

目領域の切り出しの不定性を考慮した低解像度画像からの視線方向推定

小野 泰弘[†] 岡部 孝弘[†] 佐藤 洋一[†]

本論文では、見えに基づく手法の枠組みで低解像度の目画像から視線方向を推定する手法を提案する。低解像度画像の入力を前提とすることで、被計測者をカメラの近傍に拘束しないという利点がある一方、目の領域を安定に切り出すのが困難になるという欠点が存在する。そこで提案手法では、様々な切り出しの目画像の学習パターンに対して、通常のSVD (Singular Value Decomposition) を複数のモードを取り扱えるように拡張した N モード SVD を適用することにより、目領域の切り出しの不定性に対処する。この N モード SVD を用いることにより、視線方向の変動のモードと切り出しの変動のモードを注意深く分離し、視線方向の変動を反映する特徴量を抽出する。実画像を用いた評価実験を行うことにより、提案手法がモードの分離を行わない従来の PCA (Principal Component Analysis) および、クラスごとに基底を準備する部分空間法よりも優れていることを確認した。

Gaze Estimation from Low Resolution Images Insensitive to Segmentation Error

YASUHIRO ONO,[†] TAKAHIRO OKABE[†] and YOICHI SATO[†]

We propose an appearance-based method for estimating gaze directions from low resolution images. In estimation of gaze directions from low resolution images, there exist inevitable errors in segmentation of eye regions. To improve the accuracy of gaze estimation, two key ideas are introduced in our method: using a set of training images of eye regions with artificially added segmentation error, and using N -mode SVD (Singular Value Decomposition) in order to separate image variation due to gaze directions from that due to segmentation errors. By using N -mode SVD, the feature vectors of the gaze direction can be extracted. In this paper, we describe the details of our proposed method and report experimental results demonstrating the advantage of our method over the conventional PCA (Principal Component Analysis)-based method and the subspace method in which a subspace is constructed for each class.

1. ま え が き

近年、画像処理に基づいて人間の行動を計測・理解する研究が活発に行われている。これは、人間の行動を計測・理解することにより、実世界における人間の活動を計算機などが積極的に支援することが可能になるためである。

人物の位置、頭部姿勢、手指動作などの計測対象の中でも視線方向は、人間の興味の対象を強く反映していると考えられるために特に重要である。計測した視線情報から興味の対象を推測することにより、様々な状況において人間の活動を支援することが可能になる。たとえば、駅などの公共の場所においては対象人物の

興味に応じて、電子掲示板あるいは電子広告などにより情報を提供することができる。また、商店においては客の視線の遷移履歴に基づいて、客の行動を理解したり商品の最適な配置を設計したりすることが考えられる。上記のような応用では、人間と計測装置であるカメラがある程度離れた状況において視線を計測する技術が必要になる。そこで本研究では、被計測者をカメラの近傍に拘束することなく視線を計測することを目的とする。

画像処理に基づいて視線を計測する手法は、モデルに基づく手法と見えに基づく手法の2つに大別することができる。眼球モデルを利用するモデルに基づく手法には、瞳孔-角膜反射法^{2),5),10),22)} およびエッジ検出や楕円あてはめにより虹彩あるいは瞳孔中心位置を検出する手法^{6)-8),20)} がある。しかしながら、モデルに基づく手法では眼球の3次元モデルを利用して視線

[†] 東京大学生産技術研究所

Institute of Industrial Science, The University of Tokyo

の計測を行うため、高精度に計測するためには一般に目の高解像度画像を必要とする。したがって、モデルに基づく手法では被計測者はカメラの近傍に拘束されてしまう。

一方、見えに基づく手法としては目の領域を切り出した画像を入力として、ニューラルネットワークを適用した手法^{1),12),21)} および最近傍探索を適用した手法¹⁴⁾ が提案されている。これらの手法では眼球の3次元モデルを使用しないため、高解像度画像の代わりに低解像度画像を利用してもある程度の精度で視線方向を推定することができると思われる。ところが、一般に低解像度画像を入力とした場合には画像から目の領域を安定に切り出すことは困難である。したがって、目画像の切り出しの不定性のために視線方向の推定精度が低下することが予想される。しかしながら、上記の先行研究では目画像の切り出しの不定性について考慮されていない。

そこで本論文では、被計測者をカメラの近傍に拘束しないという要求を満たすために、見えに基づく手法の枠組みで低解像度の目画像から視線方向を推定するための手法を提案する。特に、学習画像として一定の切り出しの目画像だけでなく様々な切り出しの目画像を利用することにより、目領域を正確に切り出すことができない状況においても安定に視線方向を推定することを目指す。

見えに基づく手法において、計測や認識の精度を向上させるために変動（本研究では様々な切り出しに相当）を加えた学習画像を用いるという着想は、すでに提案されている。たとえば、村瀬ら⁹⁾ は、文字認識の分野で、変動吸収特性核と呼ばれる手法を提案している。具体的には、学習画像を線形変換することで学習パターンの種類と量を増やしたうえで、部分空間法を適用している。しかしながら、切り出しの不定性をもなう視線方向の推定においては、学習画像に変動を加えるのみで安定した推定を実現するのは困難であると考えられる。たとえば、目画像の左端に黒目と思われる領域があった場合に、これが視線方向の変動に起因するのか、あるいは、切り出しの変動に起因するのかを区別するのは容易ではない。したがって、低解像度画像を用いた視線方向の推定においては、切り出しの変動と視線方向の変動をより注意深く取り扱う必要がある。

近年、顔認識、動作認識および画像合成の分野において、 N モード SVD (Singular Value Decomposition) を利用した手法が提案されている¹⁵⁾⁻¹⁸⁾。通常の SVD では、人物と画素などのように2つの変動の

モードを扱う。様々な人物の画像データによって構成される行列に対して SVD を適用すると、人物および画素の2つのモードの基底行列を抽出することができる。それに対して、 N モード SVD では、複数の変動のモードを扱うことが可能である。Vasilescu らは、顔認識に N モード SVD を適用することで、視点変化、光源変化、表情変化なども変動のモードと見なすことにより、これらの変動の影響を受けにくい人物同定を実現している¹⁶⁾。

本研究では、低解像度画像を入力として視線方向を推定するために、 N モード SVD に基づいた手法を提案する。具体的には、視線方向と画素だけでなく切り出しの変動も1つのモードとして明示的に扱うことにより、視線方向の変動のモードと切り出しの変動のモードを注意深く分離し、視線方向の変動を反映する特徴量を抽出する。具体的手順としては、あらかじめ準備した学習用の目画像から視線特徴量を抽出しておく。次に、視線方向の特徴量および切り出しの特徴量に関して双線形性を持つ評価関数を定義し、その評価関数を最小化する問題を交互法で解くことにより、1枚のテスト用の目画像から唯一の視線特徴量を抽出する。そして、これに近い学習画像の視線特徴量を探索することにより視線方向を推定する。提案手法の有効性を確認するために実画像を用いた評価実験を行い、視線方向と切り出しの変動という2つのモードを区別しない通常の PCA (Principal Component Analysis) および、クラスごとに基底を準備する部分空間法と比較した。

本論文の構成は以下のとおりである。まず、2章において、 N モード SVD を用いた提案手法について詳しく説明する。次に、3章で、実画像を用いた実験の環境および結果を報告する。最後に、4章で結論と今後の課題を述べる。

2. 提案手法

提案手法は学習段階およびテスト段階からなり、そのフローチャートを図1に示す。学習段階では、まず様々なところを注視したときの顔画像を取得する(1-a)。各々の顔画像から様々な領域で目画像を切り出すことにより、視線方向の変動および切り出しの変動を与えた学習画像を取得する(1-b)。次に、これらの学習画像に対して N モード SVD¹⁵⁾ を施すことにより、視線方向の変動を表現する特徴ベクトルおよび、切り出しの変動を表現する特徴ベクトルを抽出する(1-c)。この視線方向の変動を表現する特徴ベクトル、切り出しの変動を表現する特徴ベクトルをそれぞれ視線特徴

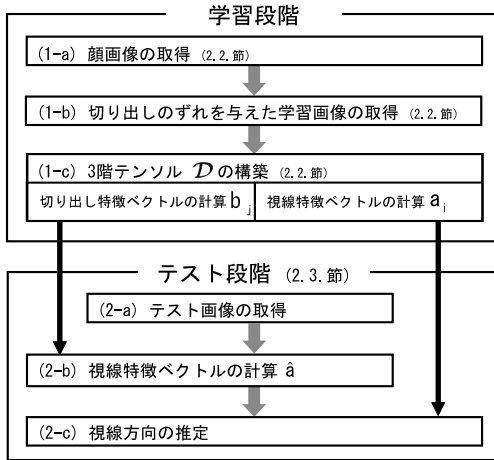


図 1 提案手法のフローチャート

Fig. 1 Flowchart of our proposed method.

ベクトル, 切り出し特徴ベクトルと呼ぶものとする.

テスト段階では, テスト画像より視線特徴ベクトルを抽出する (2-b). このようにして得られたテスト画像の視線特徴ベクトルの近傍に位置する, 学習画像の視線特徴ベクトルを近い順に 3 つ探索する. 次に, これらの特徴ベクトルを補間することによりテスト画像の視線方向を推定する (2-c).

以下において N モード SVD を説明した後, N モード SVD の枠組みにおいて目画像から視線特徴ベクトルを抽出する方法を述べる. 次に, 本提案手法を詳述する.

2.1 N モード SVD

提案手法において使用する N モード SVD¹⁵⁾ について述べる. N モード SVD とは, N 個のモードによって構成される対象データに対して, 各モードごとに基底を求め, これらの基底を利用して対象データを表現する手法である. ここでのモードとは対象データに対して変動を与える要因であり, 顔画像を例にあげると人, 視点, 照明, 表情および画素の変化のそれぞれをモードと見なすことができる. 本節では 3 モード SVD を説明する. なお, 対象データは目の画像であり, これらは様々なところを注視したときの顔画像の各々から様々な領域で切り出した目画像とする. すなわち, これらの目画像は視線方向, 切り出しおよび画素という 3 つの変動のモードからなる.

(1) テンソルの構築

取り扱う目画像は視線方向, 切り出しおよび画素の 3 つのモードからなるため, Vasilescu らに倣って 3 階テンソル \mathcal{D} を利用して各目画像のデータを記述する. ここで, 3 階テンソルの要素表現は D_{ijk} であり,

$i(1 \leq i \leq I)$, $j(1 \leq j \leq J)$, $k(1 \leq k \leq K)$ はそれぞれ, 視線方向, 切り出し方, 画素のインデックスとする.

(2) テンソルの展開

3 階テンソル \mathcal{D} に対して SVD を直接適用することは困難であるため, テンソル \mathcal{D} を行列の形に展開する. また, 各モードの基底行列を獲得するために, テンソルをモードごとに展開する. ここでは, 視線方向のモードに関してテンソル \mathcal{D} を展開する方法を述べる. まず, テンソル \mathcal{D} の要素表現 D_{ijk} において, k をある値に固定することにより行列 $F_k \in \mathbb{R}^{I \times J}$ を構成する. このとき, 視線モードに関してテンソル \mathcal{D} を展開することにより構成される行列 $D_{gaze} \in \mathbb{R}^{I \times K}$ は,

$$D_{gaze} = [F_1 F_2 \dots F_K] \quad (1)$$

と書くことができる. また, 切り出しモードおよび画素モードに関してテンソル \mathcal{D} を展開する方法を付録 A.1 に示す.

(3) 各モードの基底の獲得

視線に関する基底を獲得するために, 前手順で構成した行列 D_{gaze} に対して SVD を適用すると,

$$D_{gaze} = U_{gaze} \Sigma_{gaze} V_{gaze}^T \quad (2)$$

となる. ここで, $U_{gaze} \in \mathbb{R}^{I \times I}$ は視線方向の変動に関する基底行列である. 同様にして, 切り出しの変動に関する基底行列 $U_{cut} \in \mathbb{R}^{J \times J}$, 画素の変動に関する基底行列 $U_{pixel} \in \mathbb{R}^{K \times K}$ を計算することができる. その方法を付録 A.1 に示す.

上で求めた U_{gaze} , U_{cut} を用いて視線方向 i ($1 \leq i \leq I$) の視線特徴ベクトル $a_i \in \mathbb{R}^I$ および, 切り出し番号 j ($1 \leq j \leq J$) の切り出し特徴ベクトル $b_j \in \mathbb{R}^J$ を下式で定義する.

$$[a_1, a_2, \dots, a_I] \stackrel{\text{def}}{=} U_{gaze}^T, [b_1, b_2, \dots, b_J] \stackrel{\text{def}}{=} U_{cut}^T. \quad (3)$$

次に, N モード SVD を利用して各モードの基底を求めるために必要な計算量を考える. ここでは, 視線モードを含む N 個のモードを想定し, 視線モードの基底を抽出するために必要な計算量を考える. また, 視線モードをモード 1 とする. 視線モードの基底を計算するためには, N 階テンソル \mathcal{D} を視線モードに関して展開することにより得られる行列 $D_{gaze} \in \mathbb{R}^{I \times L}$ ($I_1 \leq L$, $L \stackrel{\text{def}}{=} I_2 I_3 \dots I_N$) に対して SVD を適用する必要がある. これに必要な計算量は $2LI_1^2 + 4I_1^3$ である⁴⁾. I_n ($1 \leq n \leq N$) はモード n の全要素数である. よって, 全要素数が I_n であるモード n を追加することにより計算量は I_n 倍になる.

(4) 基底を用いたテンソルの表現

前節で求めた各モードにおける基底行列 U_{gaze} , U_{cut} , U_{pixel} のそれぞれの要素表現 $(U_{gaze})_{il}$, $(U_{cut})_{jm}$, $(U_{pixel})_{kn}$ を用いて 3 階テンソルの要素 D_{ijk} を下式のように書くことができる.

$$D_{ijk} = \sum_{l=1}^I \sum_{m=1}^J \sum_{n=1}^K \mathcal{Z}_{lmn} (U_{gaze})_{il} (U_{cut})_{jm} (U_{pixel})_{kn} \quad (4)$$

ここで, $\mathcal{Z} \in \mathbb{R}^{I \times J \times K}$ はコアテンソルと呼ばれ各モードの基底ベクトルの相互作用を記述する. 上式を下式のように表現する.

$$\mathcal{D} = \mathcal{Z} \times_1 U_{gaze} \times_2 U_{cut} \times_3 U_{pixel} \quad (5)$$

ここで, たとえば ${}_1U_{gaze}$ の前の添え字 1 は, 式 (4) において ${}_1U_{gaze}$ の il 成分が \mathcal{Z} に作用するとき \mathcal{Z}_{lmn} の 1 次元目の添え字 l に関して和をとることを示す. また, コアテンソル \mathcal{Z} は

$$\mathcal{Z} = \mathcal{D} \times_1 U_{gaze}^\top \times_2 U_{cut}^\top \times_3 U_{pixel}^\top \quad (6)$$

と書くことができる.

また, 視線特徴ベクトル a および, 切り出し特徴ベクトル b を持つ画像 d は, 式 (7) で定義される 3 階テンソル \mathcal{B} を用いて式 (8) のように表現することができる.

$$B_{ijk} \stackrel{\text{def}}{=} \sum_{l=1}^K \mathcal{Z}_{ijl} (U_{gaze})_{kl} \quad (7)$$

$$d_k = \sum_{i=1}^I \sum_{j=1}^J B_{ijk} a_i b_j = \sum_{i=1}^I \sum_{j=1}^J B_{k(ij)} a_i b_j. \quad (8)$$

ここで, $B_{k(ij)}$ はテンソル \mathcal{B} を行列の形式で表現したものである. 具体的には, $B_{k(ij)}$ は行列 B_{pixel} の $(k, I * (j-1) + i)$ 要素であり, 行列 B_{pixel} はテンソル \mathcal{B} を画素モードで展開したものである.

2.2 学習段階

提案手法は大きく分けて学習段階およびテスト段階によって構成される. 本節では下記の手順をとる学習段階について述べる.

- (1) 目画像の取得および切り出し
- (2) 視線特徴ベクトルおよび切り出し特徴ベクトルの抽出
- (3) 行列要素 $B_{k(ij)}$ の獲得

以下において各手順を具体的に説明する.

目画像の取得および切り出し

本研究では, 視線方向の推定として PC の液晶画面上の注視点位置を推定する. 画面上の様々なところを

見た際の顔画像を取得し, 各顔画像に対して様々な領域で目画像を切り出すことにより学習画像を準備する. また, 被験者ごとに学習画像を取得する.

視線特徴ベクトルおよび切り出し特徴ベクトルの抽出

前手順で取得した目の学習画像を利用して視線特徴ベクトルを抽出する手順を説明する. まず, 前節で準備した学習画像を利用して 3 階テンソル \mathcal{D} を構成する. 次に, テンソル \mathcal{D} を視線モードに関して展開することにより行列 D_{gaze} を構成する (2.1 節参照). この行列 D_{gaze} に対して SVD を適用することにより, 視線モードの基底行列 U_{gaze} を抽出する (式 (2) 参照). そして, 式 (3) に基づいて視線特徴ベクトル a_i ($1 \leq i \leq I$) を求める. また, 式 (19) より切り出しモードの基底行列 U_{cut} を抽出する. そして, 式 (3) に基づいて切り出し特徴ベクトル b_j ($1 \leq j \leq J$) を計算する.

行列要素 $B_{k(ij)}$ の獲得

式 (7) に基づいて, 行列要素 $B_{k(ij)}$ ($1 \leq i \leq I$, $1 \leq j \leq J$, $1 \leq k \leq K$) を計算する.

2.3 テスト段階

学習段階で求めた視線特徴ベクトル, 切り出し特徴ベクトルおよび行列要素 $B_{k(ij)}$ を利用することにより, テスト段階では以下の手順によってテスト画像の視線方向を推定する.

- (1) テスト用の目画像の取得および切り出し
- (2) テスト画像の視線特徴ベクトルの抽出
- (3) 視線方向の推定

以上の各手順について説明する.

テスト用の目画像の取得および切り出し

学習用の目画像と同様にして, テスト用の目画像を取得し切り出す.

テスト画像の視線特徴ベクトルの抽出

N モード SVD の枠組みにおいて, テスト画像から特徴ベクトルを抽出するための手法が 2 つ提案されている. 第 1 は顔認識のために Vasilescu ら¹⁶⁾ によって提案されたものであり, 複数の射影行列をテスト画像に作用させることにより, テスト画像を人識別ための特徴空間に射影する手法である. 顔画像が個人, 視点, 照明および, 表情のモードからなる場合には人の総数 P , 視点の総数 V , 照明の総数 L および, 表情の総数 E を掛け合わせた数である ($P \times V \times L \times E$) 個の射影行列を準備してテスト画像に作用させるため, 人識別のための特徴ベクトルの候補が ($P \times V \times L \times E$) 個抽出される.

第 2 は Vasilescu ら¹⁹⁾ によって近年提案された手

法であり、各モードの要素数によらずにつねに唯一の射影行列を使用する．この唯一の射影行列をテスト画像に作用させることにより、すべてのモードに関して唯一の特徴ベクトルを抽出することができる点で、前述の第 1 の手法¹⁶⁾より優れている．

しかしながら、第 2 の手法¹⁹⁾を利用して低解像度画像から唯一の特徴ベクトルを抽出することはできない．それは、総画素数が、画素モードを除いたモードの要素数の積以上であるという条件を満たす場合のみ、唯一の射影行列を獲得することができるからである．この条件を本研究の目画像（総画素数 K 、視線方向 I 、切り出し数 J ）の場合にあてはめると $K \geq IJ$ となる．しかしながら、本研究では低解像度画像（ $K = 48$ ）を使用し、 $I = 20$ 、 $J = 25$ であるためこの条件を満たすことはできない．

そこで、本研究では 3 モード SVD の枠組みで低解像度画像から唯一の特徴ベクトルを抽出する手法を提案する．視線特徴ベクトル a' 、切り出し特徴ベクトル b' を用いて式 (8) に基づいて表現される画像および、テスト画像 \hat{d} を考える．このとき、評価関数 $f(a', b')$ を式 (9) によって定義する．

$$f(a', b') \stackrel{\text{def}}{=} \sum_{k=1}^K \left(\hat{d}_k - \sum_{i=1}^I \sum_{j=1}^J (B_{k(ij)} a'_i b'_j) \right)^2. \quad (9)$$

そして、式 (10) に基づき評価関数を最小化することによって、テスト画像の特徴ベクトル (\hat{a}, \hat{b}) を計算する．

$$(\hat{a}, \hat{b}) = \arg \min_{a' \in \mathbb{R}^I, b' \in \mathbb{R}^J} f(a', b'). \quad (10)$$

式 (10) の評価関数は双線形性を持つ．この評価関数では、変数 b' を固定すると、変数 a' に関しては線形であり、逆に a' を固定すると、 b' に関しては線形である．そこで、交互法^{3), 13)}を利用して、双線形性を持つ評価関数を最小化することによって唯一の特徴ベクトルを抽出する．

本提案手法を利用することにより、総画素数に関する条件を $K \geq IJ$ （前述した Vasilescu らによる第 2 の手法¹⁹⁾の条件）から $K \geq (I + J)$ に緩和することができる．つまり、解像度 (K) を固定して考えると、本提案手法では Vasilescu らによる第 2 の手法¹⁹⁾に比べて学習画像のモードの要素数（視線方向と切り出し方の和 $I + J$ ）をより大きくとることができる．

双線形性を持つ評価関数（式 (9)）を最小化する問題（式 (10)）の解法を説明する．1 変数の線形関数 f が最小となるためには、その 1 回微分がゼロである必

要がある．これより、 $\partial f / \partial a'_i = 0$ ($1 \leq i \leq I$) および $\partial f / \partial b'_j = 0$ ($1 \leq j \leq J$) が必要であり、これらの式に式 (9) を代入すると、

$$a' = M^+ \hat{d}, \quad (M)_{ki} \stackrel{\text{def}}{=} \sum_{j=1}^J B_{k(ij)} b'_j \quad (11)$$

および、

$$b' = N^+ \hat{d}, \quad (N)_{kj} \stackrel{\text{def}}{=} \sum_{i=1}^I B_{k(ij)} a'_i. \quad (12)$$

となる．ここで、 M^+ の $+$ は擬逆行列を意味する． b' の初期値を $b^{(0)}$ として、式 (11) および式 (12) に基づいて 2 つの特徴ベクトルを収束するまで交互に更新していく．そして、 $\Delta f(n) \stackrel{\text{def}}{=} f(a^{(n)}, b^{(n)}) - f(a^{(n-1)}, b^{(n-1)})$ があらかじめ定義した閾値以下になったとき反復計算を終了する．ここで、 n 回目の反復計算後の特徴ベクトルを $a^{(n)}$ および $b^{(n)}$ としている．この反復計算の後、視線特徴ベクトルの大きさは初期値に比べて変化するため、 L_2 ノルムを用いて a' を規格化しテスト画像の視線特徴ベクトル \hat{a} を得る．なお、式 (13) に基づいて b' の初期値 $b^{(0)}$ を決定している．

$$(a^{(0)}, b^{(0)}) = \arg \min_{a' \in \{a_1, a_2, \dots, a_I\}, b' \in \{b_1, b_2, \dots, b_J\}} \sum_{k=1}^K \left(\hat{d}_k - \sum_{i=1}^I \sum_{j=1}^J (B_{k(ij)} a'_i b'_j) \right)^2. \quad (13)$$

すなわち、最小二乗の意味でテスト画像に最も近い画像を与える視線特徴ベクトルおよび切り出し特徴ベクトルの組合せを探索する．

視線方向の推定

まず、下式に基づいてテスト画像の視線特徴ベクトルに最近傍の学習画像の視線特徴ベクトルを探索する．

$$i(1) = \arg \min_{i \in \{1, 2, \dots, I\}} \|\hat{a} - a_i\|^2. \quad (14)$$

上式は、テスト画像に対して最近傍の学習画像の視線特徴ベクトルが $a_{i(1)}$ であり、その視線方向が $i(1)$ となることを意味する．次に、2 番目、3 番目に近い学習画像の視線特徴ベクトルを探索する．これらの 2 番目、3 番目に近い視線方向がそれぞれ $i(2)$ 、 $i(3)$ であるとす．すなわち、テスト画像の視線特徴ベクトルに対して、最近傍の学習画像の視線特徴ベクトルおよび 2 番目、3 番目に近い学習画像の視線特徴ベクトルはそれぞれ $a_{i(1)}$ 、 $a_{i(2)}$ 、 $a_{i(3)}$ である．

このようにして探索した 3 つの学習画像の視線特徴

ベクトルを補間することによりテスト画像の視線方向を推定する．具体的には以下のように計算する．探索した3つの視線特徴ベクトルそれぞれに重みを付けて線形和をとることにより，テスト画像の視線特徴ベクトルを表現する．次に，最小二乗法を利用することによりテスト画像の視線特徴ベクトルとの誤差が最小になるような重みを計算する．ここで，誤差の式は

$$\epsilon = \left\| \hat{\mathbf{a}} - \sum_{p=1}^3 w_p \mathbf{a}_{i(p)} \right\|^2 \quad (15)$$

であり，重みの条件式は $\sum_{p=1}^3 w_p = 1$ かつ $0 \leq w_p \leq 1$ ($p = 1, 2, 3$) である．この重みを下式を利用してテスト画像の注視点座標 g を表現する．

$$\mathbf{g} = \sum_{p=1}^3 w_p \mathbf{g}(p) \quad (16)$$

ここに， $\mathbf{g}(p)$ ($1 \leq p \leq 3$) は探索した3つの視線特徴ベクトル $\mathbf{a}_{i(p)}$ に対応する注視点の2次元座標である．

2.4 提案手法の利点

本提案手法の利点および，クラスごとに基底を準備する部分空間法との差異を2点あげる．なお，クラスごとの部分空間法を利用して視線方向を連続量として推定するために，各視線方向でなく各切り出しを各クラスとして扱う．このとき，切り出しごとに基底を計算し，視線方向の変動を記述する部分空間を構築するため，切り出し領域にずれを含む場合にも視線方向を推定することができると予想される．よって，本提案手法とこの部分空間法とを比較することにより本提案手法の利点を明確にすることができると思われる．

第1の利点および差異

本提案手法を利用することにより，モードごとに特徴量を抽出することができる．ここで，各モードの特徴量は他のモードからの影響を受けにくい特徴量である．具体的には，視線モードおよび切り出しモードの特徴量を抽出することが可能であり，視線特徴量は切り出しの変動からの影響を受けにくい特徴量である．

それに対して，クラスごとに基底を準備する部分空間法を利用しても，1つのモードの特徴量しか抽出することができない．具体的には，切り出しごとに視線方向の変動を記述する部分空間を構成するため，テスト画像を各切り出しの部分空間に射影することにより視線モードの特徴量のみを抽出することができる．さらに，各部分空間を張る基底は切り出しごとに異なるため，視線モードの特徴量は切り出しの変動によって影響を受ける．

第2の利点および差異

視線モードおよび切り出しモードからなる目画像が視線特徴量および切り出し特徴量の双線形結合によって表される(式(8)参照)という仮定が成り立つものとする．この仮定のもとでは本提案手法を利用して，連続的に変化する切り出しのテスト画像から視線特徴ベクトルを抽出することができる．

これに対して，クラスごとに基底を準備する部分空間法では離散的な切り出しごとに部分空間を構成するため，連続的に変化する切り出しのテスト画像を適切な切り出しのクラスに識別することが困難であると考えられる．その結果，テスト画像から正確な視線特徴量を抽出することが原理的に難しいと予想される．

ただし，視線方向を推定することに対して本提案手法が適しているか否かを，目の実画像を利用した実験で確かめる必要がある．そこで，本提案手法による視線の推定結果と，クラスごとに基底を準備する部分空間法による結果とを比較する(3.3節の図11参照)．

3. 評価実験

提案手法の評価実験の概要および実験結果を述べる．

3.1 実験条件

実験条件および実験環境に関して説明する．目画像を切り出すためには顔画像において何らかの基準となる点が必要である．そこで，本実験では目画像の切り出しの基準となる点として目の端点を利用した．この目の端点として尤もしい位置を顔画像において検出するために，Okaらの開発した頭部姿勢追跡システム¹¹⁾を使用した．このシステムが検出する目の端点を切り出しのための基準点と見なし，目画像を切り出す方法を3.2節に詳しく説明している．この頭部姿勢追跡システムへの入力画像を取得するために3台のIEEE1394カメラ(Point Grey Research社製Flea)を使用した．これらの3台のカメラのうち1台のカメラを，視線方向の推定のための学習画像およびテスト画像を取得する目的としても利用した．被験者は5名であり，5名の各々について学習用の目画像を取得した後に視線方向を推定した．また，画像処理を行うために一般的な汎用PC(OS: Windows XP, CPU: Intel Pentium4 3.0 GHz)を使用した．そして，表示装置としてナナオ社製，1,280 × 1,024画素，18インチの表示装置を使った．目の位置からカメラまでの距離は50cmであり，学習画像およびテスト画像を取得する際の実験条件を整えるために顎のせ台を用いて頭部を固定した．

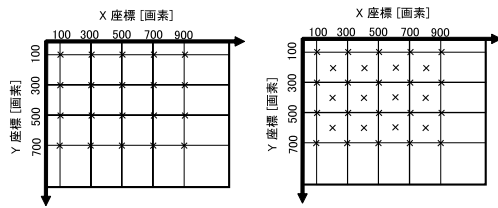


図 2 (左)学習画像用の×印表示点(右)テスト画像用の×印表示点

Fig. 2 Positions where crosshairs are displayed to grab training images (left) and test images (right).



図 3 取得した顔画像の例

Fig. 3 An example of grabbed face images.

3.2 実験手順

被験者 5 名の各人に対して、以下の手順に従い評価実験を実施した。

- (1) 顔画像の取得
- (2) 目画像の切り出し
- (3) 視線方向の推定

次に各手順を説明する。

顔画像の取得

顔画像の取得の手順を説明する。まず、液晶画面上のあらかじめ登録された複数の位置の中から 1 つのランダムな位置に × 印を表示する。学習画像を取得する目的で登録された位置を図 2 左の × 印によって示す。その × 印は画面上において正角格子状に等間隔にランダムな順番で 1 回ずつ表示され、× 印の表示点は合計 20 点である。また、テスト画像を取得する目的で登録された位置は図 2 右の × 印に示す位置であり合計 32 点である。次に、その × 印を被験者がマウスでクリックすると総画素数 144×144 画素の顔画像を保存する。被験者には × 印を見ているときに × 印をクリックするように依頼してある。取得した顔画像の例を図 3 に示す。それから、別のランダムな位置に 1 つの × 印を表示する。上記の手順を繰り返すことにより、画面上の様々な点を注視したときの顔画像を取得した。また、被験者ごとにこれらの学習画像を取得した。

目画像の切り出し

前項で取得した顔画像から目画像を切り出す手順を説明する。まず、前項で取得した顔画像 (144×144

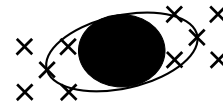


図 4 目の端点の候補

Fig. 4 Candidates of eye corners.

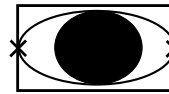


図 5 切り出された目画像

Fig. 5 A schematic diagram of segmented eye images.

画素)の一定領域の画素値の平均値を、低解像度画像において対応する画素の画素値とするという手法で、高解像度の顔画像 (144×144 画素)を 72×72 画素および 36×36 画素に低解像度化した。この方法を用いて低解像度画像を準備する理由は、低解像度の目画像に対する視線方向の推定精度と高解像度の目画像に対する精度とを比較するためである。

次に、下記の方法をとることにより顔画像から様々な領域で左目の画像を切り出した。目領域を切り出す際の基準点は目の端点とした。Okamotoらの開発した頭部姿勢追跡システム¹¹⁾が検出する目の両端点の位置を中心にして、縦横方向に ± 1 ステップずらした点(図 4 の × 印で示す位置)を目の端点の候補とした。ここで、1 ステップとは解像度が 36×36 画素, 72×72 画素, 144×144 画素の顔画像のそれぞれに対して 1 画素, 2 画素, 4 画素とした。これらの候補から 2 つの端点を選択し、下記の条件を満たすように画像内アフィン変換を施すことにより様々な領域で目画像を切り出した。アフィン変換を施す際に各画素の画素値を双線形補間により求めた。ここで、目の 2 つの端点の組合せは 25 通りである。また、学習用およびテスト用の目画像の切り出し方法は同等とした。

- 条件 1: 両端点の重心が切り出した目画像の重心に一致する。
- 条件 2: 目の両端点は図 5 の × 印に示す位置に来る。
- 条件 3: 切り出し目画像のアスペクト比は 3 である。
- 条件 4: 144×144 画素, 72×72 画素, 36×36 画素の顔画像から切り出した目画像の総画素数はそれぞれ 48×16 , 24×8 , 12×4 とする。

最後に、画像間の明るさのばらつきを抑制するために、各画像の画素値の平均値をひき、各画像の画素値の標準偏差で割ることにより各画像を規格化した。

次に、学習画像およびテスト画像の合計枚数につ

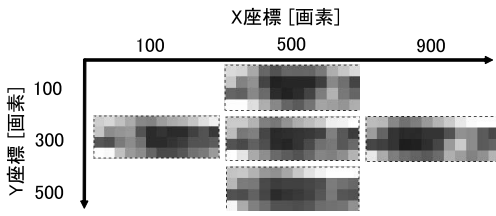


図 6 様々な注視点に対する学習用の目画像の例

Fig. 6 Examples of training images for different gaze points.

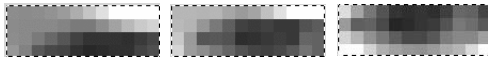


図 7 様々な切り出し方をしたときの学習用の目画像の例

Fig. 7 Examples of training images for different segmentations.

いてまとめる．学習用の注視点は合計 20 点であるため合計 20 枚の学習用の顔画像を取得した．各顔画像に対して 25 通りの切り出しを行うため学習画像は合計 $20 \times 25 = 500$ 枚である．そして，テスト用の注視点は合計 32 点であるためテスト画像の合計枚数は $32 \times 25 = 800$ 枚となる．

また，学習用に準備した目画像の例を図 6 に示す．これらは画面上の各注視点に対する目画像であり，総画素数は $12 \times 4 = 48$ 画素である．図 6 に示す座標は x, y 軸をそれぞれ液晶画面の横縦方向にとり，画面の左上を原点とするときの注視点の座標 [画素] とする．図 6 をみると，この解像度ではエッジ検出などの手法を用いて，虹彩あるいは瞳孔を検出することが困難であると予想される．

そして，注視点位置を固定して様々な領域で切り出したときの学習用の目画像の例を図 7 に示す．図 7 をみると切り出し方によって目画像の見えが大きく異なることが分かる．

視線方向の推定

視線方向を推定する手順は学習段階およびテスト段階よりなる．学習段階では学習用の目画像を利用し 2.2 節に基づいて学習画像の視線特徴ベクトル a_i ($1 \leq i \leq I$)，切り出し特徴ベクトル b_j ($1 \leq j \leq J$) および行列要素 $B_{k(ij)}$ ($1 \leq i \leq I, 1 \leq j \leq J, 1 \leq k \leq K$) を計算した．

テスト段階では，テスト用の目画像を利用して 2.3 節に基づいてテスト画像の視線特徴ベクトル \hat{a} を計算し視線方向を推定した．

3.3 実験結果

まず 3 モード SVD を利用して実画像より抽出した視線特徴ベクトルについて述べる．次に 3 モード SVD による視線方向の推定誤差を述べ，切り出しの不定性

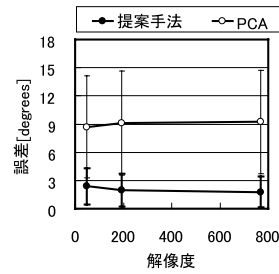


図 8 目画像の解像度の変化に対する注視点の推定誤差の変化

Fig. 8 Gaze estimation error for different image resolutions.

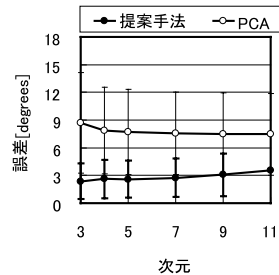


図 9 視線特徴ベクトルの次元の変化に対する注視点の推定誤差の変化

Fig. 9 Gaze estimation error against the dimension of gaze coefficients.

を別モードとして扱わない PCA による方法 (付録 A.2 参照) および，クラスごとに基底を準備する部分空間法 (付録 A.3 参照) の推定結果と比較する．

解像度の変化に対する推定誤差の変化

目画像の解像度を变化させた際の 3 モード SVD および PCA による注視点の推定誤差の変化を図 8 に示す．グラフの横軸は目画像の総画素数である．総画素数は $12 \times 4 = 48$ 画素， $24 \times 8 = 192$ 画素， $48 \times 16 = 768$ 画素の 3 種類としている．グラフ中の縦軸は 5 名の被験者のテスト画像に対して注視点を推定したときの平均誤差および標準偏差を表す．図 8 をみると，すべての解像度において 3 モード SVD による推定精度の方が PCA のそれより高いことが分かる．また，低解像度画像の推定精度は高解像度画像のそれと同程度である．これは，3.2 節で説明したように目の端点の候補 (図 4 の \times 印で示す位置) を高解像度画像では低解像度画像に比較して大きな画素で振っているためであると考えられる．なお，注視点の推定誤差は，液晶画面上における画素数の単位から視野角に変換している．

特徴量の次元の変化に対する推定誤差の変化

視線特徴ベクトルを表現するために使用する次元を变化させた際の 3 モード SVD および PCA による注視点の推定誤差の変化を図 9 に示す．グラフ中の縦軸

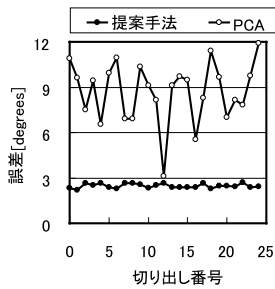


図 10 切り出し方の変化に対する注視点の推定誤差の変化，本提案手法および PCA に基づく手法による推定精度

Fig. 10 Gaze estimation error for each segmentation, estimated by our method and the PCA-based method.

は 5 名の被験者のテスト画像に対して注視点を推定したときの平均誤差および標準偏差を表す．テスト画像の解像度は 12×4 画素のみである．図 9 をみると，3～11 次元において 3 モード SVD による推定精度の方が PCA のそれより高い．

切り出し方の変化に対する推定誤差の変化

切り出し方を様々に変化させた際の 3 モード SVD および PCA による注視点の推定誤差の変化を図 10 に示す．グラフ中の縦軸は 5 名の被験者のテスト画像に対して注視点を推定したときの平均誤差および標準偏差を表す．図 10 のグラフの横軸は切り出し方のインデックスである．ここで，テスト画像の切り出し方は 3.2 節で説明した方法と同等である．なお，テスト画像の解像度は 12×4 画素のみである．図 10 のグラフをみると 3 モード SVD による推定誤差は PCA のそれと比較して切り出し方からの影響を受けにくいことが分かる．また，3 モード SVD による推定精度の方が PCA のそれより高い．

次に，本提案手法による注視点の推定精度と，クラスごとに基底を準備する部分空間法による精度とを比較する．切り出し方を様々に変化させた際のそれぞれの手法による注視点推定精度を図 11 に示す．図 11 をみると，本提案手法による推定誤差は，クラスごとに基底を準備する部分空間法のそれと比較して切り出し方からの影響を受けにくくかつ推定精度が高いことが分かる．これは，視線方向の変動および切り出し方の変動に対して視線特徴量および切り出し特徴量が線形に変化するという仮定が成り立っていることを示唆する結果であると考えられる．

各被験者の注視点の推定誤差

3 モード SVD および PCA による各被験者 (A, B, C, D, E) の注視点の推定誤差を図 12 左に示す．また，3 モード SVD および PCA を用いて注視点を推定したときの全被験者の平均誤差および標準偏差を

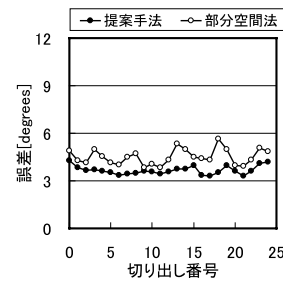


図 11 切り出し方の変化に対する注視点の推定誤差の変化，本提案手法およびクラスごとに基底を準備する部分空間法による推定精度

Fig. 11 Gaze estimation error for each segmentation, estimated by our method and the subspace method.

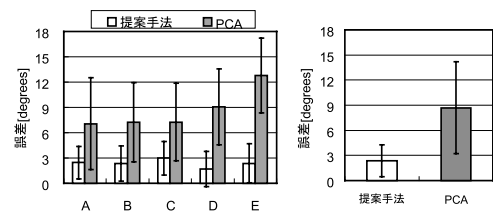


図 12 (左) 各被験者の注視点の推定誤差 (右) 全被験者の注視点の平均推定誤差

Fig. 12 Gaze estimation error for each individual (left), gaze estimation error averaged by all individuals (right).

図 12 右のグラフに示す．なお，テスト画像の解像度は 12×4 画素のみである．図 12 のグラフをみるとすべての被験者において 3 モード SVD による推定精度の方が PCA のそれより高いことが分かる．

4. む す び

本論文では，見えに基づく手法の枠組みで低解像度の目画像から視線方向を推定する手法を提案した．低解像度画像を入力とする視線計測は，被計測者をカメラの近傍に拘束しないという利点と，目の領域を安定に切り出すのが困難であるという欠点が存在する．そこで提案手法では，様々な切り出しの画像の学習パターンに対して N モード SVD を適用することにより目領域の切り出しの不定性に対処した．この N モード SVD を用いることにより，視線方向の変動のモードと切り出しの変動のモードを注意深く分離し，視線方向の変動を表現する特徴量を抽出した．実画像を用いた評価実験を行い，提案手法がモードの分離を行わない従来の PCA および，クラスごとに基底を準備する部分空間法よりも優れていることを確認した．

今後の課題として，人の違いおよび照明条件の違いをそれぞれ 1 つのモードとして扱う．また，頭部を自

由に動かすことができる状況に対応する。さらに、時系列フィルタを利用した視線追跡を検討する。

謝辞 本研究の一部は、文部科学省科学研究費補助金特定領域研究(C)(2)「人間の意図・行動理解に基づく柔軟なヒューマン・マシン・インタラクションの実現」(課題番号:13224051)の助成により行われた。

参 考 文 献

- 1) Baluja, S. and Pomerleau, D.: Non-intrusive gaze tracking using artificial neural networks, CMU CS Technical Report, CMU-CS-94-102 (1994).
- 2) Beymer, D. and Flickner, M.: Eye Gaze Tracking Using an Active Stereo Head, *Proc. IEEE CVPR 2003*, pp.II-451-458 (2003).
- 3) Buchanan, A. and Fitzgibbon, A.: Damped Newton Algorithms for Matrix Factorization with Missing Data, *Proc. IEEE CVPR 2005*, pp.316-322 (2005).
- 4) Chan, T.: An Improved Algorithm for Computing the Singular Value Decomposition, *ACM Trans. MS*, Vol.8, No.1, pp.72-83 (1982).
- 5) Hutchinson, T., White, Jr. K., Martin, W., Reichert, K. and Frey, L.: Human-Computer Interaction Using Eye-Gaze Input, *IEEE Trans. SMAC*, Vol.19, No.6, pp.1527-1534 (1989).
- 6) Ishikawa, T., Baker, S., Matthews, I. and Kanade, T.: Passive Driver Gaze Tracking with Active Appearance Models, *Proc. WCITS 2004* (2004).
- 7) 北川洋介, 加藤丈和, 呉 海元, 和田俊和: アイモデルを用いた ConDensation による視線推定, 情報処理学会研究報告, CVIM 2005-150-3, pp.17-24 (2005).
- 8) Matsumoto, Y. and Zelinsky, A.: An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement, *Proc. IEEE FG 2000*, pp.499-504 (2000).
- 9) 村瀬 洋, 木村文隆, 吉村ミツ, 三宅康二: パターン整合法における特性核の改良とその手書き平仮名文字認識への応用, 電子通信学会論文誌, Vol.J64-D, No.3, pp.276-283 (1981).
- 10) Ohno, T. and Mukawa, N.: A Free-head, Simple Calibration, Gaze Tracking System That Enables Gaze-Based Interaction, *Proc. ACM ETRA 2004*, pp.115-122 (2004).
- 11) Oka, K., Sato, Y., Nakanishi, Y. and Koike, H.: Head pose estimation system based on particle filtering with adaptive diffusion control, *Proc. IAPR MVA 2005*, pp.586-589 (2005).
- 12) Stiefelhagen, R., Yang, J. and Waibel, A.: Tracking Eyes and Monitoring Eye Gaze, *Proc. WPUI*, pp.98-100 (1997).
- 13) Shum, H., Ikeuchi, K. and Reddy, R.: Principal Component Analysis with Missing Data and Its Application to Polyhedral Object Modeling, *IEEE Trans.PAMI*, Vol.17, No.9, pp.854-867 (1995).
- 14) Tan, K., Kriegman, D. and Ahuja, N.: Appearance-based Eye Gaze Estimation, *Proc. IEEE WACV*, pp.191-195 (2002).
- 15) Vasilescu, M.A.O. and Terzopoulos, D.: Multilinear Analysis of Image Ensembles: Tensor-Faces, *Proc. ECCV 2002*, pp.447-460 (2002).
- 16) Vasilescu, M.A.O. and Terzopoulos, D.: Multilinear Image Analysis for Facial Recognition, *Proc. IAPR ICPR 2002*, pp.II-20511-20514 (2002).
- 17) Vasilescu, M.A.O.: Human Motion Signatures: Analysis, Synthesis, Recognition, *Proc. IAPR ICPR 2002*, pp.III-30456-30460 (2002).
- 18) Vasilescu, M.A.O. and Terzopoulos, D.: TensorTextures: Multilinear Image-Based Rendering, *Proc. ACM SIGGRAPH 2004*, Vol.23, No.3, pp.336-342 (2004).
- 19) Vasilescu, M.A.O. and Terzopoulos, D.: TensorTextures: Multilinear Independent Components Analysis, *Proc. IEEE CVPR 2005*, Vol.1, pp.547-553 (2005).
- 20) Wang, J., Sung, E. and Venkateswarlu, R.: Eye gaze Estimation from a Single Image of One Eye, *Proc. IEEE ICCV 2003*, pp.I-136-143 (2003).
- 21) Xu, L., Machin, D. and Sheppard, P.: A Novel Approach to Real-time Non-intrusive Gaze Finding, *British Machine Vision Conference*, pp.428-437 (1998).
- 22) Yoo, D. and Chung, M.: Non-intrusive Eye Gaze Estimation without Knowledge of Eye Pose, *Proc. IEEE FG 2004*, pp.785-790 (2004).

付 録

A.1 切り出しモード, 画素モードの基底ベクトル
切り出しモードおよび画素モードの基底ベクトルの
計算方法を説明する. テンソル \mathcal{D} の要素表現 D_{ijk}
($1 \leq i \leq I, 1 \leq j \leq J, 1 \leq k \leq K$) において, i をある値
に固定することにより構成される行列を $G_i \in \mathbb{R}^{J \times K}$
とする. また, D_{ijk} において, j をある値に固定する
ことにより構成される行列を $H_j \in \mathbb{R}^{K \times I}$ とする. こ
こで, 行列 G_i および H_j の要素表現は, それぞれ
(G_i) $_{jk}$ および (H_j) $_{ki}$ となる. このとき, 3 階テンソ
ル \mathcal{D} を切り出しモードに関して展開することにより
構成される行列 $D_{cut} \in \mathbb{R}^{J \times IK}$ は,

$$D_{cut} = [G_1 G_2 \dots G_I] \quad (17)$$

と表すことができる．また，画素モードに関して展開することにより構成される行列 $D_{pixel} \in \mathbb{R}^{K \times JI}$ は，

$$D_{pixel} = [H_1 H_2 \dots H_J] \quad (18)$$

と表すことができる．

次に，切り出しに関する基底を計算する． D_{cut} に対して SVD を適用すると，

$$D_{cut} = U_{cut} \Sigma_{cut} V_{cut}^T \quad (19)$$

となる．ここで， $U_{cut} \in \mathbb{R}^{J \times J}$ は，切り出しに関する基底行列である．

同様に，画素に関する基底を計算する． D_{pixel} に対して SVD を適用すると，

$$D_{pixel} = U_{pixel} \Sigma_{pixel} V_{pixel}^T \quad (20)$$

となる．ここで， $U_{pixel} \in \mathbb{R}^{K \times K}$ は，画素に関する基底行列である．

A.2 PCA を利用した注視点の推定手法

PCA を利用した注視点の推定手法について述べる．この手法では視線方向を表現する特徴ベクトルを求めるときに，視線方向の変動による影響と切り出しの変動による影響とを分離して考えない．まず，ある固定された視線方向の顔画像から様々な領域で切り出した目画像を固有空間に射影することにより各切り出しの目画像の特徴ベクトルを抽出する．次に，あらゆる切り出しに関して特徴ベクトルの重心をとる．重心をとった特徴ベクトルをその視線方向の特徴ベクトルとする．

以下に，この視線方向を表現する特徴ベクトルの計算方法を説明する．視線方向 i ，切り出し j の規格化された目画像 $d_{ij} \in \mathbb{R}^K$ ($1 \leq i \leq I$, $1 \leq j \leq J$) を並べて行列 D を構成する．

$$D = [d_{11}, d_{21}, \dots, d_{I1}, d_{12}, d_{22}, \dots, d_{I2}, \dots, d_{1J}, d_{2J}, \dots, d_{IJ}] \quad (21)$$

上式に対して SVD を適用すると $D = U \Sigma V^T$ となる．このとき，射影行列は U^T であるから目画像 d_{ij} を射影してできる特徴ベクトル c_{ij} は $c_{ij} = U^T d_{ij}$ と書くことができる．よって，視線方向 i を表現する特徴ベクトルは，あらゆる切り出しについての重心をとることにより $c_i = \left(\sum_{j=1}^J c_{ij} \right) / J$ となる．テスト画像 d に対して射影行列 U^T を作用させることによりテスト画像の視線特徴量 c を抽出する．そして，テスト画像の視線特徴量 c の近傍に存在する学習画像の視線特徴量を $\{c_1, c_2, \dots, c_I\}$ から 3 つ探索する．その後は 2.3 節の手順と同様にして注視点を推定する．

A.3 クラスごとの基底を持つ部分空間法による注視点の推定手法

切り出しの数だけ部分空間を準備することによって，切り出しのずれを考慮している点が上記の PCA に基づく手法に対して異なる．ここでは，視線方向を連続量として推定するために各切り出しを各クラスとして考え，切り出しごとに基底を計算し視線方向の変動を表現する部分空間を構成する．

この手法の計算方法を簡単に説明する．視線方向 i ，切り出し j の目の学習画像 $d_{ij} \in \mathbb{R}^K$ ($1 \leq i \leq I$, $1 \leq j \leq J$) を利用して行列 $D_j \in \mathbb{R}^{K \times I}$ を構成する．

$$D_j = [d_{1j}, d_{2j}, \dots, d_{Ij}] \quad (22)$$

この行列 D_j に対して SVD を適用して切り出し j の部分空間 L_j を張る基底 $U_j \in \mathbb{R}^{K \times I}$ を計算する．ただし， $K > I$ とする．

$$D_j = U_j \Sigma_j V_j^T, \quad U_j = [u_j^{(1)}, u_j^{(2)}, \dots, u_j^{(I)}] \quad (23)$$

次に， K 次元の元画像空間から K' ($\leq I < K$) 次元の部分空間 L_j への変換を行う射影行列 $P_j \in \mathbb{R}^{K \times K}$ を計算する．

$$P_j = U'_j U_j'^T, \quad U'_j \stackrel{\text{def}}{=} [u_j^{(1)}, u_j^{(2)}, \dots, u_j^{(K')}] \quad (24)$$

学習画像 d_{ij} に対して，各切り出し j の射影行列 P_j を作用させることにより，学習画像の視線特徴ベクトル $a_{ij} = P_j d_{ij}$ を計算する．

テスト画像 d と各部分空間 L_j との類似度 $S_j(d) = d^T P_j d$ を計算する．類似度が最大となる切り出し \hat{j} の部分空間 $L_{\hat{j}}$ を探索することにより，テスト画像の属する切り出し \hat{j} の部分空間 $L_{\hat{j}}$ を識別する．

そして，部分空間 $L_{\hat{j}}$ への変換を行う射影行列 $P_{\hat{j}}$ を作用させることによりテスト画像の視線特徴ベクトル $\hat{a} = P_{\hat{j}} d$ を抽出する．部分空間 $L_{\hat{j}}$ において，テスト画像の視線特徴ベクトル \hat{a} に近い学習画像の視線特徴ベクトルを近いものから 3 つ探索する．その後は 2.3 節の手順と同様にして注視点を推定する．

(平成 17 年 9 月 20 日受付)

(平成 18 年 3 月 20 日採録)

(担当編集委員 尺長 健)



小野 泰弘

1998年東京工業大学大学院理工学研究科物理学専攻修士課程修了。同年日本電気(株)半導体事業部(現NECエレクトロニクス(株))に入社。2003年10月より、東京大学大学院情報理工学系研究科博士課程に在学中。コンピュータビジョンに関する研究に従事。



岡部 孝弘(正会員)

1997年東京大学理学部物理学学科卒業。1999年同大学大学院理学系研究科物理学専攻修士課程修了。2000年同博士課程中退。2001年より東京大学生産技術研究所技官(技術職員)。コンピュータビジョン、コンピュータグラフィックスに関する研究に従事。2005年度電子情報通信学会論文賞、MIRU2004、MIRU2005優秀論文賞、2004年度PRMU研究奨励賞を受賞。電子情報通信学会、IEEE各会員。



佐藤 洋一(正会員)

1997年カーネギーメロン大学計算機科学部ロボティクス学科博士課程修了。同年より東京大学生産技術研究所研究機関研究員、講師、助教授を経て、現在同大学大学院情報学環助教授。Ph.D. in Robotics。コンピュータビジョン、ヒューマン・コンピュータ・インタラクション、コンピュータグラフィックスに関する研究に従事。MIRU2000最優秀論文賞、MIRU優秀論文賞(MIRU2004、MIRU2005)、1999年度山下記念研究賞、1999年度日本バーチャルリアリティ学会論文誌論文賞、2005年度電子情報通信学会論文賞等を受賞。電子情報通信学会、日本バーチャルリアリティ学会、ACM、IEEE各会員。