(2008年1月7日受付)

参考文献

- 1) "Requirements and Application Description on Multi-view Video Coding", ISO/IEC JTC1/SC29/WG11 N9543 (Oct. 2007)
- 2) 木全英明：“多視点映像符号化MVCの国際標準化動向”，映情学誌 (Apr. 2007)
- 3) ISO/IEC 14496-10: "Information Technology -Coding of Audio-visual Objects- Part 10: Advanced Video Coding", IS 3rd edition (Dec. 2005)
- 4) "Test of ISO/IEC 14496-10:200x/PDAMI Multiview Video Coding", ISO/IEC JTC1/SC29/WG11 N9213 (Oct. 2007)
- 5) 藤原洋監修：“最新MPEG教科書”，アスキー (1994)
- 6) ISO/IEC 13818-2: "Information Technology -Generic Coding of Moving Pictures and Associated Audio Information: Video", IS Second Edition (Dec. 2000)
- 7) "Applications and Requirements on FTV", ISO/IEC JTC1/SC29/WG11 N9466 (Oct. 2007)



やまぐち ちかあき
山口 宗明 1987年、東京工業大学金属工学科卒業。1989年、同大学院修士課程修了。同年、(株)日立製作所入社。映像符号化に関する研究・開発に従事。特に、MPEG-4 AVC/H.264映像符号化の実用化研究・開発を中心に活動。当学会編集企画幹事、正会員。