

多面体表現を用いた全方位ビデオの圧縮

波部 斉^{†a)} 山澤 一誠^{††} 野村 敏男^{†††} 曾良 洋介^{††††}
松山 隆司^{††††}

Compression Method for Omnidirectional Video using Polyhedral Representations

Hitoshi HABE^{†a)}, Kazumasa YAMAZAWA^{††}, Toshio NOMURA^{†††}, Yosuke KATSURA^{††††}, and Takashi MATSUYAMA^{††††}

あらまし

人の住む 3 次元世界をそのまま映像化する 3 次元映像メディアに関する研究が活発に行われているが、大量のデータを効率的に扱い、放送・配信・アーカイブに広く利用するためには、統一されたデータ圧縮手法の確立が不可欠になる。そこでは従来の映像情報に含まれない 3 次元幾何情報を同時に効率よく扱うことが課題となる。本論文では、3 次元映像メディアの一つである全方位ビデオに着目し、多面体表現を利用した圧縮手法を提案する。そこでは、1) 全方位ビデオの多面体表面への投影、2) 多面体の平面への展開、3) 通常の 2 次元平面映像に対する圧縮法の適用、によって、データ表現の効率が良く、既存の技術との親和性の高い圧縮手法を実現する。本論文においては、全方位ビデオの撮影・伝送・表示のシステム全体を考慮したときの提案手法の優位性を明らかにし、実映像およびシミュレーション映像を用いた実験によって、既存の全方位ビデオの表現法との定量的な比較を行う。さらに、効率良い表現のための多面体の選択指針を明らかにする。

キーワード 3 次元映像メディア、全方位ビデオ、圧縮、符号化、多面体表現

1. ま え が き

映像メディアは人間が視覚を通じて得る情報を、伝送・蓄積するための情報メディアである。映画やテレビなどによって我々は多くの情報を得ることができるが、それらは我々の住む 3 次元世界を平面で捉えたものであるため、そこから得られる情報と 3 次元世界そのものとの間には大きなギャップが存在している。それらは臨場感・写実性の欠如として顕在化してくる。それに対し、人間の住む 3 次元世界をそのまま忠

実に映像化して、エンターテイメント、教育用コンテンツ、文化資産のデジタルアーカイブなどへ応用する研究が盛んに行われている [11]。それら 3 次元映像メディアは新たな視覚情報メディアとして期待されているが、従来の映像情報に含まれない 3 次元幾何情報を同時に取り扱う必要があるため、その獲得・伝送・提示に新たな枠組が必要になる。

3 次元映像メディアとして、様々なものが提案されているが、本論文ではある視点を中心とする周囲の状況を一度に撮影した全方位ビデオ [17] に着目する。全方位ビデオは、自律動作するロボットのナビゲーション [16] や、広範囲の監視システム [14] に用いられており、近年では仮想現実感技術への応用も盛んである [12]。これらのアプリケーションは全方位ビデオがカメラの周囲の 3 次元世界を余すところなく記録しているという特徴を利用している。

昨今では、ハイビジョンカメラや IEEE1394 デジタルビデオカメラなどにより、高精細な全方位ビデオが実時間で撮影可能になっている。その結果、より臨場感の高い映像が得られ、アプリケーション領域のさらなる拡大が期待されるが、データ量が膨大になるため

[†] 京都大学大学院工学研究科, 京都市
Graduate School of Engineering, Kyoto University,
Yoshida-Hommachi, Sakyo-ku, Kyoto-shi, 606-8501 Japan
^{††} 奈良先端科学技術大学院大学情報科学研究科, 生駒市
Graduate School of Information Science, Nara Institute of
Science and Technology, Takayama 8916-5, Ikoma-shi, 630-
0192 Japan
^{†††} シャープ株式会社, 千葉市
SHARP Corporation, Nakase 1-9-2, Mihama-ku, Chiba-shi,
261-8520 Japan
^{††††} 京都大学大学院情報学研究科, 京都市
Graduate School of Informatics, Kyoto University, Yoshida-
Hommachi, Sakyo-ku, Kyoto-shi, 606-8501 Japan
a) E-mail: habe@media.kyoto-u.ac.jp

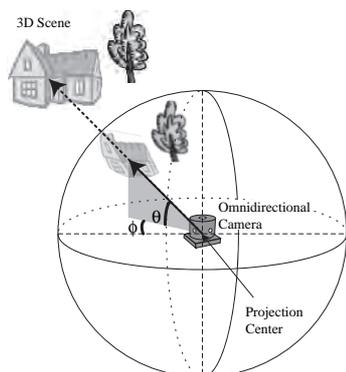


図 1 全方位ビデオの球面上へのマッピング
Fig. 1 Mapping omnidirectional video onto a sphere

その圧縮が大きな課題となる。特に、ネットワーク伝送やアーカイブでは効率的な圧縮手法が不可欠である。国際標準規格においても、近い将来の全方位ビデオ放送サービスなどの実現・普及を念頭に置き、全方位ビデオ圧縮手法の標準化に向けた取り組みが行われている [7], [15]。

全方位ビデオは、ある視点を中心とした周囲の映像を一度に撮影したものであるため、視点を中心とする球面上に視覚情報がマッピングされたものと考えることができる。このように幾何的特徴を併せ持った映像メディアを効果的に圧縮するためには、従来の 2 次元映像の圧縮とは異なり、幾何情報を取り扱う手法そのものの効率と、他の処理との親和性が非常に重要になる。すなわち、撮影時の幾何変換・情報圧縮・鑑賞時の幾何変換など全方位ビデオに対するすべての処理を含めたシステムとして、圧縮方式の議論を行う必要がある。例えば、世界地図の作成法と同様に、球面上の映像データを図 4 (a) のように平面展開して圧縮すれば、平面の上下で不要なデータが多くなり効率が悪い。また、全方位ビデオを効率的に表現する手法として、球面上に六角形の画素を考えることで直線検出やオプティカルフロー計算を効率的に行う手法 [9], [10] や、球面調和関数や球面ウェーブレットを用いて画像データを直交展開して効率的に記述するもの [5], [6] が提案されているが、2. に述べるように、伝送された映像を鑑賞する際の処理との親和性が低く、システムとして考えたときの効率が悪いといえる。

このような観点から、本論文では多面体表現を用いた全方位ビデオの圧縮手法を提案する。本手法におい

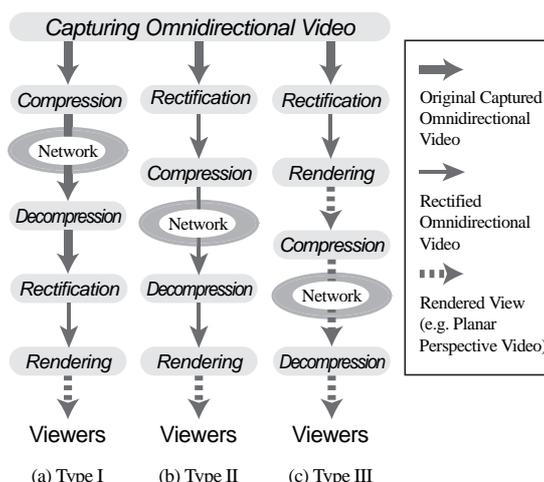


図 2 全方位ビデオの撮影・伝送・表示
Fig. 2 Capturing, transmission and rendering of omnidirectional video

ては、まず、球面上にマッピングされた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くことで平面展開図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圧縮手法を適用する。この手法は、1) 全方位ビデオのもつ視覚情報を表現するときのデータ量の効率が良い、2) デコーダ側で任意視点の映像を生成するときの幾何歪みが少なく、その処理も簡便である、3) 既存の映像コーデックが利用できることで汎用性が高い、という特徴をもち、実用面でも非常に有用な手法であると期待される。

本手法の位置付けと有効性を明確にするために、本論文においては、まず 2. において全方位ビデオの圧縮手法が満たすべき要件を整理し、3. において提案手法について詳細に論じる。次いで、提案手法の有効性を検証するための評価実験を 4. で述べる。最後に 5. で本論文のまとめと今後の課題について述べる。

2. 全方位ビデオとその圧縮

本節では全方位ビデオの特徴を示し、その圧縮に求められる要件を整理する。

2.1 全方位ビデオの撮影・伝送・表示

全方位ビデオとは、ある視点から見た全周囲の視覚情報を同時に観測した映像である。そこには奥行き情報は含まれないので、視線方向 ϕ, θ を決定すれば対応する視覚情報（画素値）は一意に定まる。すなわち、全方位ビデオの視覚情報は任意の半径の球面状にマッピングすることが可能である（図 1）。

全方位ビデオの撮影から表示までに必要な処理とその構成は図 2 のようになる。ここでは機能を、全方位ビデオの撮影 (Capturing)、魚眼レンズ [2]、曲面ミラー [18]、多眼カメラ [4] などといった撮影方法の違いを正規化し、標準的な形式の全方位ビデオを作成する処理 (Rectification)、圧縮、伸長 (Compression/Decompression)、観賞者が要求する形式での映像描画 (Rendering) に分け、考えられる構成として Type I, II, III の 3 種類を示している^(注1)。我々はこの中で Type II のシステム構成に注目する。この構成は以下の特徴をもつ。

1) 観賞者側の処理の汎用性が高い。全方位ビデオを撮影するカメラには様々な種類のものが存在する。撮影した映像に必要な幾何変換は個々のカメラシステムに依存しており、複雑な演算になることが多い。Type I ではこれらの演算を観賞者側で行う必要があるが、Type II, III ではその必要がなく、観賞者側は汎用的で簡便な処理でよいので広く一般に普及させるときに有利である。

2) 臨場感の高い映像を提供できる。1. で述べたとおり、全方位ビデオの特徴として、あたかもその場にいるような感覚を与えることが挙げられる。その効果を発揮する没入型ディスプレイ [3], [13] に表示する際には観賞者側で全方位ビデオのすべての情報を保持する必要がある。さらに、通常の平面ディスプレイに表示する場合でも、臨場感を高めるためには観賞者が行う視線方向変更の要求に対する応答を可能な限り速く行う必要がある。Type III では視線方向変更などの要求を撮影・送信側で処理するために時間遅れは避けられないが、Type I, II では速やかな応答が可能となる。また、システムのスケラビリティを考えたときには Type III では受信者の数が増えると破綻する可能性が高いが、Type I, II は柔軟に対応可能である。

2.2 全方位ビデオ圧縮手法の要件

Type II のシステムにおいて、圧縮手法に必要な条件は以下ようになる。

a) 全方位の視覚情報を効率的に表現すること

Type II では、全方位ビデオのもつすべての情報を圧縮する必要がある。そこでは、球面上にマッピングされた全方位ビデオの視覚情報を効率的に表現した手法でなければ、データ量に対する情報量の効率が悪く

なる。特に、高精細な全方位ビデオではデータ量が膨大になるためこの点が重要になる。

b) 鑑賞者側の描画処理との親和性が高いこと

前節で述べたインタラクティブな映像表示の実現には、グラフィック技術を用いた描画処理を行うことになる。十分なインタラクティブ性を実現するためには、グラフィックスハードウェアでの高速描画が容易なデータ構造である必要がある。

2.3 圧縮のためのアプローチ

以上の考察をふまえ、全方位ビデオの圧縮手法を議論する。そのアプローチとしては、以下の 3 つが考えられる。

(1) カメラで撮影したそのままのデータに既存の映像圧縮手法を適用する。

(2) 球面上のデータをそのまま直交展開して圧縮する。

(3) 球面上のデータを平面上のデータに変換し、平面上のデータに対し既存の映像圧縮手法を適用する。

(1) は図 2 に示した Type I に相当し、2.1 で述べたとおり、ここでは適切でないとと言える。

(2) によるアプローチは球面上のデータを記述する手法として自然な発想である。樋口らは球面調和関数や Haar 変換を用いた球面ウェーブレットを利用して全方位画像を記述している [5], [6] が、この拡張としての圧縮法を考えたとき、圧縮・伸長アルゴリズムが独自のものとなる点が、実用面で問題となると考えられる。また、鑑賞者が要求する映像を描画する処理を考えると、任意の視線方向の映像情報 (画素値) を抽出するための演算が複雑になるため計算コストが高く、インタラクティブな表示には不向きである。さらに、4.1 の実験結果が示すように (3) の考え方による我々の提案手法との効率の差も少ないため、これらの手法を積極的に採用する必要はないと考えられる。

それに対し、後述するように (3) のアプローチは 2.2 で述べた要件を満たすことができ、全方位ビデオに最適な圧縮手法であると考えられる。次節でその詳細を述べる。

3. 提案手法

2. での議論を踏まえ、本論文で提案する多面体表現を用いた全方位ビデオの圧縮手法について述べる (図 3)。提案手法は大きく分けて、1) 全方位ビデオの多面体表面への投影と平面上への変換、2) 変換された平面ビデオに対する既存の圧縮手法の適用、の 2 ステッ

(注1): ここでは全方位ビデオの伝送システムを念頭に置いているが、ネットワーク伝送の部分をアーカイブに置き換えれば、蓄積システムに対しても同様の議論が可能である。

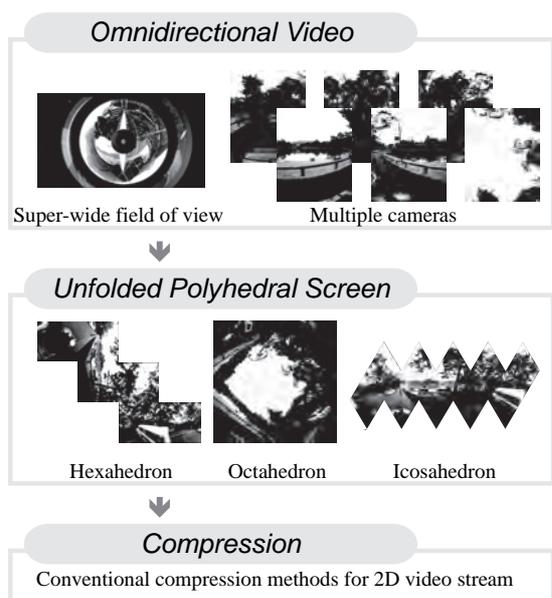


図 3 全方位ビデオの圧縮手法

Fig. 3 Compression method for omnidirectional video

ブから構成される。以下では、全方位ビデオを平面ビデオに変換する手法の中での多面体表現の優位性と、多面体表現の中の優劣について論じる。

3.1 地図投影法による全方位ビデオの表現

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、そのスクリーンを平面に展開する変換と考えることができる(図4)。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。図4(a) (b)にその一例を示す[1]。

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2.2で述べたとおり、1) 同一データ量で表現できる情報量、2) 鑑賞者側での映像描画、の2点での効率の良さが求められる。1)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2)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。

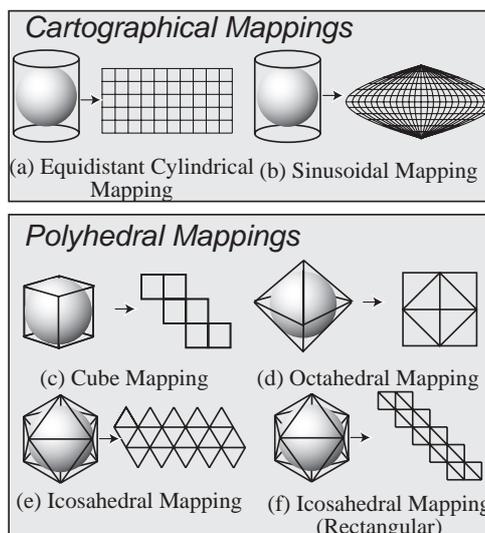


図 4 全方位ビデオの平面への変換

Fig. 4 Transform omnidirectional video onto a planar screen

3.2 全方位ビデオの多面体表現

これに対し、我々は多面体表現(図4(c)~(f))の適用を提案する。

3.2.1 多面体表現の利点と計算コスト

多面体は球をよく近似するため1)に関して効率が良い。2)についても多面体の各頂点の平面展開図上の座標と3次元空間中の座標の関係をあらかじめ与えておけば、他の対応関係は線形補間で正確に計算できる。このように全方位ビデオデータを表現すれば、汎用PCのグラフィック処理ライブラリで、リアルタイムに任意の視点からの映像を生成可能である。よって、2)についても要件を満たしているといえる。さらに、多面体展開図の作成を行う際にも、展開図上のある点に相当する方位 ϕ, θ の計算が線形補間で求まるので、非常に高速な演算が可能である。これらの処理時間は平面上に展開したビデオ映像に対して行うMPEG-4などの圧縮処理に比べて十分短い。

3.2.2 多面体や平面展開図の形状と圧縮効率

多面体表現には各面の形状および平面展開の手法によって様々な方法が考えられる。以下でそれらが圧縮効率に及ぼす影響を議論する。

図5は球面から多面体のある面への投影を2次的に示している。球面上で同じ角度 α に相当する領域A, Bでも多面体表面に投影したときの面積 a, b が多面体表面の接線となす角 θ に依存して変化することが

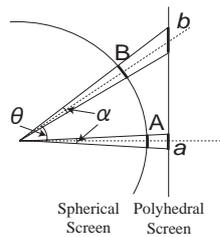


図 5 球面から多面体表面への投影

Fig. 5 Projection from spherical surface onto polyhedral surface

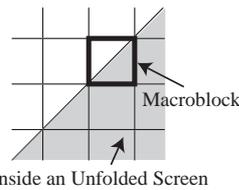


図 6 マクロブロックと展開した多面体の関係

Fig. 6 Macroblock and unfolded polyhedral screen

分かる．その間には近似的に $b/a = 1/\cos^2\theta$ の関係が成立する（詳細な導出は付録 1. 参照）ので θ が大きくなるに従い，多面体表面の 1 画素に対応する球面上の面積が減少していく．よって， θ の最大値が大きい多面体では，その表面の 1 画素に対応する球面上の面積が大きく変化するため分解能にばらつきが生じる．そのとき，多面体表面の接線方向で十分な分解能を確保しようとするれば θ が大きな領域で必要以上のデータが必要になる．よって， θ の最大値が小さい，面数が多い多面体のデータ効率が良いことが分かる．データ効率については 4.3 においてシミュレーション結果を示す．

次に，平面ビデオに変換した後の既存の圧縮手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表現が望まれる．一般に，多面体の面数が多くなるほどその展開図は複雑な形状になるため，以上 2 つの条件はトレードオフの関係にある．

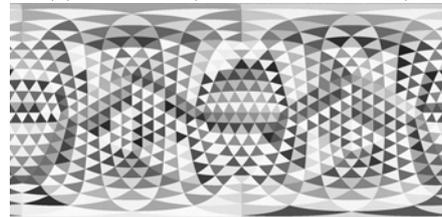
また，MPEG に代表される 2 次元映像の圧縮手法は矩形の映像を前提としているが，多面体表現を用いた場合，その展開図の輪郭は必ずしも矩形にはならない（図 4 (e))．MPEG-4 の任意形状オブジェクト符号化を利用すれば輪郭が矩形でない展開図にも対応で



(a) Data Set 1 (15fps, About 4.7Mppf)



(b) Data Set 2 (30fps, About 0.7Mppf)



(c) Data Set 3 (15fps, Synthetic Data)

図 7 実験に使用した全方位映像（正距円筒図法で表示）
Fig. 7 Test data sets (Depicted with the Equidistant-cylindrical projection)

きるが，図 6 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（図 4 (f))．この場合には正多面体でなくなるために先述の分解能の偏りが増大し，効率低下が発生するトレードオフが考えられる．

以上をまとめると，多面体表現を用いた場合の圧縮効率を決定する要因には，1) 球面に対する分解能の偏り（面数を増やすほど良い），2) 平面展開図の形状（切れ目が少なく，輪郭が矩形の方が良い）がある．これらの相互の影響を評価するために，4. の評価実験では図 4 (c) ~ (f) に示す 4 種類の表現方法を比較する (c) は正六面体であり (d) は正八面体の各面を変形させ正方形に展開したもの (e) は正二十面体 (f) は正二十面体の各面を変形させ輪郭が矩形になるように平面展開したものである．

4. 評価実験

本節では、評価実験によってこれまでに述べた議論を確認する。評価実験に用いる映像として、実環境で撮影した2種類の解像度が異なる全方位ビデオデータと、シミュレーションによって生成したデータを用意した。その例を図7に示す。Data Set 1は池の周囲を走行する車両の上に設置した多眼式の全方位カメラで撮影した映像であり、Data Set 2は橋の上を歩行する人が持った曲面ミラー式の全方位カメラで撮影したものである。Data Set 3は空間的に一様な模様の中に設置された全方位カメラが回転している様子を模擬している。なお、図中のppfは1フレーム当りの画素数を示している。それぞれのデータの時間長は10秒間に統一した。

4.1 実験1—多面体表現と球面直交展開法の比較

まず、過去に全方位ビデオの表現手法として提案されている、球面調和関数や球面ウェーブレットを用いて、球面上の映像情報を直接直交展開して表現する手法[5],[6]と本論文の提案手法との比較を行う。2.3で述べたように、球面直交展開による手法は鑑賞者側での描画処理の計算コストが高く、全方位ビデオの特徴を生かしたインタラクティブな表示を行う際に不利になる。従って、そもそも球面直交展開法は全方位ビデオの圧縮手法として不相当であると考えられるが、ここでは球面直交展開法の表現効率を我々が提案する多面体表現と比較し、表現効率の意味からも球面直交展開法を積極的に採用する必要がないことを示す。この趣旨および、先行研究で提案されている球面直交展開法が1フレームの全方位画像を表現するものであることから、ここでの比較は1フレームの全方位画像データを用いることとした。

4.1.1 球面調和関数による直交展開との比較

Data Set 1および3を利用して、球面調和関数を利用した表現法と多面体表現との比較を行う。球面調和関数によって全方位ビデオ $f(\theta, \phi)$ は、

$$f(\theta, \phi) = \sum_{n=0}^{n_{\max}} \sum_{m=-n}^n Y_{nm}(\theta, \phi), \quad (1)$$

$$Y_{nm}(\theta, \phi) = \begin{cases} A_{nm} \cos m\phi P_n^m(\cos \theta), \\ B_{n|m|} \sin |m|\phi P_n^{|m|}(\cos \theta), \end{cases} \quad (2)$$

と直交展開される[5],[6]。ここで $P_n^m(x)$ は Legendre

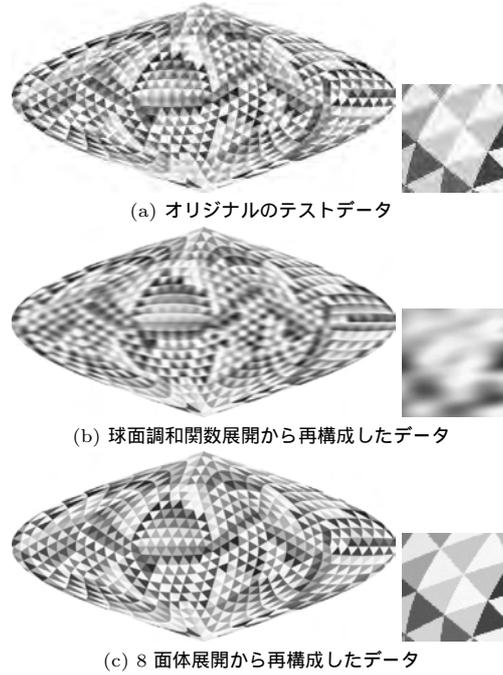


図8 球面調和関数による直交展開との比較

Fig. 8 Comparison to spherical harmonics functions

表1 球面調和関数による直交展開との画質比較 (dB)
Table 1 Comparison to spherical harmonic functions

	Data Set 1	Data Set 3
球面調和関数 (b)	22.1	18.7
八面体 (c)	33.6	30.0

陪関数である。Data Set 1 および 3^(注2)を経度方向に1100点、緯度方向に550点サンプリングしたデータから上式を用いて直交展開 ($n_{\max} = 150$) し、再度全方位画像に再構成した。Data Set 3を再構成したものを図8(b)に示す。図はSinusoidal 図法(図4(b))を用いて示しておりサイズは1000 × 500である。

比較対象として、八面体展開(図4(d))を行ったあとで再構成した全方位画像を(c)に示す。八面体展開図の画素数は球面調和関数展開に利用したサンプリングデータの数とほぼ同数となるようにした。

これらの結果をみると、球面調和関数展開を行った場合には詳細なエッジが再現できていないことが分かる。元のデータから直接生成した全方位画像(a)との間で(b)と(c)のPSNRを求めたところ、表1の

(注2): Data Set 2は有効視野が水平方向に限られているため、実験1ではData Set 1と3を使用することとした。

表 2 球面ウェーブレットによる直交展開との画質比較 (dB)

Table 2 Comparison to spherical wavelet functions

圧縮率	Data Set 1		Data Set 3	
	10%	50%	10%	50%
球面ウェーブレット	25.6	29.4	17.8	25.4
多面体 + DCT	26.9	31.4	24.3	26.9

ような結果が得られた．定量的に画質の差は明らかである．これは球面調和関数展開を行う際の計算誤差に由来するものであり，文献 [6] でもその問題点が指摘されている．計算誤差は計算法を改良して解決できる可能性があるが，同時に計算コストが高くなる．よって，2.3 で述べたようにインタラクティブな描画処理との親和性が低くなり，全方位ビデオの圧縮手法として不適當であると考えられる．

4.1.2 球面ウェーブレットによる直交展開との比較

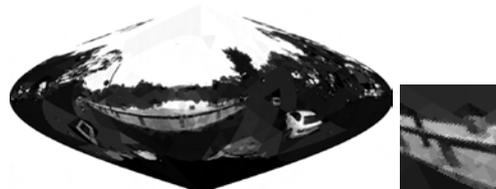
次に，球面ウェーブレットを用いた表現との比較を行う．ここでは，球面を測地ドームで近似し，球面 Haar 基底を用いた手法 [6] を比較対象とする．評価方法は以下のとおりである．

- (1) 正二十面体から作成したレベル 7 の測地ドーム (面数 163842) で Data Set 1 および 3 の中の 1 フレームをサンプリングしたデータに対してウェーブレット展開を行い，スケーリング関数とウェーブレット関数の係数を得る．
- (2) ウェーブレット関数の係数のうち，有意な情報を持たないものを削除し，データ量を圧縮する．
- (3) 圧縮後のデータから全方位画像を再構成する．
- (4) 元データから八面体展開 (図 4 (d)) を行い平面展開図を得る．平面展開図の画素数はウェーブレット変換に用いたデータ数とほぼ同数にする．
- (5) 展開図から得られた DCT 係数から有意な係数を抽出し，データ量を圧縮する．
- (6) 圧縮した DCT 係数から全方位画像を再構成する．
- (7) (3) で得られた画像と (6) で得られた画像のオリジナルデータに対する画質劣化を PSNR で評価する．

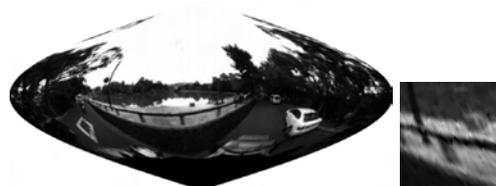
通常の映像圧縮手法ではこのような係数の展開・圧縮以外にも様々な処理が行われるが，4.1 の冒頭に述べたように，球面ウェーブレットの表現効率を多面体表現と比較することがここでの趣旨であるため，このような簡便な手法で比較を行った．



(a) オリジナルのテストデータ



(b) 球面ウェーブレット展開後に再構成したデータ (10% 圧縮)



(c) 8 面体表現から DCT 展開し再構成したデータ (10% 圧縮)

図 9 球面ウェーブレットと 8 面体表現から再構成された全方位ビデオの例

Fig. 9 Reconstructed images from spherical wavelet functions and octahedral representation

評価結果を表 2 に示し，圧縮率 10% のときの画像を図 9 に示す．この結果をみると，球面ウェーブレットと多面体表現に大きな差は認められない．先述のとおり，この実験では映像圧縮に不可欠なフレーム間の冗長度削減も考慮していないため，この実験のみで優劣を結論付けることはできない．しかし，2.3 で述べたように，多面体表現には 1) 描画側での処理が高速・簡便になる，2) 既存の映像コーデックをそのまま利用できる，という利点がある．これと本節の結果を考えると，多面体表現が球面ウェーブレットに比べて十分有用な手法であると言える．

4.2 実験 2—多面体表現と地図投影法の比較

実験 2 では 3.1 に述べた多面体表現と地図投影法を比較する．評価手法は図 10 に示すように以下の 3 ステップから構成される．

- (1) 各展開方法による展開図の作成

実験 2 では Data Set 1 と Data Set 2 を使い，これまでに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 θ ，経度 ϕ を平面上の x, y 座標に対応づける正距円筒図法 (図 4 (a)) と正距円筒図法の経度方向のサンプリング間隔を緯度

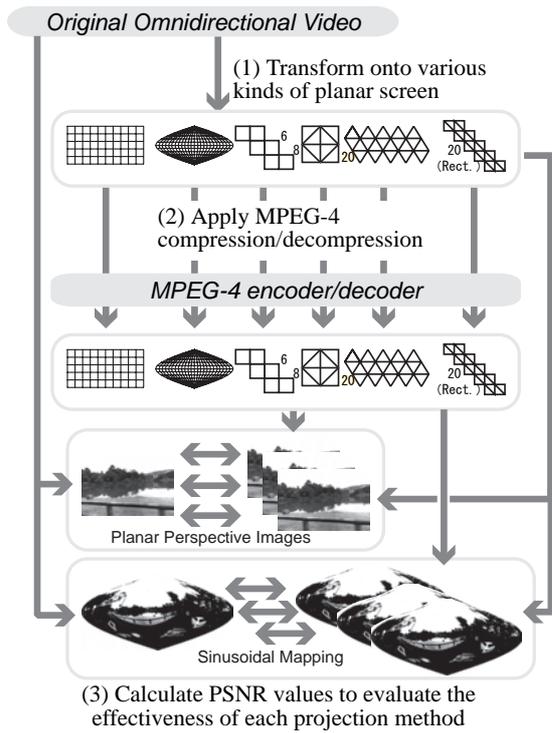


図 10 実験 2 と 3 で用いた実験手法
Fig. 10 Experimental Method 2 and 3

の余弦 $\cos \theta$ に反比例させることで球面上で均等にサンプリングする Sinusoidal 図法 (図 4 (b)) に着目し、図 4 の (c) ~ (f) で示す多面体表現と比較した。

(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、デコードを行う。任意形状オブジェクトを扱えるように、MPEG-4 コアプロファイルを用いた。

ここまでのステップの中で、展開図上の画素数は高解像度・低解像度の 2 通りのものを作成し、それらを一定のビットレートになるように MPEG-4 エンコードした。これにより、圧縮後のデータ量を一定とする条件の中での、展開図上の画素数や展開方法の画質への影響が比較可能となる。

(3) 画質評価

上記の手順で生成された各画像の画質評価を行う。画質評価の際には、元の全方位ビデオデータをリファレンスとし、各々の展開手法について、上記ステップ (1) で得られた展開図、もしくはステップ (2) の結果で得られたデコード画像の画質を評価した。これにより、展開方法そのものによる画質劣化と MPEG エ

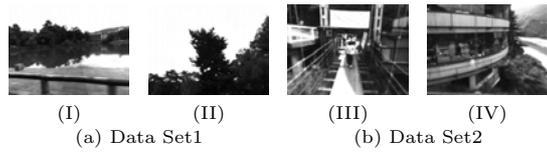


図 11 画質評価に使用した平面透視投影映像
Fig. 11 Planar perspective images for quality evaluation

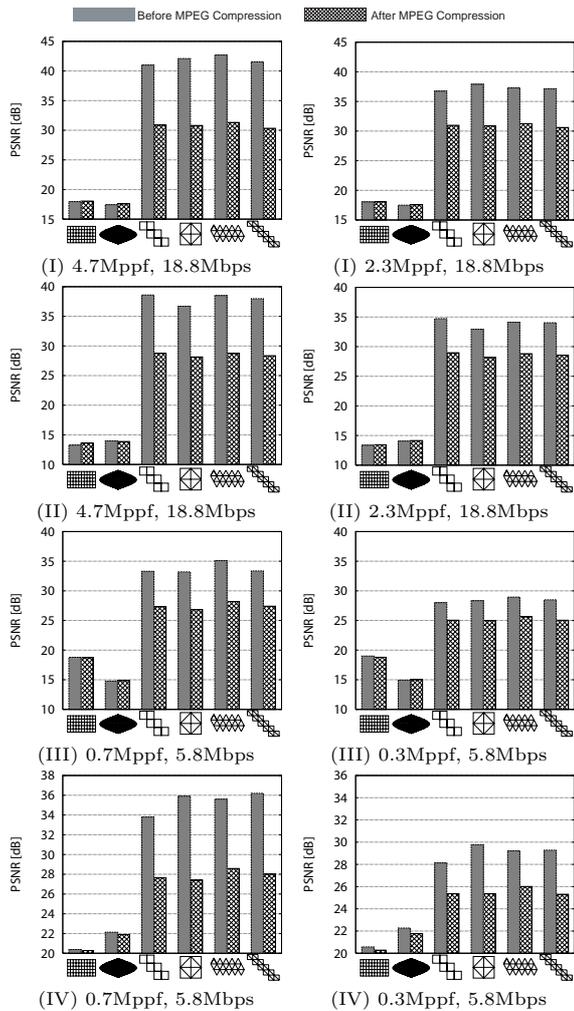


図 12 評価実験 2 の結果 — 多面体表現と地図投影法の比較

Fig. 12 Experimental results 2 — Comparison between polyhedral model and cartographical mapping

ンコードによる画質劣化を詳細に評価できる。

評価の際には全方位ビデオから図 11 に示すような、特定視野の平面透視投影映像を再生成し、その画質を

表 3 評価実験 3 の結果 (1) — 多面体表現のデータ効率の比較

Table 3 Experimental results 3(1) — Comparison of data efficiency of polyhedral mappings

手法	正六面体	八面体	二十面体	二十面体 (矩形)
標準偏差	2223	1999	1040	1050

PSNR で評価する手法を採用した。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って生成した。

実験 2 の結果を図 12 に示す。図中の (I) ~ (IV) はそれぞれ図 11 の透視投影映像を用いて行った評価結果、ppf は展開図の 1 フレームあたりの画素数、bps は MPEG 圧縮のビットレートを示す。また、図中にはステップ (1) で得られた MPEG 圧縮前の映像を使って評価した結果とステップ (2) で得られた MPEG 圧縮後の映像を使って評価した結果を並べて示している。

結果をみると、多面体表現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、3.1 で述べたように、有限個のパッチ に対するテクスチャマッピングを行った場合、地図投影法では幾何的な歪みが見られるためである。このように、地図投影法に比べて多面体表現が有効であることが確認できた。また、例外もあるが多面体表現の中では正 20 面体の効率が最も良い傾向を示している。

4.3 実験 3—多面体表現同士の比較

次に、多面体表現の中でどの表現手法が良いかを検証する実験を行った。比較対象は実験 2 と同じ 4 種類の多面体表現手法である。

まず、3.2 で議論した、球面上の映像情報に対する分解能の偏りをシミュレーションで定量的に評価する。シミュレーションでは、球面上に隙間なく微小領域を設定し、その領域を表現するために割り当てられる平面展開図上の画素数を求めた。分解能の偏りがなく、この画素数は球面全体で等しくなる。具体的手法は以下ようになる、1) 正 20 面体を再帰分割して得られる測地ドームの各面を球面上の微小領域とする、2) それぞれの手法の展開図上の画素を再度球面へ投影する^(注3)、3) 測地ドームの各面毎に投影された画素数を求め、その標準偏差を計算する。

このシミュレーションの結果を表 3 に示す。この結果は 3.2 で述べた MPEG 圧縮の影響を含まない平面変換の効率を表している。結果からは 3.2 の議論のと

(注3) : 各展開図上の画素数は 4.7M ピクセルにほぼ等しくなるようにした

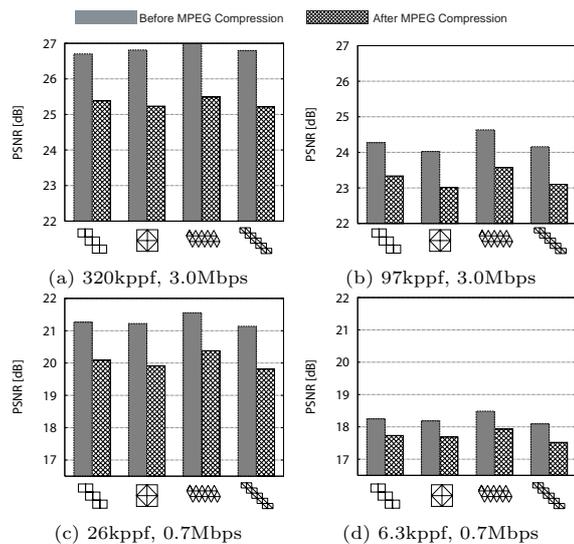


図 13 評価実験 3 の結果 (2) — 多面体表現同士の比較

Fig. 13 Experimental results 3(2) — Comparison of Polyhedral Mappings

おり、多面体の面数が多い二十面体が最も偏りが少なく、また、二十面体でも輪郭を矩形にする変換を行うことで偏りが増していることが分かる。

次に実験 2 と同様の画質比較を行う。手法は基本的に実験 2 と同じであるが、全方位ビデオデータとして Data Set 3 を用いた。実映像は方位によってテクスチャが異なるために画質への影響に方位依存性が生じるが、Data Set 3 は一様な模様としているためにより方位依存性のない評価が可能となる。さらに、実験 2 では透視投影映像を作成する際の視野の選択が結果に影響を及ぼすことも考えられるので、全方位ビデオから空間的に均等な間隔で映像情報をサンプリングし、それを比較に用いる手法を採用した。全方位ビデオからの均等な間隔のサンプリングは、図 4(b) に示す Sinusoidal 図法を再生成することと同等になる。したがって、この手法は各展開図から再生成した Sinusoidal 図法による展開図を比較することと等価である。

図 13 に実験 3 の結果を示す。Data Set 3 は任意の解像度とできるが、ここでは 4 種類の解像度のデータを用いた。実験 2 と同様に 1 つの MPEG-4 ビットレートに対して、2 種類の解像度の展開図を用い、それらの画質の差を比較することとした。

結果をみるとどの解像度でも実験 2 と同様に正二十面体が最も画質が良くなっており、概ね正二十面体、

正六面体，二十面体（矩形），八面体の順になっている傾向が見られる。しかし，正六面体以降は差が少なく，逆転する場合もある。また，平面展開図の画素数を大きくした方が画質が良いことや，MPEG 圧縮の前後で画質の優劣の傾向が変化しないことがわかる。

この結果では，平面展開図の輪郭を矩形としたことや，平面展開図の画像端での物体の出現・消失を少なくしたことの効果が見られない。すなわち，3.2 で MPEG 圧縮との親和性を考慮して導入した展開手法の効果がないことがわかる。全方位ビデオは元々データ量が大きいため，これらの要因によるデータのロスが相対的に少なく，表 3 に示した，球面上の映像情報に対する分解能の偏りの影響が大きかったものと考えられる。すなわち，多面体の選択においては，分解能の偏りを少なくすることが最も重要であることがわかる。

5. むすび

本論文では，多面体表現による全方位ビデオの圧縮手法を提案した。多面体表現は球面上の全方位ビデオを表現する際の分解能の偏りが少ないため，同一のデータ量で表現できる情報量が多くなる。また，鑑賞者側で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。さらに，通常の MPEG などの動画圧縮手法との親和性が高いことも大きな特徴である。今回の実験では，多面体表現の中で正二十面体が最も効率的であるという結果が得られた。

これらの結果を踏まえ，我々は提案手法を動画圧縮手法の国際標準を審議する ISO/IEC JTC1/SC29/WG11(MPEG) に提案し，国際標準化に向けての活動を行っている [8]。

今後の課題としては，鑑賞者の注目領域の変化などに応じて，領域毎の圧縮率を動的に制御可能な方式の実現などがあげられる。

謝辞

本研究の一部は科学研究費補助金特定領域研究 13224051，文部科学省「知的資産の電子的な保存・活用を支援するソフトウェア基盤技術の構築」事業における「大型有形・無形文化財の高精度デジタル化ソフトウェアの開発」プロジェクトの補助を受けて行った。実験に使用したデータおよび評価ソフトの一部は JEITA3DMM 委員会並びに奈良先端科学技術大学院大横矢研究室から提供されたものである。本研究に関

して貴重なご意見をいただいた JEITA3DMM 委員会の委員各位に感謝する。

文 献

- [1] 地図学用語辞典（増補改訂版），日本国際地図学会地図用語専門部会（編），技報堂出版，東京，1998。
- [2] Z. L. Cao, S. J. Oh, E. L. Hall, "Dynamic Omnidirectional Vision for Mobile Robots," J. Robotics Systems, vol.3, no.1, pp.5-17, 1986.
- [3] C. Cruzz-Neira, D. J. Sandin, T. A. DeFanti, "Surround-screen Projection-based Virtual Reality: The Design and Implementation of the CAVE," Proc. ACM SIGGRAPH '93, pp.135-142, 1993.
- [4] C. Fermuller, Y. Aloimonos, P. Baker, R. Pless, J. Neumann, B. Stuart, "Multicamera Networks: Eyes from Eyes," Proc. IEEE Workshop on Omnidirectional Vision, pp.11-18, 2000.
- [5] 樋口正明, 高野孝英, 苗村健, 原島博, "パノラマ全周画像の直交展開に関する基礎検討," 映像メディア処理シンポジウム (IMPS'98), I-1.14, pp.29-30, 1998。
- [6] 樋口正明, 高野孝英, 苗村健, 原島博, "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, 1-6, pp.31-36, 1999。
- [7] ISO/IEC JTC1/SC29/WG11(MPEG), "Applications and Requirements for 3DAV," WG11 document N5326, Awaaji, Japan, 2002.12.
- [8] ISO/IEC JTC1/SC29/WG11(MPEG), "Report on 3DAV Exploration," WG11 document N5878, Trondheim, Norway, 2003.7.
- [9] 木室義彦, 長田正, "球面六角形画像ピラミッドを用いた球面投影による全方位画像処理," 日本ロボット学会誌, vol.14, no.6, pp.809-819, 1996。
- [10] 木室義彦, 長田正, "球面画像上でのオプティカルフロー計算," 信学論, vol.J79-D-II, no.11, pp.1862-1870, 1996。
- [11] T.Matsuyama, X.Wu, T.Takai, S.Nobuhara, "Real-Time 3D Shape Reconstruction, Dynamic 3D Mesh Deformation, and High Fidelity Visualization for 3D Video," Computer Vision and Image Understanding, vol.96, no.3, pp. 393-434, 2004.12.
- [12] Y.Onoe, K.Yamazawa, H.Takemura, N.Yokoya, "Telepresence by real-time view-dependent image generation from omnidirectional video streams," Computer Vision and Image Understanding, vol.71, no.2, pp.154-165, 1998.
- [13] 柴野伸之, 畑中智行, 中西弘泰, 星野洋, 長濱龍一郎, 澤田一哉, 野村淳二, "都市環境ヒューマンメディアにおける球面型没入ディスプレイ," 日本バーチャルリアリティ学会論文誌, vol.4, no.3, pp.549-554,1999.
- [14] 十河卓司, 石黒浩, モーハン M. トリベティ, "複数の全方位視覚センサによる実時間人間追跡システム," 信学論, vol.J83-D-II, no.12, pp.2567-2577, 2000。
- [15] A.Smolic, D.McCutchen, "3DAV Exploration of Video-Based Rendering Technology in MPEG," IEEE Trans. Circuits and Systems for Video Tech-

- nology, vol.14, no.3, pp.348-356, 2004.3.
- [16] Y.Yagi, W.Nishii, K.Yamazawa, M.Yachida, "Stabilization for Mobile Robot by using Omnidirectional Optical Flow," Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp. 618-625, 1996.
- [17] 八木康史, 横矢直和, "全方位ビジョン: センサ開発と応用の最新動向," 情処学論: コンピュータビジョンとイメージメディア, pp.1-18, 2001.
- [18] 山澤一誠, 八木康史, 谷内田正彦, "移動ロボットのナビゲーションのための全方位視覚系 HyperOmni Vision の提案," 信学論, vol.J79-D-II, vol.5, pp.698-707, 1996.

付 録

1. 球面の分解能の変化

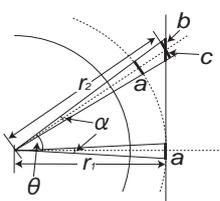


図 A.1 球面から多面体への投影 (2)

Fig. A.1 Projection from spherical surface onto polyhedral surface(2)

図 A.1 に示すように球の中心から多面体平面までの距離 r_1, r_2 をとると $r_1/r_2 = \cos \theta$ であるので, 図中の長さ c と a の間には $a/c = \cos \theta$ が成立する. ここで α が十分小さいとすると $c/b = \cos \theta$ とおける. 以上より $b/a = 1/\cos^2 \theta$ が成立する.

(平成 xx 年 xx 月 xx 日受付)

波部 齊 (正員)

1997 京大・工・電気工学第二卒. 1999 同大学院・工・電子通信工学専攻修士過程修了. 同年三菱電機(株)入社. 2002 京大大学院情報メディアセンター助手. 現在, 同大学院工学研究科電子工学専攻助手. コンピュータビジョン, 3次元映像メディア

の撮影・伝送・表示に関する研究に従事. IEEE Computer Society, 情報処理学会会員

山澤 一誠 (正員)

1992 年阪大・基礎工・情報工卒. 1996 年同大学院博士後期課程中退. 同年奈良先端科学技術大学院大学・情報科学研究科

助手. 2002 年奈良先端科学技術大学院大学・情報科学研究科助教授. コンピュータビジョン, 複合現実感の研究に従事. 1997 年信学会論文賞受賞. 2002 年信学会学術奨励賞受賞. 博士(工学). 情報処理学会, 日本バーチャルリアリティ学会, IEEE 各会員.

野村 敏男

1991 年, 京都大学大学院工学研究科電子工学専攻修士課程修了. 同年, シャープ(株)入社. 以来, 画像圧縮・画像処理に関する研究に従事. 1997-1998 年, 米カリフォルニア大学パークレー校客員研究員.

曾良 洋介 (学生員)

2003 年京大・工・電気電子工学科卒業. 2005 年同大学院情報学研究科修士課程修了. 現在, 同博士後期課程在学中. コンピュータビジョン, 3次元ビデオの研究に従事.

松山 隆司 (正員)

1976 年京大大学院修士課程修了. 京大助手, 東北大助教授, 岡山大教授を経て 1995 年より京大大学院電子通信工学専攻教授. 現在同大学院情報学研究科知能情報学専攻教授. 2002 年学術情報メディアセンター長, 京都大学評議員. 2005 年情報環境機構長. 工博. 画像理解, 人工知能, 分散協調視覚, 3次元ビデオの研究に従事. 最近は「人間と共生する情報システム」の実現に興味を持っている. 1980 年情報処理学会創立 20 周年記念論文賞, 1990 年人工知能学会論文賞, 1993 年情報処理学会論文賞, 1994 年電子情報通信学会論文賞, 1995 年第 5 回国際コンピュータビジョン会議 Marr Prize, 1996 年国際パターン認識連合 Fellow, 1999 年電子情報通信学会論文賞, 2000 年画像センシングシンポジウム優秀論文賞, 2004 年 FIT 優秀論文賞. 情報処理学会元理事.