

自由視点画像合成に基づく光線空間符号化手法の提案と評価方法の検討

Proposal of Ray-Space Coding Based on Free-Viewpoint Image Synthesis and Its Evaluation Method

田口 裕一†
Yuichi Taguchi

苗村 健†
Takeshi Naemura

1 はじめに

実空間を光線情報の集合として記述し、自由視点画像を合成する手法として、光線空間の概念が提案されている [1-3]。4次元光線空間のデータは、被写体空間を多数の視点から撮影した画像群をもとに構築される。これを補間処理することにより、その被写体空間を様々な視点位置から見た画像を自由に合成することができる。ここで、合成画像の品質を向上させるためには、被写体空間をより多数の視点から、より密な間隔で撮影することが必要となる。このため、カメラアレイ [4] やレンズアレイ [5] を利用したシステムは近年ますます大規模化し、そのデータ量も膨大なものとなってきている。

このような光線空間データの効率的な蓄積、伝送のためには、圧縮符号化の技術が必要不可欠となる。筆者らは、圧縮効率も勿論のこと、画像合成時のコストや、アプリケーションの利用形態などを総合的に考慮した、機能的な符号化が必要であると考えている。そこで、これまでに自由視点画像合成技術を前処理として利用することで、新たな機能性を付加する圧縮符号化技術について検討を進めてきた [6, 7]。本稿では、これまでの検討を踏まえ、新たな符号化方式の枠組みを提案するとともに、その枠組みに適した評価手法について検討する。さらに、インテグラルフォトグラフィから構築される光線空間データを対象として、実証実験を行い、提案手法の有効性を検証する。

2 自由視点画像合成に基づく光線空間符号化

2.1 提案符号化手法の概要

提案する符号化手法の概要を図1に示す。提案手法では、まず入力多視点画像 (a) を利用して、ある与えられた視点位置 (b) における合成画像および被写体の視点依存幾何モデル (c) を求める。この視点位置には、被写体空間を俯瞰する代表的な視点位置や、ユーザが被写体空間を眺めている視点位置などが考えられる。圧縮率を最大化する観点から自動的に与えることや、製作者の意図を反映するために手動で与えることも可能であろう。筆者らの提案する枠組みは、何らかの方法で与えられた視点位置に対応する合成画像を、光線空間符号化の出発点として位置づけるものである。以下では、このような視点位置を「代表視点位置」、それに対応する合成画像を「代表視点画像」と呼ぶことにする。代表視点画像は、復号時に最初に復元される画像となり、これは被写体空間のサムネイルと見なすこともできる。次に、代表視点位置における視点依存幾何モデルと代表視点画像から、各入力画像の予測画像 (d) を合成する。この予測画像と入

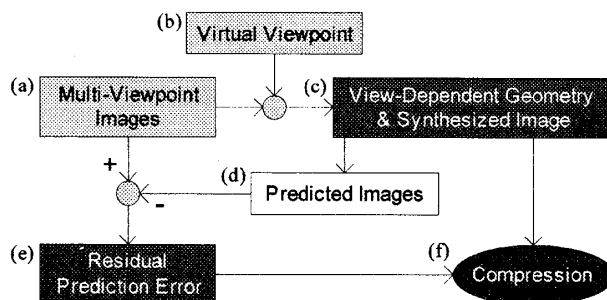


図1: 提案符号化手法の概要

力画像との差分を取り、予測残差 (e) を求める。

以上の手順により、代表視点画像、視点依存幾何モデル、予測残差が求められる。これらをこの順番で符号化し、符号ストリームを作成する (f)。この代表視点画像を被写体空間のサムネイルとして、従来の2次元画像ビューアなどで閲覧できるような符号ストリームを設計することにより、従来技術との親和性の高い符号化方式を提供することが可能となる。

2.2 提案手法の位置づけ

光線空間の符号化手法は、以下の3つに分類して考えることができる。

- (1) 光線空間を多次元空間として扱う手法
- (2) 光線空間を入力多視点画像群のまま処理する手法
- (3) 被写体の幾何モデルを合成利用する手法

(1) は画像内の相関と画像間の相関を同等に扱う手法であり、ベクトル量子化 [3] や多次元 DCT [8]、フラクタル符号化 [9] などが利用されている。(2) は画像間の予測を利用した手法であり、一般に視差補償予測符号化 [10] と呼ばれる。動き補償予測と組み合わせることにより、これを動空間に拡張する手法 [11] などが存在する。(3) はポリゴンなどの幾何モデルを用いる3次元CG技術との親和性が高い手法である [12, 13]

提案手法は、代表視点画像を最初に復号できることおよび、2次元画像向けシステムとの親和性が高い符号ストリームを作成することを目標とし、代表視点位置における合成画像と視点依存幾何モデルを符号化の出発点とする。これは、(3) の手法のように被写体の完全な幾何モデルを求めるのではなく、代表視点位置における2次元情報のみを光線空間から抽出し、モデル化を行うことに相当する。また、(2) の手法では参照画像として入力多視点画像を利用するが、提案手法では合成画像を参照し入力画像を予測する。つまり、光線空間を代表視点画像に対応する2次元断面と、その他の付加情報に分離して符号化を行うことになる。

†東京大学大学院情報理工学系研究科, Graduate School of Information Science and Technology, The University of Tokyo

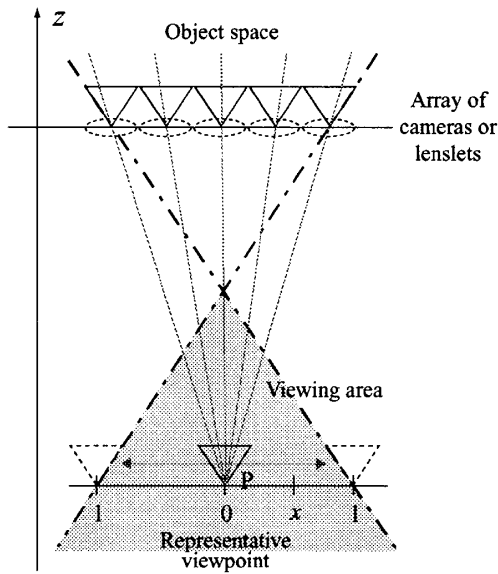


図2: 提案手法の評価方法

2.3 提案手法の評価方法

従来の符号化手法は、圧縮後の光線空間から合成される自由視点画像が、あらゆる視点位置において一様な品質を持つように設計されてきた。一方、提案手法は、代表視点位置を中心に据えた手法であり、この視点位置から離れるにしたがって、予測が難しくなり、伝送すべき残差成分の情報量が増えることになる。例えば、遠隔地

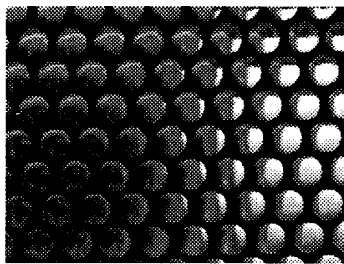
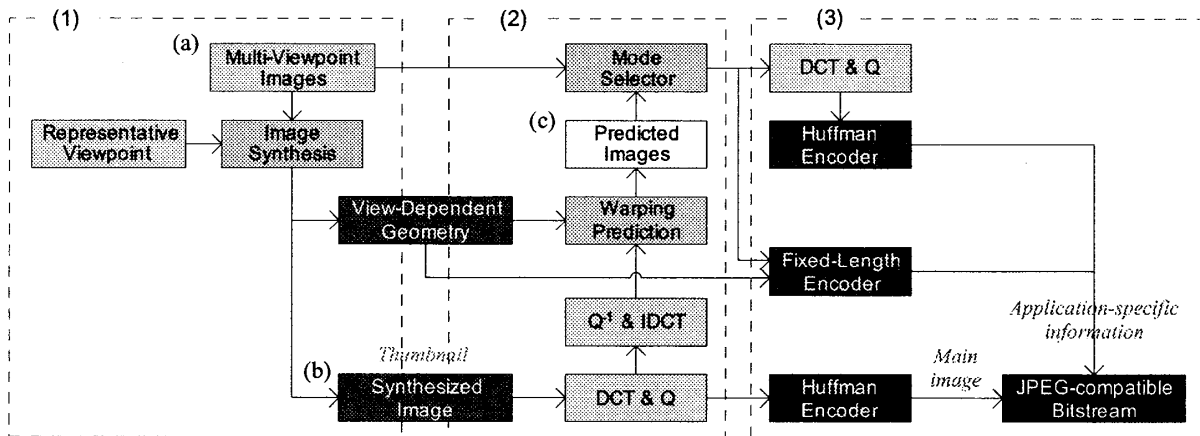
のユーザがある視点位置で被写体空間を眺めているとき、そのユーザの視点動き得る範囲を重点的に符号化して情報を送るような動空間伝送方式においては、合成品質の均一性よりも、重要な部分をいかに少ない情報量で伝送するかが重要な課題になってくる。このような観点から提案手法を評価するには、入力多視点画像ではなく自由視点合成画像の品質を評価する必要がある、また、品質の平均値だけではなくその分布が重要な意味を持つてくる。そこで本稿では、提案する光線空間符号化の枠組みに適した評価手法について検討する。

具体的には、図2のように、符号化時に利用した代表視点位置を中心として、距離 x を変化させながら合成画像の品質を評価することを考える。距離 x の円周上において、圧縮前の光線空間から合成した画像と圧縮後の光線空間から合成した画像の PSNR 値を算出し、その平均や分散を求めることで、合成画像品質の視点位置依存性を評価することができる。

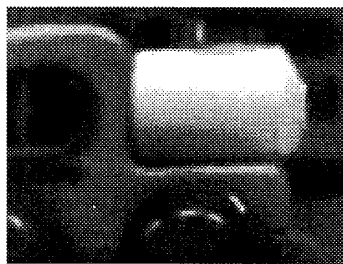
なお、図2における視域(ユーザ視点位置の可動範囲)は、画像合成の際に全ての入力視点の情報を利用できる視点位置という条件で定義している。これは、今回利用する LIFLET [5] により取得されたインテグラルフォトグラフィに対して妥当な条件であり、視域の定義は入力画像の種類によりそれぞれ異なる。

3 インテグラルフォトグラフィに対する提案手法の実装

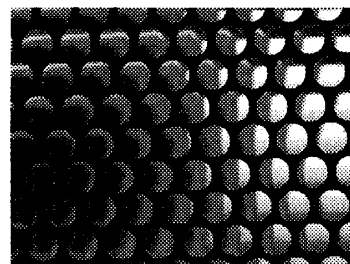
2章で提案した符号化の枠組みにおける、1つの具体的な実装例を図3に示す。ここでは、実装の詳細を3つのステップに分けて説明する。



(a) A part of multi-viewpoint images



(b) A synthesized image (thumbnail)



(c) A part of predicted images

図3: 符号化手順の詳細とインテグラルフォトグラフィを利用した画像例

(1) 代表視点位置における画像の合成と幾何モデルの構築

まず、多視点画像と代表視点位置を入力とし、代表視点位置における画像と幾何モデルの合成を行う。ここで入力多視点画像として用いたインテグラルフォトグラフィは、図3(a)のような、レンズアレイを利用して取得された画像である。円領域の画像は要素画像と呼ばれ、それぞれのレンズ位置から見た被写体空間の情報を記録している。

ここでは、文献[14]のアルゴリズムに従って画像合成を行った。これは、ステレオマッチングを用いて、要素画像1つに対し1つの奥行きを推定し画像合成を行う手法である。これにより、画像合成に利用する要素画像の個数と同じ数の頂点を持つ、視点依存のポリゴンモデルが構築される。画像合成は、推定した奥行き情報を利用して要素画像の一部分を切り出し、ポリゴンモデルにテクスチャマッピングすることにより行われる。

(2) 予測画像の合成と符号化モードの選択

代表視点画像と幾何モデルを利用して、予測画像を合成する。ここでは、ワーピングを利用して入力多視点画像の予測を行った。符号器と復号器で生じる量子化ひずみの差異を吸収するため、予測の際には局所復号した合成画像を利用する。予測画像の品質にはばらつきがあるため、 16×16 画素のマクロブロックごとに、符号化モードを以下の3つに分けて符号化する。

- 予測のみ
- 予測および残差符号化
- イントラ符号化

予測画像におけるマクロブロックの画質が、あらかじめ決められた q_{min} より高ければ、そのマクロブロックの付加情報は符号化されない。一方、品質が低い場合には、予測残差ともとの入力画像で符号化評価値の比較を行い、評価が高い方を符号化する。入力画像が選択された場合には、予測画像は利用されないためイントラ符号化となる。なお、この評価値の比較には、MPEG-2 Test Model 5 [15] で利用されている基準と同等のものを利用している。

また、符号化に先立って、円形の入力画像や予測残差画像を矩形にするために、水平パディング処理を行う。

これは、無効画素によって発生する情報量を最小限に抑えるための措置である。

(3) 符号ストリームの作成

最後に、代表視点画像、幾何モデル、マクロブロック情報をまとめて符号ストリームを生成する。この際、代表視点画像をJPEG画像として符号化し、それ以外を付加情報としてアプリケーション拡張領域に格納することで、通常のJPEGビューアでも代表視点画像を再現することが可能な符号ストリームを作成することができる[7]。

なお、幾何モデルやマクロブロック情報は等長符号化、符号化するマクロブロックの画像にはDCTベースの符号化を行っている。専用のビューアでは、これらの付加情報を利用した自由視点画像合成が可能となる。

4 実験と評価

4.1 実験の概要

実験で利用した画像は、図3で示したインテグラルフォトグラフィである。これは、筆者らのグループが検討を行っているLIFLET [5]により取得されたものであり、XGAの解像度を持つカメラ1台でレンズアレイを撮影している。ここでは、画像合成の際に598個(横26×縦23)の要素画像を利用したため、この要素画像群を符号化の対象とした。

本稿では、マクロブロックの画質を管理するパラメータである q_{min} を変化させながら符号化を行った。代表視点位置は、撮影に用いたレンズアレイ面(幅28mm)の中心から200mm引いた位置とした。また、比較対象として、以下の2つの方式を利用した場合の特性を評価した。

- インテグラルフォトグラフィ画像を1枚の画像とみなし、JPEG圧縮を行う(Intra)。
- それぞれの要素画像を切り出し、パディング処理を行った後にJPEG圧縮を行う(Padding + Intra)。

なお、圧縮率は以下のようにして求めた。すなわち、提案手法では代表視点画像・幾何モデル・マクロブロック情報の符号量の合計を、比較対象の2つの方式ではイントラ符号化された画像の符号量の合計を、XGAの解像度を持つインテグラルフォトグラフィ画像のデータ量で割ることにより求めた。

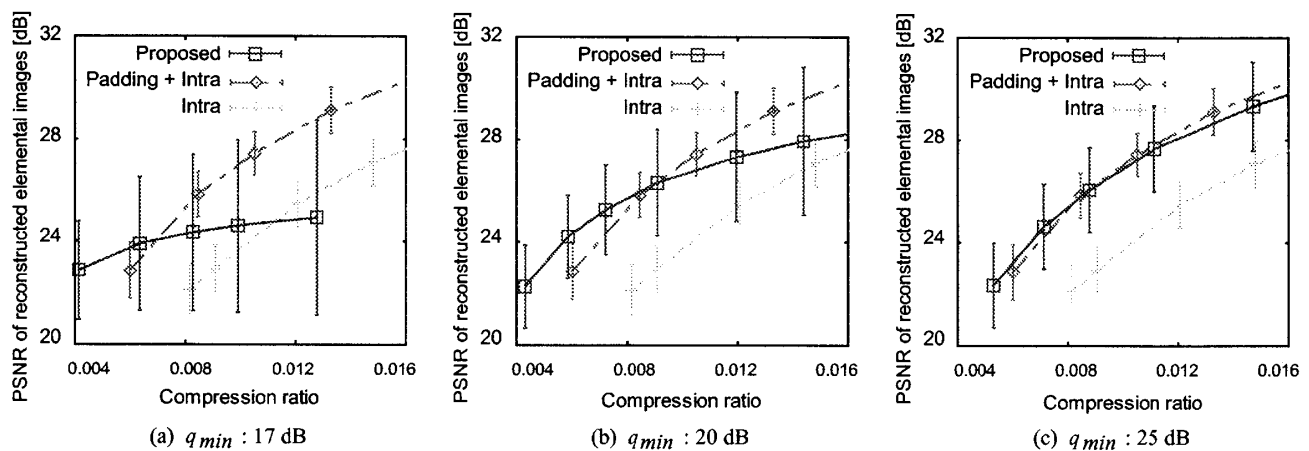


図4: 再構成される要素画像のレート歪特性

4.2 復号される要素画像のレート歪特性

符号化したデータから要素画像を復号した場合のレート歪特性を図4に示す。図4では、それぞれの要素画像に対して輝度値のPSNRを求め、その平均値と標準偏差が示されている。PSNRは要素画像の有効画素領域、つまり円形の内部の領域で計算した。

提案手法の特性は、他の2つの方式に比べると、平均値では低ビットレート(高圧縮率)において良好であり、全体を通して分散が大きいことが分かる。これは、要素画像ごとの品質のばらつきが大きいことを意味しており、高品質の部分(グラフ中の標準偏差を表すバーの上部)で比較すると、視点依存型の圧縮方式である提案手法が勝っている部分もあり得ることが推察される。そこで、2.3節で述べたような、視点位置に依存した合成画像の品質評価を行う。

4.3 視点位置に依存した合成画像品質の評価

図2で定義した視域において、合成画像の品質を評価する。符号化時に利用した代表視点位置(点P)を0、奥行き方向を固定した場合の最大可動距離を1として、距離が x の円周上18箇所の視点に対応した画像を合成し、PSNRを測定した。圧縮率は0.007で一定とし、距離 x における合成画像のPSNRの平均値と標準偏差を示したものが図5である。

これより、提案手法では再現される合成画像の品質に視点位置依存性があることが分かる。代表視点位置付近の品質は高く、そこから離れるにしたがって品質が下がる。つまり、提案手法は代表視点位置に対して重要度を設定できる符号化方式であると言える(圧縮率を下げることで一定の品質を保証することも可能であるが、特に高圧縮率の場合に視点依存性が顕著になる)。また、 q_{min} のパラメータを制御することにより、同じ符号量でも重要度の重み付けを変化させられることが分かる。一方、全ての要素画像をイントラ符号化した場合には、合成画像の品質は視点位置によらずほぼ一定である。

このような提案手法の特徴は、例えば3次元動空間通信において、ユーザの現在の視点位置情報をもとに、次のフレームではその周辺の画質を重視した符号化を行うなどのアプリケーションにおいて有効であると考えられる。この場合は、サムネイルにはJPEG画像ではなく、MPEG動画などを利用することが可能である。

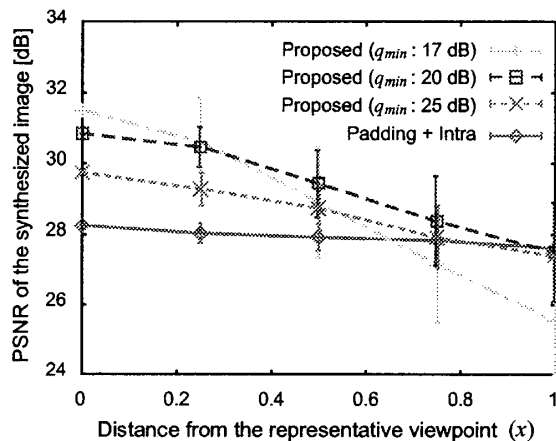


図5: 視点位置による合成画像の歪特性(圧縮率: 0.007)

5 まとめ

本稿では、自由視点画像合成に基づく光線空間符号化の枠組みを提案し、その評価方法の検討を行った。また、インテグラルフォトグラフィ入力に対する具体的な実装を行い、再構成される合成画像の品質が視点依存性を持つことを示した。今後は、代表視点位置の選択手法を検討し、動空間通信システムの構築を進めていく。

謝辞

本研究を進めるにあたり有益な御助言を頂いた、東京大学の原島博教授、高橋桂太氏に深く感謝致します。

参考文献

- [1] 柳澤 健之, 苗村 健, 金子 正秀, 原島 博, “光線空間における3次元物体の操作,” 信学ソ大, D-169, 1995.
- [2] 藤井 俊彰, 金子 正秀, 原島 博, “光線群による3次元空間情報の表現とその応用,” テレビ誌, vol. 50, no. 9, 1996, pp. 1312–1318.
- [3] M. Levoy, P. Hanrahan, “Light Field Rendering,” Proc. ACM SIGGRAPH’96, Aug. 1996, pp. 31–42.
- [4] B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, M. Levoy, “High Performance Imaging Using Large Camera Arrays,” Proc. ACM SIGGRAPH 2005.
- [5] T. Yamamoto, M. Kojima, T. Naemura, “LIFLET: Light Field Live with Thousands of Lenslets,” ACM SIGGRAPH’04 Emerging Technologies, etech.0130, Aug. 2004.
- [6] 田口 裕一, 苗村 健, “自由視点画像合成に基づく光線空間符号化方式の基礎検討,” 信学総大, D-11-144, Mar. 2005.
- [7] 田口 裕一, 苗村 健, “光線空間符号化のための自由視点画像合成に関する検討,” 3次元画像コンファレンス, July 2005.
- [8] 高野 孝英, 苗村 健, 原島 博, “仮想オブジェクト面を用いた空間符号化,” 信学論, D-II, vol. J82-D-II, no. 10, 1999, pp. 1804–1815.
- [9] 苗村 健, 原島 博, “自己相似モデリングによる多眼3次元画像の補間と情報圧縮,” テレビ誌, vol. 48, no. 10, 1994, pp. 1215–1221.
- [10] M. Lukacs, “Predictive Coding of Multi-Viewpoint Image Sets,” Proc. IEEE ICASSP, 1986, pp. 521–524.
- [11] 岡 慎也, 圓道 知博, 藤井 俊彰, 谷本 正幸, “Multi-directional Picture を用いた動的光線空間情報圧縮,” 信学技報, IE2004-94, Dec. 2004, pp. 15–20.
- [12] M. Magnor, P. Ramanathan, B. Girod, “Multi-View Coding for Image-Based Rendering Using 3-D Scene Geometry,” IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 11, Nov. 2003, pp. 1092–1106.
- [13] M. Kitahara, H. Kimata, K. Kamikura, Y. Yashima, “Progressive Compression of Surface Light Fields,” NTT Tech. Review, vol. 2, no. 8, 2004, pp. 27–34.
- [14] S. Mitsuda, T. Yamamoto, K. Takahashi, T. Naemura, H. Harashima, “Interactive View Synthesis from Integral Photography using Estimated Depth Information,” Proc. SPIE Three-Dimensional TV, Video, and Display II, vol. 5243, Sept. 2003, pp. 116–124.
- [15] MPEG-2 Test Model 5, <http://www.mpeg.org/MPEG/MSSG/tm5/>.