

Fuzzy Domain Knowledge を用いた無校正画像からの 3次元構造の復元：顔画像生成への適用

磯野 勝宣[†] 張 正友^{†,††,★} 赤松 茂[†]

無校正画像の応用には新規画像生成のような多くのアプリケーションがある。任意方向画像の生成を行う場合、射影空間またはアフィン空間において、方向を指定することは容易ではない。しかし、対象物体のユークリッド構造が得られれば、ユークリッド空間における、任意の姿勢の指定が可能となる。本論文では、距離や角度といったカテゴリーに基づく知識 (FDK) を用いて無校正画像からのユークリッド構造の復元を行う。我々は通常、特定の物体に対しては正確ではないが、その物体のカテゴリーに対する知識を持っている。これらの知識のあいまいさはガウス変数としてモデル化される。本手法を新規顔画像の生成に適用し、顔のユークリッド構造の復元を行い、任意方向の新規顔画像の生成を行った。

Euclidean Structure from Uncalibrated Images Using Fuzzy Domain Knowledge: Application to Facial Image Synthesis

KATSUNORI ISONO,[†] ZHENGYOU ZHANG^{†,††,★}
and SHIGERU AKAMATSU[†]

Use of uncalibrated images has found many applications such as image synthesis. However, it is not easy to specify the desired position of the new image in projective or affine space. This paper proposes to recover Euclidean structure from uncalibrated images using domain knowledge such as distances and angles. The knowledge we have is usually about an object category, but not very precise for the particular object being considered. The variation (fuzziness) is modeled as a Gaussian variable. Six types of common knowledge are formulated. Once we have a Euclidean description, the task to specify the desired position in Euclidean space becomes trivial. The proposed technique is then applied to synthesis of new facial images.

1. はじめに

無校正 (カメラパラメータが未知) のカメラで撮影された画像 (無校正画像) を用いた研究は、新規画像生成などの多くのアプリケーションに応用することができる。また、無校正画像からは、距離情報の抽出はできないが、ピンホールカメラを仮定することにより、射影構造 (Projective structure)¹⁾ を得ることは可能である。得られた射影構造は、同一平面性や同一直線性、複比といった豊富な情報を持っており、物体認識などに利用することができる²⁾。しかしながら、パーチャリリティなどの、他の多くのアプリケーシ

ンに対しては、射影構造だけではなく、通常の直交座標系での3次元構造 (3次元ユークリッド構造) が必要な場合が多い。

本論文では、カテゴリーに基づく知識 (Fuzzy Domain Knowledge: **FDK**) を用いて、2枚の無校正画像から3次元ユークリッド構造を復元する手法を提案する。そして、その応用例として、顔の3次元ユークリッド構造の復元を行う。

FDK は、点の位置や2点間の距離といった簡単なものから、距離の比率といった複雑なものまである。“Fuzzy” とは、ある特定の対象物の正確な測定値を持たないことを意味しており、対象物のカテゴリーについての一般的な知識を FDK として用いる。そのため、FDK を用いて復元された3次元ユークリッド構造は、必ずしもその物体の正確な3次元ユークリッド表現ではない。しかし、3次元ユークリッド構造を得ることにより、様々な方向からの画像を取得することができ、新規画像生成などの多くのアプリケーションには有用

[†] 株式会社 ATR 人間情報通信研究所
ATR Human Information Research Laboratories

^{††} 仏国 INRIA
INRIA Sophia-Antipolis, France

[★] 現在、米国マイクロソフト
Presently with Microsoft Research, Redmond, USA

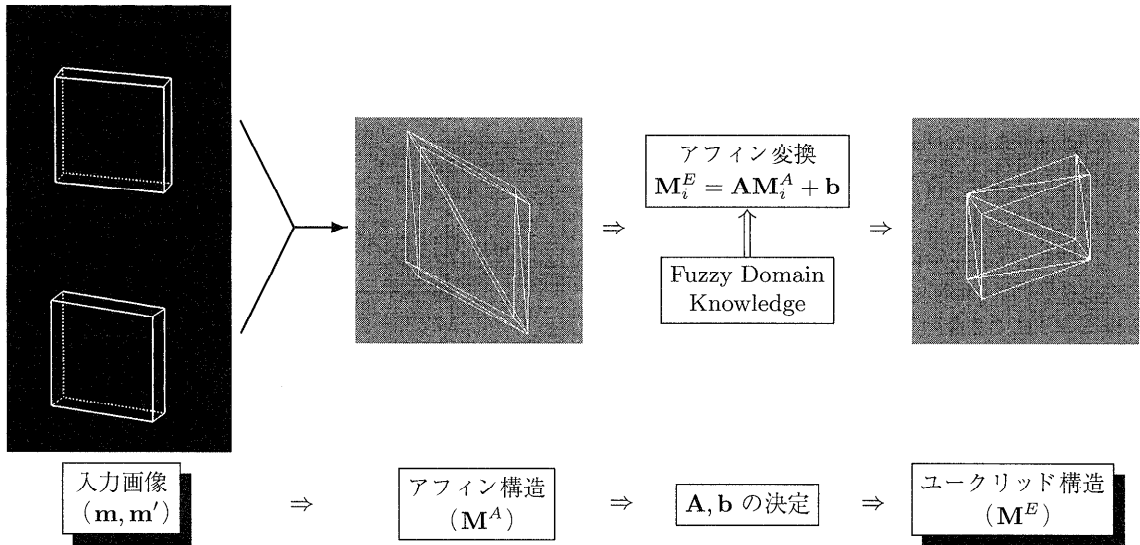


図 1 3次元ユークリッド構造復元の概要. 2枚の無校正画像から, 3次元アフィン構造を復元する. アフィン構造からユークリッド構造へのアフィン変換 (A, b) を FDK を用いて求め, 3次元ユークリッド構造を復元する.

Fig. 1 Overview of Euclidean reconstruction.

である (平行性, 平行な直線の長さの比などの, すべての正確なアフィン情報は保存されている).

また, 今回の対象物の奥行きは, カメラから対象物までの距離と比較して十分に小さいとし, カメラモデルをアフィンカメラ (弱中心射影) で近似することとする.

以下, 2章では, 2枚の無校正画像からの, FDK を用いた 3次元ユークリッド構造の復元手法について概説する. 3章では, 提案手法を顔画像に適用し, 顔の 3次元ユークリッド構造の復元を行い, 復元された 3次元ユークリッド構造から任意方向顔画像を生成する方法について述べる.

2. Fuzzy Domain Knowledge を用いた 3次元形状の復元

2.1 概要

本論文では, カメラパラメータが未知のカメラで撮影された, 同一物体で姿勢の異なる 2枚の画像 (2枚の無校正画像) から, その物体の 3次元構造を復元する手法について提案する.

一般に同一物体で姿勢の異なる 3枚の画像が与えられれば, structure-from-motion の原理を用いて, その物体の 3次元構造を復元することができる. 一方, Koenderink ら³⁾は, 入力画像が 3枚ではなく 2枚だけのときには, 3次元アフィン変換の自由度を残しての構造復元が可能であることを示した. このような 3

次元構造を 3次元アフィン構造と呼ぶ. ここで, 最終的に求めたい 3次元構造は, アフィン変換の自由度をなくした, 通常の直交座標系における 3次元構造であり, これを 3次元アフィン構造と区別して, 3次元ユークリッド構造と呼ぶ.

3次元アフィン構造 M^A , 3次元ユークリッド構造 M^E , 2枚の画像 m, m' を特徴点の集合として表し, 特徴点の 3次元アフィン座標, 3次元ユークリッド座標, 画像座標をそれぞれ M_i^A, M_i^E, m_i, m'_i (i は特徴点番号) とする. 3次元アフィン構造 M^A から 3次元ユークリッド構造 M^E へのアフィン変換は,

$$M_i^E = AM_i^A + b, \quad (1)$$

であり, A は 3×3 の正則行列, b は 3次元ベクトルである (12 自由度). つまり, (A, b) が得られれば, 2枚の無校正画像から得られる 3次元アフィン構造より, 3次元ユークリッド構造を復元することができる.

2枚の無校正画像からの 3次元ユークリッド構造復元の概要を図 1 に示す. ある物体の 2枚の無校正画像が与えられ, 画像間で対応がとれている特徴点が見られるとき, 弱中心射影を仮定すると, 文献 4) の手法を用いることにより, 2枚の画像間のエピソード幾何学が復元され, アフィン基本行列 F_A を求めることができる. F_A が得られれば, その物体の 3次元アフィン構造を復元することができる⁵⁾. すなわち, 画像間の特徴点のペア (m_i, m'_i) から, それらが対応する 3次元アフィン構造上の点の座標 M_i^A を求めるこ

とができる。 M^A が得られれば、式(1)の A と b を決定することによって、対象物体の3次元ユークリッド構造 M^E を復元することができる。

(A, b) は、たとえば、少なくとも4点の3次元アフィン座標とそれらに対応する3次元ユークリッド座標が与えられれば、線形最小自乗法を使って簡単に求めることができる。しかし、2枚の無校正画像のみを用いた場合、それらを撮影したカメラパラメータは未知であるので、画像のみから3次元ユークリッド座標を得ることはできない。レンジファインダなどの特殊な計測装置を用いて対象物体の3次元ユークリッド座標を得る方法もあるが、対象物体によっては計測ができない場合もある。また、このような計測装置は手軽に利用できるものではないため、利用範囲が制限されてしまう。そこで、今回は、3次元ユークリッド座標以外の知識として、対象物体のある部分の長さや、長さの比、角度などの知識も用いて (A, b) の推定を行う。このような知識は、その物体の一般的な知識から比較的容易に得ることができる。たとえば、立方体を考えた場合、「各辺の長さはすべて等しく、各頂点部でそれらの辺は直交している」といった知識をこれまでの経験により持っている。このようなカテゴリーについての知識を Fuzzy Domain Knowledge (FDK) と呼ぶこととする。Fuzzy とは、ある特定の対象物の正確な知識ではなく、その対象物が属するカテゴリーについての知識であることを意味する。

2.2 基準座標系

3次元アフィン構造から3次元ユークリッド構造へのアフィン変換の自由度は12である。ここでは、3次元ユークリッド構造復元の基準となる基準ユークリッド座標系 $O-XYZ$ の設定を行う。基準ユークリッド座標系の設定を行うことにより、アフィン変換の自由度は6となる。

基準ユークリッド座標系は、位置、方向に関して、それぞれ3自由度、合計6自由度を持つので、同一直線上にない3点を任意に選ぶことにより設定することができる。

M_0^A, M_1^A, M_2^A を任意の3次元アフィン構造上の3点とし、これらのアフィン変換(式(1))後の3次元ユークリッド構造上の点を M_0^E, M_1^E, M_2^E とする(図2)。ここで、以下の3つの条件を考え、基準ユークリッド座標系の設定を行う。

(1) M_0^A は、原点 $O = [0, 0, 0]^T$ に変換される。

$$M_0^E = AM_0^A + b = O. \quad (2)$$

これは、3つの拘束条件を与える。

(2) M_0^E と M_1^E を結ぶベクトルは、 X 軸と平行に

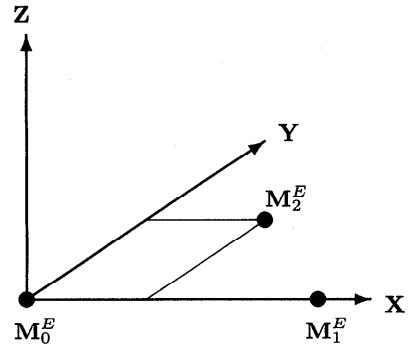


図2 3点指定による基準ユークリッド座標系の設定。同一直線上にない3点を任意に選ぶことにより基準ユークリッド座標系を設定する。

Fig. 2 Fixing a Euclidean reference frame by choosing three points.

なる。

$$(M_1^E - M_0^E) = A(M_1^A - M_0^A) = k[1, 0, 0]^T.$$

ここで、 k は任意の非零値である。 $[v_{11}, v_{12}, v_{13}]^T \equiv A(M_1^A - M_0^A)$ とすると、拘束条件は、

$$v_{12} = 0 \text{ and } v_{13} = 0, \quad (3)$$

となる。

(3) M_0^E, M_1^E, M_2^E を通る平面は、 XY 平面となる。これは、 M_0^E と M_2^E を結ぶベクトルが、 Z 軸と垂直であるということと等価である。

$$\begin{aligned} [0, 0, 1]^T (M_2^E - M_0^E) \\ = [0, 0, 1]^T A(M_2^A - M_0^A) \\ = 0. \end{aligned}$$

$[v_{21}, v_{22}, v_{23}]^T \equiv A(M_2^A - M_0^A)$ とすると、拘束条件は、

$$v_{23} = 0, \quad (4)$$

となる。

以上より、6つの拘束条件が得られ、基準ユークリッド座標系を設定することができる。

2.3 Fuzzy Domain Knowledge (FDK)

基準ユークリッド座標系を設定することにより、アフィン変換の残りの自由度は6となる。したがって、3次元ユークリッド構造を復元する(A, b を推定する)ためには、少なくとも、さらに6つの拘束条件が必要となる。また、基準ユークリッド座標系の設定を行わない場合は、12個以上の拘束条件が必要である。そこで、対象物体のFDKとして、表1に示す6種類の知識を考え、それらを組み合わせる用いることにより、(A, b)を推定する。

一般的に、これらの知識は完全には正確ではない。そこで、そのあいまいさをガウス変数によってモデル

表1 Fuzzy Domain Knowledge
Table 1 The fuzzy domain knowledge.

	FDK	拘束条件数
(1)	点のユークリッド座標	3
(2)	2点間の距離 (ベクトルの長さ)	1
(3)	2つのベクトルの成す角度	1
(4)	ベクトルの向き	2
(5)	ベクトルの同一性	(2)+(4)=3
(6)	2つのベクトルの長さの比率	1

化し、マハラノビス距離の二乗を最小にすることで、6種類の拘束条件を表す。

2.3.1 点のユークリッド座標

3次元アフィン座標 M_i^A に対応する 3次元ユークリッド座標の期待値は \hat{M}_i^E であり、共分散行列 Λ_i を持つものとする。式 (1) より得られる 3次元ユークリッド座標 M_i^E とその期待値 \hat{M}_i^E とのマハラノビス距離の二乗は次式で表される。

$$r_p^2 = (M_i^E - \hat{M}_i^E)^T \Lambda_i^{-1} (M_i^E - \hat{M}_i^E). \quad (5)$$

2.3.2 2点間の距離 (ベクトルの長さ)

M_i^E, M_j^E 間のユークリッド距離 d_{ij} の期待値は \hat{d}_{ij} であり、標準偏差 $\sigma_{d_{ij}}$ を持つものとする。 d_{ij} と \hat{d}_{ij} とのマハラノビス距離は、

$$r_d = (d_{ij} - \hat{d}_{ij}) / \sigma_{d_{ij}}, \quad (6)$$

となる。ここで、

$$d_{ij} = \|M_i^E - M_j^E\| = \|A(M_i^A - M_j^A)\|,$$

である。

2.3.3 2つのベクトルのなす角

\hat{v}_i と \hat{v}_j をアフィン座標系における 2つの 3次元ベクトル (アフィンベクトル) とする。2つのアフィンベクトルをユークリッド座標系に変換した 2つのベクトル (ユークリッドベクトル) のなす角の期待値は $\hat{\theta}$ であり、標準偏差 σ_θ を持つものとする。ユークリッドベクトルは $v_i = A\hat{v}_i$ と $v_j = A\hat{v}_j$ であり、2つのベクトルのなす角 θ を v_i, v_j で表すと、cosine 関数または sine 関数を用いて次のように表すことができる。

$$\cos \theta = \frac{v_i^T v_j}{\|v_i\| \|v_j\|} \quad \text{or} \quad \sin \theta = \frac{\|v_i \times v_j\|}{\|v_i\| \|v_j\|}.$$

いま、 $\theta = \hat{\theta} + \delta\theta$ とし、 $\delta\theta$ は小さいとすると、 $\delta\theta$ は、

$$\begin{aligned} \cos \theta &= \cos(\hat{\theta} + \delta\theta) \approx \cos \hat{\theta} - \sin \hat{\theta} \delta\theta, \\ \delta\theta &\approx \frac{\cos \hat{\theta} - \cos \theta}{\sin \hat{\theta}}, \end{aligned} \quad (7)$$

または、

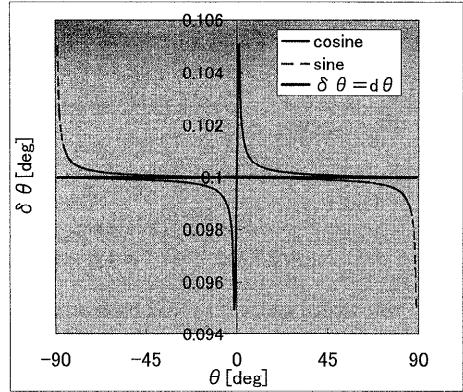


図3 $\delta\theta$ の近似。 $\delta\theta$ を sine 関数または cosine 関数を用いて近似する。 $-\pi/2 \leq \hat{\theta} \leq \pi/2$ の範囲において、 $-\pi/4 < \hat{\theta} < \pi/4$ では、 sine 関数、それ以外では、 cosine 関数を用いた方が、 $d\theta$ をより近似できる。

Fig.3 Approximation of $\delta\theta$ by using sine or cosine function.

$$\sin \theta = \sin(\hat{\theta} + \delta\theta) \approx \sin \hat{\theta} + \cos \hat{\theta} \delta\theta,$$

$$\delta\theta \approx \frac{\sin \theta - \sin \hat{\theta}}{\cos \hat{\theta}}, \quad (8)$$

となる。ここで、 $\delta\theta = d\theta$ (一定) とし、 $\hat{\theta}$ を変化させたときの、式 (7)、式 (8) の右辺の値の変化を図 3 に示す。図 3 から、 $-\pi/2 \leq \hat{\theta} \leq \pi/2$ の範囲において、 $-\pi/4 < \hat{\theta} < \pi/4$ では、式 (8) の方が、それ以外では、式 (7) の方が、 $d\theta$ をより近似していることが分かる。

以上より、 $-\pi/2 \leq \hat{\theta} \leq \pi/2$ の範囲において、 $-\pi/4 < \hat{\theta} < \pi/4$ のとき、 sine 関数を用い、それ以外では、 cosine 関数を用いるものとする。それぞれのマハラノビス距離は以下のとおりである。

i) cosine 関数を用いた場合：

$$\begin{aligned} r_\theta &= \frac{\delta\theta}{\sigma_\theta} \\ &= \frac{\cos \hat{\theta} - v_i^T v_j / (\|v_i\| \|v_j\|)}{\sigma_\theta \sin \hat{\theta}}. \end{aligned} \quad (9)$$

ii) sine 関数を用いた場合：

$$\begin{aligned} r_\theta &= \frac{\delta\theta}{\sigma_\theta} \\ &= \frac{\|v_i \times v_j\| / (\|v_i\| \|v_j\|) - \sin \hat{\theta}}{\sigma_\theta \cos \hat{\theta}}. \end{aligned} \quad (10)$$

2.3.4 ベクトルの向き

\hat{v}_i をアフィンベクトルとし、 $v_i = A\hat{v}_i$ をユークリッドベクトルとする。今、あるユークリッドベクトル \hat{v} と v_i との向きが等しい、つまり、2つのベクトルが平行であるという FDK が与えられたとする。これは、 v_i と \hat{v} とのなす角の期待値は 0° であるとい

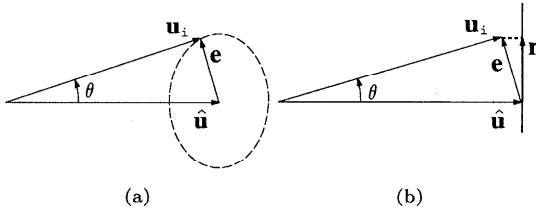


図4 ベクトルの向き。2つのベクトルが平行であるということは、それらのベクトルのなす角度が 0° であるということと等価である。

Fig. 4 Knowledge about the orientation of a vector.

うことと等価である。

$\mathbf{u}_i = \mathbf{v}_i / \|\mathbf{v}_i\|$ と $\hat{\mathbf{u}} = \hat{\mathbf{v}} / \|\hat{\mathbf{v}}\|$ を単位ベクトルとし、エラーベクトル $\mathbf{e} = \mathbf{u}_i - \hat{\mathbf{u}}$ は単位球面上にあるものとする (図4(a))。ここで、 \mathbf{e} の $\hat{\mathbf{u}}$ に垂直な面への投影を考える (図4(b)は垂直断面を示す)。 \mathbf{r} は、簡単に

$$\mathbf{r} = \mathbf{u}_i - (\mathbf{u}_i^T \hat{\mathbf{u}}) \hat{\mathbf{u}},$$

によって表すことができる。 θ が、標準偏差 σ_θ で、ガウス分布に従うとすると、 \mathbf{r} は、 $\hat{\mathbf{u}}$ 周りの平面上の2次元ガウス分布に従う (共分散行列 $\text{diag}(\sin^2 \sigma_\theta, \sin^2 \sigma_\theta)$)。これを3次元に拡張すると、 $\hat{\mathbf{u}} = [0, 0, 1]^T$ とすると、 \mathbf{r} の共分散行列 $\mathbf{\Lambda}_r$ は、 $\text{diag}(\sin^2 \sigma_\theta, \sin^2 \sigma_\theta, 0)$ で与えられる。一般的な $\hat{\mathbf{u}}$ に対しては、 $\hat{\mathbf{u}} = \mathbf{R} [0, 0, 1]^T$ となるような回転行列 \mathbf{R} を考えることができ、 \mathbf{r} の共分散行列は

$$\mathbf{\Lambda}_r = \mathbf{R} \text{diag}(\sin^2 \sigma_\theta, \sin^2 \sigma_\theta, 0) \mathbf{R}^T,$$

となる。 \mathbf{R} を定義する方法としては、たとえば、 $\hat{\mathbf{u}}$ に垂直なベクトル $\hat{\mathbf{u}}_0$ を考えると、回転行列は以下のようになる。

$$\mathbf{R} = [\hat{\mathbf{u}}_0 \times \hat{\mathbf{u}} \quad \hat{\mathbf{u}}_0 \quad \hat{\mathbf{u}}].$$

以上より、マハラノビス距離の二乗は次式で表される。

$$\begin{aligned} r_u^2 &= \mathbf{r}^T \mathbf{\Lambda}_r^{-1} \mathbf{r} \\ &= \mathbf{r}^T \mathbf{R} \text{diag}(1/\sin^2 \sigma_\theta, 1/\sin^2 \sigma_\theta, 0) \mathbf{R}^T \mathbf{r}. \end{aligned} \quad (11)$$

2.3.5 ベクトルの同一性

ユークリッドベクトル $\mathbf{v}_i = \mathbf{A} \tilde{\mathbf{v}}_i$ は $\hat{\mathbf{v}}_i$ と等しいというFDKが与えられたとき、このFDKは、ベクトルの長さ (式(6)) とその向き (式(11)) の2つのFDKの組合せで表すことができる。

2.3.6 2つのベクトルの長さの比率

長さの比率はユークリッド構造の復元において、新たな拘束として利用することができる。 $\tilde{\mathbf{v}}_i$ と $\tilde{\mathbf{v}}_j$ を2つのアフィンベクトルとする。これらに対応するユークリッドベクトルは、 $\mathbf{A} \tilde{\mathbf{v}}_i$ 、 $\mathbf{A} \tilde{\mathbf{v}}_j$ であり、その長さはそれぞれ、 $\|\mathbf{A} \tilde{\mathbf{v}}_i\|$ 、 $\|\mathbf{A} \tilde{\mathbf{v}}_j\|$ で表される。よって、こ

れら2つのベクトルの長さの比は、 $\|\mathbf{A} \tilde{\mathbf{v}}_i\| / \|\mathbf{A} \tilde{\mathbf{v}}_j\|$ となる。

ユークリッド座標系における、これらのベクトルの長さの比率の期待値は \hat{r} であり、標準偏差 σ_r を持つものとする、マハラノビス距離は次式で与えられる。

$$r_r = \frac{\|\mathbf{A} \tilde{\mathbf{v}}_i\| / \|\mathbf{A} \tilde{\mathbf{v}}_j\| - \hat{r}}{\sigma_r}. \quad (12)$$

2.4 ユークリッド構造の復元

上述の6種類のFDKを用いて、3次元アフィン構造を3次元ユークリッド構造に変換するアフィン変換 (\mathbf{A}, \mathbf{b}) を推定する。6種類のFDKは独立ではなく、互いに影響し合っている。そこで、各FDKを別々に扱うのではなく、FDKの和をとることによって、FDK間の相互関係をより記述できると考え、まず、次元の異なるFDKをマハラノビス距離で表し、次元の統一 (無次元化) を行い、得られたマハラノビス距離の二乗和を最小にすることで (\mathbf{A}, \mathbf{b}) の推定を行う。

$$\begin{aligned} J &= \sum r_p^2 + \sum r_d^2 \\ &+ \sum r_\theta^2 + \sum r_u^2 + \sum r_r^2 \rightarrow \min. \end{aligned} \quad (13)$$

これは、非線形の最小化問題であり、その解法としての標準的な方法である、Levenberg-Marquardt法⁶⁾を用いて推定を行った。なお、初期値は $\mathbf{A} = \text{diag}(1, 1, 1)$ 、 $\mathbf{b} = [0, 0, 0]^T$ とした。

ここで、アフィン構造からユークリッド構造へのアフィン変換の自由度は12であり、基準座標系の設定を行うことで、自由度は6減るので、残りの自由度は6となる。したがって、少なくともあと6個の拘束条件があれば、アフィン変換 (\mathbf{A}, \mathbf{b}) を推定することができる。

基準座標系 (各軸の方向) が設定されていれば、あとは座標系における長さが定義できれば (\mathbf{A}, \mathbf{b}) を推定することができる。長さを定義するためには、たとえば、簡単なものでは、同一平面上にない、いくつかのベクトルの長さ各軸との角度が分かればよい。拘束条件が最低6個必要であることを考慮し、これをFDKで表すと、「ベクトルの長さ $\times 3$ 」と「2つのベクトルのなす角度 $\times 3$ 」となる。以上で、拘束条件数は、合計12となり、 (\mathbf{A}, \mathbf{b}) を推定することができる。なお、基準座標系の設定を行う場合は、2.2節で述べた手法を用いる必要があり、式(2)、(3)、(4)で与えられる6つの拘束条件が、式(13)に追加される。

基準系の設定を行わずに、FDKのみを用いて復元を行う場合は、2.2節で与えられる条件と同様な条件をFDKにより与えればよい。たとえば、基準座標系の設定には、原点に変換される点と、各軸の方向が分

表 2 復元結果
Table 2 Result of Euclidean reconstruction.

point	Truth			Estimated			Affine		
	x	y	z	x	y	z	x	y	z
0	0	0	0	0.0	0.0	0.0	101.0	403.5	-106.2
1	100	0	0	100.0	0.0	0.0	333.0	440.5	-451.9
2	100	100	0	100.8	99.9	0.0	333.0	199.5	-482.6
3	0	100	0	0.8	99.9	0.0	101.0	163.5	-136.6
4	0	0	25	0.4	0.0	25.0	129.0	383.0	-132.7
5	100	0	25	100.0	0.0	24.9	360.0	420.0	-477.1
6	100	100	25	100.8	100.0	25.0	360.0	179.0	-507.7
7	0	100	25	0.4	99.5	25.1	129.0	142.0	-163.3

かればよい. さらに, ユークリッド座標系での長さが定義できるような FDK を与えれば, (\mathbf{A}, \mathbf{b}) を推定することができる. つまり, 3 つのベクトルの方向 (各軸に平行) とその長さを与えればよい. これらを FDK で表すと,

- (1) 3次元アフィン座標 \mathbf{M}_i^A は, 3次元ユークリッド座標 $[0, 0, 0]^T$ に変換される (点のユークリッド座標).
- (2) ベクトル $\mathbf{A}\hat{\mathbf{v}}_i$ は, X 軸に平行なベクトル $\hat{\mathbf{v}}_i$ と等しい (ベクトルの同一性).
- (3) ベクトル $\mathbf{A}\hat{\mathbf{v}}_j$ は, Y 軸に平行なベクトル $\hat{\mathbf{v}}_j$ と等しい (ベクトルの同一性).
- (4) ベクトル $\mathbf{A}\hat{\mathbf{v}}_k$ は, Z 軸に平行なベクトル $\hat{\mathbf{v}}_k$ と等しい (ベクトルの同一性).

となる. ここで, $\mathbf{A}\hat{\mathbf{v}}_i$ は, それぞれ, ユークリッドベクトルの推定値と期待値である. 以上より, 拘束条件は 12 となり, アフィン変換 (\mathbf{A}, \mathbf{b}) を推定することができる.

2.5 復元実験

図 1 の入力画像を用いて, 3次元ユークリッド構造の復元実験を行った. 特徴点は, 直方体の 8 個の頂点とし, その中の 3 点を用いて, 基準ユークリッド座標系の設定を行った. FDK としては, 同一平面上にない 3 つのベクトルの長さとして 3 つの角度を用いた.

表 2 に復元された各特徴点のアフィン座標値, 3次元ユークリッド座標値を実際の座標値とともに示す. 各特徴点のユークリッド座標値は, 実際の座標値とはほぼ等しい結果が得られている. なお, このときのアフィン変換 (\mathbf{A}, \mathbf{b}) は,

$$\mathbf{A} = \begin{bmatrix} -0.5405 & 0.0788 & -0.6445 \\ -0.7885 & -0.3427 & -0.5657 \\ 2.068 & -0.1741 & 1.369 \end{bmatrix},$$

$$\mathbf{b} = \begin{bmatrix} -45.94 \\ 157.93 \\ -18.55 \end{bmatrix},$$

であった.

3. 顔画像生成への適用

任意方向・任意表情の顔画像を生成する技術^{7)~9)}は, テレビ会議システムなど, 様々なアプリケーションに応用できる. たとえば, 少ない実画像から様々な顔画像を生成する技術は, ビデオ会議などの情報転送量を少なくすることができる. つまり, 会議の前に一度だけ実画像を送っておけば, 後は, 姿勢や表情のパラメタのみを送るだけでよい. また, バーチャルリアリティなどの多くのアプリケーションでは, 画像だけではなく, 3次元ユークリッド構造が必要とされる場合が多い. いったん 3次元ユークリッド構造が得られれば, 任意方向画像も容易に得ることができる.

3次元ユークリッド構造の復元を行わない場合, 任意方向の新規顔画像の生成は, たとえば, 次のように行うことができる.

2.1 節で示したように, 画像間のアフィン基本行列が得られれば, 画像間で対応のとれている特徴点のペア $(\mathbf{m}, \mathbf{m}')$ から, それらに対応するアフィン構造 \mathbf{M}^A を求めることができる. アフィン構造 \mathbf{M}^A と新規画像 \mathbf{m}'' の間には, 次の投影関係が成り立つ.

$$\mathbf{m}_i'' = \mathbf{B}\mathbf{M}_i^A + \mathbf{c}. \quad (14)$$

ここで, \mathbf{B} は 2×3 の行列であり, \mathbf{c} は 2次元ベクトルである. つまり, \mathbf{B} と \mathbf{c} が得られれば, 新規画像 \mathbf{m}'' を得ることができ, \mathbf{B} と \mathbf{c} を任意に変化させることにより様々な方向からの画像を得ることができる. ここで問題となってくるのは, ある特定の (望みの) 方向の画像を得るための \mathbf{B} と \mathbf{c} をどのように求めるかということである. ここで, たとえば, ある方向から撮影された画像が与えられたとする. これを参照画像として, 同一平面上にない, 少なくとも 4 点の参照画像上における点と, それらに対応する, 3次元アフィン構造上の点を得られれば, \mathbf{B} と \mathbf{c} を簡単に求めることができる. アフィン構造だけからでは, ユー

クリッド空間, つまり通常の3次元空間での特定の方向を指定するための \mathbf{B} と \mathbf{c} を求めるのは困難である.

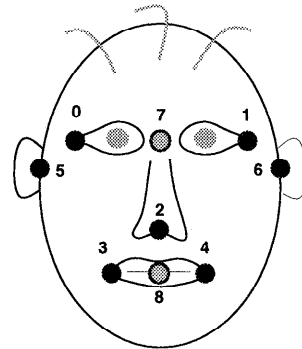
以上の方法を用いることにより, 任意方向の顔画像の生成を行うことができる. しかし, この方法では, 異なる方向の顔画像の生成を行う際には, それに応じた参照画像を用意しなくてはならない. 容易に任意方向の顔画像を得るためには, 顔の3次元ユークリッド構造が得られればよい. そこで, 本論文では, 2章で述べた手法を顔画像に適用し, 顔の3次元ユークリッド構造の復元を行い, 任意方向顔画像の生成を行った.

3.1 顔の FDK

我々は, 顔の形状に対する豊富な知識を持っている. たとえば, 両目を結ぶ線は頭部の垂直対称線とほぼ垂直に交わる. このような, 顔という対象物が属するカテゴリの中で共通する, 距離や角度といった一般的な知識を FDK として用いる. 前述のとおり, FDK は完全には正確ではなく, そのあいまいさをガウス変数でモデル化している. このあいまいさは, 同一カテゴリ内におけるデータのばらつきを表しており, たとえば, 顔の場合, 顔立ちの違いといった, カテゴリ内の個人差を表している. したがって, 同一カテゴリの物体において, 形状にばらつきがあっても, その違いは, すでに FDK に含まれているので, 同一カテゴリの物体であれば, 同一の FDK を用いて3次元ユークリッド構造の復元を行うことができる.

ここで, たとえば, 同じ顔カテゴリであっても, 欧米人の FDK を用いて日本人の顔を復元するなどといった場合は, 同じ顔カテゴリではあるが, 人種のカテゴリの不整合による影響が復元結果に現れてしまう. つまり, どのようなカテゴリに分類し, それらをいかに適切に適用していくかによって, 復元結果が影響を受けてしまう. 上の例の場合は, 日欧, 両方の人物から得られる FDK を用いるか, または, カテゴリとして, 性別, 人種などによる分類を行って FDK を取得し, より近いカテゴリの FDK を用いることで, カテゴリの不整合による影響を軽減することができる.

今回, FDK として, 2点間の距離, 2つのベクトルのなす角度, 2つのベクトルの長さの比率を用い, それらを36名の女性の頭部3次元レンジデータと31名の男性の頭部3次元レンジデータから統計分析によって取得した. FDK を指定するための基準点を図5に示す. 表3には, 今回用いた顔の FDK の一覧を示す. 表3(a), (b), (c) は, それぞれ, 2点間の距離, 2つのベクトルのなす角, 距離の比率の統計分析によって得られた平均値と標準偏差を示している.



- 0: 右目尻
- 1: 左目尻
- 2: 鼻頂点
- 3: 口右端
- 4: 口左端
- 5: 右耳付け根
- 6: 左耳付け根
- 7: 0と1の中点
- 8: 3と4の中点

図5 Fuzzy Domain Knowledge の基準点. 顔の FDK の指定に用いた基準点

Fig. 5 Feature points selected for the domain knowledge base.

表3 顔における Fuzzy Knowledge

Table 3 A subset of the fuzzy knowledge base for faces.

(a) Distances [mm]				
points	female faces		male faces	
	distance	s.d. (σ_d)	distance	s.d. (σ_d)
0 \leftrightarrow 1	122	5	125	6
0 \leftrightarrow 2	84	4	89	5
0 \leftrightarrow 3	95	4	102	6
3 \leftrightarrow 2	63	3	66	4
3 \leftrightarrow 4	58	4	59	5
5 \leftrightarrow 0	99	5	106	5
5 \leftrightarrow 2	159	6	171	6
5 \leftrightarrow 3	132	6	144	7

(b) Angles [degrees]				
vectors	female faces		male faces	
	angle	s.d. (σ_θ)	angle	s.d. (σ_θ)
$\angle(01, 02)$	44	2	46	2
$\angle(01, 34)$	3	2	3	2
$\angle(02, 03)$	41	3	40	3
$\angle(32, 34)$	63	2	65	4
$\angle(50, 03)$	94	4	93	3
$\angle(53, 03)$	44	3	44	2

(c) Ratio of distances				
distances	female faces		male faces	
	ratio	s.d. (σ_r)	ratio	s.d. (σ_r)
d_{02}/d_{32}	1.33	0.08	1.35	0.09
d_{50}/d_{53}	0.75	0.04	0.74	0.03

3.2 3次元ユークリッド構造復元による任意方向画像の生成

3.2.1 ユークリッド構造の復元

2枚の無校正画像を図6に示す. 今回用いた画像は, FDK を取得するために用いた人物データとは異なる人物の画像を用いた.

初めに, 画像間で対応のとれている特徴点を取得する. 特徴点の取得は, 顔の内部については, FUTON



(a) 入力画像-I

(b) 入力画像-II

図 6 入力画像。カメラパラメータが未知のカメラで撮影された同一人物で姿勢の異なる 2 枚の画像

Fig. 6 Input images.



(a) 入力画像-I

(b) 入力画像-II

図 7 特徴点。特徴点の取得は目視による手入力で行う。(a)と(b)の各特徴点に対応している。

Fig. 7 Matched points between two facial images.

システム¹⁰⁾のサンプリングツールを用いて、目視による手入力で行い(83点)、さらに、髪の方に、いくつかの特徴点を手入力により取得した。なお、特徴点は、特徴点どうしを結んだ直線が、顔の造作(眉、目、口など)の境界(エッジ)をなぞるように設定されている。今回用いた特徴点(114点)を図7に示す。ここで、頬の部分のように、オクルージョンが起きている部分では、隠れているエッジをそれに一番近い、見えているエッジに割り当てることとする。つまり、図7に示すように、明らかに見えているエッジ間で対応付けを行う: 図7(b)の右頬のエッジは図7(a)の見えている右頬のエッジに対応するものとする(実際の対応とは異なる)。

次に、顔のFDKとの対応をとるため、図5に示した基準点を入力画像(図6)に対して手入力で指定する。基準点6は入力画像-IIでは見えないので、今回、この点は使用しない。基準点が得られれば、2章で述べられた方法と、3.1節で得られた女性の顔に対するFDKを用いて、顔の3次元ユークリッド構造の復元を行うことができる。今回、基準ユークリッド座標系の設定には、基準点0, 1, 3を用いた。

表4(a)に、本手法により復元された、(a)基準点

表 4 ユークリッド構造復元結果

Table 4 Results on the recovered Euclidean structure.

(a) Euclidean coordinates [mm]						
	Truth			Estimated		
	x	y	z	x	y	z
0	0.0	0.0	0.0	0.0	0.0	0.0
1	122.3	0.0	0.0	122.6	0.0	0.0
2	60.2	49.3	-30.2	62.7	47.1	-37.1
3	28.6	85.7	-1.4	31.4	88.3	0.0
4	91.9	88.4	1.6	91.5	92.8	-3.9
5	-35.5	16.3	91.7	-26.4	19.4	91.3

(b) Distances [mm]

points	Truth	Estimated
0 ↔ 1	122.3	122.6
0 ↔ 2	83.5	86.8
0 ↔ 3	90.4	93.7
3 ↔ 2	56.2	63.7
3 ↔ 4	63.4	60.4
5 ↔ 0	99.7	97.0
5 ↔ 2	158.5	158.7
5 ↔ 3	132.6	128.2

(c) Angles [degrees]

vectors	Truth	Estimated
$\angle(01, 02)$	43.9	43.7
$\angle(01, 34)$	3.4	5.7
$\angle(02, 03)$	37.5	41.1
$\angle(32, 34)$	59.2	61.4
$\angle(50, 03)$	91.7	95.6
$\angle(53, 03)$	48.7	48.9

(d) Ratios of distances

distances	Truth	Estimated
d_{02}/d_{32}	1.49	1.36
d_{50}/d_{53}	0.75	0.76

のユークリッド座標値, (b) 2点間の距離, (c) 2つのベクトルのなす角度, (d) 2つのベクトルの長さの比率を実測値とともに示す。なお、実測値は、入力画像を撮影したときとほぼ同時期に、3次元スキャナ(Cyberware3030)を用いて測定された値である。

本手法では、単純にFDKの値に近づけているわけではなく、その人物固有のアフィン情報を保持した状態でユークリッド構造の復元を行っている。そのため必ずしも推定値がFDKと近い値を示すとは限らない。また、今回の人物は、FDKを抽出するために用いたデータの中に含まれていない人物である。しかし、「日本人女性」という同一カテゴリーに属している。今回、推定値が多少のばらつきがあるものの、実測値に近い値を示したことは、カテゴリーの知識を用いることの有効性、つまり、本手法の有効性を示しているといえる。

得られた3次元ユークリッド構造の特徴点を用い

て三角形パッチを生成し、ワイヤーフレームモデル (WFM) としたものを図 8 に示す。これにテクスチャをマッピングし、任意方向に回転させることにより、任意方向顔画像を容易に作成することができる。

3.2.2 テクスチャ生成

WFM における三角形パッチは、2 枚の入力画像に対して、それぞれに対応する三角形パッチがある (オクルージョンの起こっている部分は除く)。そこで、テクスチャマッピングを行うとき、どちらのテクスチャ

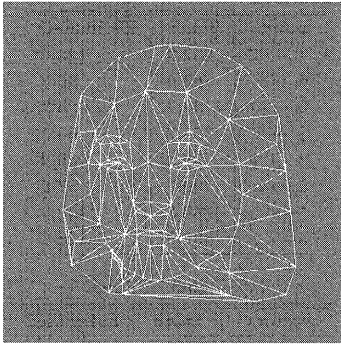


図 8 ワイヤーフレームモデル。本手法により復元された顔の 3 次元ユークリッド構造の特徴点を用いて三角形パッチを生成し、ワイヤーフレームモデルとした。

Fig. 8 Wire-frame-model.

を用いるか決定しなくてはならない。通常、2 枚の画像は同じ色ではなく、対応するテクスチャの大きさも同じでない。テクスチャとして 1 つの画像のみを用いた場合、画像の小さな三角形パッチを WFM の大きな三角形パッチにマッピングしなければならない場合など、生成画像の質が悪くなってしまう。また、今回の例では、入力画像-II では、左耳が見えないといった問題もある。

そこで、本論文では、三角形パッチの面積に応じた重みによって、両方の画像を用いてテクスチャを生成することとする。 Δ_1 と Δ_2 を入力画像の 2 つの対応する三角形パッチとし、 Δ_3 を WFM の対応する三角形パッチとする。これらの面積はそれぞれ S_1, S_2, S_3 で表される。 $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ を対応する三角形パッチ $\Delta_1, \Delta_2, \Delta_3$ の内部のピクセルとする。 $I(\mathbf{p})$ をピクセル \mathbf{p} の RGB 値とすると、 \mathbf{p}_3 の RGB 値は以下の式で計算される。

$$I(\mathbf{p}_3) = \frac{S_1 I(\mathbf{p}_1) + S_2 I(\mathbf{p}_2)}{S_1 + S_2}.$$

WFM に上述の手法で得られたテクスチャをマッピングし、顔モデルとして仕上げる。得られた顔モデルを任意方向に回転させ、それを画像面上に投影することによって、任意方向顔画像を得ることができる。図 9

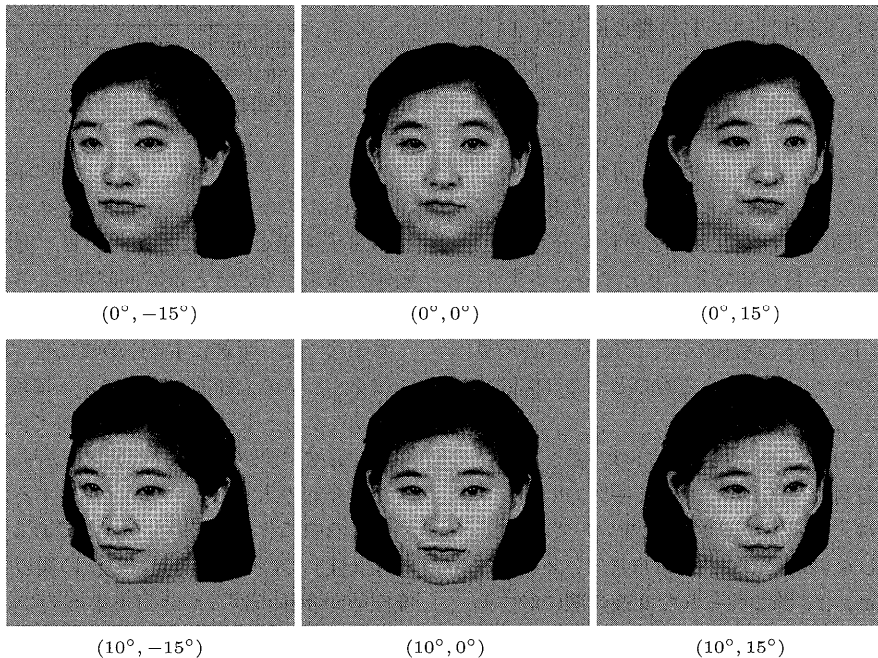


図 9 復元された 3 次元ユークリッド構造からの任意方向顔画像の生成。回転角 (θ, ϕ) で回転させて得られた画像である (3 次元構造を水平軸周りに θ 、鉛直軸周りに ϕ 回転)。

Fig. 9 Synthesized images from the recovered Euclidean shape.

に、本手法によって得られた顔モデルを異なる回転角で回転させ、それらを正射影することによって取得した6枚の生成画像を示す。図中の回転角(θ, ϕ)は、水平軸周りに θ 、鉛直軸周りに ϕ 回転させたことを示している。顔を回転させたとき、頬の部分において不自然さが見られるが、これは、現在用いている特徴点は髪の毛の周りを除いて83点と少なく、頬の部分に特徴点が設定されていないためである。

4. おわりに

本論文では、Fuzzy Domain Knowledge (FDK) を用いて、カメラパラメータが未知のカメラで撮影された2枚の画像から、対象物体の3次元ユークリッド構造を復元する手法について述べた。提案手法を顔画像に適用し、顔の3次元ユークリッド構造の復元を行った。いったん、3次元ユークリッド構造を得ることができれば、任意の方向を簡単に指定することができ、任意方向顔画像も容易に取得することができる。

今後は、今回得られた顔の3次元ユークリッド構造を基に、表情に関するFDKを用いて、任意表情を持つ顔の3次元ユークリッド構造の生成を行っていく予定である。

参考文献

- 1) Zhang, Z. and Xu, G.: A Unified Theory of Uncalibrated Stereo for Both Perspective and Affine Cameras, *Journal of Mathematical Imaging and Vision*, Vol.9, No.3, pp.213-229 (1998).
- 2) Shashua, A.: Projective Structure from Uncalibrated Images: Structure from Motion and Recognition, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.16, No.8, pp.778-790 (1994).
- 3) Koenderink, J. and van Doorn, A.: Affine structure from Motion, *Journal of Optical Society of America*, Vol.8, No.2, pp.377-385 (1991).
- 4) Shapiro, L., Zisserman, A. and Brady, M.: 3D Motion Recovery via Affine Epipolar Geometry, *The International Journal of Computer Vision*, Vol.16, pp.147-182 (1995).
- 5) Xu, G. and Zhang, Z.: *Epipolar Geometry in Stereo, Motion and Object Recognition: A Unified Approach*, Kluwer Academic Publishers (1996).
- 6) Press, W., Teukolsky, S., Vetterling, W. and Flannery, B.: C言語による数値計算のレシピ, 技術評論社 (1993).
- 7) Mukaigawa, Y., Nakamura, Y. and Ohta, Y.: Synthesis of arbitrarily oriented face views

from two images, *Proc. 2nd Asian Conference on Computer Vision*, Singapore, Vol.III, pp.718-722 (1995).

- 8) Seitz, S. and Dyer, C.: View Morphing, *Computer Graphics, Annual Conference Series*, pp.21-30, ACM SIGGRAPH (1996).
- 9) Beymer, D., Shashua, A. and Poggio, T.: Example Based Image Analysis and Synthesis, A.I. Memo No.1431, MIT AI Lab. (1993).
- 10) 蒲池みゆき, 向田 茂, 吉川左紀子, 加藤 隆, 尾田政臣, 赤松 茂: 顔・表情認知に関する心理実験のための顔画像合成システム-FUTON System, 信学技報, HIP97-39, pp.73-80 (1998).

(平成10年12月25日受付)

(平成11年6月3日採録)



磯野 勝宣

昭和45年生。平成4年慶應義塾大学理工学部機械工学科卒業。平成6年同大学院理工学研究科生体医工学専攻修士課程修了。同年(株)インテック入社。平成8年より(株)

ATR人間情報通信研究所に出向。顔イメージの計算機による表現と生成に関する研究に従事。平成10年度電子情報通信学会学術奨励賞受賞。電子情報通信学会会員。



張 正友

昭和40年生。昭和62年仏国ナンシー大学・コンピュータサイエンス修士。平成2年仏国パリXI大学・コンピュータサイエンス博士・工学博士。平成8年から平成9年にかけて(株)ATR人間情報通信研究所客員研究員。仏国

INRIA主任研究員。現在、米国マイクロソフト。コンピュータビジョン、移動ロボット、動的情景解析、センサフュージョン等に関する研究に従事。IEEE会員。

**赤松 茂 (正会員)**

昭和 27 年生。昭和 50 年東京大学工学部計数工学科卒業。昭和 52 年同大大学院工学系研究科計数工学専攻修士課程修了。同年日本電信電話公社（現 NTT）入社。以来、文字認識、画像処理の研究に従事。昭和 60 年から昭和 61 年にかけて米国カリフォルニア大学 Irvine 校客員研究員。平成 4 年より（株）ATR 人間情報通信研究所第二研究室長、現在に至る。平成 11 年より東京工業大学大学院総合理工学研究科客員教授を兼務。工学博士。平成 11 年 5 月より電子情報通信学会ヒューマン情報処理研究専門委員会委員長。パターン認識のヒューマンインタフェースへの応用、感性情報処理、特に人間とコンピュータによる顔情報処理に関心を持つ。IEEE, ACM, 電子情報通信学会, 日本バーチャルリアリティ学会各会員。
