

固有空間法による学習に基づく カラー画像からの距離画像の推定

川中 昇平 日浦 慎作 佐藤 宏介

大阪大学大学院 基礎工学研究科

560-8531 大阪府豊中市待兼山町 1-15

Tel +81-6-6850-6372 heisho@sens.sys.es.osaka-u.ac.jp

概要: 対象物体の詳細な形状情報はリアルな CG の描画に有用であるが、現在ではそれらを取得するには単なるカラー画像の取得に比べて高価・大型で特殊な機器や多くの労力・時間を必要とする。また一方で、レンガや石敷などの、CG において表面属性として扱われることの多い物体は色情報が形状情報や他の場所の色情報と相関を持つと仮定できる場合が多い。そこで、本研究では固有空間を用いて画像のもつ自己相関性を学習することでカラー画像の超解像度化を行い、また、2次元カラー画像と距離画像との相関性を学習することでカラー画像から距離画像を推定する実験を行なった。その結果、本研究で提案した手法を用いると、うまく学習データを選べば一般的な補間技術では復元できないような情報を復元でき、また、学習によってカラー画像から距離画像のバンプ値を推定できることが確認された。

キーワード: 距離画像, 学習, 超解像度, 自己相関, 固有空間法

Estimation of Range Images from Color Images Using Learning on the Eigenspace Method

Shohei KAWANAKA Shinsaku HIURA Kosuke SATO

Graduate School of Engineering Science, Osaka University

Machikaneyama-cho 1-15, Toyonaka, Osaka, 560-8531, Japan

Tel +81-6-6850-6372 heisho@sens.sys.es.osaka-u.ac.jp

Abstract: In order to show real computer graphics, shape information on objects is effective, but expensive measurement system and a lot of labors are needed. On the other hand, a image of brickwork or stone paving can be assumed that its color information is correlate with its shape or texture in different location. Therefore, in this research, we make super-resolution color images using autocorrelation leaned in color images based on the eigenspace method, and estimate range images from 2D color images using correlation learned between color images and range images. We showed that our proposal method can interpolate the detail of the texture which general interpolation cannot reconstruct if the leaning data was chosen well, and that same system make it possible to estimate bump values of range images from color images.

Keywords: Range Image, Learning, Super-resolution, Autocorrelation, Eigenspace Method

1 はじめに

我々は日常、カメラによる物体やシーンの撮影をよく行なうが、これはカメラを用いてその物体や環境そのものを記録・保存しその再現をしようとしているのではなく、その見た目だけ、すなわち視覚情報のみを記録・保存・再現することを目的をしている。また、人間は単なる2次元の写真からでも物体・環境の形状をある程度想像することができるため、この見た目だけの記録・保存・再現であっても、過去の状況を確認したり、他人に遠隔地の状況を伝達したり、撮影時には確認する時間がなかった詳細を再確認したりすることが可能になる。

しかし、通常のカメラによって撮影される画像は、撮影点から見た物体表面の明るさの情報だけを記録しているにすぎず、視点の変更ができないという問題が生じる。もし視点の変更が可能となれば、当時の状況を視点を変えながら確認したり、VRシステムによりウォークスルーして環境を体験することが可能となる。このような3次元形状提示は物体・環境の情報を多視点から表示させることができるため、2次元カラー写真のような1視点のみの画像と比べてよりリアルで正確な情報提示技術の一つとして注目されている。

この視点の変更を可能とするためには、シーンの形状情報が必要となるが、レンジファインダは非常に高価で計測時間が長いなど、欠点も多い。そのため、これまでカメラの運動により物体の見え方が変わることを用いて形状計測を行なう Shape from Motion(SFM)[1]や、画素の陰影から物体表面の法線ベクトルを復元する Shape from Shading(SFS)[2][3]、Photometric Stereo[4]等の手法が提案されている。これらはいずれも演繹的に計算して物体表面の法線ベクトルやバンプを求めているが、光源の位置や物体の反射関数等の情報が明確である必要がある。

一方、画像には連続した模様や、同じ材質で構成されている部分が数多く存在し、そういった画像におけるある1点の色度や明度といった特徴はその周りの領域と非常に高い自己相関性を持つ。特にレンガや石敷などの、CGにおいて表面属性として扱わ

れることの多い物体は、色情報が形状情報や他の場所の色情報と高い相関性を持つと仮定できる場合が多い。

そこで本研究ではビルの壁面や石敷、レンガなど、画像の持つテクスチャと各点の法線ベクトル、またはバンプ値との相関性が高い対象に関して固有空間法を用いて学習することで、カラー画像から形状を推定する手法を提案する。また、木目やレンガ、石畳などのカラー画像がもつ自己相関性を学習することで、それを低解像度画像に反映させて画像の超解像度化を実現する方法を提案する。

2 固有空間を用いた画像推定法

2.1 固有空間の生成

フラクタル的な性質を含む画像、つまり冗長な情報を含んでいる画像から、画像を記述するルールを抽出するために画像サイズ($W \times H$)より十分小さなウィンドウ($w \times h: w < W, h < H$)で画像の局所領域を複数切り出し、学習サンプル列を得る。その後、学習サンプルをラスタスキャンして画像ベクトルを生成する。

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_R], \quad \mathbf{x}_i = [x_{1,i}, x_{2,i}, \dots, x_{w \cdot h,i}]^T \quad (1)$$

ただし、 R は切り出した局所領域の数であり、この学習サンプルでは欠損領域を含まないようにする。また、各値は最大を1に正規化してあるものとする。さらに、各画素における平均ベクトルを $\mathbf{c}$ として、今後

$$\mathbf{x}_i = \mathbf{x}_i - \mathbf{c} \quad (2)$$

と置き直すことにする。この画像ベクトル列 $\mathbf{X}$ の共分散行列は次のように書ける。

$$\mathbf{Q} = \mathbf{X} \mathbf{X}^T \quad (3)$$

この共分散行列の固有ベクトルを基底に選ぶことで固有空間を生成する。実際の計算においては特異値

分解 [5] を用いて固有空間を生成する.

$$E_D = [e_1, e_2, \dots, e_D], \quad e_i = [e_{i1}, e_{i2}, \dots, e_{i(w \cdot h)}]^T \quad (4)$$

ただし, 各固有ベクトル e_i に対応する固有値を λ_i とすると,

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{w \cdot h} \quad (5)$$

である. この固有空間の次元 D は累積寄与率 (cumulative proportion) などを参考に決定する. 累積寄与率は固有ベクトル e_i に対応する固有値 λ_i を用いて

$$c.p.(D) = \sum_{i=1}^D \lambda_i / \sum_{i=1}^R \lambda_i \quad (6)$$

と定義される. 十分な累積寄与率が得られているならば, その固有ベクトルは画像のルールを抽出していると考えられることができ, このとき, 欠損領域と学習した領域との相関性が高ければこれらの固有ベクトルにより欠損箇所を復元することができる.

2.2 BPLP を用いた欠損画素の補間

学習により得られた固有ベクトル列を画像を記述するルールとして補間を行なう. 学習で設定したウィンドウサイズで切り出した画像から画像ベクトルを生成し, その i 番目のものを x_i をする. この画像ベクトルを固有空間 E へ投影すれば投影点

$$p_i = E^T x_i \quad (7)$$

へ投影される. ここで, Σ を対角行列とし, 画像ベクトル内に含まれる欠損画素に相当する要素を 0, それ以外を 1 とすると, 欠損領域を含む画像ベクトルは,

$$\begin{aligned} \hat{x}_i &= \Sigma x_i \\ &= [x_{i1}, x_{i2}, x_{i3}, 0, 0, 0, x_{i7}, \dots, 0]^T \end{aligned} \quad (8)$$

のようにかける. このとき, 欠損画素を含む場合の投影点 $\hat{p}_i$ と含まない場合の投影点 p の対応関係を

$$\hat{p}_i \cong E^T \Sigma E p_i = T p_i \quad (9)$$

で求めることができる. よって, 行列 T の逆行列を求めることができれば式 (9) の逆写像と 1 次結合を組み合わせることにより,

$$x_i^* \cong E T^{-1} E^T \hat{x}_i \quad (10)$$

で欠損した領域の画像を近似的に補間することができる [6]. しかし, T^{-1} が存在するには T のランクがフルである必要があり, 欠損画素数 (Lost Pixels) と固有空間の次元数との関係は

$$D \leq w \cdot h - (\text{Lost Pixels}) \quad (11)$$

を満たさなければならない. 画像補間の再現性は式 (10) の近似の妥当性に影響されるが, その指標としては累積寄与率を用いることができる.

補間段階では学習段階と同様の切り出しウィンドウを用い, ウィンドウサイズより小さなステップ幅で画像の左上から等間隔に切り出してウィンドウをずらし, 各ウィンドウごとに欠損箇所の補間を行なう.

以下の図 1, 2 に固有空間を用いた BPLP の学習と補間の流れを表した概念図を示す.

3 カラー画像の超解像度化

固有空間法をカラー画像に適応させる方法としては, 各画素には RGB 値が情報として格納されているので, RGB の成分毎に固有空間を生成することも考えられるが, 本研究では

$$x_i = [x_{R1}, x_{G1}, x_{B1}, \dots, x_{Gw \cdot h}, x_{Bw \cdot h}]^T \quad (12)$$

のように RGB 値を順に並べる手法をとる.

3.1 画像の拡大による冗長画素の出現

天野らは局所領域にテロップなどの画像欠損を含む画像に対して BPLP を用いた欠損データの補間 [6] を行なっているが, 本研究では画像を拡大したときに生じる冗長画素を欠損画素と見なして画像補間を行なう. 例えば, 図 3(a) のように画像を単純に 2 倍

に拡大すると、灰色部分の画素ような冗長画素が生じる。この画像は特別な補間をしなければ最近傍法により、図 3(b) のようにこれらの 4 画素内に格納されている RGB 値は全てそのウィンドウ内の左上の画素の RGB 値と同じ値を取るため、左上の画素以外の 3 画素の RGB 値は冗長な情報と見なすことができる。

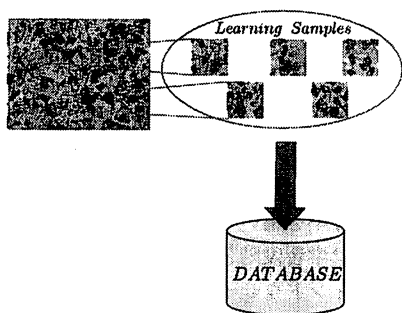


図 1: 学習段階

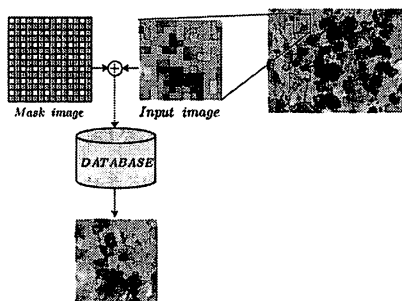


図 2: 補間段階

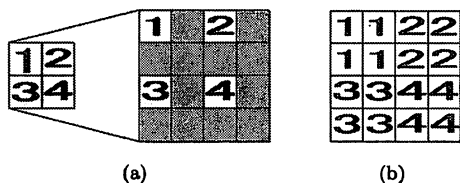


図 3: 冗長画素の出現 (a) と最近傍法による補間 (b)

3.2 超解像度化

本研究で用いる手法が適応できる画像は自己相関性が高いものである必要があるため、図 4 のような鉱物テクスチャ画像 (RGB 各 8bit, 640 × 480 画素) を用いて補間実験を行なった。まずはじめに大きさ 4 × 4 画素のウィンドウサイズを考え、学習サンプルを多数切り出した。また、獲得した学習サンプル列から固有空間を生成した。

図 4 の元の鉱物テクスチャ画像を 1/2 倍に縮小し、さらにその画像を 2 倍に拡大した (図 5)。このとき、特別な画像補間を行なわなかったため、最近傍法による補間がなされたので冗長画素が出現しているのが分かる。この画像を入力画像とし、ウィンドウ内の冗長画素を欠損と見なした。

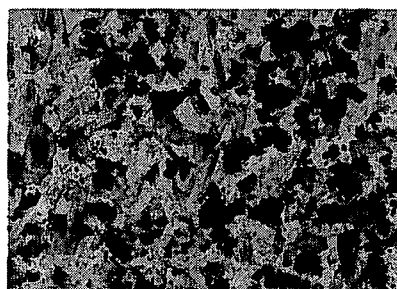


図 4: 鉱物テクスチャ画像

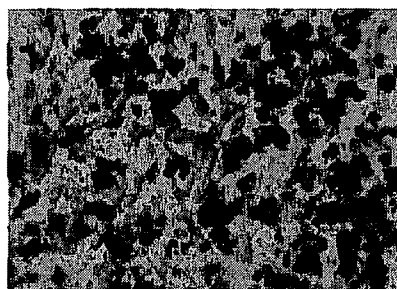


図 5: 縮小拡大画像

補間段階では、入力画像の左上から右下まで 4×4 画素のウィンドウサイズで、ウィンドウのステップ幅をウィンドウサイズの半分として、本研究の手法による画像補間を行なった。ただし、学習サンプル作成時に切り出したウィンドウと補間段階で切り出したウィンドウとは異なる部分を使用した。補間結果を図6に示す。次に、学習ウィンドウを 64×48 にし、先ほどと同様に左上から1画素ずらして学習サンプルを取得し、固有ベクトルを生成後、画像補間を行なった。結果を図7に示す。また、以下に元画像(図8(a))、入力画像(図8(b))、学習サイズ 4×4 での補間結果(図8(c))、学習サイズ 64×48 での補間結果(図8(d))の拡大図、ならびに同画像をバイリニア補間(図8(e))、バイキュービック補間(図8(f))した結果を示す。

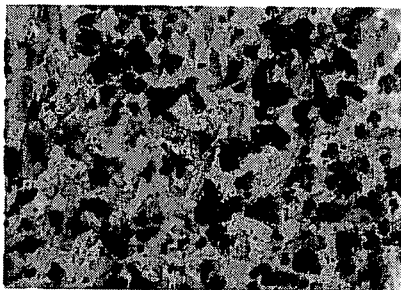


図6: 学習ウィンドウサイズ 4×4 での画像補間結果

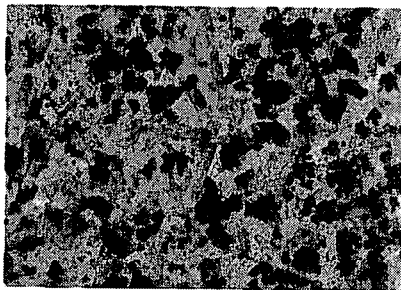


図7: 学習ウィンドウサイズ 64×48 での画像補間結果

また、本研究では誤差指標として、画像の劣化の尺度の一つであるPSNR (peak signal-to-noise ratio) を用いた。式(13)に示す。

$$PSNR = 10 \log_{10} \frac{W \times H \times 3 \times T^2}{\sum_i (x_i - x_i^*)^2} \quad (\text{dB}) \quad (13)$$

ここで T は画像の階調数 $-1(8 \text{ ビット/ピクセル画像では } T = 255, \text{ 情報量の最大値が } 1 \text{ に正規化してある画像では } T = 1)$ である。表1は各補間結果のPSNR、及び使用した固有空間の次元数と累積寄与率を表したものである。表1を見ると、 4×4 サイズのウィンドウで学習、補間した結果はバイリニア、バイキュービック補間とほぼ同程度のPSNR値を出していることが分かる。一方、 64×48 サイズのウィンドウで学習、補間した結果はPSNR値が最も低く、ノイズが多いことを示しているが、図8を見ると、バイリニア、バイキュービック補間と違い高周波成分も復元しているので、知覚的には元画像と近い結果となった。

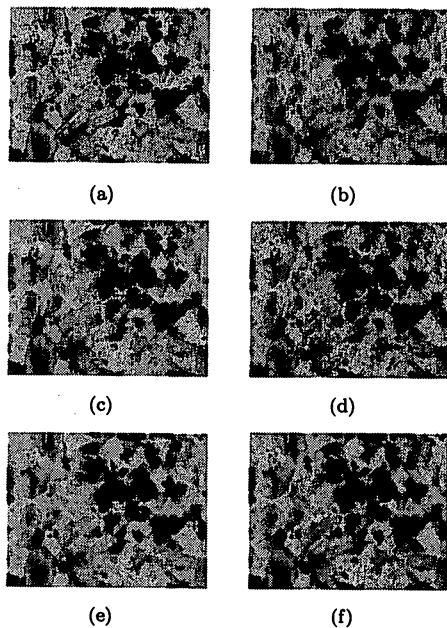


図8: 各補間結果の拡大図

表 1: 各補間結果の元画像との誤差

	PSNR	固有空間 次元数	累積 寄与率
学習サイズ:4×4	23.200	4	92.7
学習サイズ:64×48	20.431	300	99.1
バイリニア	23.641		
バイキュービック	23.843		

4 距離画像推定

カラー画像の各画素に距離画像の奥行き情報(バンプ)を付加した画像における画像ベクトルを,

$$\mathbf{x}_i = [x_{R,1}, x_{G,1}, x_{B,1}, x_{Z,1}, \dots, x_{G,w \cdot h}, x_{B,w \cdot h}, x_{Z,w \cdot h}]_i^T \quad (14)$$

とする。このベクトルを用いて固有空間を生成する。ここで $x_{Z,i}$ はカラー画像を左上からラスタ走査して i 番目の画素に対応する距離画像の奥行きの値である。ただし値は RGB 値と同様、大きさ 1 に正規化してある。つまり、 i 番目の画素の情報は $x_{R,i}, x_{G,i}, x_{B,i}, x_{Z,i}$ となる。このように RGB 値とバンプ値を同じ尺度にし、同じベクトルに組み込むことで、同一固有空間上で表現できる。カラー画像から各画素に対応するバンプ値を推定する段階では、入力画像ベクトルにおいてバンプ値をマスクするようなマスク行列を設定し、2.2 と同様の手法で値を補間することとする。

本研究の手法を実現するのに必要となるカラー画像とそれに対応する距離画像として、コニカミノルタ社の非接触式デジタイザ VIVID910 で撮影されたデータを用いた。同一 CCD からカラー画像と距離画像を得ることができるため、カラー画像とその各画素に対応した距離画像を得ることができる。本研究では RGB 値とバンプ値とが高い相関性を持っている画像が実験の対象となるので、今回はある一つの平面上に凹凸が生じているような物体を対象とした。

しかし、VIVID910 で得られた距離画像は視点に対して傾き・歪みを生じている場合が多く、この傾きを補正しないと、同じ明るさでもバンプ値が全く異なる部分が多数出てくるので、カラー画像の RGB 値

と距離画像のバンプ値との相関性が得られない。そこで、この傾きを補正するために主成分分析を用いる。画像ベクトル列を

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_R], \quad \mathbf{x}_i = [X_i, Y_i, Z_i]^T \quad (15)$$

と定義する。ここで X_i, Y_i, Z_i は i 番目の画素の距離データであり、式 (2) を満たしているものとする。この画像ベクトル列の共分散行列 $\mathbf{Q}$ の固有ベクトル列 $\mathbf{E}$ は、

$$\mathbf{E} = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3], \quad \mathbf{e}_i = [e_{1,i}, e_{2,i}, e_{3,i}]^T \quad (16)$$

となり、得られた $\mathbf{e}_1, \mathbf{e}_2$ は図 9 のようにデータの第 1 主軸、第 2 主軸を表しているのので、これを用いて以下のようにデータの傾きを補正することができる。

$$\mathbf{X}' = \mathbf{E}^T \mathbf{X} \quad (17)$$

図 10 は VIVID910 で撮影した、エジプトの遺跡の壁面に描かれたヒエログリフのカラー画像である。この図 10 のカラー画像と同時に取得した距離画像を先の手法によって補正したものが、図 11(a) である。今回の実験ではウィンドウサイズを 2×2 とし、3 の超解像度化と同様の手法で学習サンプル、固有空間を作成し、補間段階では使用する固有空間の次元を 1 とし、画像ベクトルのバンプ値を全てマスクして BPLP によるバンプ値の推定を行なった。推定結果を図 11(b) に示す。推定結果を見ると、カラー画像の RGB 情報からバンプ値が学習され、真の値ではないが概形の特徴を自然に見える程度に推定していることが分かる。次に、図 12 の今城塚古墳の石敷

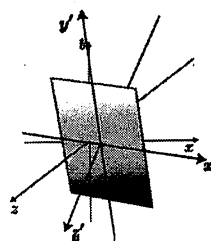


図 9: 主軸変換

(補正済み)に対して同様に学習し、推定を行なった。先の実験と同様、学習ウィンドウサイズを 2×2 、固有空間の次元を2とした。図13(a)はオリジナルの距離画像、図13(b)は図12のカラー画像から推定した結果である。また、図14(a)、図14(b)はそれぞれ図13(a)、図13(b)の拡大図を示した。先のヒエログリフと同様、概形の特徴はとらえているが、図14(a)、図14(b)から分かるとおり、細かい特徴までは復元できていないことが確認できた。これは、学習サイズが小さいために、画像の陰影とバンプ値との相関性程度しか学習できず、石敷の汚れ具合に影響して、表面ががたがたした結果になったと考えられる。表2は、ヒエログリフ、石敷のバンプ値の推定結果と元データとの誤差とPSNR、及び使用した固有空間の次元数と累積寄与率を表したものである。

表 2: 推定結果と元データとの誤差

	PSNR	固有空間 次元数	累積 寄与率
ヒエログリフ	31.623	1	84.1
石敷	34.506	2	99.3

5 考察

カラー画像補間においては、大きなウィンドウを用いて学習した場合には知覚的にはよい結果が得られたが、ノイズが生じ、PSNRが非常に低くなった。

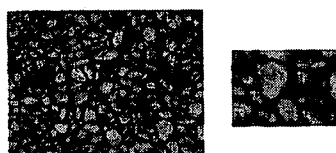
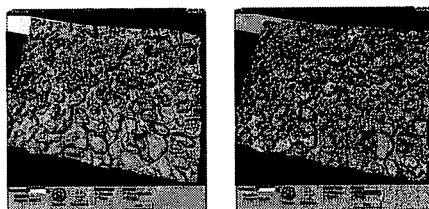


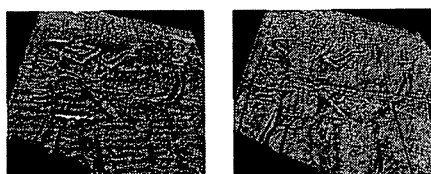
図 12: 石敷カラー画像とその拡大図



(a)

(b)

図 13: 石敷距離画像 (a) とその推定結果 (b)



(a)

(b)

図 11: ヒエログリフ距離画像 (a) とその推定結果 (b)



(a)

(b)

図 14: 石敷距離画像 (a) とその推定結果 (b) の拡大図

そのため、類似性のある別の画像で学習して、それを用いて対象画像の超解像度化をはかるには階層的に補間を行なうなどの工夫が必要となる。

また、距離画像推定においては、ヒエログリフにおいても石敷においても概形特徴は本研究の手法により推定できることが確認できた。しかし、以下のような問題点が挙げられる。まず初めに、使用できる固有空間の次元の低さである。固有空間内の非常に値の大きい固有値に対応した固有ベクトルのうち、パンプ値と強い相関を持ったベクトルがどのような学習ウィンドウを用いても必ず表れる。つまり、補間しようとしてもすぐにパンプ値に大きく依存している固有ベクトルを使用してしまうので、式(9)の行列 T が Singular になってしまい補間結果の値がオーバーフローしてしまう。次に、本研究の提案システムは単純に RGB 値のみから距離画像を推定しているため、画像において全体の明度が一定でない場合はうまく学習・推定ができないという点が挙げられる。この問題を解消するには画像の明度成分を抽出し、学習用カラー画像と入力カラー画像どちらの画像も明度を一定にするなどの対策が考えられる。

6 今後の課題と展望

本研究ではカラー画像の超解像度化を行なうにあたって、1枚の画像を拡大縮小することによって得た低解像度画像を対象に評価した。しかし本研究は遠くから撮影した画像に対して近くから撮影した画像を用いて学習したものを反映させて超解像度化をはかる、ということが本来の目的である。そのため、今後は実際に同じ物体を遠距離、近距離で撮影したものを用意して、本研究と同様の結果が得られるかどうかの検証を行なう。

距離画像推定においては、よりカラー情報と距離情報との相関性が高くなるようなパラメータの設定や学習の仕方を検討していく。また、ヒエログリフのような凹凸の小さい画像に対しては、同様の手法を用いて画像のパンプ値ではなく物体表面の傾きを推定することで、画像を回転させることによる陰影

の変化が生まれ、よりリアルな CG の提示ができることが期待される。

参考文献

- [1] D.D. Morris, K. Kanatani, and T. Kanade: "Uncertainty Modeling for Optimal Structure from Motion Vision Algorithms Theory and Practice", Springer LNCS, September, 1999.
- [2] B. K. P. Horn: "Shape from Shading: A Method for Obtaining the Shape of a Smooth Object from One View", MIT Project MAC Internal Report TR-79 & MIT AI Laboratory Technical Report 232, 1970
- [3] K. Ikeuchi and B. K. P. Hone: "Numerical Shape from Shading and Occluding Boundaries", Artificial Intelligence, Vol. 17, pp. 141 ~ 184, 1981
- [4] R. J. Woodham: "Photometric Method for Determining Surface Orientation from Multiple Images", Optical Engineering, Vol. 19, pp. 139 ~ 144, 1980
- [5] 金谷健一: "これなら分かる応用数学教室-最少二乗法からウェーブレットまで-", 6-3, 共立出版 (2003-6)
- [6] 天野敏之, 佐藤幸男: "固有空間法を用いた BPLP による画像補間", 電子情報通信学会論文誌 D-II, Vol. J85-D-II, No. 3, pp. 457-467 (2002)