

多面体モデルを用いた全方位ビデオの圧縮

Compression of Omnidirectional Video with Polyhedral Mapping

波部 斉

山澤 一誠[†]

野村 敏男^{††}

松山 隆司^{†††}

Hitoshi HABE Kazumasa YAMAZAWA Toshio NOMURA Takashi MATSUYAMA

京都大学大学院工学研究科

Graduate School of Engineering, Kyoto University

^{††} シャープ (株)

SHARP Corporation

[†] 奈良先端科学技術大学院大学情報科学研究科

Graduate School of Information Science, NAIST

^{†††} 京都大学大学院情報学研究科

Graduate School of Informatics, Kyoto University

Abstract When we transmit or store omnidirectional video, because of its wide field of view, the amount of data becomes very large. Therefore, it is necessary to compress omnidirectional video to achieve practical usability. However, when we apply conventional compression methods to omnidirectional video, their efficiency is limited. This is because the geometry of omnidirectional video is not taken into account in conventional compression methods. In this paper, we propose a novel method for compressing omnidirectional video. In our method, omnidirectional video is mapped onto a polyhedral screen, and MPEG-4 compression is applied to the unfolded (planar) polyhedral screen.

1 はじめに

複数のカメラまたは超広角光学系を利用することで、ある視点を中心とした全方位の視覚情報を同時に獲得することができる。このようにして撮影された映像（全方位ビデオ）を利用者に提示することで、あたかも撮影された場所にいるような感覚（没入感）を与えることが可能であるが、臨場感が高い映像を実時間で伝送して利用者に提示するためにはその圧縮が不可欠である。

全方位ビデオは、視点を中心とする球面上に映像情報がマッピングされたものと考えることができ、従来の2次元のビデオにはない幾何構造をもっている。そのため、既存の映像に対する圧縮手法では、データ量などの効率が悪くなってしまふ。また、全方位カメラで撮影したそのままの映像を圧縮・伝送すると受信側でカメラに応じた幾何変換を行なう必要が生じ、効率的ではない。

これに対し、球面直交展開によって1枚の全方位画像を記述する手法 [1] が提案されているが、本稿では全方位ビデオを対象とし、多面体モデルによる圧縮手法を提案する。本手法においては、まず、全方位ビデオの視覚情報を多面体上に投影する。次いで、その多面体の辺を切り開いて展開図とすることで平面上のビデオ映像を得る。最後に、平面ビデオ映像に対して通常の動画圧縮法を適用する。この手法は、全方位ビデオの撮影媒体によらない手法であり、次節以降で述べるように球面上の視覚情報を効率良く表現することができる。さらに、多面体モデルを用いることで、復号して任意視野の映像を生成するときの幾何歪みを少なくすることで、既存のグラフィックアクセラレータとの親和性が高いために高速描画が可能になる、という利点がある。

2 全方位ビデオの平面への変換

前述のとおり、提案手法は大きく分けて、1) 全方位ビデオのもつ情報の平面への変換、2) 変換された平面ビデオに対する既存の圧縮法の適用、のステップから構成される。本節ではまず、1) のステップについて述べる。

2.1 スクリーンの形状

球面上の映像データの平面への変換は、球を囲むスクリーンに球面上のデータを投影し、そのスクリーンを展開する変換と考えることができる (Fig. 1)。このような変換は、世界地図の描画方法 (地図投影法) として使用されており、主に円柱や円錐がスクリーンとされてきた。

多くの地図投影法は、直観的な見やすさや距離・方位の測りやすさなどに主眼がおかれている。一方、全方位

ビデオの表現方法としては、a) データ量、b) 受信した側での透視投影映像の生成、の2点での効率の良さが求められる。a) に関しては地図投影法の中でも球面上のデータを過不足なく表現する方法があるが、そのような手法では、平面上の座標と球面上の座標の間の変換に非線形な演算が含まれるため b) の面での計算効率が悪く、透視投影映像上の幾何的な歪みも避けられない。よって、地図投影法は全方位ビデオの表現法として適切ではないといえる。

これに対し、我々は多面体モデル (Fig.1(c),(d)) の適用を提案する。多面体モデルでは球をよく近似できるため a) に関しても効率が良く、座標変換に非線形演算が含まれないため b) についてもグラフィックアクセラレータによる高速描画が可能になる。次節では圧縮に適した多面体モデルについて考察する。

2.2 多面体モデルの比較

平面ビデオに変換した後の既存の圧縮法の適用を考えると、フレーム間の冗長度の削減の面から、スクリーンを展開した平面 (展開図) 上で画像端における物体の消失・出現が少ない方が望ましい。すなわち、平面展開の際に切れ目が少ない単純な形状の多面体モデルが望まれる。

また、データ量に関しても効率が良い多面体を選択する必要がある。Fig.2 は球面から平面への投影を示しており、球面上で同じ立体角 α に相当する領域でも球に接する領域 a と離れた領域 b では平面スクリーン上での面積が変化することを示している。a, b は多面体の面の中心と稜線の部分に相当する。a と b との差は平面上での解像度の偏りに相当し、これが少なくなる、つまり面数が多い多面体のデータ量の効率が良いことになる。

以上2つの条件は相容れないため、本稿では評価実験によってその影響を検証する。

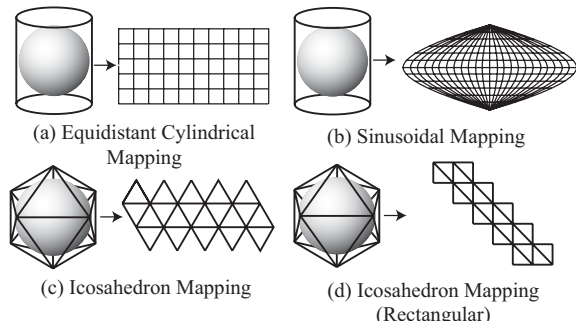


Fig.1: Mapping onto Unfolded Screens

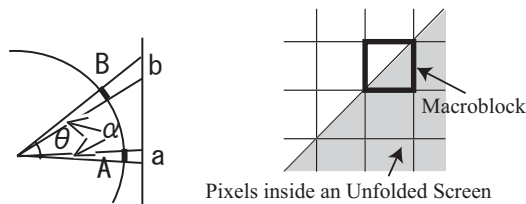


Fig.2: Resolution Distribution

Fig.3: Macroblock and Unfolded Screen

3 平面上の全方位ビデオの圧縮

前節の手法で平面上に変換された全方位ビデオは、通常の映像圧縮法によって圧縮される。MPEG に代表される 2 次元映像の圧縮法は矩形の映像を前提としているが、多面体モデルを用いた場合、その展開図の輪郭は必ずしも矩形にはならない。MPEG-4 の任意形状オブジェクト符号化を利用すれば輪郭が矩形でない展開図にも対応できるが、Fig.3 に示すように斜めの輪郭を含むマクロブロックには展開図外の不必要な領域が含まれてしまうため、効率の低下が起こる可能性がある。これに対し、平面展開図の輪郭に斜めの部分が含まれないように変換することも可能である (Fig.1(d))。この場合には逆に、展開図上の解像度の歪みによる画質低下も考えられるので、評価実験によってその影響を明らかにする。

4 評価実験

本節では、評価実験によってここまで述べた議論を確かめていく。評価実験は以下の 3 ステップから構成される。本実験では全方位カメラで撮影した実映像とシミュレーションで生成した全方位ビデオを用いて評価を行った。

(1) 各展開方法による展開図の作成：与えられた全方位ビデオからここまで述べた各展開方法による展開図を作成する。ここでは地図投影方法の中で、緯度・経度を平面上の xy 座標に対応づける正距円筒図法 (Fig.1(a)) と正距円筒図法の経度方向の解像度を緯度に応じて変化させることで球面上で均等にサンプリングする Sinusoidal 図法 (Fig.1(b))、多面体モデルの中で正六面体、八面体、正二十面体 (Fig.1(c)) および Fig.1(d) に示した輪郭を矩形にした正二十面体展開図の間で比較を行なった。

(2) MPEG-4 によるエンコード、デコード：各展開図に対して、MPEG-4 エンコード、デコードを行なう。

(3) デコード後の展開図の画質評価：デコードした展開図から、ある視野での透視投影映像を生成しその画質を評価する。透視投影映像は 3 次元空間に配置したポリゴンにテクスチャマッピングを行なって生成した。評価のリファレンスとするのは、オリジナルの全方位ビデオから直接生成した透視投影映像である。

なお、上記のステップの中で、展開図上の画素数と MPEG-4 のビットレートは比較対象同士で同一とした。これにより、最後の画質評価で最良のものが、一定のデータ量の中で最も効率的に表現できるものであるといえる。

4.1 実験結果

まず地図投影法と多面体モデルとの比較を行なった。Tab.1(a) に画質評価の結果の PSNR を示す。表中の ppf は 1 フレーム辺りの画素数であり、I~IV は Fig.4 上段に示す 4 種類の視野の透視投影画像について比較した結果である。結果をみると、多面体モデルに比べて地図投影法の画質が著しく低下していることがわかる。これは、2.1 節で述べたように地図投影法では幾何的な歪みが現れるためであると考えられる。Tab.1(a) をみると多面体モデルの中では正 20 面体が最も良い値になっている。

Tab.1(a) の結果は、実映像によるものであるため、透視投影映像を作成する視野の選び方によって結果に偏りが生じる恐れがある。そこで、球面に均一にテクスチャが分布するシミュレーション映像を用い、デコード後の展開図から全ての方位の画素情報を球面上で等間隔に取

Tab.1: Comparison of Mapping Methods

(a) Cartographical vs. Polyhedral (dB)

	4.7Mppf		0.7Mppf	
	I	II	III	IV
<i>Cartographical Mappings</i>				
Equidistant Cylindrical	18.0	13.6	18.7	20.3
Sinusoidal	17.6	13.9	14.9	21.9
<i>Polyhedral Mappings</i>				
Cube	30.9	28.8	27.3	27.6
Octahedron	30.8	28.1	26.8	27.4
Icosahedron	31.3	28.7	28.2	28.6
Icosahedron(Rect.)	29.9	27.9	26.3	26.3

(b) Comparison of Polyhedral Mappings (dB)

	4.7Mppf		0.7Mppf	
	V	VI	V	VI
Cube	39.5	29.7	38.1	28.6
Octahedron	39.6	29.8	37.9	28.2
Icosahedron	40.0	29.8	38.5	28.4
Icosahedron(Rect.)	39.6	29.0	37.9	27.3

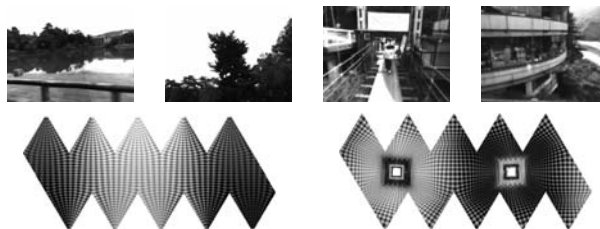


Fig.4: Compared Images

り出して比較を行なうことにした。Tab.1(b) にその結果を示す。表中 V~VI は Fig.4 下段に示す 2 種類のシミュレーション映像の結果を示している。これを見ても、やはり正 20 面体が最も良いという傾向が分かる。

また、Tab.1(b) からは、3 で述べた平面展開図の輪郭を矩形とした効果が見られないことがわかる。全方位ビデオは元々データ量が大きいため、任意形状符号化におけるデータのロスが相対的に少なく、輪郭を矩形としたことによる平面上の解像度の歪みの影響が大きかったものと考えられる。

5 まとめ

本稿では、多面体モデルによる全方位ビデオの圧縮法を提案した。多面体モデルは球面上の全方位ビデオを表現する際の解像度の偏りが少ないため、データ量の効率が良い。また、透視投影映像を生成する際に非線形な演算が必要ないため、受信側でも効率良く高速に全方位ビデオを表示できる。今回の実験では、多面体モデルの中でも正 20 面体が最も効率的であるという結果が得られたが、限られた組合せでの実験であったので、評価データやパラメータの組合せを増やして詳細な実験を行っていくことが今後の課題である。特に、多面体の 3 次元空間内での配置は画質に影響を及ぼすことが考えられるので、詳細に評価を行なう必要がある。

謝辞

本研究は科学研究費補助金基盤研究 A 13308017 の補助を受けて行った。実験に使用したデータは JEITA3DMM 委員会並びに奈良先端科学技術大学院大学横矢研究室から提供されたものである。

参考文献

[1] 樋口他: "球面直交展開を用いたパノラマ全周画像の記述", 3 次元画像コンファレンス '99, pp.31-36, 1999.

京都大学大学院工学研究科電子工学専攻

〒 606-8501 京都市左京区吉田本町

Tel:075-753-7465

e-mail:habe@media.kyoto-u.ac.jp