

EbC 画像対を用いた三次元物体の高速線形姿勢推定

天野 敏之^{†a)} 玉木 徹^{††b)}

A Fast Linear Pose Estimation Method of 3D Object Using EbC Image Pair

Toshiyuki AMANO^{†a)} and Toru TAMAKI^{††b)}

あらまし 本論文では三次元物体の姿勢パラメータを二次元画像から高速に推定する方法として、EbC (Estimation-by-Completion) 法を提案する。EbC 法は、アピアランスベース姿勢推定を、画像に埋め込まれた情報トラックの復元という問題としてとらえ、固有空間法による学習結果をもとに BPLP 法で画像補完することで情報トラックを復元し、パラメータ推定を実現する。また、画像補完とパラメータ推定の計算を 2 枚の画像 (EbC 画像対) に集約し、各パラメータの推定を画像の内積演算と簡単な三角関数のみで実現するため、パラメータ推定を高速に行うことができる。実験では、鉛直軸周りの回転と画像面での並進の 3 自由度の姿勢パラメータ推定結果を示し、EbC 法の精度と計算コストについて述べる。

キーワード 姿勢推定, パラメトリック固有空間法, BPLP, EbC 画像対

1. ま え が き

三次元物体の位置や姿勢などのパラメータを二次元画像から推定する問題は、コンピュータビジョンの最も基本的かつ重要な課題である。この課題の解決方法は、幾何学的な形状モデルと画像を照合することにより認識を行うモデルベースの方法論と、あらかじめ多数学習していた見た目の変化との照合により認識を行うアピアランスベースの方法論に大別できる。パラメトリック固有空間法 [1], [2] で広く知られているアピアランスベースの方法は、形状モデル生成を必要としないため簡便に認識のための学習を行うことができる。また、認識方法も容易であり実用的な方法である。

パラメトリック固有空間法は固有顔 [4] を用いた顔画像認識をパラメトリックに変化する画像列に応用した手法であり、三次元物体の姿勢や照明方向の変化を固有空間上の多様体でパラメトリックに表現する手法である。このパラメトリック固有空間法は再現性が高く

簡便であるため、動画像認識 [7] や視覚サーボ [8], 照明設計 [2], shape from shading [9] など様々な応用がなされている。しかし、アピアランスベースの物体認識ではパラメータの自由度に対して積算的に増加する学習サンプルを必要とし、計算コストや記憶容量が膨大となるところに問題がある。天野ら [10] は距離画像を入力として用いることで、見た目の大きさ変化や照明条件による画像の変化をキャンセルして学習サンプルの増加を抑える工夫を示したが、濃淡画像を用いる場合には見た目の大きさや照明条件をパラメータとして学習を行わざるを得ない。しかし、学習はパラメータ推定を行う前にあらかじめ行われるため、工業応用や視覚ナビゲーションなどへの応用では、学習は現実的な時間の計算で行われればよく、認識のみ高速に行うことができればよい。

一方、岡谷ら [11] は線形モデルによる物体の姿勢推定手法を提案している。この研究では、パラメータ推定を一般化逆行列により求められた係数ベクトルと入力画像との内積演算で表現しており、学習データの記憶コストとパラメータ推定コストの大幅な削減を実現している。

本研究では、固有空間による特徴抽出と画像補完を用い、岡谷らのアプローチとは別の視点から線形モデルによるパラメータ推定を行う EbC (Estimation-by-Completion) 法を提案する。具体的には、物体の姿勢

[†] 奈良先端科学技術大学院大学情報科学研究科, 生駒市
Graduate School of Information Science, Nara Institute of
Science and Technology, 8916-5 Takayama-cho, Ikoma-shi,
630-0192 Japan

^{††} 広島大学大学院工学研究科情報工学専攻, 東広島市
Graduate School of Engineering, Hiroshima University, 1-4-
1 Kagamiyama, Higashihiroshima-shi, 739-8527 Japan

a) E-mail: amano@is.naist.jp

b) E-mail: tamaki@hiroshima-u.ac.jp

情報などを表す情報トラックを学習サンプルに付加して学習を行う。パラメータ推定は、入力される画像中の情報トラックを解読することで実現する。もちろん、カメラで取得される画像には情報トラックは存在しないため、学習で得られた固有ベクトルとの射影関係から入力画像の情報トラックを復元することでパラメータ推定を実現する。この復元のためには、学習に基づく画像補間手法である BPLP [12] を用いる。これらのプロセスは線形演算で実現できるため、画像補完とパラメータ推定の計算を 2 枚の画像 (EbC 画像対) に集約でき、各パラメータの推定を 2 回の画像の内積演算と 1 回の簡単な三角関数演算のみで実現できる。

2. 復元によるパラメータ推定

本章では、EbC 法における情報トラックを埋め込んだ学習サンプルの生成と、その情報トラックの復元について述べる。

2.1 BPLP 法の概要

ここでは天野らが提案した画像補間のための BPLP 法 [12] について簡単に説明する。各画素 x_1, \dots, x_N を要素にもつ N 次元の画像ベクトル

$$\chi = [x_1, x_2, \dots, x_N]^T \quad (1)$$

を一つのサンプルとし、 M 枚の画像 $\{\chi_1, \chi_2, \dots, \chi_M\}$ を学習サンプルとして用意する。このような学習サンプルに対して共分散行列を計算し、 D 個 ($D \leq M$) の固有ベクトル $\{e_1, e_2, \dots, e_D\}$ を求める。それらを表す固有空間を

$$E = [e_1, e_2, \dots, e_D] \quad (2)$$

とおく。ここで、補間するべき失われた画素 (欠損画素) をもつ画像 χ' が与えられたとき、 $N \times N$ 次元単位行列のうち欠損画素に対応する部分だけが 0 である欠損行列

$$\Sigma = \text{diag}(1, 1, \dots, 0, \dots, 1) \quad (3)$$

を定義する。そして画素の欠損を

$$\chi' = \Sigma \chi \quad (4)$$

とみなし、以下の式で欠損画素値を推定した画像 $\hat{\chi}$ を求める。

$$\hat{\chi} = E(E^T \Sigma E)^{-1} E^T \chi' \quad (5)$$

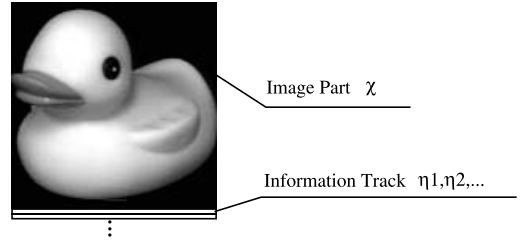


図 1 学習画像 χ と情報トラック η_1, η_2, \dots の例
Fig. 1 An example of learning sample χ and information tracks η_1, η_2, \dots

2.2 BPLP 法を用いた情報トラック復元

本研究では、パラメトリック固有空間法などの教師あり学習と同様に、パラメータの変化とともに画像ベクトル空間で滑らかな多様体を形成する画像群が得られる物体を対象とし、多様体の形状を近似し得る程度に密にサンプリングされた画像群 χ_i を学習サンプルとする。また、学習において物体のパラメータ θ は一つに限らず、多自由度のパラメータ推定を扱うものとする。以下では、物体の変化が 3 自由度で三つのパラメータ θ_j ($j = 1, 2, 3$) が得られると仮定して説明する。本研究では、パラメータ推定を行うために図 1 に示すように、画像の 1 ライン分 (幅 w) の情報トラックベクトル

$$\eta_j = [y_{j1}, y_{j2}, \dots, y_{jw}]^T \quad (6)$$

を生成する。この情報トラックベクトル η_j とパラメータ θ_j は、一意に対応するものとする。そして、各学習画像サンプル χ にこれらのトラック情報トラックベクトルを付け加え、拡張した学習サンプルベクトル

$$\zeta = [\chi^T, \eta_1^T, \eta_2^T, \eta_3^T]^T \quad (7)$$

を生成する。このようにして得られた M 枚の学習サンプル $\{\zeta_1, \zeta_2, \dots, \zeta_M\}$ に対して、BPLP 法と同様に固有空間 E を求める。パラメータ推定では、推定すべきパラメータをもつ新たな画像 χ' が与えられたとき、

$$\zeta' = [\chi'^T, \mathbf{0}^T, \mathbf{0}^T, \mathbf{0}^T]^T \quad (8)$$

なる拡張ベクトルを考え、これを情報トラック部分が欠損したベクトル $\zeta' = \Sigma \zeta$ であると考えて、以下の式で BPLP 法により補完されたベクトル $\hat{\zeta}$ を求める。

$$\hat{\zeta} = E(E^T \Sigma E)^{-1} E^T \zeta' \quad (9)$$

ただし欠損行列は

$$\Sigma = \text{diag}(\overbrace{1, \dots, 1}^M, \overbrace{0, \dots, 0}^{3w}) \quad (10)$$

である。

情報トラックの推定値を得るには、 $\hat{\zeta}$ から対応する部分を取り出せばよい。しかしここで、固有空間 E を画像部分とそれぞれの情報トラック部分に分けて

$$E = [E_I^T, E_{O_1}^T, E_{O_2}^T, E_{O_3}^T]^T \quad (11)$$

と分割して記述しておく、式 (9) を変形して、各情報トラックの推定値 $\hat{\eta}_j$ は χ' から直接次式で求めることができる。

$$\hat{\eta}_j = E_{O_j}(E_I^T E_I)^{-1} E_I^T \chi' \quad (12)$$

3. EbC 画像対によるパラメータ推定

3.1 パラメータ復元と情報トラック

前述のように、パラメータを一意に決定する情報トラックの記述の仕方には様々な方法が考えられるが、本研究では物体の姿勢角度のように周期性があるパラメータ^(注1)を記述するために、以下のような正弦波を用いる。

$$y_{ji} = K \cos\left(\frac{2\pi}{w}(i-1) - \theta_j\right) + C, \quad i = 1, \dots, w \quad (13)$$

ここで K, C は任意定数である。また w は画像の幅であるので、各情報トラック η_j は、画像の幅を周期とし、与えられたパラメータ θ_j を位相にもつ正弦波である。正弦波以外にも、三角波やデルタ関数などを用いることが考えられるが、正弦波を用いると BPLP による情報トラックは必ず正弦波に復元される^(注2)ため、パラメータを一意に決定できる利点がある。

また、正弦波の位相を求める方法としては、離散フーリエ変換が考えられるが、情報トラックは正弦波であるため、 \cos と \sin による同一角周波数のベクトル

$$\omega_c = \left[\cos 0, \cos\left(\frac{2\pi}{w}\right), \dots, \cos\left(\frac{2(w-1)\pi}{w}\right) \right] \quad (14)$$

$$\omega_s = \left[\sin 0, \sin\left(\frac{2\pi}{w}\right), \dots, \sin\left(\frac{2(w-1)\pi}{w}\right) \right] \quad (15)$$

を用意し、それらと情報トラックとの内積

$$c_j = \omega_c^T \hat{\eta}_j \quad (16)$$

$$s_j = \omega_s^T \hat{\eta}_j \quad (17)$$

を求める。すると、位相は逆正接関数

$$\hat{\theta}_j = \tan^{-1}\left(\frac{s_j}{c_j}\right) \quad (18)$$

により求めることができる。これが復元されたトラック情報から得られるパラメータの推定値である。このような計算の簡単さも、情報トラックに正弦波を用いる一つの利点である。

3.2 2枚の画像によるパラメータ推定

式 (16), (17) の計算は、式 (12) を用いると、

$$c_j = \omega_c^T \hat{\eta}_j = \omega_c^T E_{O_j}(E_I^T E_I)^{-1} E_I \chi = \{\omega_c^T E_{O_j}(E_I^T E_I)^{-1} E_I\} \chi = \Omega_{c_j}^T \chi \quad (19)$$

$$s_j = \{\omega_s^T E_{O_j}(E_I^T E_I)^{-1} E_I\} \chi = \Omega_{s_j}^T \chi \quad (20)$$

とできる。ただし、

$$\Omega_{c_j}^T = \omega_c^T E_{O_j}(E_I^T E_I)^{-1} E_I \quad (21)$$

$$\Omega_{s_j}^T = \omega_s^T E_{O_j}(E_I^T E_I)^{-1} E_I \quad (22)$$

である。つまり、情報トラックを正弦波で記述することにより、パラメータ推定は、 Ω_c と Ω_s との内積演算と、逆正接関数の計算のみで表現することができる。本研究ではこのパラメータ推定方法を EbC 法と呼ぶ。また、 Ω_c, Ω_s は、画像ベクトルと同じ N 次元のベクトルであるため、2枚の画像とみなすことができる。これらの画像を EbC 画像対と呼ぶ。

4. アルゴリズム

4.1 学習過程

Step 1. 回転ステージによる物体の回転や移動などを行い、各パラメータを変化させて物体を撮影し、 M 枚の学習サンプル $\{\chi_1, \chi_2, \dots, \chi_M\}$ を得る。

Step 2. 撮影した各学習サンプル χ に情報トラック η_j (1 自由度当り 1 ライン) を付加した画像 ζ を作成する。

Step 3. 情報トラックが付加された M 枚のサンプル $\{\zeta_1, \zeta_2, \dots, \zeta_M\}$ に対して、固有空間 E を求める。

(注1): ただし、並進移動などの周期性をもたないパラメータに $-\pi \leq \theta \leq \pi$ を割り当てると $\theta = \pm\pi$ で矛盾が生じる。そのため、並進移動の最小値を位相角 $\theta = -\pi/2$ 、最大値を位相角 $\theta = \pi/2$ に割り当てる。

(注2): この性質については付録で説明する。

$M < N$ の場合, $Z = [\zeta_1, \zeta_2, \dots, \zeta_M]$ に対して特異値分解を用いて E を計算する.

Step 4. 得られた固有空間 E を画像部分 E_i と情報トラック部分 E_{oi} に分離し (式 (11)), 各パラメータごとに EbC 画像対 Ω_{cj}, Ω_{sj} を生成し, これらの画像対を学習結果として保持する.

4.2 推定過程

Step 1. 新たな画像ベクトル χ' が与えられた場合, 各パラメータを二つの変数に対する逆正接関数により位相角を推定する.

$$\hat{\theta}_j = \tan^{-1} \left(\frac{\Omega_{sj}^T \chi'}{\Omega_{cj}^T \chi'} \right) \quad (23)$$

ただし, $\Omega_{cj}^T \chi' < 0$ のときは, 推定値に π 加算する^(注3). また, 適宜 2π を足し引きして推定値が $-\pi \leq \hat{\theta}_j < \pi$ の範囲となるようにする.

Step 2. パラメータが姿勢角度であれば位相角がそのまま推定値となる. 並進移動など非周期的なものであれば, パラメータの最大値・最小値から得られるゲインを推定位相角に掛けて推定結果とする.

5. パラメータ推定実験

5.1 EbC 画像対の生成

実験を行うにあたり, 学習サンプルとして図 2 に示す Columbia Object Image Library (COIL-20)[13] を用いた. これらの画像は解像度 128×128 画素, 256 階調のグレースケール画像である. 学習のための三次元物体の運動のパラメータは, 物体の鉛直軸周りの回転 R_θ [deg] と, 画像面に対する横方向 v_x 及び縦方向 v_y [pixel] の並進移動の 3 自由度とした. COIL-20 で提供されている画像は, R_θ は 0 [deg] より 5 [deg] 刻みであり, 1 周 72 ステップである. このうち $R_\theta = 0, 10, \dots, 350$ [deg] の 36 枚のサンプルを学習サンプル, $R_\theta = 5, 15, \dots, 355$ [deg] の 36 枚を検証用のサンプルとした. 並進 v_x, v_y は, 画像を縦横に並進移動させることで仮想的に三次元物体が並進移動したサンプルを生成した. その範囲はそれぞれ $-12 \leq v_x \leq 12, -12 \leq v_y \leq 12$ とし, パラメータの刻み Δv を 12, 6, 4, 3 [pixel] の 4 種類^(注4) に変化させて, それぞれについて実験を行った. これらのパラメータに対して, 対応する情報トラックの位相 θ_j を

$$\theta_1 = R_\theta, \quad \theta_2 = \pi \frac{v_x}{24}, \quad \theta_3 = \pi \frac{v_y}{24} \quad (24)$$

とした. また, 画像では負の値を表現することはでき



図 2 学習サンプル
Fig.2 Learning samples.

ないので, 正弦波を $K = 127, C = 127$ に設定し, η の最小値と最大値をそれぞれ 0 と 254 に抑え, これを学習サンプル画像 χ に追加して学習用画像 ζ を生成した. そして, これらの学習サンプル画像に対して特異値分解を行い, 固有ベクトル列 E を得た. 固有空間の次元は高めれば高いほど忠実な射影を行うことが期待できるが, 次元を高くすると計算コストが高くなる^(注5). 本研究では累積寄与率を目安にして固有空間の次元を, 累積寄与率 99.0% を超える最小の次元と設定した (表 1 参照). 図 3 と図 4 に EbC 画像対の一例を示す. ただし, 最小値を黒, 最大値を白として可視化している. また, 学習サンプル数, 生成した固有空間の次元, 計算時間 (3.0 GHz CPU) を表 1 に示す.

5.2 パラメータ推定

前節で説明した検証用のサンプル 36 枚 ($R_\theta = 5, 15, \dots, 355$ [deg]) について, 図 5 のように画像を ± 12 [pixel] の範囲で 1 画素ずつ縦横に並進移動させたテスト画像 ($25 \times 25 = 625$ 枚) を生成してパラメータ推定を行った^(注6). つまり, $\Delta v = 12$ の学習であれば, 縦横の並進移動量 v_x, v_y が $-12, 0, +12$ [pixel] のときにのみ学習サンプルと一致することになり, それ以外の場合は学習からパラメータが推定されたことになる. また, 姿勢角度, 横方向及び縦方向のパラメータ推定すべての計算時間 (3.0 GHz CPU) は 1 回当たりおよそ 533 [μ s] であった.

姿勢角度推定結果の一例として, $R_\theta = 60$ [deg] 及び

(注3): 実際に作成したプログラムでは, ANSI C 標準ライブラリ中の atan2() 関数を用いて計算している.

(注4): 例えば $\Delta v = 12$ のとき, 学習する並進パラメータは $v_x = -12, 0, 12$ であり, $\Delta v = 3$ のときは $v_x = -12, -9, -6, -3, 0, 3, 6, 9, 12$ である. つまり, $\Delta v = 12$ のとき, 学習サンプル数は $M = 36 \times 3 \times 3 = 324$ 枚となる (表 1 参照).

(注5): 固有空間の次元は計算コストのほかに, 汎化性能に対しても影響を与える. 詳細については 6.2 で述べる.

(注6): 横軸はテスト画像の番号, 縦軸は移動量 (v_x, v_y) = $(-12, -12), (-11, -12), \dots, (0, 0), \dots, (+12, +12)$ を意味する. テスト画像の番号の小さい順に, v_x, v_y ともに小さい値であり, まず v_y (太線) を固定して v_x (細線) を増やしている.

表 1 画像サンプル数 M , 固有空間の次元 D と計算時間
 Table 1 The results of computational times at number of samples M , eigenspace dimension D .

Δv	M	D	Comp. time
12	324	131	19 s
6	900	252	2 min. 18 s
4	1764	337	8 min. 46 s
3	2916	388	23 min. 14 s

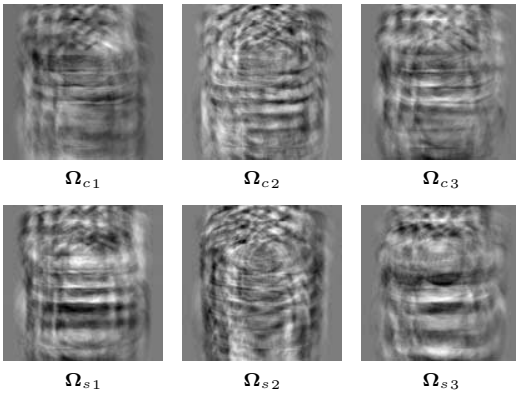


図 3 EbC 画像対 ($\Delta v = 12$)
 Fig. 3 EbC image pairs ($\Delta v = 12$).

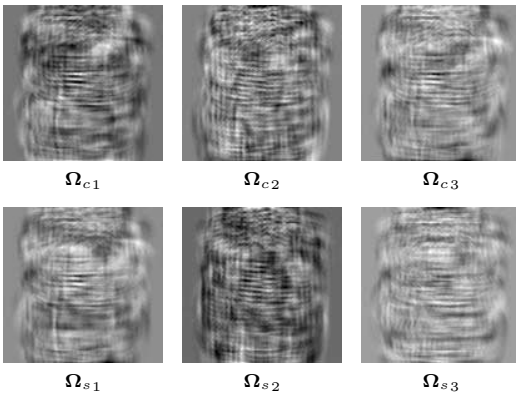


図 4 EbC 画像対 ($\Delta v = 4$)
 Fig. 4 EbC image pairs ($\Delta v = 4$).

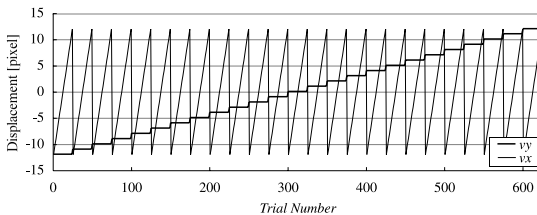


図 5 各テスト画像における並進移動量設定値
 Fig. 5 Set up values of translation at each test image.

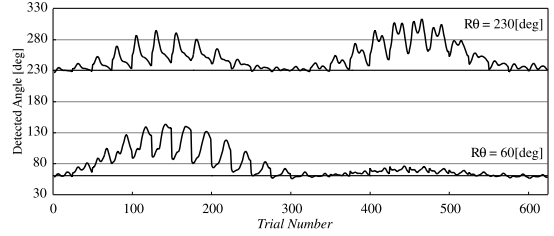


図 6 姿勢角推定結果 ($\Delta v = 12$ の場合)
 Fig. 6 Pose estimation results (In case of $\Delta v = 12$).

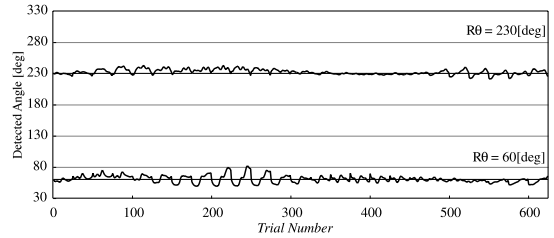


図 7 姿勢角推定結果 ($\Delta v = 6$ の場合)
 Fig. 7 Pose estimation results (In case of $\Delta v = 6$).

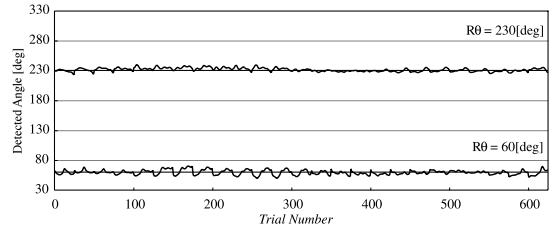


図 8 姿勢角推定結果 ($\Delta v = 4$ の場合)
 Fig. 8 Pose estimation results (In case of $\Delta v = 4$).

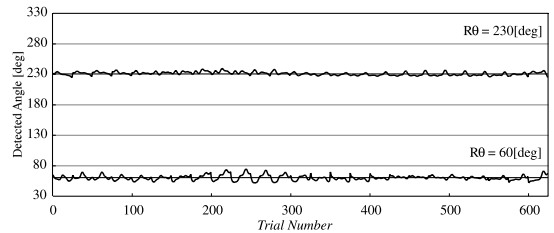


図 9 姿勢角推定結果 ($\Delta v = 3$ の場合)
 Fig. 9 Pose estimation results (In case of $\Delta v = 3$).

230 [deg] について, 各テスト画像を与えた際に推定された姿勢角度を図 6 ~ 図 9 に示す. 横軸はテスト画像番号であり, 縦軸は図 5 の対応する番号の変位量 (v_y, v_x) の画像を与えたときに推定された姿勢角を表している. これらの結果を見ると, 全体的な傾向としては Δv が小さくなるにつれて推定値が真値に近づいており, 学習サンプルを密にすることで精度が向上していることが分か

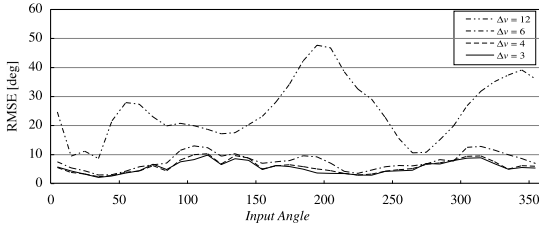


図 10 姿勢角 R_θ の最小二乗残差
Fig. 10 RMSE of estimated pose angles.

る．また理想的には，並進移動量にかかわらず，真値に一樣に近い推定値が得られることが期待されるが，特に $\Delta v = 12$ の結果では大きく湾曲した鋸状の推定結果が得られている．この理由は， $\Delta v = 12$ では $(v_x, v_y) = (-12, -12), (-12, 0), \dots, (+12, +12)$ [pixel] の 9 種類の並進移動量でしか学習を行っていないためである．つまり，テスト画像番号の 1, 313, 625 付近では， v_y について学習している画像に近い画像がテスト画像として与えられており推定精度がよく，それ以外では学習済みの画像とは異なるため，推定結果が真値から離れて大きく湾曲している．またのこぎり歯状になっているのは，各 v_y に対して $v_x = -12, 0, +12$ [pixel] が学習済みのため真値に近くなり，それ以外では推定結果が真値から離れるためである．

図 10 に，各姿勢角のテスト画像 625 枚についての推定値の最小二乗残差を示す．この結果を見ると $\Delta v = 12$ では全体的に誤差が大きく，テスト画像の姿勢角が 195 [deg] のとき残差が最大となり 49.4 [deg] であった． Δv が小さく学習が密になるに従って，全体的に誤差が小さくなる傾向にある．しかし $\Delta v = 6$ 以下では，推定精度にあまり大きな変化は見られない．

図 11 と図 12 に並進移動量 v_x 及び v_y の最小二乗残差を，表 2 にすべて入力に対する推定精度の平均した結果を示す．図 10 と同様に， $\Delta v = 12$ のときが最も残差が大きく， Δv が小さくなるに従って残差が小さくなる傾向がある．また，表 2 より，学習において Δv を細かくすることにより，並進移動量のみならず，姿勢角度の推定誤差も小さくなっていることが確認できる．

その他の物体での姿勢パラメータ推定誤差を図 13 に示す．これらは， $\Delta v = 3$ [pixel] として学習を行い，前述と同様の方法によりパラメータ推定誤差の最小二乗残差を求めて横軸を姿勢角，縦軸を並進移動量 (v_x 及び v_y の最小二乗残差の平均) として各物体につい

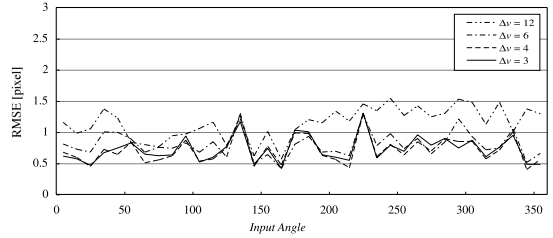


図 11 横方向並進移動量 v_x の最小二乗残差
Fig. 11 RMSE of estimated horizontal translations.

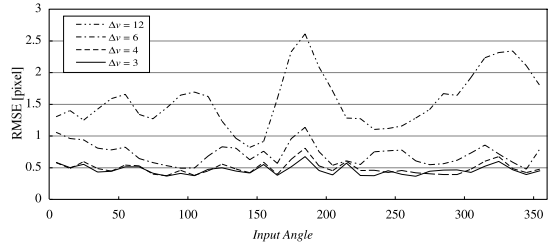


図 12 縦方向並進移動量 v_y の最小二乗残差
Fig. 12 RMSE of estimated vertical translations.

表 2 姿勢角及び並進移動量の平均推定精度
Table 2 The average of pose angle and translation estimation accuracy.

Δv	12	6	4	3
R_θ [deg]	24.8	7.76	5.86	5.31
v_x [pixel]	1.11	0.833	0.692	0.696
v_y [pixel]	1.53	0.724	0.490	0.450

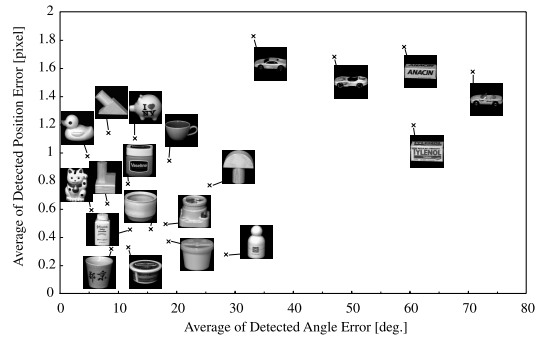


図 13 他の物体のパラメータ推定誤差
Fig. 13 The parameter estimation error of other objects.

て推定誤差をプロットした結果である．この結果を見ると，物体の形状が回転体に近いものについてはパラメータ推定誤差が小さく，自動車の模型などのように姿勢角の変化により見た目の大きさが大きく変化する物体はパラメータ推定誤差が大きかったことが確認できる．本研究で用いた画像ライブラリでは，取得され

た画像から物体領域を切り出し大きさの正規化を施しており^(注7)、これらの処理がパラメータ推定の精度に影響を与えていると予想される．具体的には、姿勢角の変化と比べて画像の拡大による画像ベクトルの変化が大きいため、固有ベクトルの生成で特徴量として姿勢角変化による画像ベクトルの変化に重みが置かれなかったことが予想される．この仮説については今後の研究で検証する予定である．

6. 考 察

6.1 パラメータ推定が可能となる条件

EbC 法は BPLP 法による線形射影により学習サンプル $\zeta = [\chi^T, \eta_1^T, \eta_2^T, \eta_3^T]$ の画像ベクトル部分 χ と情報トラック部分 η_j を対応づけている．情報トラック部分 η_j は w 次元部分空間で $C = [C, C, \dots, C]$ を中心とする円上の点に投影され^(注8)、3.1 で示したように位相角と学習サンプルのパラメータ θ_j が対応している．パラメータ推定では、まず BPLP 法による線形射影で入力画像 χ' の情報トラック η_i を推定する．その後学習で定められた η_i と θ_i の関係よりパラメータを推定していると考えられることができる．したがってパラメータの値が正しく推定される条件の一つとして、まず学習サンプル ζ の画像ベクトル部分 χ が式 (12) により対応する情報トラック部分 η_j へ誤差なく射影される必要がある．この射影が成立するためには、学習サンプル数 M が画像ベクトルの次元 N に満たないことが必要である^(注9)．つまり、画像の解像度が 128×128 画素であれば、学習サンプル数は 16384 未満でなければならない．この条件を満たす場合、学習サンプルと同一の入力に対して誤差なく情報トラックが推定でき、正しい姿勢パラメータが得られる．

もう一つの条件としては、パラメータ推定において BPLP 法により推定された入力画像 χ' の情報トラック $\hat{\eta}$ が正しい位置に投影されることが必要である．以下では、ある一つのパラメータ θ_j について考える．入力画像 χ' のパラメータの真値を θ' とする．また画像ベクトル空間において χ' と二乗距離が近い二つの学習サンプルを χ_1, χ_2 、これらのパラメータの値をそれぞれ θ^1, θ^2 とするとき、

$$\chi' \approx \frac{|\theta^2 - \theta'| \chi_1 + |\theta^1 - \theta'| \chi_2}{|\theta^1 - \theta'| + |\theta^2 - \theta'|} \quad (25)$$

が成り立ち、学習サンプルが十分多いことが必要である．ただし $\chi_1 \neq \chi_2$ とする．この条件は、 χ' が画像ベクトル空間上で χ_1 と χ_2 を結ぶ線分上の点とみな

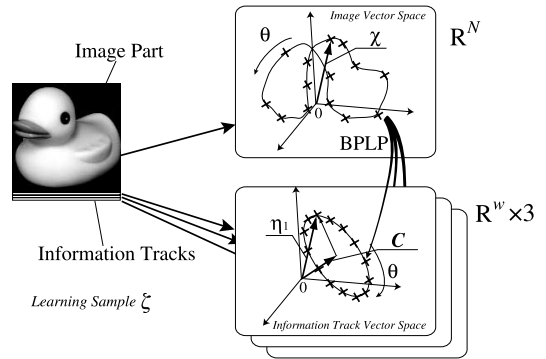


図 14 BPLP による画像ベクトルの射影
Fig. 14 Projections of Image vector to information tracks by the BPLP method.

すことができ、更に χ' が χ_1 と χ_2 間で $|\theta^2 - \theta'|$ と $|\theta^1 - \theta'|$ の比率に応じた場所に位置することを意味している．EbC 法で用いられている射影は線形射影であるため、 χ_1 と χ_2 及び BPLP 法で推定された χ' の情報トラックをそれぞれ $\eta^1, \eta^2, \hat{\eta}$ とするとき、 $\hat{\eta}$ は η^1 と η^2 を結んだ線分の上に $|\theta^2 - \theta'|$ と $|\theta^1 - \theta'|$ の比率に応じて投影される．そのため、 $\hat{\eta}$ の偏角 $\hat{\theta}$ は $\theta^1 < \hat{\theta} < \theta^2$ を満たす．このとき、学習サンプルが十分多ければ $\theta^1 \approx \theta^2$ となり、 $\hat{\eta}$ の偏角は真値 θ' に近づくため正確なパラメータ推定が可能となる．

6.2 固有空間の次元と姿勢検出精度について

また、固有空間を用いた次元圧縮は、一般的に学習サンプルの詳細が省略され、パラメータ推定精度が低下することが懸念されるが、パターン認識における次元の呪い [14] と同様に必ずしも高次元の固有空間の次元数が高ければパラメータ推定精度が向上するとは限らない．本論文で提案した EbC 法は、最終的に学習結果が画像対に集約され、入力画像との内積演算によりパラメータが推定されている．しかし、その演算は画像とパラメータの変化を共通する固有空間で記述して射影関係よりパラメータを求めることを意味しており、提案手法でも固有空間の次元によりパラメータ推

(注7): coil-20 には、切出し及び大きさの正規化を施したもののほかに処理を施していない画像ライブラリも用意されている．

(注8): $\sum_{j=1}^w (y_{ji} - C)^2 = \frac{1}{2} w K^2$ であるため η_i はすべての偏角 ($0 < \theta_j \leq 2\pi$) で C からの距離が一定である (図 14)．また、適切な α と β を用いて $\eta_i = \alpha \omega_c + \beta \omega_s + C$ と記述できるため、 η_i は C を通り、 ω_c と ω_s により張られる平面の上の円上の点である．

(注9): $\eta_j = A\chi$ と記述するとき、 η_j の要素 y_{ji} は行列 A の行ベクトル a_i を用いて $y_{ji} = a_i \chi$ と表すことができる．これは χ を変数、 a_i を係数とし、原点を通過する N 次元平面の方程式である．したがって、サンプルが N 点未満であれば y_{ji} を通る超平面が存在し、1 対 1 の射影が実現できる．

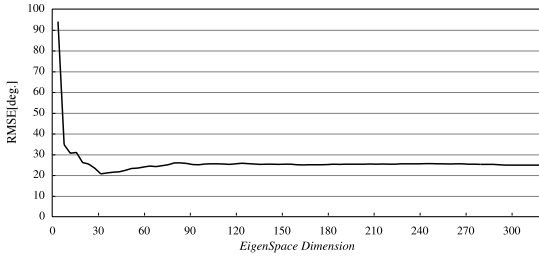


図 15 各次元での姿勢角度推定誤差

Fig. 15 Pose estimation error at each eigenspace dimension.

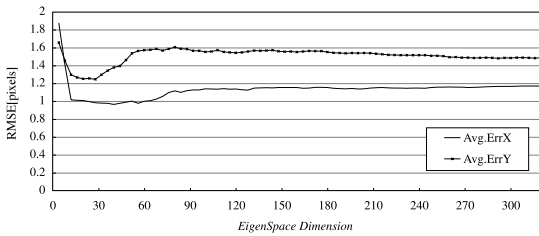


図 16 各次元での位置推定誤差

Fig. 16 Translation estimation error at each eigenspace dimension.

定精度が変化する．固有空間の次元に対するパラメータ推定精度の関係を調査した結果を示す．図 15 は姿勢パラメータ，図 16 は位置パラメータの推定誤差を示しており，縦軸は縦横の並進移動量を ± 12 [pixel] の範囲で 1 画素ずつ並進させた 625 サンプルでパラメータ推定誤差の RMSE を求めた後に，すべての画像での平均値を求めた結果を示している．また，汎化性能を評価するために，学習では COIL-20 の奇数番の画像を，パラメータ推定では偶数番の画像を用いた．いずれのパラメータも固有空間の次元を 30 次元から 40 次元付近で推定誤差が最小となっていることが分かるが，この理由としては，次元削減による汎化が関係しているものと思われる．例えば画像ノイズのようにパラメータ変化と無関係な特徴量が少数の学習サンプルに共通して含まれる場合，その特徴は高次の主成分に反映される．このような場合，固有空間の次元を高く設定すると非学習データに含まれる同様なパラメータ変化と無関係な画像ノイズによりパラメータ推定に誤差を与えることが考えられる．最適な固有空間の次元を定める方法としては，個々の応用についてはチューニングにより固有空間の次元を定めることが考えられる．しかし，一般的な議論として固有空間の次元とパラメータ推定精度との間にはどのような関係があるか

ということは説明が困難である．この関係については今後の研究で解明していきたい．

6.3 岡谷らの手法との違いについて

本研究では，情報トラックを付加して固有空間を生成し，画像補完によりパラメータ推定を行うアプローチを考えた．しかし，岡谷らが提案した手法 [11] のように，固有空間を用いずに直接画像対を算出することも可能である．式 (12) で表される線形方程式を，ある行列 A_j を用いて

$$\eta_j = A_j \chi \quad (26)$$

とすれば，学習サンプル $\{\chi_1, \chi_2, \dots, \chi_M\}$ とその情報トラックをムーア-ペンローズ型一般逆行列

$$A_j = Z_{O_j} (Z_I^T Z_I)^{-1} Z_I^T \quad (27)$$

を用いた連立方程式で結び付けることができる．ただし， Z_I と Z_{O_j} は学習サンプルの画像部分と j 番目の情報トラック列である．

しかし，式 (12) と一般逆行列を比べると，行列計算のための要素の積和回数は提案手法が $D^2 N + D^2 w + DNw$ に対して一般逆行列では $M^2 N + M^2 w + MNw$ である．また，計算すべき逆行列は式 (27) に対して式 (12) であり，大きさはたかだか $M \times M$ である．もちろん，提案手法では固有ベクトルを生成する計算コストを必要とするが，本論文で示した実験の条件では行列の積の計算コストが圧倒的に大きく，全体の計算コストを半分程度にまで削減できる．また岡谷ら [11] は，周期的ではないパラメータは一つの N 次元ベクトルで記憶しているが，本手法ではすべてのパラメータを二つの N 次元ベクトル（すなわち画像対）として記憶している．そのため，本手法によるパラメータ推定の計算コストは上述の 2 倍であるが，入力画像を二次元のパラメータ空間 [15] へ投影して偏角でパラメータを求めることが可能であり，画像の明度変化に対して頑強である上に，姿勢角度のように周期的なパラメータを矛盾なく表現することができる．

精度の点から比較すると，適切な次元を選ぶことができれば本手法の方が有利である．本手法は，すべての固有ベクトルを用いると（つまり $D = M$ のとき）岡谷らの手法と同等になるが，図 15 と図 16 で示しているように，固有ベクトル数が多いほど推定誤差が小さいとはいえない．岡谷らの手法と同等である最大次元 $D = M = 324$ の誤差よりも， $D = 30$ 程度としたときの誤差の方が明らかに小さく，適切な次元を選

んだ方が推定誤差が小さいことが分かる。本手法では累積寄与率が 99.0%になる最小の次元を D として用いるが、最初の実験では(表 1 参照) $D = 131$ であり(平均誤差である図 15 と図 16 で見る限り)、次元最大のときの誤差よりも小さいとはいえない。そのため、しきい値とする累積寄与率の値を変更するなど、誤差が小さくなる適切な次元を選ぶ方法は今後の検討課題である。

6.4 学習データの記憶容量について

本論文で提案した方法に対する従来手法としては、岡谷らの方法とパラメトリック固有空間法が挙げられる。

岡谷らの方法と比べ、提案手法は固有空間を生成する、情報トラックという概念を用いてパラメータを記述する、などの違いはあるものの、学習により数枚の画像が得られるところは同じである。提案手法では、各パラメータについて 1 対の画像が学習データとして作成される。これに対して、岡谷らの方法では並進など周期的でないパラメータについては 1 枚の画像として学習データが生成される。そのため、提案手法の学習データの記憶容量は岡谷らの方法と比較して 1.0 倍から 2.0 倍となる。本論文では、鉛直軸周りの回転と画像面での並進 2 軸をパラメータとして設定している。そのため、この設定では提案手法の記憶容量は岡谷らの方法と比べ 1.5 倍となる。また、画像サイズを 128×128 画素、4 バイトの実数型で記憶をするものとすれば、岡谷らの方法は 262 k バイトに対して提案手法は 393 k バイトである。

パラメトリック固有空間法では、学習データとして固有ベクトルと固有空間上での軌跡が得られる。この際、固有空間の次元をいくりにするかということは、認識性能に影響する上、記憶容量も変化する。そのため一概に比較することは困難であるが、仮に固有空間の次元を八次元 [6] とすると、固有ベクトルの記憶に 524 k バイト必要である。また、軌跡を投影点群として記憶すると仮定すれば、本論文の設定において $\Delta v = 6$ では学習サンプル数は 900 であり、投影点群の記憶に 29 k バイト必要である。そのため、記憶容量は合計 553 k バイトとなり、提案手法の 1.4 倍である。ただし、本論文で扱ったパラメータは並進移動であり、学習サンプルは照明条件の変化 [6] と比較すると大きな分散をもつことが予想される。そのため、パラメトリック固有空間法で本論文と同様のパラメータ推定を行うためには固有空間を八次元よりも高くする必要が

あり、記憶容量はこの試算よりも大きくなる。

7. む す び

本論文では三次元物体の姿勢パラメータを二次元画像から高速に推定する EbC 法を提案した。EbC 法はアピアランススペースのパラメータ推定手法に、画像への情報トラックの概念を導入し、BPLP 法による画像補完を用いることで復元された情報トラックよりパラメータの推定を実現した。この際に、計算過程を EbC 画像対で表現して内積演算と三角関数計算のみでパラメータ推定を実現した。今後の課題としては、回転パラメータに対するより詳細な推定実験、学習サンプル数・自由度・画像と情報トラックの次元数とパラメータ推定精度の関係の検討、EbC 画像対生成コストの省力化などが挙げられる。

謝辞 本研究に対して、有益なコメントや議論をして頂いた九州大学大学院システム情報科学研究院内田誠一氏、NTT データ坂野鋭氏、筑波大学大学院システム情報工学研究科福井和広氏に感謝致します。

文 献

- [1] S.K. Nayar, H. Murase, and S.A. Nene, "Parametric appearance representation," in *Early Visual Learning*, Chapter 6, pp.131–160, Oxford University Press, 1996.
- [2] H. Murase and S.K. Nayar, "Illumination planning for object recognition using parametric eigenspaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.16, no.12, pp.1219–1227, 1994.
- [3] L. Sirovich and M. Kirby, "Low dimensional procedure for the characterization of human face," *J. Optical Society of America*, vol.4, pp.519–524, 1987.
- [4] M. Turk and A.P. Pentland, "Face recognition using eigenface," *CVPR91*, pp.586–591, 1991.
- [5] A.P. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigenspace for face recognition," *CVPR94*, pp.84–91, 1994.
- [6] 村瀬 洋, シュリー ナイヤー, "2 次元照合による 3 次元物体認識—パラメトリック固有空間法," *信学論 (D-II)*, vol.J77-D-II, no.11, pp.2179–2187, Nov. 1994.
- [7] H. Murase and R. Sakai, "Moving object recognition in eigenspace representation: Gait analysis and lip reading," *Pattern Recognit. Lett.*, vol.17, pp.155–162, 1996.
- [8] 出口光一郎, 野口 崇, "固有空間法による視覚サーボにおける制御に適した固有部分空間の再構成," *信学論 (D-II)*, vol.J80-D-II, no.6, pp.1522–1529, June 1997.
- [9] 岡谷貴之, 出口光一郎, "固有空間法を利用した陰影からの曲面の形状復元," *情処学研報 CVIM*, vol.95, no.34, pp.1–7, 1995.
- [10] 天野敏之, 井口征士, "距離画像の固有空間法による物体

- 探索 ; 信学論 (D-II), vol.J83-D-II, no.2, pp.584-592, Feb. 2000.
- [11] T. Okatani and K. Deguchi, "Yet another appearance based method for pose estimation based on a linear model," MVA2000, pp.258-261, 2000.
- [12] 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," 信学論 (D-II), vol.J85-D-II, no.3, pp.457-465, March 2002.
- [13] S.A. Nene, S.K. Nayar, and H. Murase, "Columbia object image library (COIL-20)," Technical Report CUCS-005-96, Columbia University, 1996.
- [14] G.F. Hughes, "On the mean accuracy of statistical pattern recognizers," IEEE Trans. Inf. Theory, vol.14, no.1, pp.55-63, 1968.
- [15] 玉木 徹, 天野敏之, "マルチポート固有空間法," 部分空間研究会 Subspace2006 予稿集, pp.7-15, July 2006.
- [16] 天野敏之, 玉木 徹, "Estimation-by-Completion : 3次元物体の線形姿勢推定手法," MIRU2006 画像の認識・理解シンポジウム予稿集, pp.460-465, July 2006.
- [17] S. Mika, G. Rätsch, J. Watson, B. Schölkopf, and K.-R. Müller, "Fisher discriminant analysis with kernels," in Neural Networks for Signal Processing IX, pp.41-48, IEEE, 1999.

付 録

復元された情報トラックについて

固有ベクトルは学習サンプル ζ_j の分布に応じた重み $w_{i,j}$ を用いて一次結合により

$$e^i = w_{i,1}\zeta_1 + w_{i,2}\zeta_2 + \dots + w_{i,M}\zeta_M \quad (\text{A}\cdot 1)$$

と記述できる [17]. つまり, 第 i 固有ベクトルの情報トラック部分

$$e_{o^i} = [e_{o^i_1}, e_{o^i_2}, \dots, e_{o^i_w}] \quad (\text{A}\cdot 2)$$

は振幅と位相が異なる複数の正弦波の合成となる. また, 振幅が異なる二つの正弦波の合成は加法定理及び合成公式により

$$k_1 \cos(\omega t + \phi_1) + k_2 \cos(\omega t + \phi_2) \quad (\text{A}\cdot 3)$$

$$= -\sqrt{a^2 + b^2} \cos(\omega t + \theta_0) \quad (\text{A}\cdot 4)$$

$$\theta_0 = \frac{\pi}{2} - \tan^{-1} \frac{b}{a} \quad (\text{A}\cdot 5)$$

$$a = k_1 \cos \phi_1 + k_2 \cos \phi_2 \quad (\text{A}\cdot 6)$$

$$b = k_1 \sin \phi_1 + k_2 \sin \phi_2 \quad (\text{A}\cdot 7)$$

であり, 二つ以上の正弦波の合成についても同様に位相と振幅が変化した角周波数 ω の正弦波になることは明らかである. したがって e_{o^i} の要素は

$$e_{ok}^i = \sum_{j=1}^M w_{i,j} \left\{ K \cos \left(\frac{2\pi}{w} k + \theta_j \right) + C \right\} \quad (\text{A}\cdot 8)$$

$$= K' \cos \left(\frac{2\pi}{w} k + \theta' \right) + C' \quad (\text{A}\cdot 9)$$

となる.

BPLP による欠損の推定は, 欠損により生じた投影誤差を固有空間上で補正し, その後固有ベクトルとの一次結合により行われる [12]. そのため, BPLP により推定された欠損のない状態での固有空間上の投影点を $p = [p_1, p_2, \dots, p_D]$ とすると, 情報トラックは

$$\hat{\eta} = p_1 e_o^1 + p_2 e_o^2 + \dots + p_D e_o^D \quad (\text{A}\cdot 10)$$

と表すことができる. また, 先に示したように, 角周波数 ω で振幅が異なる複数の正弦波の合成は角周波数 ω の正弦波となるため, $\eta = [\eta_1, \eta_2, \dots, \eta_w]$ の要素 η_i は適当な定数 K'', θ'', C'' により

$$\eta_i = K'' \cos \left(\frac{2\pi}{w} i + \theta'' \right) + C'' \quad (\text{A}\cdot 11)$$

と表すことができる. したがって, 情報トラックに正弦波を用いると BPLP による情報トラック復元で正弦波が復元される.

(平成 18 年 10 月 2 日受付, 19 年 2 月 5 日再受付)



天野 敏之 (正員)

平 7 阪大・基礎工・システム卒. 平 12 同大大学院博士後期課程了. 同年, 名古屋工業大学電気情報工学科助手. 平 19 奈良先端科学技術大学院大学情報科学研究科助教. 距離画像計測応用, 固有空間法を用いた画像処理等に従事. 博士 (工学). 平 12 年度本会学術奨励賞, MIRU2006 インタラクティブセッション優秀賞受賞. 芸術科学会会員.



玉木 徹 (正員)

平 8 名大・工・情報卒. 平 10 同大大学院工学研究科情報工学専攻博士課程前期課程了. 平 13 同専攻博士課程後期課程了. 同年, 新潟大学工学部情報工学科助手. 平 15 同大学大学院自然科学研究科助手. 現在, 広島大学大学院工学研究科情報工学専攻助教. 画像処理, コンピュータビジョンの研究・教育に従事. 博士 (工学). MIRU2006 インタラクティブセッション優秀賞受賞. 情報処理学会, 映像情報メディア学会, IEEE 各会員.