視点間の射影幾何的關係を利用した多視点動画のデータ圧縮

Data Compression of Multiple-View Movies via Projective Geometry of the Viewpoints

秋山 修[†], 正会員 斎藤 英雄[†]

Osamu Akiyama[†] and Hideo Saito[†]

Abstract In this paper, we propose a method for compressing the data of multiple-view movie considering redundancy of the viewpoints. This is achieved by dividing multiple-view movies of dynamic scenes captured by several cameras into foreground and background, and applying projective geometry to the viewpoints for each region. Specifically, a viewpoint, called the base viewpoint of compression, is selected from multiple viewpoints and another viewpoint, called the target viewpoint of compression, is predicted by applying projective geometry between the base viewpoint and the target viewpoint. We also propose a method to decide the optimal viewpoint for compressing multiple-view movies efficiently. The results of compressed multiple-view movies and obtaining an optimal viewpoint for compressing them are evaluated in experiments.

キーワード：多視点, データ圧縮, 射影幾何, 視点間冗長性

1. ま え が き

近年, コンピュータの処理能力やビデオ機器の発達に伴って, 同一シーンを多数のカメラで撮影した多視点画像が利用されるようになってきた. 例えば, 米国 CBS テレビが 2001 年のスーパー・ボウル中継で実際に放送に用いた“Eye-Vision”システム¹⁾が有名である. また, 新たな映像表現方法の一つとして, 映画“Matrix”²⁾では, 時間が静止した状態で視点が動くような, コンピュータグラフィックスが, 実画像を元にして生成されている. また, 複数のカメラで撮影された画像から任意視点画像を生成する研究³⁾や, 大規模な空間で行われる動きのあるイベントを対象として, 複数のカメラで撮影された映像をもとに, 自由に視点を変化させた映像を提示する研究⁴⁾⁵⁾などが挙げられる.

一方で現在, インターネット上で画像のストリーミング配信などが行われている. このような映像データをネットワークで伝送したり, 蓄積するには画像の圧縮技術が重要になってくる.

このように, 多視点画像の利用や画像のストリーミング

配信といったことを考えると, 将来的に上記で述べた任意視点画像を含めた多視点画像のストリーミング配信, また放送といったことが行われると予想され, 多視点動画の圧縮技術が重要になってくる. 画像の圧縮として, MPEG-1,2 や新しい画像符号化標準 MPEG-4 などが知られている. また, 動画圧縮の研究として, MPEG-4 スプライト符号化を利用した超低レート動画圧縮の研究⁶⁾や三角メッシュを利用した動画の圧縮研究⁷⁾⁸⁾などが報告されている. しかし, これらの画像圧縮の対象は, 1 視点で撮影された映像である. ステレオ画像で, 片方の画像と視差情報を利用して, 他方の画像も復元する研究⁹⁾も報告されているが, 2 視点の動画圧縮である. また, 多視点画像の圧縮として多視点画像間のエピソード幾何拘束を用いて, 多視点画像間の相関を利用した予測符号化¹⁰⁾が提案されているが, 密な視点間隔で撮影された剛体を対象としており, また静的なシーンを対象としている. 多視点画像のストリーミング配信や放送といったことを考えると, 動きのあるシーンを対象とした多視点動画圧縮技術の重要性は, 高まってくるものと予想される. 近年のブロードバンドの普及によるネットワークの回線の高速化が進んでいるとは言え, 多視点動画のデータ量は膨大であるために, 多視点動画の圧縮技術は必要不可欠なものとなってくる.

多視点動画の圧縮として, 1 視点動画の圧縮, 例えば, MPEG で視点数分を独立して行う方法と視点間の冗長な部分を考慮した方法が考えられる. しかし, この方法

2004 年 2 月 9 日受付, 2004 年 9 月 22 日最終受付, 2004 年 10 月 14 日採録

[†]慶應義塾大学 大学院 理工学研究科

(〒 223-8522 神奈川県横浜市港北区日吉 3-14-1. TEL 045-566-1753)

[†]Graduate School of Science and Technology, Keio University

(3-14-1 Hiyoshi, Kouhoku-ku, Yokohama-shi, Kanagawa, 223-8522, Japan)

図 1 提案手法の流れ
Flowchart of proposed method.

だと異なる視点で同じ物が見えているにもかかわらず、視点毎に独立して圧縮することになり、効率が悪くなる。そこで本論文は、射影幾何を利用することにより、視点間冗長性を考慮して、数台程度のカメラで撮影された動きのあるシーンを対象にした、多視点動画を圧縮する手法を提案する。視点間の冗長な部分を考慮する際に、視点間の射影幾何を利用するのは、視点間射影幾何の関係が画像だけから容易に得られるためである。提案手法は、多視点動画の中で、ある視点 i をもとに他の視点 j との間の射影幾何的关系を利用して、視点 j の画像を圧縮するものである。ここで視点 i をどの視点にするかで、圧縮効率が変わってくる。また視点によっては、視点間の冗長な部分を考慮した圧縮が非効率になることも考えられる。そこで本論文では、最も圧縮効率の良い視点を決定する方法や視点間の圧縮がどのような場合に非効率であるかについても提案する。

本論文では、まず第 2 章で提案手法のおおまかな流れについて述べる。次に第 3 章で視点間の射影幾何的关系について述べる。そして第 4 章で提案手法の各過程の詳細について述べる。さらに第 5 章で、提案手法で最も圧縮効率の良い視点を決定方法について示し、第 6 章では実験とその結果を、第 7 章で今後の展望について示す。

2. 処理の概要

提案手法では、多視点動画の中で、ある視点を基準として圧縮を行う。この視点を圧縮基準視点と呼ぶことにする。次に圧縮基準視点とそれ以外の視点(別視点)との射影幾何関係を利用して、別視点の画像を圧縮する。この圧縮される別視点を圧縮対象視点と呼ぶことにする。この圧縮では、対象シーンを前景領域と背景領域に分ける。そしてこれらの領域が、平面近似される複数の領域より構成されるものと考え、領域毎に圧縮基準視点と圧縮対象視点間で平面射影変換を適用すれば、圧縮基準視点画像から圧縮対象視点画像を予測できることを利用する。このように予測され生成される画像を、視点間予測画像と呼ぶことにする。

この視点間予測画像と圧縮対象視点画像の差分画像を符号化することにより、圧縮対象視点画像を圧縮する。このような提案手法のおおまかな流れについて図 1 に示す。

このように圧縮の際に、領域毎に考慮するということは、オブジェクト毎に圧縮するオブジェクトベース符号化と同じである。オブジェクトベース符号化では、動画中のオブジェクト抽出が重要であり、そのため動画中のオブジェクト抽出といった研究が行われている¹¹⁾。提案手法でも、平面近似する領域として、大きく分けて前景領域と背景領域が考えられるが、この平面近似する際に、前景領域の抽出といった問題がある。本論文では、このようなオブジェクト抽出が対象の多視点動画に対して、既に行われていることを前提として、視点間の冗長性を利用した圧縮法を提案するものである。そこで本論文では、簡単のために、前景と背景を切り分けるために、あらかじめ撮影しておいた背景画像を利用した背景差分を用いている。

圧縮基準視点の決定については、提案手法による多視点動画の圧縮効率が最大となる時、つまりより高い圧縮率でより高画質な復元画像が得られる時の視点を圧縮最適視点とし、この圧縮最適視点を前景の大きさと視点間で重複している背景の大きさを利用して求める。

3. 平面射影変換

平面射影変換の様子を図 2 に示す。3 次元空間中で同一平面上にある全ての点は、2 枚の画像 I_1, I_2 間で次式の関係を満たす。

$$\tilde{p}_2 = H \tilde{p}_1 \quad (1)$$

なお、ここで $\tilde{p}_1, \tilde{p}_2$ は、 I_1, I_2 間で対応点の取れている点の斉次座標である。行列 H は平面射影行列、あるいは Homography と呼ばれ、3 次元空間中の各平面毎に異なる値をもつ。 H は 3×3 の行列で、9 要素をもつが、スケールを任意に設定できるため、自由度は 8 である。よって 2 枚の画像間で対応点を 4 組以上から H を算出できること

ができる。平面射影変換を用いた場合、片方の画像である点を与えると、他方の画像でそれに対応する1点が直ちに求まる。提案手法では、圧縮基準視点の画像から圧縮対象視点の画像を予測し、視点間予測画像を作成するのにこの平面射影変換を用いる。

図2 平面射影変換
Homography.

4. 提案手法

4.1 前景領域と背景領域の分離

提案手法では、多視点動画を圧縮する際、対象シーンが複数の平面近似領域から成り立っていると考えるために、まず背景領域と前景領域を分離する。このため、提案手法

図3 前景領域の抽出
Extraction of foreground region.

図4 背景領域の分割
Division of background region.

では背景差分を用いて背景領域と前景領域を分離する。この時、背景差分の閾値を大きく設定すると、前景領域であるにも関わらず背景領域と判断され、また前景領域内で穴ができてしまうことがある。このようなことを避けるために、背景差分の閾値を小さく設定することにより、真の前景領域がほぼ確実に前景領域に含まれるようにする。これは、真の前景領域が背景差分により誤って背景領域と判断されてしまうと、平面射影変換を適用して生成される視点間予測画像の精度が悪くなり、結果的に、提案手法による圧縮の効率が下がってしまうことを避けるためである。逆に、真の背景領域が前景領域と誤って判断されてしまっても、生成される視点間予測画像は、背景領域に似た前景領域となるため精度はあまり悪くならず、また圧縮の効率はほとんど悪くならない。以上から、提案手法では背景差分の閾値を小さく設定した。この様子を図3に示す。図3の(d)からも閾値を大きくすると前景領域を完全に抽出できていないことが分かり、図3の(c)から閾値を小さくすると、前景領域を完全に抽出できていることがわかる。

4.2 視点間の領域の対応付け

提案手法では、圧縮基準視点画像から平面射影変換を適用することにより、圧縮対象視点画像を予測する。第3章で述べたように平面射影行列を求めるには、4点以上の対応点の組合せが必要である。ここでは、背景領域・前景領域それぞれについて、平面射影行列を求める4点の組合せをどのように対応付けているかを述べる。また、この平面射影行列を求める対応点の組み合わせは、圧縮データから画像を復元する際にも用いる。

(1) 背景領域の対応付け

あらかじめ撮影しておいた、図3の(b)のような背景画像を平面と近似できる複数の領域に分割する。領域を分割した例を図4に示す。図4のように平面と近似できる領域を複数選び、視点間で対応する領域内での対応点4点以上から平面射影行列を求める。

(2) 前景領域の対応付け

4.1で分離した圧縮基準視点の前景領域と圧縮対象視点の前景領域の各領域で4点ずつ対応付けを行う。

図5 前景領域の対応付け
Correspondence of foreground region among viewpoints.

具体的にはまず、図3の(c)のような圧縮基準視点の前景領域と圧縮対象視点の前景領域に対して、ラベリング処理を行う。次に、背景領域として抽出した地面の平面射影行列を用いて、圧縮基準視点のラベル L_i が地面と接する1点を、圧縮対象視点に射影変換した点を含むラベル L_j を求める。このラベル L_i とラベル L_j を同じ領域であるに対応付けをする。最後に、対応付けされたラベル L_i とラベル L_j から4点の対応点を求め、この4点の組合せから前景領域の平面射影行列を求める。なお、ここでの4点の組合せは、図5に示すようなラベル L_i とラベル L_j の外接長方形を構成するものである。これらの処理を圧縮基準視点の全てのラベルに対して行う。

4.3 視点間予測画像と視点間差分画像

図6 視点間予測画像と視点間差分画像
Predict image and difference image among viewpoints.

4.2で対応付けされた背景領域では、背景領域を構成する複数の平面近似領域毎に、平面射影行列が求まっている。各領域を平面近似しているため、圧縮基準視点からそれぞれの領域に対して、平面射影変換をそれぞれ行うことにより、圧縮対象視点における背景領域の画像を予測できる。

前景領域でも同様に、前景領域の各領域を一つの平面であると近似して、圧縮基準視点からそれぞれの領域に対して、平面射影変換をそれぞれ行うことにより、圧縮対象視点における前景領域の画像を予測できる。

このように、圧縮対象視点の背景領域を予測したものと前景領域を予測したものを合成すれば、視点間予測画像を生成でき、圧縮対象視点の画像を予測できる。この視点間予測画像と圧縮対象視点画像との差分である視点間差分画像を圧縮すれば、圧縮対象視点の原画像を視点間の冗長性

を考慮せずに圧縮するよりも、多視点動画のデータ量をより効率良く圧縮することができる。これらの背景・前景予測や視点間予測・視点間差分画像の結果を図6に示す。

4.4 圧縮方法と復元方法

(1) 圧縮方法

提案手法では、ある視点を圧縮基準視点として、他の視点である圧縮対象視点の画像を圧縮するものであり、圧縮基準視点の動画データは、他の視点とは独立して圧縮する。また、圧縮対象視点の動画データに対しては、背景・前景の各領域毎の平面射影行列を求めるための視点間の対応点の組合せデータと図6(f)のような視点間差分画像の動画データを圧縮する。ここで圧縮データの視点間の対応点の組合せデータは、動画データに比べ、無視できる程度少なくなっている。

(2) 復元方法

圧縮基準視点の動画データは、上記のように他の視点とは独立して圧縮されているため、この圧縮データだけから復元できる。圧縮対象視点の動画データに対しては、まず、復元した圧縮基準視点の動画データと視点間の対応点の組合せデータから図6(e)のような視点間予測画像を生成する。次に、この視点間予測画像と圧縮データである視点間差分画像から、圧縮対象視点の動画データを復元する。このように視点間の射影幾何学的関係を利用して、圧縮基準視点と圧縮対象視点の動画データを復元、つまり多視点動画のデータを圧縮データから復元する。

5. 視点間圧縮最適視点決定方法

提案手法では、多視点のうち、どの視点を圧縮基準視点とするかによって、圧縮の効率が変わってくると考えられる。提案手法による多視点動画の圧縮において、より圧縮効率が良く、さらに復元画像の画質が良くなる圧縮基準視点を、視点間圧縮最適視点と定義する。ここでは、この視点間圧縮最適視点を決定する方法について述べる。提案手法では、前景領域と背景領域を分離して多視点動画を圧縮するため、まず、背景領域・前景領域それぞれについて最適な圧縮基準視点を決定することを考える。次に、背景領域について決定された最適な圧縮基準視点と、前景領域について決定された最適な圧縮基準視点から視点間での冗長性を考慮した、多視点動画の圧縮に最適な圧縮基準視点を決定する。5.1では、背景領域について述べ、5.2では、前景領域、そして5.3では、背景と前景から視点間での冗長性を考慮した、多視点動画の圧縮に最適な圧縮基準視点の決定方法について述べる。

5.1 視点間圧縮最適視点(背景領域)

視点 i , j 間で背景領域の視点間冗長性を考えてみる。図7に視点 i を圧縮基準視点とし、視点 j を圧縮対象視点とした時の視点間冗長領域を示す。この時、視点間冗長領域は、4.3で述べたように、視点 i から平面射影変換を利用して視点 j を予測した、図6の(c)のような予測背景画像

図 7 視点間冗長領域
Region of overlap among viewpoints.

の領域である。背景領域における奥行きの変化はカメラからの距離に比べて小さいため、背景領域を平面と近似して視点 i から視点 j を予測した時の予測誤差は小さいと考えられる。したがって、平面射影変換を利用して予測される領域が大きければ大きいほど圧縮の効率が良くなる。つまり、圧縮基準視点とそれ以外の圧縮対象視点との間で予測される領域 (視点間冗長領域) の和が、最も大きくなる時の視点を圧縮基準視点として選択できれば、圧縮効率が良いことになる。そこで背景領域のみを考慮した場合、次式の値が最小となる視点 i を視点間圧縮最適視点とする。

$$V_B(i) = \sum_j \frac{1}{\alpha_j} \quad (2)$$

ここで α_j とは、視点 j の画像に対する図 7 中の *Overlap*、つまり視点間冗長領域の比である。 α_j が逆数になっているのは、 α_j が大きいほど上式の値が小さくなるようにするためである。また、後に述べる前景領域や、背景領域と前景領域を合わせて考慮する際に同様の考え、つまり上式のような視点間圧縮最適視点の評価式の値が小さくなるほうが、より最適であるという考えをするためである。

一方、逆に予測領域が小さすぎると視点間の冗長性を考慮した圧縮が非効率になる。つまり、視点間の冗長性を考慮しない圧縮よりも圧縮効率が悪くなる。なぜなら、視点間の冗長性を考慮した圧縮では、予測領域が小さすぎると、予測領域に比べて非予測領域が大きくなるため、視点間差分画像データのエントロピーが増大するからである。そこで次式のように、視点間冗長領域 α_j が閾値よりも小さい場合、視点 i を圧縮基準視点とした時の視点 i, j 間では、視点間の冗長性を考慮した圧縮が非効率であるとみなし、視点 i, j 間では、視点 i を圧縮基準視点として選ばないこととする。

$$\alpha_j < Th_1 \quad (3)$$

ここで Th_1 は実験により求めた閾値である。

5.2 視点間圧縮最適視点 (前景領域)

4.1 で切り分けた前景領域にラベリング処理をし、圧縮対象視点のラベル a での前景領域とその外接長方形を図 8 に示す。本論文では、前景領域としては人物などを想定しており、必ずしも平面近似が成立しない。しかし、前景領

図 8 ラベル a の前景領域とその外接長方形
Foreground region and circumscribed rectangle of the label a .

域のあるラベル a が小さければ、平面近似による平面射影変換を利用しても、精度良く予測できる。逆に、ラベル a が大きくなればなるほど、平面近似による平面射影変換を利用しても精度が悪い予測前景画像を生成することになり、その結果、提案手法での視点間の冗長性を考慮した圧縮が非効率になる。そこで本論文では、画像全体に対する前景領域の面積比を基準にして、視点間圧縮最適視点を決定するものとする。ただし、前景領域が小さくても、その外接長方形が大きい場合は、平面近似の精度が低下すると考え、前景領域とその外接長方形の大きさの比も考慮した、下記の評価式の値が最小となる視点 i を視点間圧縮最適視点とする。

$$V_F(i) = \sum_j \sum_a \beta_{ja} \sqrt{\frac{R_{ja}}{A_{ja}}} \quad (4)$$

ここで、 β_{ja} 、 A_{ja} 、 R_{ja} は、それぞれ、ラベル a の圧縮対象視点 j の画像に対する面積比、ラベル a の面積、ラベル a の外接長方形の面積である。

上記のように、圧縮対象視点の前景領域の大きさやその外接長方形が大きくなると、予測精度が悪くなるので、次式を満たす場合に、視点 i を圧縮基準視点とした時の視点 i, j 間では、視点間の冗長性を考慮した圧縮が非効率であるとする、つまり視点 i, j 間では、視点 i を圧縮基準視点として選ばないこととする。

$$\gamma_{jmax} = \beta_{jmax} \sqrt{\frac{R_{jmax}}{A_{jmax}}} > Th_2 \quad (5)$$

ここで、上式の左辺値は、圧縮対象視点 j の全てのラベルに対して、求めた値の中で最大値の時のものである。 Th_2 は実験により求めた閾値である。

5.3 視点間圧縮最適視点

前景領域・背景領域それぞれに対して、視点間の圧縮に最適な視点を決定した。これらから多視点動画の視点間の冗長性を考慮した圧縮に最適な視点は、(2) 式と (4) 式を合わせて考える。つまり、圧縮基準視点とそれ以外の圧縮対象視点との間で予測される領域 (視点間冗長領域) の和が最も大きくなる視点、かつ図 8 に示すような圧縮対象の前景領域とその外接長方形の大きさの和が、最も小さくなる時の視点を選択すれば、圧縮効率が良いことになる。このことから、次式を最小にする時の視点 i を視点間圧縮最適視点とする。

$$V(i) = W_b V_B(i) + W_f V_F(i) \quad (6)$$

ここで W_b , W_f は重み係数である。また、視点間の冗長性を考慮した圧縮が非効率になるのは、まず (3) 式を満たす時、次に (3) 式を満たさないが、(5) 式を満たす時である。

このように (6) 式を最小とする視点を圧縮基準視点として、提案手法による視点間の冗長性を考慮した多視点動画の圧縮をすれば、効率の良い圧縮を行うことができる。

表 1 に、本手法において視点間圧縮最適視点決定に用いる評価式と閾値を示す。

表 1 視点間圧縮最適視点決定に用いる評価式と閾値
Criteria and thresholds for deciding the optimal viewpoint for compressing multiple-view movies considering redundancy among viewpoints.

	Criterion	Thresholds
Background	$V_B(i)$	α_j
Foreground	$V_F(i)$	γ_{jmax}
Back + fore	$V(i)$	$\alpha_j \cdot \gamma_{jmax}$

6. 実験と検討

提案手法の有効性を示すために、図 9 に示す (A)Soccer, (B)Lifting, (C)Dance のような様々な環境や視点数の多視点動画画像列に対して、提案手法による多視点動画画像の圧縮を適用し、実験を行った。

表 2 視点間圧縮最適視点の評価値 ((A)Soccer)
Evaluation value of the optimal viewpoint for compressing multiple-view movies considering redundancy among viewpoints ((A)Soccer).

	$V_B(i)$	$V_F(i)$	$V(i)$
視点 1	3.334	0.0286	4.2377
視点 2	2.269	0.0245	3.0436
視点 3	2.334	0.0235	3.0786

表 3 視点間圧縮最適視点の評価値 ((B)Lifting)
Evaluation value of the optimal viewpoint for compressing multiple-view movies considering redundancy among viewpoints ((B)Lifting).

	$V_B(i)$	$V_F(i)$	$V(i)$
視点 1	9.251	0.2109	12.3731
視点 2	5.439	0.2288	8.6745
視点 3	4.974	0.2189	6.8948
視点 4	5.567	0.1701	7.9716

表 4 視点間圧縮最適視点の評価値 ((C)Dance)
Evaluation value of the optimal viewpoint for compressing multiple-view movies considering redundancy among viewpoints ((C)Dance).

	$V_B(i)$	$V_F(i)$	$V(i)$
視点 1	2.394	0.2460	4.8852
視点 2	2.770	0.2286	5.0563
視点 3	2.202	0.2400	4.6027

6.1 可逆圧縮実験

まず、多視点動画画像の視点間冗長性を利用して、データ圧縮を行う提案手法が、多視点動画画像の可逆圧縮に有効であることを確認するための実験を行った。入力多視点画像のうち圧縮対象視点画像については、圧縮基準視点画像か

ら生成された視点間予測画像との差分である視点間差分画像 (例えば、図 6(f)) を用いたエントロピー符号化 (ハフマン符号化) を用いて可逆圧縮した。この結果、圧縮基準視点画像から圧縮対象視点画像を予測するために生成される視点間差分画像をエントロピー符号化すると、24bit/pixel の原画像が 15~17bit/pixel 程度のデータ量に圧縮された。一方、視点間予測画像を利用せずに、原画像の輝度値をそのまま用いてエントロピー符号化すると、約 21~23bit/pixel となった。つまり、提案手法に基づき視点間冗長性を考慮した圧縮を行うことによって、4~6bit/pixel 程度のデータ量が削減ができたことになり、提案手法が有効であることを確認できた。なお、圧縮基準視点画像については、視点間差分画像が得られないことになるため、データ量は減少しない。

(1) 視点間圧縮最適視点の決定

この可逆圧縮による実験により、視点間圧縮最適視点の決定方法の評価を行った。可逆圧縮による実験であるのは、圧縮率と復元画像の画質の両方を考慮しなければならない非可逆圧縮とは違い、可逆圧縮は復元画像が原画像であり、多視点動画画像の圧縮率だけを容易に評価できるためである。

表 2 から表 4 に、(A)Soccer, (B)Lifting, (C)Dance のそれぞれの多視点動画画像での (2), (4), (6) 式の値を示す。また、圧縮率とこれらの関係を表したものを図 10 に示す。ここで圧縮率とは、圧縮対象視点の原画像のデータ量に対する図 6 の (f) のような差分画像のデータ量の比である。表 2~表 4 から、それぞれの多視点動画画像に対して、視点 2, 視点 3, 視点 3 の $V(i)$ の値が最小であり、図 10 から $V(i)$ の値が最小の時、多視点動画画像を最も圧縮していることがわかり、(A)Soccer, (B)Lifting, (C)Dance のそれぞれの視点間圧縮最適視点は、視点 2, 視点 3, 視点 3 といえる。また実験値を線形近似すると、単調増加の関係が得られたことにより、提案手法による多視点動画画像の圧縮で最も効率良く圧縮できる、視点間圧縮最適視点の決定方法の有効性が確認できた。

(2) 提案手法による圧縮の効率性

提案手法による多視点動画画像の圧縮が非効率になる、つまり、圧縮率が 1 を超えると考えられるのは、第 5 章で述べたように、(3), (5) 式を満たす時であり、これらの式の閾値 Th_1 , Th_2 と多視点動画画像の圧縮率の関係を表した実験結果を図 11 と図 12 に示す。図 11 で、提案手法による視点間の冗長性を考慮した圧縮が非効率である、つまり圧縮率が 1 を超えているのは、予測できる背景領域が小さくなるためである。また、(A)Soccer, (B)Lifting, (C)Dance のどの多視点動画画像でも、提案手法の圧縮率が 1 を超えているのが、 $\alpha_j = 0.7$ ぐらいであるのがわかる。よってこの実験結果より、(3) 式の閾値 Th_1 を 0.7 とした。図 12 で、 γ_{jmax} の値が大きくなるにつれて、圧縮の効率が悪くなっているのは、前景領域の平面近似の精度によるもので、画像の大きさに対する前景領域の比率が大きくなると、5.2

(a) 視点 1

(b) 視点 2

(c) 視点 3

(A) Soccer

(a) 視点 1

(b) 視点 2

(c) 視点 3

(d) 視点 4

(B) Lifting

(a) 視点 1

(b) 視点 2

(c) 視点 3

(C) Dance

図 9 多視点原画像
Multiple-view original image.

(a) Soccer

(b) Lifting

(c) Dance

図 10 圧縮率と $V(i)$
Compression rate and $V(i)$.

(a) Soccer

(b) Lifting

(c) Dance

図 11 圧縮率と閾値 Th_1
Compression and threshold Th_1 .

でも述べたように、平面近似の精度が次第に悪くなるためである。ここで (A)Soccer, (B)Lifting, (C)Dance のそれぞれの画像の大きさに対する前景領域の比率は、それぞれ 0.01~0.1%, 1~2%, 10~12% である。これらの三つの異なる前景領域の大きさと得られた圧縮率との関係から、線形近似をして、それから提案手法による圧縮が非効率になるのは、 $\gamma_{jmax} = 0.14$ であるので、(5) 式の閾値 Th_2 を 0.14 とした。

図 12 圧縮率と閾値 Th_2
Compression rate and threshold Th_2 .

6.2 非可逆圧縮実験

MPEG のように、指定されたデータ量になるような、フレーム間予測符号化を行う非可逆な動画画像圧縮を行った場合の提案手法の有効性を確認するために、動画画像に対して 5 フレームづつ片方向フレーム間予測符号化を用いた動画画像圧縮を行い、視点間の冗長性を考慮しない場合（従来法）と、冗長性を考慮する提案手法の両者において、同じデータ量になるように圧縮した場合の画質を比較した。

視点間の冗長性を考慮しない場合の圧縮データは、圧縮基準視点・圧縮対象視点ともに動画画像を視点毎に独立したものとして扱い、動画画像圧縮を行ったデータである。一方、提案手法による圧縮データは、4.4 の圧縮方法で述べたように、圧縮基準視点の動画画像を動画画像圧縮したデータ、そして、背景・前景の各領域毎の平面射影行列を求めるための視点間の対応点の組合せデータ、さらに図 6 の (f) のような視点間差分画像の動画画像を動画画像圧縮したデータとなる。

この比較実験では、各フレームに対して均等に量子化を行い、量子化のレート制御は行わなかった。また、すべての視点のデータ量を同一とし、インターネットにおけるストリーミングビデオ等における利用を想定して、256kbps から 1150kbps を各視点に割当てた場合の実験を行った。なお、すべての視点のデータ量を同一としたので、圧縮基準視点の復元画像は、視点間冗長性を考慮しない場合と、それを考慮する提案手法の間で同一となる。そこで、両者の手法の画質の比較は、圧縮対象視点画像について行った。

画質の比較には、次式に示すような、画像の輝度値に対する PSNR を用いた。

$$PSNR = 10 \log_{10} \frac{255^2 N}{\sum (p - q)^2} \quad (7)$$

ここで、 N は画像内の画素数、 p は復元した画像の輝度値で、 q は原画像の輝度値である。

(1) 復元画像の例

図 13 に圧縮対象画像に対する復元画像の例を示す。提案手法による圧縮での復元画像は、4.4 の復元方法で述べたように、フレーム間予測符号化に基づいて動画画像圧縮された視点間差分画像を利用して復元している。また図 13 の復元画像は、6.1 で求めた視点間圧縮最適視点である視点 2、視点 3、視点 3 を圧縮基準視点とし、圧縮対象視点である視点 3、視点 2、視点 1 の画像を復元した結果である。提案手法による視点間の冗長性を考慮した圧縮から復元した画像は、どの多視点画像も知覚的に原画像とほとんど変わらない画像が得られているのがわかる。

(2) 復元画像の画質

図 14 に、各視点ともに同一ビットレート (800kbps) で、提案手法に基づいて動画画像圧縮を行った場合と、視点毎に独立に動画画像圧縮した場合（従来法）のそれぞれについて、図 13 に示した圧縮対象視点における復元画像の画質を PSNR で評価し比較した結果を示す。

図 14 より、画質の差は、(A)Soccer では、2~3dB、(B)Lifting では、1~1.4dB、(C)Dance では、0.8dB~1.2dB 程度あり、提案手法による視点間の冗長性を考慮した圧縮の方が、視点毎に独立して動画画像圧縮したものよりも復元画像の画質がより良いことがわかり、提案手法による視点間の冗長性を考慮した圧縮の有効性が確認できる。また、他の圧縮対象視点についても同様な結果が得られた。

図 15 では、圧縮するビットレートを変化させたときの、図 13 に示した圧縮対象視点における復元画像の画質を比較した。図 15 の縦軸は、動画画像の全フレームで求めた PSNR を平均したものである。図 15 から、どのようなビットレートで圧縮しても、提案手法による視点間の冗長性を考慮した圧縮の方が、視点毎に動画画像圧縮したものより良い結果が得られているのがわかる。特に、(A)Soccer と (B)Lifting では、1150kbps で視点毎に独立に動画画像圧縮したものの復元動画画像と、512kbps で提案手法による視点間の冗長性を考慮した圧縮の復元動画画像の画質がほぼ同じ、またそれ以上の画質が得られている。これは、同じ画質の復元画像を得るために、提案手法による視点間の冗長性を考慮した圧縮に必要な圧縮データが、視点毎に独立して動画画像圧縮するのに必要な圧縮データの半分以下で良いということになり、データ量が膨大である多視点動画画像の圧縮に効果的であることが示唆されている。

(3) 視点間のデータ配分と復元画像の画質

ここで示した実験では、上記のように各視点について均一のビットレートを割当てたものである。しかし、多視点画像全体のビットレートは変化させなくても、視点によりビットレートの配分を変化させれば、圧縮データから復元

(a) 原画像

(b) 提案手法

(c) 従来法

(A) Soccer

(a) 原画像

(b) 提案手法

(c) 従来法

(B) Lifting

(a) 原画像

(b) 提案手法

(c) 従来法

(C) Dance

図 13 復元画像の例
Examples of reconstructed image.

(a) Soccer

(b) Lifting

(c) Dance

図 14 同一ビットレート符号化時 (800kbps) の PSNR による画質評価の例
Examples of image quality under same bitrate(800kbps).

(a) Soccer

(b) Lifting

(c) Dance

図 15 ビットレートと画質評価の例
Examples of bitrate and image quality.

される画像の画質も変化すると考えられる。

図 16 に、3 視点の動画 Soccer について、3 視点全体のビットレートは 2400kbps のままで、3 視点それぞれに配分するビットレートを変化させて提案手法により圧縮し、圧縮対象視点で復元した画像の画質を示す。この実験例からは、3 視点に同一のビットレートを割当てるよりは (Case 2)、圧縮基準視点に割当てるビットレートを高めた方が画質が向上し (Case 1)、逆に圧縮基準視点に割当てるビットレートを低くしてしまうと画質が劣化してしまう (Case 3)、という傾向がわかる。しかし、このような傾向は圧縮対象となる多視点動画の性質に依存すると考えられるので、ビット配分を最適に決定することは容易ではない。このような視点間のビットレート配分の最適化については、今後の課題としたい。

7. ま と め

視点間の射影幾何関係を利用し、視点間の冗長性を考慮した多視点動画の圧縮方法を提案した。その提案手法の圧縮方法を様々な多視点動画に対して適用し、視点毎に独立して圧縮するよりも、より高圧縮でより良い復元画像の画質を得ることができることを確かめ、提案手法による視点間の冗長性を考慮した多視点動画の圧縮方法の有効性を確認した。

また、提案手法による圧縮が最も効率が良くなる圧縮基準視点、つまり視点間圧縮最適視点の決定方法について提案し、様々な多視点画像 (前景の大きさが異なる) に対して実験により、その提案方法の有効性を確認した。

提案手法による多視点動画の圧縮方法により、本論文で示した結果よりも、より高圧縮でより画質の良い復元画像を得るためには、視点間差分画像の精度を良くすればよい、つまり視点間予測画像の精度を良くすればよいと考えられる。視点間予測画像の精度に関わってくるのは、提案手法では、対象シーンを平面近似される複数の領域より構成されるものと考えているため、平面と近似できる領域の分割方法を今後検討していきたい。また本論文で提案した視点間圧縮最適視点の決定方法を、より様々な多視点動画に対して適用し、さらにその有効性を確認していきたい。

【文 献】

- 1) Eye Vision : "http://www.pvi-inc.com/eyevision/".
- 2) Manex Entertainment, Inc., : "http://www.mvfx.com,"
- 3) S.E.Chen, L.Williams : "View Interpolation for Image Synthesis," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.20, no.11, November 1998.
- 4) N.Inamoto, H.Saito : "Fly Through View Video Generation of Soccer Scene," International Workshop on Entertainment Computing(IWEC2002), Workshop Note, pp.94-101, May 2002.
- 5) 北原格, 石川寛亨, 渡辺真生, 大田友一 : "大規模空間の多視点映像を用いた運動視差の再現可能な自由視点映像の生成・提示方式," 画電学誌, 31, 4 (2002)
- 6) 秦泉寺久美, 渡辺裕, 岡田重樹, 小林直樹 : "MPEG-4 スプライト符号化を用いた超低レート動画圧縮," 信学論, J84-D-II, 5, pp.758-768 (2001)
- 7) Pohsiang Hsu, K.J.Ray Liu, Tsuhan Chen : "A Low Bit-Rate Video Codec Based on Two-Dimensional Mesh Motion Compensation

(a) 視点 1

(b) 視点 3

	視点 1(kbps)	視点 2(kbps)	視点 3(kbps)	合計 (kbps)
		基準視点		
Case 1	600	1200	600	2400
Case 2	800	800	800	2400
Case 3	900	600	900	2400

図 16 ビットレートの割当ての違いによる画質の評価 (Soccer)

Image quality for different bitrate assignment (Soccer).

with Adaptive Interpolation," IEEE Trans.Circuits Syst. Video Technol.vol.11,Jan.2001.

- 8) Yan Yaoping, Wu Chengke : "A Novel Video Coding Scheme Using Delaunay Triangulation," Journal of Visual Communication and Image Representation, Vol.9, Mar.1998.
- 9) Haluk Aydinoglu, Monson H. Hayes : "Stereo Image Coding: A Projection Approach," IEEE Trans. Image Processing vol.7, Apr.1998.
- 10) 畑幸一, 栄藤稔, 千原國宏 : "多視点画像の高効率符号化," 信学論, J82-D-II, 11, pp.1921-1929 (1999)
- 11) 境田慎一, 苗村昌秀, 金次保明 : "背景差分法と時空間 watershed による領域成長法を併用した動画オブジェクトの抽出," 信学論, J84-D-II, 12, pp.2541-2555 (2001)

あきやま
秋山 修

2002 年、慶應義塾大学理工学部情報工学科卒業。2004 年、同大学大学院理工学研究科開放環境科学専攻修士課程修了。在学中、コンピュータビジョン・画像処理の研究に従事。現在、日本アイ・ビーエム (株) 勤務。

さいとう ひでお
斎藤 英雄

1987 年、慶應義塾大学理工学部電気工学科卒業。1992 年、同大学大学院理工学研究科開放環境科学専攻修士課程修了。同年同大学助手、1995 年、専任講師、2001 年助教授。博士 (工学)。コンピュータビジョン、画像処理等に関する研究に従事。1997 年から 1999 年まで、米国カーネギーメロン大学ロボティクス研究所訪問研究員 (日本学術振興会海外特別研究員)。2000 年から 2003 年まで、科学技術振興事業団さきがけ 2 1 研究員 (情報と知)。電子情報通信学会、情報処理学会、日本 V R 学会、画像電子学会、計測自動制御学会、IEEE 会員。