

画像の線形化に基づく 任意光源方位 HDR 画像列の可逆圧縮

武田 涼平¹ 日浦 慎作¹ 浅田 尚紀¹

概要: 本研究では, 対象物体を様々な方位から照明した画像列を可逆圧縮する方法について検討した. 平行光源により照明される拡散反射物体は, 一般に 3 枚の基底画像の線形結合により表すことができる. しかし, 実際の画像には鏡面反射や影, 相互反射などの影響が含まれる. そこで入力画像列の線形化によりこれを拡散反射成分のみの画像 (線形化画像) に変換し, 少数の基底画像の線形結合により表すとともに, 実画像と線形化画像の差分である残差画像を生成する. 残差画像は多くの画素で画素値が 0 に近く, エントロピー圧縮の効果が高いため, 入力画像列をそのまま可逆圧縮する方法に比べ情報を失うことなく小さなデータ量で原画像を表現できる. 実験では 16bitHDR 画像を入力とし, 基底画像の枚数と各画素のビット深度を変化させた場合の圧縮率の変化について検証した.

キーワード: HDR 画像, 可逆圧縮, 線形化画像

Lossless-Compression of HDR Images under Various Lighting Conditions based on Photometric Linearization

RYOHEI TAKETA¹ SHINSAKU HIURA¹ NAOKI ASADA¹

Abstract: We have investigated lossless compression method for HDR images of an object illuminated under various lighting conditions. The object consists of only diffuse reflection illuminated by directional light source can be expressed by the linear combination of three basis images. However, real images include various photometric factors such as shadow and specularity. Therefore, we convert input images into ones with only diffuse factor by photometric linearization, and residual images are calculated as the difference between input images and linearized images. Since most pixel values of residual images are close to 0, high compression ratio is achieved by popular entropy compression. For 16bit HDR input images, we examined the compression ratio with various numbers and bit depths of basis images.

Keywords: HDR image, Lossless compression, Linearized image

1. はじめに

近年, 物体が人間に与える質感に関する研究が行われている. 質感とは, 物体の素材, 手触り, 柔らかさなど様々な情報を人に与える. 実画像のみならず, CG 分野においても質感に関する研究が行われており, 近年では, 非常に

実物体に近い質感を生成することができる. しかし, これらの画像は, 反射特性や形状, 光源を変えながら反射現象をシミュレートしたものであり, 画像に含まれた情報がどのように質感認知に影響しているのかは未だによく分かっていない. そこで, 物体のどのような情報が人間にどのような質感を与えているのかという研究がなされている.

この質感研究の一環として, 本研究では, 物体の質感を画像として保存する手法について述べる. 物体の質感を画像に保存することにより, その画像を研究の素材や, 脳科

¹ 広島市立大学 大学院情報科学研究科
Graduate School of Information Sciences, Hiroshima City University

学・心理物理学などにおける視覚刺激など、多くの分野において利用できるような提供することを本研究の目的とする。

物体の質感を画像に保存するためには、主に二つの問題がある。一つ目は、輝度値のダイナミックレンジである。現在、一般に扱われている画像は 8bit である。一方、現実の輝度のダイナミックレンジは 10^{10} といわれており、8bit 画像では、現実の輝度のダイナミックレンジは収めることができない。そこで近年では、HDR 画像という輝度のダイナミックレンジを広げた画像が用られるようになっている。本研究では、この HDR 画像を直接撮影できる HDR カメラで物体を撮影し、質感を保存する。

二つ目は、記憶容量の問題である。物体の見かけの質感は、視点位置と照明環境に依存する。よって、ある物体の質感をくまなく収めるためには、様々な観測方向、照明環境下でその物体を撮った画像を保存しなければならないため、記憶容量が膨大になってしまう。そこで本研究では、物体とカメラを固定し、光源方向のみを変えた HDR 画像を任意光源方位 HDR 画像（以後、入力画像群と呼ぶ）とし、この任意光源方位 HDR 画像列を効率良く可逆圧縮する手法を提案する。

画像圧縮の研究の多くは不可逆圧縮であり、不可逆圧縮に比べ圧縮率の低い可逆圧縮の研究は乏しいと言える。また、HDR 画像の圧縮についても、近年まで広く用いられることがなかったこともあり、十分な研究が行われていない。

2. 提案手法

光源方向のみを変えた入力画像群は、輝度の分布は大きく異なるが、光源の方位に対して一定の法則に従い変化する。この法則を利用すると、入力画像群の拡散反射成分を少ない枚数の画像（以後、基底画像と呼ぶ）の線形結合で表すことができる。入力画像群を 3 枚の基底画像で表わされるように拡散反射成分のみにすることを画像の線形化と呼ぶ。入力画像群が拡散反射成分のみの場合、入力画像群を線形化することにより、記憶に要する容量を大幅に減らすことができる。しかし、実際には、鏡面反射や影があるため、少数枚の基底画像では、入力画像群の鏡面反射や影を表すことができない。よって、鏡面反射や影を残差として残すことにより可逆性を与える（図 1）。

変換後のデータ量は、入力画像の枚数が膨大になることを考慮すると、数枚の基底画像と係数組の容量は無視できるため、残差画像のデータ量にほぼ等しいといえる。また、残差画像は後に述べるように入力画像に比べ情報エントロピーが小さく圧縮率が高い。よって、残差画像を可逆圧縮することにより圧縮率を向上させる。この手法のポイントは、入力画像に比べてなるべく情報エントロピーの少ない残差画像を生成することである。以下に、提案手法の流れを示す。

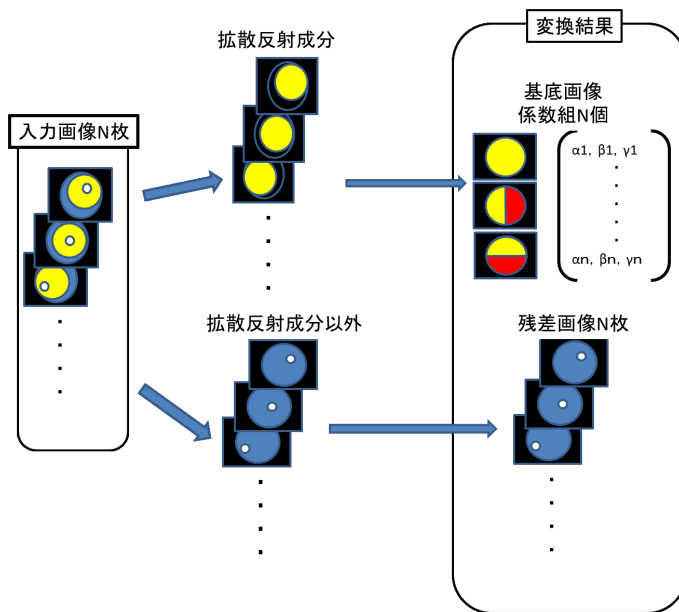


図 1 圧縮の流れ

Fig. 1 Flow of our method

(1) 入力画像群を撮影する

対象物体と HDR カメラを固定し、光源方向のみを変えて複数枚 HDR 画像を撮影することにより対象物体の質感を保存する。

(2) 入力画像群を線形化する

入力画像群を線形化することにより、入力画像群の拡散反射成分を基底画像と係数組に変換する。

(3) 残差画像を生成する

入力画像から線形化画像を用いて拡散反射成分を取り除き、残差画像を生成する。

(4) 変換結果を可逆圧縮する

基底画像、係数組、残差画像を可逆圧縮する。この可逆圧縮したときのデータ量を提案手法における圧縮結果とする。なお今回は、可逆圧縮にはハフマン圧縮を用いた。

提案手法により圧縮された画像は、基底画像、係数組、残差画像となる。展開する場合は、基底画像と係数組により入力画像の拡散反射成分を表し、残差画像を足しあわせることにより入力画像を再現する。

2.1 入力画像群の線形化

この章では、入力画像群の線形化とその手法について述べる。入力画像群を線形化することにより、拡散反射成分を少数枚の基底画像と線形結合に用いる係数組へと変換する。

Shashua[1]によれば、平行光源、完全拡散反射面を仮定することで、光源方向の異なる 3 枚の基底画像の線形結合により任意光源方向の画像を表現できる。3 枚の基底画像をベクトルで表現したものを $\hat{I}_1, \hat{I}_2, \hat{I}_3$ (互いに

独立である必要がある) とするとき任意光源方位の画像 $I_k = (i_{(k,1)} \ i_{(k,2)} \ \dots \ i_{(k,p)})^T$ は式 (1) のように線形結合で表現できる.

$$I_k = \alpha_k \hat{I}_1 + \beta_k \hat{I}_2 + \gamma_k \hat{I}_3 \quad (1)$$

ここで, $c_k = (\alpha_k \ \beta_k \ \gamma_k)^T$ を画像 I_k に対する線形結合の係数組とする. しかし, 実画像は影や鏡面反射を含むため, 式 (1) を満たさない. そこで, 式 (1) を満たすように, 拡散反射のみの画像に変換する. この変換処理を画像の線形化と呼ぶ. 式 (1) はある 1 枚の画像の輝度を表しているの, m 枚の画像群の輝度は式 (2) で表される.

$$\begin{aligned} A &= \begin{pmatrix} I_1 & I_2 & \dots & I_m \end{pmatrix} \\ B &= \begin{pmatrix} \hat{I}_1 & \hat{I}_2 & \hat{I}_3 \end{pmatrix} \\ K &= \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_m \\ \beta_1 & \beta_2 & \dots & \beta_m \\ \gamma_1 & \gamma_2 & \dots & \gamma_m \end{pmatrix} \\ A &= BK \end{aligned} \quad (2)$$

森らは, 主成分分析に光学現象の分類を行う手法を組み合わせるにより, 高速かつ高精度な画像の線形化が実現出来ることを示している [2]. 本研究では, 主成分分析と光学現象の分類に基づいた画像の線形化を用いることにより, 入力画像群を線形化する. アルゴリズムを以下に示す.

(1) 入力画像群を取得する

任意光源方位 HDR 画像列を入力画像群とする.

(2) 基底画像の画素値を求める

基底画像群行列 B を求めるため, 入力画像群行列 A の共分散行列 ($A^T A$) に主成分分析を行い, 第一, 第二, 第三主成分を 3 枚の基底画像とする.

(3) 係数組を求める

最小二乗法により, 係数組 K を求める.

(4) 線形化画像を求める

基底画像群行列と係数組の内積 BK を計算することにより, 入力画像群を線形化する.

(5) 光学現象を分類する

入力画像の画素値 $i_{(k,p)}$ とそれを線形化した画像の画素値 $i_{(k,p)}^L$ を比較して, 各画素において生じている光学現象を分類する (図 2). なお, T_s は影を除去するための閾値であり, 入力画像から数点サンプリングすることにより決定する. T_1, T_2 は本来の拡散反射成分からどれだけ離れているかを示す閾値であり, 任意に決める.

(6) 鏡面反射のにじみ領域を分類する

鏡面反射のにじみを考慮するため, 入力画像において鏡面反射領域であると分類された画素の 8 近傍が鏡面反射領域に分類されている場合は, その画素を中心に 5 ピクセル四方を鏡面反射のにじみ領域に分類する.

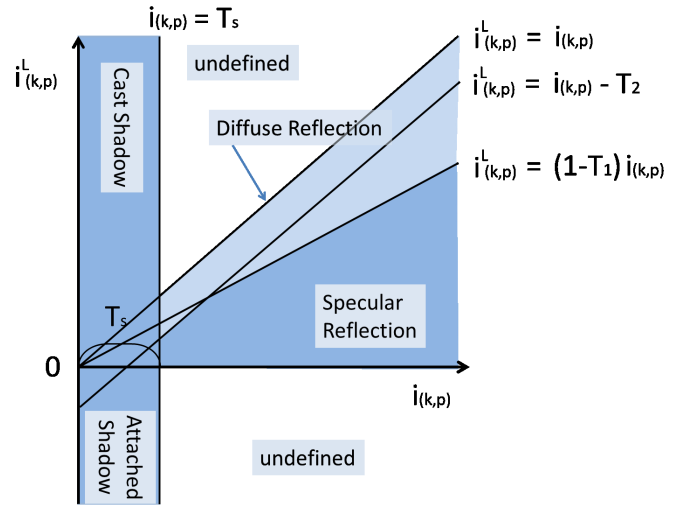


図 2 光学現象の分類

Fig. 2 Classification of photometric factors

(7) 埋め込みによりノイズ領域を除去する

鏡面反射・attached shadow・cast shadow・鏡面反射のにじみに分類された入力画像の画素に, 線形化画像の画素値を埋め込む.

(8) 閾値を少し厳しくして, 処理を繰り返す

画素値を埋め込むことによって更新された入力画像群行列をもとに, (2) から (7) の過程を繰り返す. その際, 光学現象の分類に用いる閾値 T_1, T_2 を次第に小さくしていく.

(9) 線形化画像を出力する

閾値 T_1, T_2 が最終値まで小さくなった場合は, 線形化画像を出力し, 処理を終了する. 最終値は任意に決める.

ここで, 3 枚の基底画像の線形結合で表される線形化画像を (2) ~ (4) の処理により出力している. しかし, 本研究では, 圧縮率の改善のため基底画像のビット深度と枚数を変化させる. よって, (2) において基底画像の枚数に対応して主成分の 1, 2, 3, ... と昇順に選んでいき, その基底画像のビット深度を変化させた後 (3), (4) を処理し, 線形化画像を出力する.

2.2 基底画像の変換による圧縮率の改善

本研究では, 基底画像の枚数, bit 数を変化させることにより, 圧縮結果を改善する.

理論上, 入力画像の拡散反射成分は 3 枚の基底画像の線形和により表現できる. しかし, 相互反射・環境光, ノイズなどの影響により, 3 枚では精度良く表現できない場合がある. 後に述べるように基底画像の線形和により表現された拡散反射成分は, 残差画像の生成において用いられ, 拡散反射画像の精度は, 残差画像の情報エントロピーに関係してくる. そこで, 基底画像の枚数のある範囲内で変化

させる。

基底画像のデータ量は圧縮結果に直接関係するため、なるべく小さい方がよい。そこで、基底画像の各画素のビット深度をある範囲内で変化させる。ビット深度が浅いほど基底画像のデータ量は減少するが、残差画像のエントロピーは増大すると考えられる。

2.3 残差画像の生成

入力画像から拡散反射成分を取り除くことにより、鏡面反射や影を残差画像として残す。入力画像から拡散反射成分を取り除くには、可逆性を保つため、それを線形化した画像を用いる（線形化画像）。線形化画像は負の値を含むため、正の値のみ差分をとり、負の値の部分はそのまま入力画像の画素値を用いる。

残差画像では入力画像の拡散反射成分は0になるため、拡散反射成分が画像の大部分に分布している場合、残差画像の画素値はほとんど0又は、0に近い値となる。また、影は理論的には画素値は0となるため、残差画像におけるエントロピーは鏡面反射の領域に依存する。入力画像群の鏡面反射部分が少ない程、圧縮率が高くなると考えられる。

3. 実験

実験では、提案手法により可逆圧縮率が改善されていることを確認する。また、基底画像の枚数、ビット深度を変化させることにより、圧縮結果にどのように影響するかも確認する。

3.1 シミュレーション実験

まず、ノイズの影響を受けないシミュレーション画像に対して提案手法を用いる。環境光・相互反射の影響を受けない球を対象物体とし、光源を高さ方向に4段階、水平方向に8段階動かして生成した合計32枚の画像を入力とした。画像は各画素を浮動小数で計算した後、16bit 値へ変換した解像度500×500のグレースケールである。シミュレーション画像のファイルサイズを表1に、シミュレーション画像の例とそれを線形化した画像、残差画像、基底画像3枚を図3に示す。基底画像、線形化画像は負の値を含むため、この画像では負の値を赤で表示している。残差画像（図3(c)）では、環境光・相互反射、またノイズの影響を受けないため、鏡面反射部分のみが残っており、ハフマン圧縮による効果が高いと考えられる。

次に基底画像のbit数を8~10、枚数を3~7枚で変化させた時の残差画像を可逆圧縮した結果を図4に示す。ノイズの影響を受けないため、基底画像3枚で拡散反射成分を精度良く表せ、残差画像の圧縮率は3枚から増加してもほとんど変わらないはずである。また、基底画像のビット深度は、深いほど拡散反射成分を精度よく表せるので、ビット深度が増加するにつれ残差画像の圧縮率も高くなる。

表1 シミュレーション画像のファイルサイズ

Table 1 Size of simulation image

	1枚	計32枚
ファイルサイズ (kbyte)	492	15744

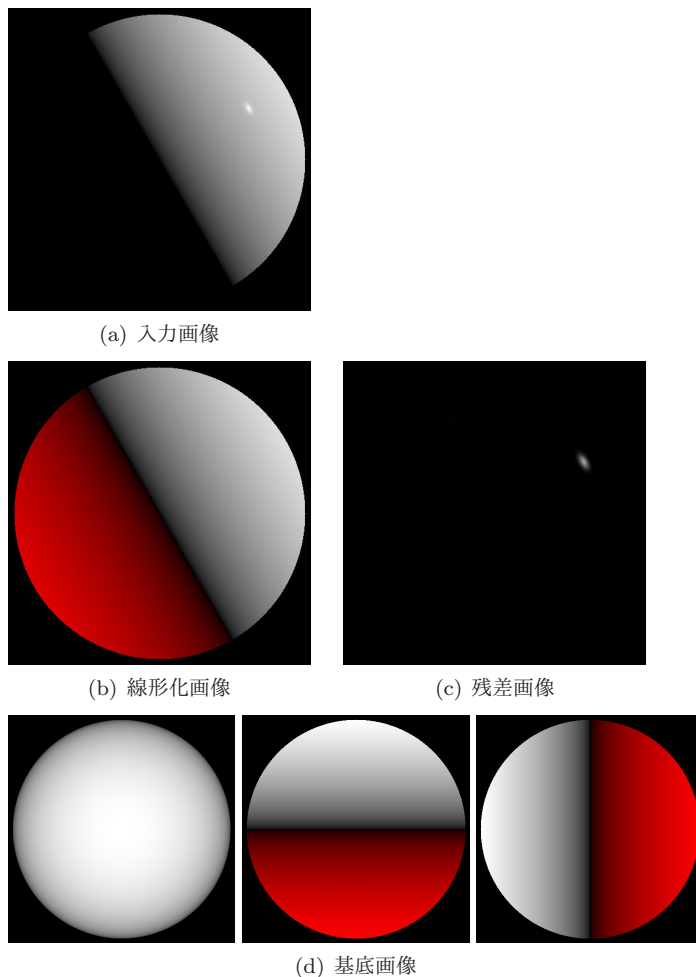


図3 シミュレーション画像の線形化の結果

Fig. 3 Result of photometric linearization for simulation images

図4を参照すると、基底画像3枚から枚数を増やしても残差画像の圧縮結果は大きく変わらないことが確認できる。また、ビット深度が深いほど圧縮率がよくなっていることも確認できる。

次に図4のファイルサイズに基底画像、係数組のファイルサイズを足したものを図5に示す。残差画像のファイルサイズは、基底画像の枚数に関わらず一定なので、基底画像が増えるにつれ圧縮結果のファイルサイズが増加していることが確認できる。そのため、基底画像の枚数が3枚のときが各ビット深度で一番良い圧縮結果となっている。最も最適な圧縮結果は、基底画像の枚数が3枚、ビット深度が16bitの時となった。

最後に、入力画像をそのままハフマン圧縮した場合と、提案手法により圧縮したサイズを表2に示す。なお提案手法について、基底画像の枚数が3枚、ビット深度が16bit

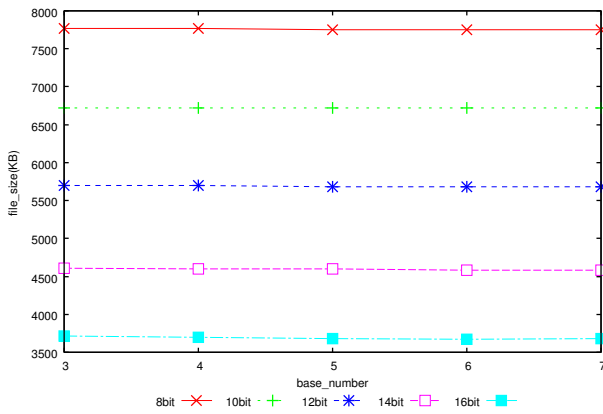


図 4 残差画像のみの圧縮結果

Fig. 4 Result of compressing only residual images

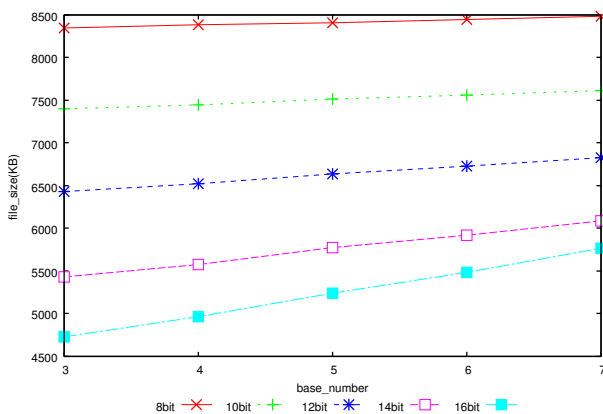


図 5 提案手法による圧縮結果

Fig. 5 Result of compressing images by our method

表 2 提案手法とハフマン圧縮の比較

Table 2 Compare with our method and Huffman compression

	ファイルサイズ (kbyte)	圧縮率 (%)
ハフマン圧縮	10820	69
提案手法	4728	30

提案手法のサイズ割合 : (残差画像:79% 基底画像: 21%)

のときの圧縮結果を示す。提案手法では、ハフマン圧縮のみに比べ、倍以上に圧縮率を改善することができた。シミュレーション実験では、入力画像に相互反射・環境光やノイズの影響はなく、残差画像に鏡面反射のみが残ったため、このような結果が得られたと考えられる。また、鏡面反射部分が図 3 のように画像に対して小さい部分に分布していたため、圧縮率が非常に高くなったと考えられる。

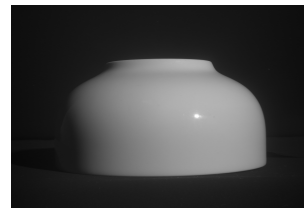
3.2 実画像実験

実画像実験では、HDR カメラとして Apogee 社の Alta16000、解像度 4872×3248 を、対象物体として、鏡面反射が起こりやすいような白いお椀を用いた。画像は、解像度 1218×812 の 16bit 画像でグレースケールとして保存した。対象物体とカメラを固定し、暗室で光源を高さ方向に 4 段階、水平方向に 8 段階動かして生成した合計 32

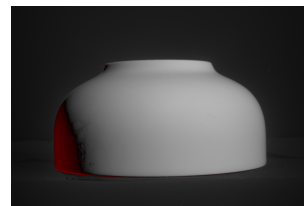
表 3 実画像のファイルサイズ

Table 3 Size of real picture

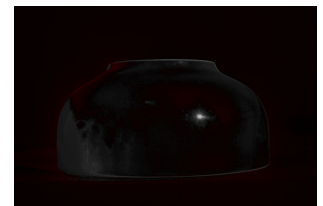
	1枚	計 32枚
ファイルサイズ (kbyte)	1932	61824



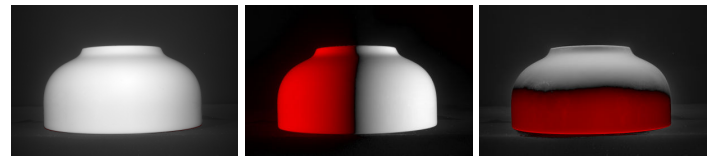
(a) 入力画像



(b) 線形化画像



(c) 残差画像



(d) 基底画像

図 6 実画像の線形化の結果

Fig. 6 Result of photometric linearization for real pictures

枚の画像を入力とした。入力画像のファイルサイズを表 3 に示す。入力画像とそれを線形化した画像、残差画像、基底画像 3 枚を図 6 に示す。基底画像、線形化画像は負の値を含むため、この画像では、負の値を赤で表示している。残差画像 (図 6 (c)) では、シミュレーション実験に比べ鏡面反射のみの画像となっていない。

次に基底画像の bit 数を 8~10、枚数を 3~7 枚で変化した時の残差画像を可逆圧縮した結果を図 7 に示す。基底画像を増加させた時の残差画像の圧縮率が、シミュレーション画像に比べ良くなっている。実画像はノイズなどを含むため、拡散反射成分を基底画像 3 枚で精度良く表現できなかったためと考えられる。また、ビット深度に関しては、12bit~16bit の間で残差画像の圧縮率はほとんど変わっていない。よって、残差画像の圧縮率は基底画像が 12bit 以上の精度でも変わらないと考えられる。

次に図 7 のファイルサイズに基底画像、係数組のサイズを加えたものを図 8 に示す。基底画像の枚数は、シミュレーション実験と同様に 3 枚の時が各ビット深度で最適な枚数となっている。最も最適な圧縮結果は、基底画像が 3 枚、bit 数が 10bit の時となった。シミュレーション画像と比較すると実画像では、線形化画像の精度は提案手法の圧縮結果に影響しにくいと考えられる。

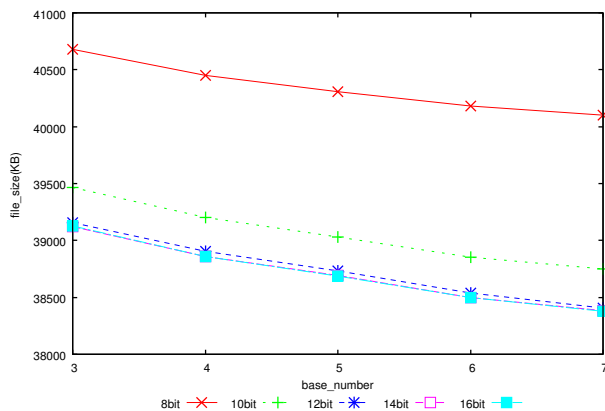


図 7 残差画像のみの圧縮結果

Fig. 7 Result of compressing only residual images

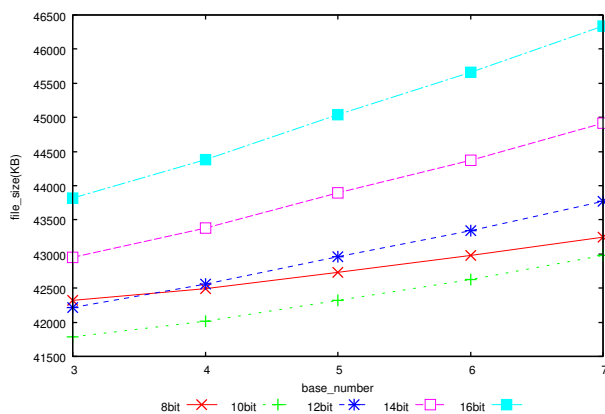


図 8 提案手法による圧縮結果

Fig. 8 Result of compression images by our method

表 4 提案手法とハフマン圧縮の比較

Table 4 Compare with our method and huffman compression

	ファイルサイズ (kbyte)	圧縮率 (%)
ハフマン圧縮	48448	78
提案手法	41783	68

提案手法のサイズ割合 : (残差画像:94% 基底画像:6%)

最後に、入力画像をそのままハフマン圧縮した場合と提案手法により圧縮したサイズを表 4 に示す。なお提案手法について、基底画像が 3 枚、ビット深度が 10bit のときの圧縮結果を示す。提案手法では、ハフマン圧縮のみよりも、約 1 割の圧縮率改善をすることができた。しかし、シミュレーション画像に比べ圧縮率が悪い。これは、実画像のノイズなどの影響により、拡散反射成分を精度良く表現することができず、鏡面反射成分のみの残差画像を生成できなかったためと考えられる。図 6 (c) ではわかりにくいですが実際の値を調べると、物体と影との境界線付近で 1000 以上の大きな値が見られた。また、残差画像において、拡散反射成分を取り除いた部分以外(影となっている部分)では、拡散反射を取り除いた部分に比べ 0 に近い値をもっておらず、それらも残差画像の圧縮率に影響したと考えられる。

4. 考察

今回は、相互反射・環境光やノイズのないシミュレーション画像と実際に撮影した実画像について提案手法を用いた。シミュレーション実験では、基底画像の枚数を 3 枚以上増やしても残差画像の圧縮率に大きな変化はなかった。また、ビット深度では、8~16bit の間で、ビット深度の向上にともない残差画像の圧縮率が向上した。実画像実験では、基底画像の枚数を 3 枚から増やすことにより、残差画像の圧縮率が向上した。ビット深度では、12~16bit の間では、残差画像の圧縮率がほとんど変わらなかった。上記のことをふまえ、実画像が相互反射・環境光やノイズを含んでいることを考慮すると、相互反射・環境光やノイズを含まない画像では、基底画像の枚数は 3 枚とし、ビット深度を変化させることにより圧縮結果を最適化するのが良いと考えられる。逆に、相互反射・環境光やノイズを含んでいる画像では、基底画像のビット深度よりも、基底画像の枚数により圧縮結果を最適化するのが良いと考えられる。また、シミュレーション実験と実画像実験では、やはりシミュレーション実験の方が良い圧縮結果を得られたため、提案手法の圧縮率は、相互反射・環境光やノイズがないほうがより良くなると考えられる。

今回、入力画像群は 32 枚としたが入力画像がさらに増えることを考慮すると、基底画像のファイルサイズは無視できるものとなってくる。表 2、表 4 から提案手法における基底画像のサイズの比率は小さいことが確認できる。また、実験において基底画像のビット深度、枚数により圧縮結果が変化することも確認している。よって、今回よりも更に枚数が多い場合は、基底画像の枚数の増加、ビット深度の増加により圧縮率の向上が見込めると考えられる。

5. おわりに

本研究は、物体の質感を保存することを目的とし、HDR カメラと対象物体を固定し光源方向のみを変化させ撮影することにより HDR 画像列として物体の質感を保存し、その際の記憶容量の問題について解決する方法を提案した。その手法は、HDR 画像列の拡散反射成分を画像の線形化に基づき少数枚の基底画像と係数組へ変換、拡散反射成分以外を残差画像として残し、それらを可逆圧縮することにより記憶容量を削減するというものであった。実験において、シミュレーション実験では 30%、実画像実験では 68% の圧縮率を得ることができた。また、基底画像のビット深度、枚数は、画像のノイズなどの状態に対して変化させることにより圧縮率の向上が見込めることが確認できた。

今後の課題としては、反射特性の異なる多様な物体に提案手法を用いた場合の検討、実画像における圧縮率の改善などが挙げられる。圧縮率の改善では、可逆圧縮法の改良

や撮影環境の改善などが考えられる。

参考文献

- [1] A.Shashua, "Geometry and photometry in 3D visual recognition", Ph.D thesis, Dept. Brain and Cognitive Science, MIT(1992)
- [2] Takahiro Mori, Shinsaku Hiura and Kosuke Sato, Shadow and Specular Removal by Photometric Linearization based on PCA with Outlier Exclusion, Int. Conf. on Computer Vision Theory and Applications, to appear. pp.221-229(2012)
- [3] 石井育規, 福井孝太郎, 向川康博, 尺長健:"光学現象の分類基準を利用した画像の線形化", 画像の認識・理解シンポジウム (MIRU2002), pp.167-176(2002)