

多面体モデルを用いた全方位ビデオの圧縮

曾良 洋介[†], 波部 齊^{††}, 山澤 一誠^{†††}, 松山 隆司[†]

[†] 京都大学大学院情報学研究科知能情報学専攻

^{††} 京都大学学術情報メディアセンター

^{†††} 奈良先端科学技術大学院大学情報科学研究科

概要 全方位ビデオが用いられる場面が増えてきている。その特長でもある撮影範囲の広さゆえにデータ量が大きく、圧縮が必要となる。しかし、全方位ビデオは球面という三次元の物体にマッピングされた視覚情報であり、既存の二次元映像に対する動画圧縮方法は適用することができない。そこで、本論文では全方位ビデオを効率良く圧縮する方法として、多面体モデルを用いて球面から平面へ全方位ビデオを変換し、その二次元映像に対して MPEG 圧縮する方法を提案する。また、評価実験において各多面体の圧縮効率を画質劣化の割合によって評価する。

Compression of Omnidirectional Video with Polyhedral Mapping

Yosuke KATSURA[†], Hitoshi HABE^{††}, Kazumasa YAMAZAWA^{†††},
and Takashi MATSUYAMA[†]

[†]Department of Intelligence Science and Technology,
Graduate School of Informatics, Kyoto Univ.

^{††} Academic Center for Computing and Media Studies, Kyoto Univ.

^{†††} Graduate School of Information Science,
Nara Institute of Science and Technology

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. We evaluate various kinds of polyhedral and cartographical screen and investigate their effectiveness.

1 序論

撮影機械の発達、コンピュータのグラフィック処理能力の向上は目覚ましい。その発達により、任意視点から対象を見ることができる三次元ビデオ映像や、ある視点からの任意方向の映像を見ることができる全方位ビデオなど、従来の二次元映像では不可能だった、三次元的なシーンを扱う映像が可能になった。

特に、全方位ビデオについては、全方位を撮影するカメラシステムが市販され、比較的容易に撮影することが可能になってきている。また、全方位ビ

デオを用いた監視システムが発表されたり、ロボットの遠隔操作のための視覚センサとして全方位ビデオを用いた研究 [1] がなされたりと、全方位ビデオへの関心が高まってきている。

全方位ビデオの特長は、視点を中心とした全球の映像が見られるほどの撮影範囲の広さである。しかし、その撮影範囲の広さゆえにデータ量が大きくなってしまいう問題がある。そのため、高解像度の場合のフレームレートはデータ転送容量の制限を受けてしまい、映像としての利点が失われてしまう。高解像度で十分なフレームレートを実現しよう

とすると、データの圧縮が必要となってくる。

現在の全方位カメラから得られるのは二次元映像であり、既存の技術を用いて圧縮をしようとすると、MPEGによって標準化されている圧縮方法を用いてそのまま圧縮してしまう方法が考えられる。しかし、この方法では、カメラシステム固有のパラメータも一緒に送信しなくてはならない。全方位ビデオをカメラシステムに依存しないデータとして扱い、その幾何学的特長に沿った符合化をする必要がある。

カメラシステムに依存しない全方位ビデオは地球上に視覚情報がマッピングされたものであると考えられる。そこで、本論文では、全球を効率よく表し、圧縮に向くデータ構造を考える必要がある。

この問題提起を踏まえ、本論文では第2節で全方位ビデオの効率的な圧縮方法について考察し、第3節では、提案された各方法を用いて圧縮をした際に起こる画質劣化についての評価実験を行なう。最後に第4節で実験のまとめと今後の課題を述べる。

2 多面体モデルを用いた全方位ビデオの圧縮

本節では全方位ビデオを効率良く圧縮する方法について考察していく。

2.1 全方位ビデオの圧縮

まず、今回扱う全方位ビデオは、視覚情報は視点を中心とした地球上にマッピングされているが、その全球の奥行きは定義されないとする。

本論文で注目しているのは動画としての圧縮であり、その時空間的な冗長度を削減することによるデータ圧縮が有効であると考えられる。そのためには動きベクトルを用いた圧縮方法が適している。

三次元空間に動きベクトルを適用することも考えられるが、インターネットでの情報配信などを考えれば、既存の技術を応用し、広く使われるような圧縮方法にする方がよりよいと考えられる。しかし、現在の標準的な動きベクトルによる動画圧縮は平面的な二次元映像を対象としている。そのため、全方位ビデオにその動画圧縮を適用するためには、全球にマッピングされたデータを二次元のデータに変

換しなければならぬ。これに対して、全方位ビデオは三次元のうち方位のみで奥行き情報がないため、経度と緯度の二つのパラメータによって表現できる。つまり、平面への展開が可能であり、既存技術を用いた圧縮もできると考えられる。

本論文と異なるアプローチで全方位の画像データを表現する方法として、球面を一次元配列で表現する方法 [2] や二次元配列で表現する方法 [3] が提案されているが、この二つの手法は静止画に対する方法であり、動画については、空間的な近傍関係を保持しないため動きベクトルの適用ができないと考えられる。

本論文では、全球のデータを平面画像に変換し、その平面画像に対して従来の動画圧縮技術を用いて圧縮をすることで全方位ビデオの圧縮を試みる。

2.2 球面から平面への展開

球面を二次元に展開する方法としては、球を囲むスクリーンに球上のデータを投影し、そのスクリーンを展開することで球面を二次元に変換する方法が考えられる。その中でも、スクリーンの形状によって以下の二つが挙げられる

円錐や円柱をスクリーンとした方法

スクリーンを円錐や円柱とした展開方法は世界地図にも使われている。投影面の形状、投影中心の場所、投影面と地表の関係、投影面と地軸の関係などによって様々な投影図が考案されている。また、投影図からさらに変形を施すことでより扱いやすくしているものもある。その変換の際に、展開図の作成や透視投影図の作成の際に非線形な計算が必要となる。

多面体をスクリーンとした方法

球上のデータを多面体のスクリーンに投影し、その多面体を展開する方法も考えられる。この方法だと、線形な計算で処理できるため、処理速度の向上が望める。

この後、上記二つの方法による展開図で、今回の圧縮に最も適した展開方法がどれなのかを評価していく。ここでの評価観点として以下のものが挙げられる。

解像度の偏り 解像度の偏りとは、つまり、球状のデータから二次元のデータに変換する際のサン

プリングの密度が場所によって異なるということである。この偏りはデータ量を不必要に大きくする原因となる。

スクリーンの形状 全方位ビデオを見る場合、まずスクリーンとなるオブジェクトを作り、そのオブジェクトにテクスチャを張るという手順によって実現されることが多い。曲面を含むオブジェクトをスクリーンとする場合、そのオブジェクトを作る段階で歪みが生じることになる。なぜならば、コンピュータが扱える数値は離散化されているため、曲面は多面体の面（パッチ）の数を増やすことによって近似されているからである。より歪みを少なくするにはパッチ数の増加をすればよいが、投影される多面体のパッチの数が多きことは、テクスチャマッピングにおける処理を複雑にする原因となる。

展開図の形状 展開図の形状は必ずしも矩形とならないため、二次元に変換した画像には無効領域ができる。この無効領域も MPEG 圧縮の対象領域としてしまうのは効率的でない。これは、MPEG-4 で採用された任意形状の VOP (Video Object Plane) を用いて圧縮の対象となる領域を選ぶことで、対応できる。しかし、対象の形状のデータが必要となるので、矩形の画像の方がよりデータ量が小さくなる。

展開図上の動き 現在の動画圧縮は対象の形状の変化が少なく、動きが直線的であることを前提としている。そのため、直線的な動きが歪んで投影される場合、圧縮の効率が悪くなる。

また、本来つながっている部分が離れてしまうと、その部分を跨ぐような動きベクトルは使えなくなってしまうので圧縮効率が下がるものと考えられる。

2.3 地図展開法による展開

直観的であり、見た目にわかりやすいという利点がある。先程も述べたように、その多くは円筒などの曲面により構成されたスクリーンに投影している。そのため、非線形な変換を用いて球面から二次元への変換をしていることになる。これは前述した通りパッチ数の増大や計算量の増加という欠点に結び付く。

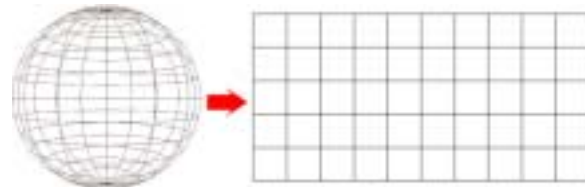


図 1: 正距円筒図法

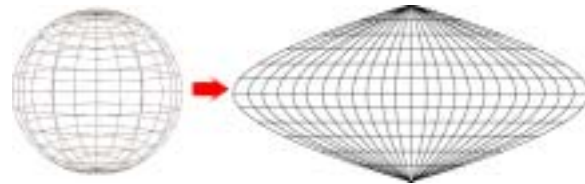


図 2: Sinusoidal 図法

従来、様々な地図展開法が提案されているが、

- 球の全体を投影することができる。
- 経緯線が単純な式で表される。

の二点を考慮すると、次に挙げる二つの地図展開法を用いて球を展開するのが良いと考えられる。

正距円筒図法 図 1 のような図法で、経緯線は垂直線と水平線で表され、それぞれが平面上で等間隔にならぶ図法である。赤道から θ [度] 離れるにつれて、経線方向に拡大される。そのためそのデータ量は元の球と比較すると、赤道の長さを画像の幅とした場合には 1.57 倍のデータ量となる。展開図は矩形となり、形状としては MPEG の圧縮に最も適している。また、スクリーンが円柱状であり、パッチを用いた映像表示をするのに向いていない。

Sinusoidal 図法 図 2 のように、平面上で緯線が水平線で正しい間隔、正しい長さで表され、経線も正しい間隔に並ぶような図法である。拡大される部分が無く、元の球と等しいデータ量で解像度の偏りなく表現できる。正距円筒図法同様、展開図に曲線を含んでおり、パッチを用いた映像表示をするのに向いていない。

2.4 多面体モデルによる展開

視点を投影中心として、球上の視覚情報を多面体に投影し、その多面体を展開することで二次元に変

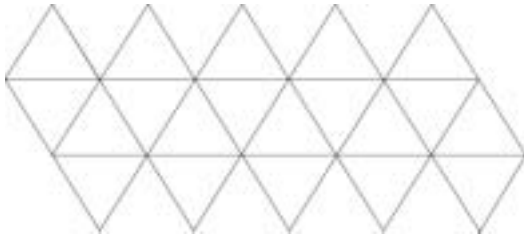


図 3: 二十面体の展開図

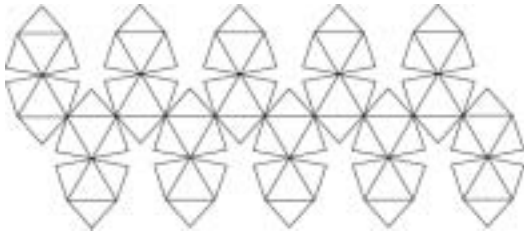


図 4: 八十面体の展開図

換する手法である．様々な多面体がある中で，どのような多面体が全方位ビデオの圧縮に適した多面体なのかを考察していく．

2.4.1 正多面体による展開の優位性

図 3 は正二十面体の展開図であり，図 4 は正二十面体の各面を 4 分割し，頂点座標を正規化することで得られる八十面体の展開図である．多面体を構成する面が全て合同であれば，展開図の継目をより少なくすると考えられる．継目の少ない単純な展開図は動きベクトルが適用しやすく，また，各面に投影された映像は同じ歪み方をするので圧縮効率が上がると考えられる．

よって，八十面体は圧縮には適していないと考えられるため，今回は正四面体，正六面体，正八面体，正十二面体，正二十面体の五つの正多面体を比較することにした．以下では，データ量の観点から五つの正多面体を比較していく．

2.4.2 ピクセルのオーバーサンプリング

球状の視覚情報を多面体に投影する際，その投影中心は球の中心となる．球面を等間隔にサンプリングして多面体に投影していくと，図 5 に見られるように，球の中心からの距離の違いが原因となり，引

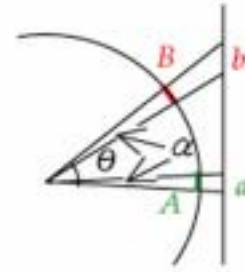


図 5: 球面からの投影

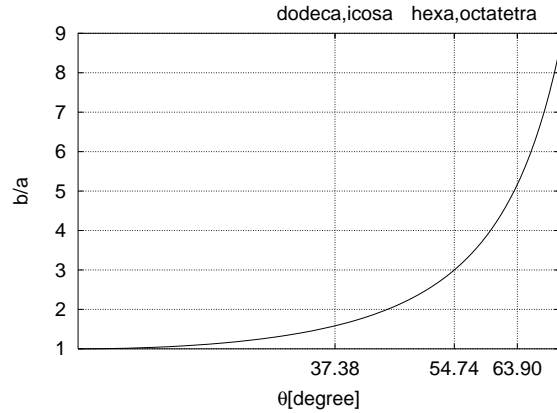


図 6: オーバーサンプリング

き延ばされる部分が出てくる．図 5 中の A, B はある角度 α によって球面から切り取られる面積で， a, b はそれぞれ A, B がスクリーン上に投影された際のスクリーン上の面積で $A = B$ である．球上からの視覚情報の損失を防ぐためには対応領域が最も小さい部分を多面体の 1 画素に対応させる必要がある．(図で $a = 1$ とすることに対応) その場合，引き延ばされる部分は球上の 1 画素に対して 1 画素以上の画素が割り当てられることになる．(図で $b/a > 1$ となる) このオーバーサンプリングはデータ量を不必要に大きくしてしまう原因になる． b/a は見込み角 θ の関数であり，

$$\frac{b}{a} = \frac{1}{\cos^2 \theta} \quad (1)$$

で表され，図 6 のような単調増加をする．

データ量を少なくするためには θ の最大値が小さくなるが必要となる． θ の最大値は，各面の重心とその面の頂点がなす角度である．つまり，多面体の一つの面が小さい方が圧縮に向いていると考えられる．各多面体ごとに球上の 1 画素に対する多面

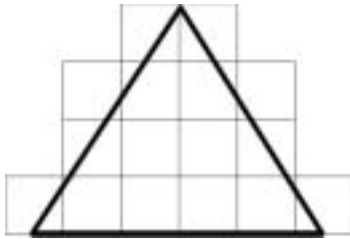


図 7: VOP の適用

体の平均画素数,つまり,表 1 に b/a と,平均画素数 \times 面積,つまり多面体全体での画素数の関係を示す.この表から,面の数を増やして球への近似を良くするほど,オーバーサンプリングの点でも面積の点でもより望ましい値をとっており,無駄なデータが少なくなることがわかる.

2.5 展開図の形状と MPEG の親和性

2.2 節で述べたように,展開図形状が矩形にならない場合には任意形状の VOP を用いることで効率的な圧縮が可能になる.しかし,VOP は 16×16 の正方形を基準として領域を選択することになっているので,二十面体のように三角形によって構成されている場合などは図 7 のようにどうしても無効領域を選択範囲に含めなければならない.

そこで,図 8 のような変換を行なうことを考える.この変換によって,展開図は矩形の組み合わせとなり,VOP による領域選択により過不足なく圧縮の対象となる.しかし,この変換により,サンプリング密度分布が変化してしまう.そのことにより,球状の 1 画素に対する多面体の平均画素数が変化し,変換する前は 1.14 画素だったのが,1.33 画素まで増えてしまう.

すなわち,展開図を矩形にすることによる MPEG 圧縮時の効率化と,データ量の増大はトレードオフの関係にある.3 節でこの変換の効果を定量的に評価し,この変換の有効性を検証する.

3 評価実験

本実験では展開方法に起因する画質劣化を定量的に評価していくことで,どの展開方法が最も効率良く全方位ビデオを圧縮できるのかを調べていく.

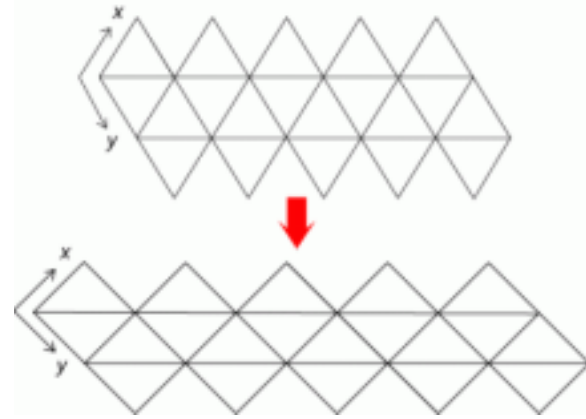


図 8: 正二十面体の変形

3.1 実験の概要

本実験には以下の 3 つのステップがある.

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,2.3 節で述べた正距円筒図法と Sinusoidal 図法,2.4 節で述べた多面体モデルの中で正六面体,正八面体,正二十面体および図 8 に示した変形した正二十面体展開図(以下,正二十面体(変形)と呼ぶ)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
作成した各展開図に対して,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,ある視線方向での透視投影画像を生成し,その画質を評価する.視線方向の決定と評価の方法については後述する.評価のリファレンスとするのは,与えられた全方位ビデオから直接生成した透視投影画像である.

なお,上記の実験の流れの中ではいくつかのパラメータを定めなくてはならない.以下では,本実験の目的である,展開方法の違いによる画質劣化の評価,という観点でのパラメータの決定について述べる.

有効画素数について ステップ 1 で作成する展開図には画像情報をもつ画素(有効画素と呼ぶ)とそう

表 1: 多面体の比較

| | 正四面体 | 正六面体 | 正八面体 | 正十二面体 | 正二十面体 | 八十面体 |
|-----------|---------|---------|-------|---------|-------|-------|
| 見込み角 | 63.90 ° | 54.74 ° | | 37.38 ° | | - |
| 平均画素値 | 2.04 | 1.67 | 1.50 | 1.22 | 1.14 | 1.04 |
| 面積 × 平均画素 | 44.26 | 40.0 | 31.18 | 20.39 | 17.37 | 13.91 |

でない画素が存在する．本実験では有効画素数がほぼ等しくなるようにして，各展開方法による展開図を作成する．このことにより，2.2 節で述べた解像度の偏りによる画質劣化を評価することができる．

その有効画素数は二つの場合について実験を行なう．まず，元の全方位データと同等の有効画素数を持つ展開図についての実験を行なう．さらに，画質劣化による影響が大きくなるような，有効画素数を少なくした展開図についても実験を行なう．表 2 に本実験におけるそれぞれの展開図の有効画素数を示す．

エンコード時の圧縮率について 全ての展開図に対して圧縮率が等しくなるようにエンコードを行なう．このことにより，2.2 節で述べた展開図上の動きと圧縮効率との関係进行评估することができる．

展開図の画質評価について 展開図の画質評価は，次式に従って計算される PSNR R を比較して行なう．

$$R = 10 \log_{10} \left(\frac{255^2}{S/(N \cdot T)} \right),$$

$$S = \sum_{p,t} \left(i(p,t) - \hat{i}(p,t) \right)^2.$$

ここで， $i(p,t)$ は評価対象の透視投影画像の位置 p ，時刻 t での画素値を示し， $\hat{i}(p,t)$ はリファレンスとする映像を示す． N は 1 フレームの画素数， T は評価する透視投影画像のフレーム数を示す．

いくつかの視線方向に対して，VGA サイズの透視投影画像を作成し，その上で PSNR を計算する．ここで選択する視線方向によって PSNR も変化してくる．これは，展開の際に起こる解像度の偏り，または本来つながっている場所が切り離されることが原因である．特に頂点付近では歪みが顕著であると考えられる．よって，いくつかの視線方向における透視投影画像を作成することで，その歪みの影響を評価する．

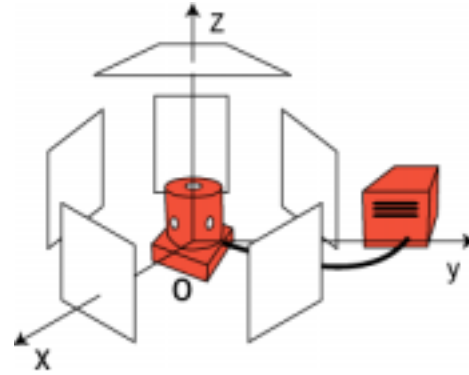


図 9: Ladybug カメラシステム

3.2 実験に用いた全方位ビデオ

このデータは，Point Grey Research 社製の Ladybug という多眼カメラを用いて撮影されたものである．この Ladybug は，小型カメラヘッド部と蓄積ユニットにより構成されている．小型カメラヘッド部は図 9 のように，水平方向には五つのカメラが等間隔に並んでおり，鉛直上方向きには一つのカメラがある．この六つのカメラによって全球の 75% 以上が撮影可能である．それぞれのカメラは 1024×768 の解像度で，最大 15fps の映像が撮影可能である．今回の実験では，走行車の屋根にカメラを取り付けられて撮影された，10 秒間 (150 フレーム) の映像を使用した．

3.3 実験結果

本実験で得られた結果を表 3(a) に示す．表中では，透視投影画像を作成した視線方向を図 10 に示す角度 (θ, ϕ) で表現している．各視線方向での映像の例を図 11 に示す．実験の結果から以下のことがわかる．

まず，地図展開法の中では，Sinusoidal 図法が優れている傾向がわかる．これは，2.3 節で述べた通

表 2: 展開図のサイズ

| | 有効画素数 [画素] |
|---------------|------------|
| 高解像度 | |
| 正距円筒図法 | 4718592 |
| Sinusoidal 図法 | 4732094 |
| 正六面体 | 4707277 |
| 正八面体 | 4668002 |
| 正二十面体 | 4765687 |
| 正二十面体 (変換) | 4785077 |
| 低解像度 | |
| 正距円筒図法 | 1174386 |
| Sinusoidal 図法 | 1176748 |
| 正六面体 | 1145029 |
| 正八面体 | 1166674 |
| 正二十面体 | 1198386 |
| 正二十面体 (変換) | 1146072 |

り, Sinusoidal 図法では展開図上の解像度の偏りが少ないことによると思われる。また, 多面体モデルの中では, 正二十面体を用いた展開方法が他の展開方法よりも画質劣化が少ないことが分かった。これは, 2.4.2 節で述べた考察に合致する結果である。また, 変換を施した正二十面体の展開図はその有効性が認められなかった。

なお, 視線方向 ($0^\circ, 90^\circ$) での評価結果の傾向は他の結果と異なるが, これは図 11(c) に示しているように, 今回用いたデータでは鉛直上方でのテクスチャが少なかったためであり, 各手法の性能を正しく反映していないと思われる。

次いで, 視線方向の選択によらず, 全方位でどれだけ画質劣化しているかを評価するために, デコード後の展開図から, 球面上で等間隔にサンプリングして画素値を得て, 元のデータより同じように得られた画素値と比較した。その結果を表 3(b) に示す。この結果によると, 今回用いた全方位ビデオの場合は, 視線方向の選択によらず, 表 3(a) と同じ傾向が見られることがわかる。

以上の実験の結果, Sinusoidal 図法および正二十面体による展開法を用いた場合の画質劣化が少ないことが明らかになった。しかし, 2.3 節に述べたように, Sinusoidal 図法は球面から二次元展開図への変換の際に非線形な演算が必要になり, 実用面では正二十面体が優れているといえる。

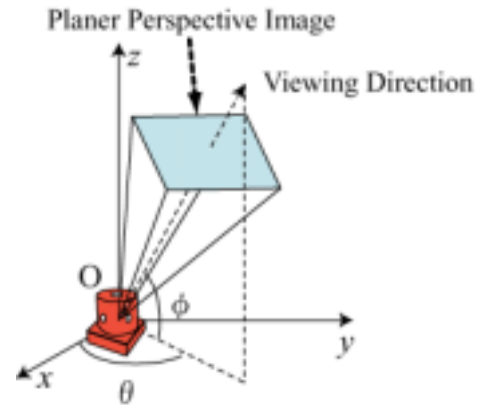


図 10: 透視投影画像の生成

4 結論

本論文では, 全方位ビデオを圧縮する方法として, 全方位ビデオを多面体モデルを用いて二次元平面へ変換してから MPEG 圧縮する方法を提案した。そして, いくつかの展開方法の中でも解像度の偏りが少ないため, データの冗長性が低い, 正二十面体を用いた展開方法と Sinusoidal 図法による展開方法が画質劣化が少ないことを示した。両者を比較すると, 透視投影画像作成の際に非線形な変換を用いない, 正二十面体を用いた展開方法がより良い方法であると考えられる。

しかし, 我々が使用した評価手法の結果は圧縮される映像に依存するものであるため, 今後多種の映像で実験を行なう必要がある。

その他の課題としては, 世界座標系に対する多面体の向きに関する評価が挙げられる。多面体の向き次第で, 展開図上での動きベクトルの向きが変わってくる。このとき, 動きベクトルが途切れないように多面体の向きを決めることで, 圧縮効率が上がると考えられる。よって, 今後, 全方位ビデオ上で見られる動きと多面体の向きが圧縮の効率に与える影響についても評価していく必要がある。



(a) $(\theta, \phi) = (-135^\circ, 0^\circ)$

(b) $(\theta, \phi) = (135^\circ, 35^\circ)$

(c) $(\theta, \phi) = (0^\circ, 90^\circ)$

図 11: 透視投影画像の例

表 3: 画質評価のために計算した PSNR(dB)

(a) 特定の視線方向について計算した PSNR

| 視線方向 (θ, ϕ) (°) | (-135, 0) | (135, 35) | (0, 90) |
|---------------------------|-----------|-----------|---------|
| 高解像度 | | | |
| 元データ | 30.6 | 27.9 | 28.6 |
| 正距円筒図法 | 30.8 | 28.5 | 29.9 |
| Sinusoidal 図法 | 31.2 | 28.5 | 28.6 |
| 正六面体 | 30.9 | 28.8 | 28.9 |
| 正八面体 | 30.8 | 28.1 | 29.5 |
| 正二十面体 | 31.3 | 28.7 | 29.4 |
| 正二十面体 (変換) | 29.9 | 27.9 | 28.9 |
| 低解像度 | | | |
| 正距円筒図法 | 26.6 | 24.5 | 26.0 |
| Sinusoidal 図法 | 27.1 | 24.5 | 24.5 |
| 正六面体 | 26.7 | 24.8 | 24.8 |
| 正八面体 | 26.6 | 24.0 | 25.5 |
| 正二十面体 | 27.1 | 24.7 | 25.2 |
| 正二十面体 (変換) | 25.4 | 23.6 | 24.6 |

(b) 全方位について計算した PSNR

| | |
|------------|------|
| 高解像度 | |
| 正六面体 | 30.6 |
| 正八面体 | 30.4 |
| 正二十面体 | 30.7 |
| 正二十面体 (変換) | 29.5 |
| 低解像度 | |
| 正六面体 | 25.5 |
| 正八面体 | 25.1 |
| 正二十面体 | 25.6 |
| 正二十面体 (変換) | 24.7 |

謝辞

本研究は科学研究費補助金基盤研究 A 13308017 および特定領域研究 13224051 の補助を受けて行った。

本研究を行うにあたり有益なご助言をいただいたシャープ (株) の野村敏男氏に感謝する。

今回の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横矢研究室から提供されたものである。ここに感謝の意を表す。

参考文献

- [1] 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センサ. 電子情報通信学会論文誌 D-II, Vol.J79-D-II, No.5, pp.698-707, (1996.5).
- [2] 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, (1999.6).
- [3] 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. 電子情報通信学会論文誌 D-II, Vol.J79-D-II, No.11 pp.1862-1870, (1996.11).