

Title	多重解像度解析による顔濃淡・距離画像の特徴解析に関する研究
Author(s)	金森, 証
Citation	
Issue Date	2002-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/1524">http://hdl.handle.net/10119/1524</a>
Rights	
Description	Supervisor:小谷 一孔, 情報科学研究科, 修士

修 士 論 文

多重解像度解析による  
顔濃淡・距離画像の特徴解析に関する研究

北陸先端科学技術大学院大学  
情報科学研究科情報処理学専攻

金森 証

2002年3月

修 士 論 文

多重解像度解析による  
顔濃淡・距離画像の特徴解析に関する研究

指導教官 小谷一孔 助教授

審査委員主査 小谷一孔 助教授

審査委員 宮原誠 教授

審査委員 阿部亨 助教授

北陸先端科学技術大学院大学  
情報科学研究科情報処理学専攻

910035 金森 証

提出年月: 2002 年 2 月

## 概要

顔画像解析手法は、画像の輝度値や色情報を特徴空間へ写像するなどして画素情報から顔画像特徴を抽出する appearance ベースの手法と、顔画像から輪郭や特徴点を抽出して顔の形状特徴を抽出する model ベースの手法の二つに大別できる。この二つの解析手法は、大局的情報または局所的情報を捉える手法である。

画像解析において、大局的、局所的情報を異なる解像度から抽出、解析する手法に、多重解像度解析がある。顔画像の特徴を多重解像度解析することにより、顔の大局的特徴と局所の特徴が分離、抽出、解析できれば、appearance ベースの手法と model ベースの手法の問題点を軽減しそれぞれの長所を有する顔画像解析が期待できる。

本研究では、顔画像の特徴解析に多重解像度解析を用い、多重解像度による大局的、局所的情報を捉え、各解像度が顔のどのような特徴情報を含んでいるのかを明らかにすることを試みた。解析に用いる画像は、照明による顔の濃淡変化を顔の形状特徴として捉えている顔濃淡画像と、顔の直接的な形状特徴を捉えている顔距離画像を用いた。

顔濃淡・距離画像を離散コサイン変換、離散 Wavelet 変換により多重解像度解析し、顔画像の特徴情報を解析した。顔濃淡・距離画像の多重解像度解析より得られた特徴情報が、個人を特定する情報を含んでいるかについて、個人識別実験を行い評価した。

# 目次

第1章	はじめに	1
1.1	背景	1
1.2	目的	2
第2章	画像の多重解像度解析	3
2.1	画像の多重解像度解析とスペクトル	3
2.1.1	周波数スペクトル	3
2.1.2	パターンスペクトル	7
2.2	画像の多重解像度解析の研究例	8
2.2.1	多重解像度解析に周波数スペクトルを用いた研究例	8
2.2.2	多重解像度解析にパターンスペクトルを用いた研究例	10
第3章	顔画像の多重解像度解析	11
3.1	顔画像データ	11
3.1.1	顔濃淡画像	11
3.1.2	顔距離画像	14
3.1.3	顔濃淡画像と顔距離画像の正規化	17
3.2	顔濃淡・距離画像の多重解像度解析によって得られた各係数の特徴	18
3.3	顔濃淡画像の多重解像度解析	26
3.3.1	離散コサイン変換による多重解像度解析と個人識別特性	26
3.3.2	離散 Wavelet 変換による多重解像度解析と個人識別特性	32
3.3.3	離散コサイン変換と離散 Wavelet 変換による多重解像度解析の個人 識別特性の比較	38
3.4	顔距離画像の多重解像度解析	38
3.4.1	離散コサイン変換による多重解像度解析と個人識別特性	38
3.4.2	離散 Wavelet 変換による多重解像度解析と個人識別特性	42
3.4.3	離散コサイン変換と離散 Wavelet 変換による多重解像度解析の個人 識別特性の比較	49
3.5	顔濃淡画像と顔距離画像の多重解像度解析結果による個人識別特性の比較	50

第4章	特徴量を用いた顔画像の多重解像度解析	52
4.1	顔画像の特徴量による解析	52
4.2	特徴量の定義	52
4.3	特徴量を用いた個人識別実験	54
4.3.1	実験条件と識別方法	54
4.3.2	各特徴量を単独に用いた識別実験	55
4.3.3	各特徴量を組み合わせた個人識別実験結果	57
4.3.4	笑い顔画像の識別実験	57
4.3.5	辞書に無表情と笑い顔画像を組み合わせた場合の識別実験	60
第5章	結言	65
5.1	まとめ	65
5.2	今後の課題	66
	謝辞	68
	付録A JPEG2000 システムにおける顔画像認識システム	69
	参考文献	70

# 第1章 はじめに

## 1.1 背景

顔画像解析手法は、その基本的なアプローチによって次の二つに大別できる。

- appearance ベースの手法
- model ベースの手法

前者は顔画像の輝度値や色情報を特徴空間へ写像するなどして画素情報から顔画像特徴を抽出するものであり、後者は顔画像から輪郭や特徴点を抽出して顔の形状特徴を抽出するものである。それぞれの特徴の概要をまとめると、

### 1. appearance ベースの手法の特徴

画素値を直接解析に用いるので顔画像輪郭や特徴点抽出の精度や特性の影響がない。顔画像の画素数が多くなると計算コストが急激に増加するが、画素数が少なくなると特徴抽出の分解能が低下する。画素の値 (例えば輝度値) はその画素の位置における値なので、画像中の顔の位置や大きさなどの正規化がされていないと特徴抽出精度が著しく低下する。

### 2. model ベースの手法の特徴

顔画像輪郭や特徴点抽出することにより正規化誤差の影響が小さい。また、特徴点の数は画素数よりもずっと少ないので計算コストは appearance ベースの手法よりも少ない。ただし、輪郭検出や特徴点抽出を完璧に行うのは不可能であり、これらの精度や特性の影響が大きい。

上記2手法のいずれが優れているかを結論づけることは難しく、顔画像の研究の現状を見ると、問題点を除外して議論するものや、適用する目的に特化して手法を選択するなどの研究手法がとられている [1]-[10]。

一方、画像が表現できる細かさの程度を表すものとして解像度がある。画像は解像度変換することで、同一の画像を異なる解像度で表現できる。このような一つの画像を複数の解像度に分解することを多重解像度解析という [11]。画像の多重解像度解析によって、画像の大局的・局所的な特徴を異なる解像度の画像から抽出、解析することができる。顔画像を多重解像度解析することにより、顔の大局的特徴と局所的特徴が分離、抽出、解析できれば、

- 大局的特徴は顔画像の大まかな特徴を捉えているからデータ量が少なく処理コストが小さい。位置づれなどの正規化の影響は小さい。
- 局所の特徴は顔画像輪郭や特徴点の情報を有しているので分解能の高い特徴抽出と解析が行える。

上記は顔画像から特徴点抽出などの前処理が行うことなく画素値から直接計算できるので、appearance ベースの手法と model ベースの手法の問題点を軽減し、それぞれの長所を有する顔画像解析が期待できる。

更に別の視点から顔画像解析について考えてみる。顔の特徴の大部分は形状に現れている（顔色というものもあるが、ここでは議論しない）。appearance ベースの手法は、顔の形状特徴を顔画像の画素情報として抽出、解析しようとするものである。これは、照明光が顔にあたり、その反射特性が顔形状特徴を与えると考え、反射光を撮影して画像化することにより形状特徴を画素値として間接的に抽出、解析する。一方、model ベースの手法は顔画像の輪郭や特徴点により、より直接的に形状特徴を抽出、解析するものであるが、輪郭や特徴点に形状特徴を代表させている。

レーザ光を当て測距して、3次元形状を距離画像として直接的に画像化する装置がある[12]。これを顔画像に適用して距離画像（顔距離画像と呼ぶ）を得れば、appearance ベースの手法と model ベースの手法を融合した中間的な手法あるいは第3の手法を実現できるであろう。

これまでの顔画像研究では、上記した顔画像の多重解像度解析や顔距離画像の解析は、まだ十分な研究が行われていない。

## 1.2 目的

本研究では、多重解像度解析により顔画像を多重解像度に分離し、各解像度において顔のどのような特徴情報が含まれているのかを明らかにする。

もし、解像度ごとに得られた特徴情報が、顔による表情の解析（表情解析）や個人の特定（個人識別）の情報として用いることができるならば、解像度ごとに得られた顔画像の特徴情報より表情解析や個人識別などの顔画像解析が期待できる。そこで、多重解像度に分離した顔画像の特徴情報を解析し、個人識別を試みる。

顔画像の特徴解析に用いる顔画像は、2次元の濃淡パターン情報を表す顔濃淡画像と、3次元の形状情報を表す顔距離画像の2種類の画像に対して解析を試みる。



## 第2章 画像の多重解像度解析

### 2.1 画像の多重解像度解析とスペクトル

画像が表現できる細かさの程度を示すものとして解像度がある。画像は、異なる解像度で複数の画像を表現できる。このような一つの画像を複数の解像度に分離する手法に、多重解像度解析がある。画像の多重解像度解析は、画像の大局的・局所的な特徴情報を異なる解像度の画像により捉えることが可能である。画像を多重解像度解析する手法には、周波数スペクトルを用いる手法とパターンスペクトルを用いる手法がある。周波数スペクトルによる多重解像度解析には、離散フーリエ変換、離散コサイン変換、離散 Wavelet 変換等が用いられ、パターンスペクトルによる多重解像度解析には、モルフォロジーが用いられている。

#### 2.1.1 周波数スペクトル

周波数スペクトルは、画像の周波数と位相に関する特徴を表す。画像のすべての領域に対して基本波形により分解するため、画像の大局的・局所的な特徴を捉えることができる。周波数スペクトルによる多重解像度解析には、離散フーリエ変換 (DFT)、離散コサイン変換 (DCT) および離散 Wavelet 変換がある。以下に、それぞれの解析手法の特徴について述べる。

##### 離散フーリエ変換

離散フーリエ変換は、非周期的な画像信号を正弦波と余弦波に分解、あるいは非周期的な画像信号を正弦波と余弦波の合成として表現する手法である。離散フーリエ変換は、異なる周波数の正弦波と余弦波により多重解像度解析を行う。この変換により、画像信号中に、どのような周波数成分をもつかを位相と振幅の情報によりを知ることができる。この位相や振幅は、複素数を用いて一つの数として表す。画像の周波数情報を複素数を用いて表現するため、計算量が膨大となる問題がある。また、画像中の画素の位置情報を保有していないため、画像のどこに、どの様な周波数成分が含まれているのかについて視覚的に捉えることができない。以下に、離散フーリエ変換の定義を示す。

離散フーリエ変換の定義式を式 2.1 に示す。  $h(m, n)$  は画像信号、  $H(e^{j\omega_1}, e^{j\omega_2})$  は  $h(m, n)$

のスペクトルを表す.  $M, N$  は, それぞれ画像の縦・横のサイズを表す.  $m$  は水平方向,  $n$  は垂直方向の画像の座標を表す.  $\omega_1, \omega_2$  は, それぞれ水平方向, 垂直方向の周波数を表す.  $j$  は,  $j = \sqrt{-1}$  とする.

$$H(e^{j\omega_1}, e^{j\omega_2}) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} h(m, n) e^{-j\omega_1 m} e^{-j\omega_2 n} \quad (2.1)$$

## 離散コサイン変換

離散コサイン変換は, 非周期的な画像信号を余弦波のみを用いて分解する, あるいは非周期的な画像信号を余弦波のみの合成として表現する手法である. 2次元の離散コサイン変換は, 画像信号を図 2.1 に示した基底画像の分解, 合成により表現する. すなわち, 離散コサイン変換により変換された画像は, 元の画像を周波数成分別に分解したときの成分の大きさを表している. 離散コサイン変換は, 基底画像により多重解像度解析を行う. この解析手法によって, 離散フーリエ変換と同様, 画像信号がどのような周波数成分をもつかを知ることができる. 離散コサイン変換は, 画像情報の表現において, 離散フーリエ変換と異なり, 画像の位相や振幅の情報を複素数を用いて表さない.

このような特徴を持つ離散コサイン変換は, 実世界での技術応用として JPEG の圧縮方式に用いられている. 圧縮方式では, 変換の対象画像を複数のブロックに分割し, 各ブロックに対して, 離散コサイン変換を行う. 各ブロックごとに離散コサイン変換を行うため位置情報を保有した解析が可能である. しかし, 離散コサイン変換は, 各ブロックごとに分けて変換を行うため, ブロック間で画像信号の関係が打ち切られる. そのためブロック歪みとして画像に表れるといった問題がある. 以下に, 離散コサイン変換の定義を示す.

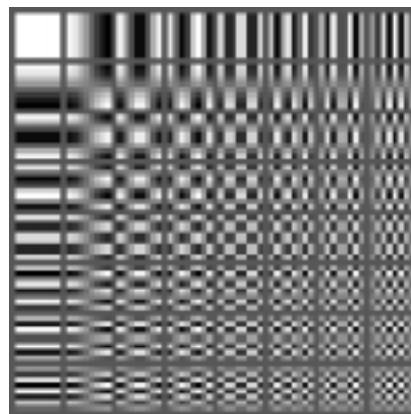


図 2.1: 2次元 DCT(8 × 8) の基底画像の例

離散コサイン変換の定義を式 2.2 に示す.  $o(m, n)$  は画像信号,  $O(u, v)$  は基底画像の変換係数を表す.  $N$  は, 画像を水平・垂直方向に小さなブロックに分割するときのサイズを表す.  $m, u$  は水平方向,  $n, v$  は垂直方向の画像の座標を表す.  $C(u), C(v)$  は, 式 2.3, 2.4 に示す.

$$O(u, v) = \left(\frac{2}{N}\right) C(u)C(v) \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} o(m, n) \cos \left[ \frac{(2m+1)u\pi}{2N} \right] \cos \left[ \frac{(2n+1)v\pi}{2N} \right] \quad (2.2)$$

$$C(u) = \begin{cases} \frac{1}{\sqrt{2}} & u = 0 \\ 1 & u \neq 0 \end{cases} \quad (2.3)$$

$$C(v) = \begin{cases} \frac{1}{\sqrt{2}} & v = 0 \\ 1 & v \neq 0 \end{cases} \quad (2.4)$$

## 離散 Wavelet 変換

Wavelet 変換は, 画像信号から, Mother Wavelet と呼ばれる波形と相似なものを抽出する解析手法である. ここで Mother Wavelet とは, 三角関数や対数関数のような特定の決まった関数ではなく, 局在する波を表す様々な関数の総称である. 画像変換の基底となる Mother Wavelet は, 空間の階層構造を利用して作られる. 離散 Wavelet 変換は, この Mother Wavelet の各階層構造により, 画像を多重解像度解析する. 画像の解析は, Mother Wavelet を伸縮, 平行移動することによって, 解析する波形中の Mother Wavelet と相似な様々なスケールの波形を画像の位置情報を保有したまま抽出することができる. 画像に離散 Wavelet 変換による分解アルゴリズムを適用することで, 画像を低域成分, 水平, 垂直, 対角方向成分の 4 つに分離することができる. 離散 Wavelet 変換による 2 次元の分解アルゴリズムの構成図を図 2.2 に示す. 図 2.2 の HPF はハイパス・フィルタ, LPF はローパス・フィルタを表す.

離散フーリエ変換に対する離散 Wavelet 変換の特徴として, 離散フーリエ変換が位相や振幅の情報を複素数を用いて表すのに対し, 離散 Wavelet 変換では, 実数で表す. また, 離散コサイン変換に対する離散 Wavelet 変換の特徴として, 離散コサイン変換では, 重複しない隣接した画像のブロックを切り出し, そのブロックに変換を行うが, 離散 Wavelet 変換では, 画像データを重複するように切り出し, 変換を行う. よって, 画像信号の隣接関係を打ち切られることがなく, 多重解像度解析が行える.

離散 Wavelet 変換は, 以下の式で与えられる.

$$w^j(x) = 2^j \sum_{k=1}^8 \overline{\phi(2^j x - k)} f(x) \quad (2.5)$$

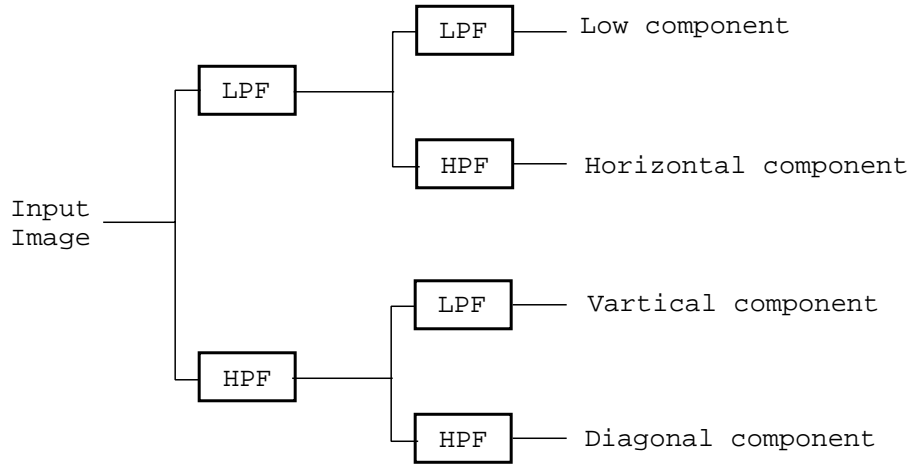


図 2.2: 離散 Wavelet 変換による 2次元分解アルゴリズムの例

ここで,  $f(x)$  は任意の信号,  $w$  は wavelet 係数,  $\phi(x)$  はスケーリング関数,  $k$  は wavelet の大きさを表す<sup>1</sup>.  $j$  は解像度を表し, ここではレベルと呼ぶ. スケーリング関数  $\phi(x)$  は, 次の Two-Scale 関係を満たす.

$$\phi(x) = \sum_{k=1}^8 p_k \phi(2^j x - k) \quad (2.6)$$

数列  $\{p_k\}$  は Two-scale 数列と呼ばれ, 数列  $\{p_k\}$  によってスケーリング関数  $\phi(x)$  は決定される. また, Mother wavelet  $\psi(x)$  は,

$$\psi(x) = \sum_{k=1}^8 q_k \psi(2^j x - k) \quad (2.7)$$

によって決定される. 数列  $\{q_k\}$  も Two-scale 数列と呼ばれる.  $\phi(x)$  が与えられると,  $\{\phi(2^j x - k)\}$  の張る空間  $V_j$  が決定する. 空間  $V_j$  の任意の関数  $f_j$  は次のように表される.

$$f_j(x) = \sum_{k=1}^8 s_k^j \phi(2^j x - k) \quad (2.8)$$

Two-Scale 関係から空間  $V_j$  の階層構造

$$\cdots \subset V_{j-1} \subset V_j \subset V_{j+1} \subset \cdots \quad (2.9)$$

が決まる. また,  $\phi(x)$  に対応する  $\psi(x)$  が与えられると  $\{\psi(2^j x - k)\}$  の張る空間  $W_j(x)$  が決まり, 任意の関数  $g_j \in W_j$  は次のように表される.

$$g_j(x) = \sum_{k=1}^8 w_k^j \psi(2^j x - k) \quad (2.10)$$

<sup>1</sup>本研究では  $k = 8$  とした

これらの関係から、スケーリング関数, Wavelet 関数により張られた空間  $V_j, W_j$  には,

$$f_j(x) = f_{j-1}(x) + g_{j-1}(x) \quad (2.11)$$

という関係が成り立つ. 各レベル  $j$  における  $f_j(x), g_j(x)$  を用いて多重解像度解析を行う. 以降では, 各レベル  $j$  の水平, 垂直, 対角方向の Wavelet 係数は, それぞれ  $w^{j,h}, w^{j,v}, w^{j,d}$  と表し, それに対応する関数を  $g_j^h(x), g_j^v(x), g_j^d(x)$  と表す.

## 2.1.2 パターンスペクトル

パターンスペクトルは, 画像のサイズの異なる構造要素と呼ばれる基本形状で分解した重ね合わせとして定義される. 画像の構造要素による分解は, 画像のすべての領域に対して行われるため, 画像の大局的な特徴を捉えることができる. これは, 特徴点などの画像の局所的な特徴を捉える画像解析に比べ, 前処理や正規化処理が必要ないことから, 特徴抽出処理を容易に行うことができるという利点がある.

パターンスペクトルは, 周波数スペクトルを抽出する正弦波の方向や振幅などの情報を構造要素と呼ばれるパターンに圧縮しているため, 構造要素次第で画像の方向性に関してモデル化することができる. また, 構造要素を回転対象図形にすることで, 画像の位相変化だけでなく回転変化に対しても不変な特徴量となる.

パターンスペクトルによる解析手法には, モルフォロジーがある. 以下に解析特徴を述べる.

### モルフォロジー

モルフォロジーは, 集合演算的操作からなる 2 値画像または濃淡画像の変形手法の一貫した理論体系で, 処理対象画像と構造要素およびそれらの集合演算で様々な画像処理を実現できる. 同一の演算であっても, 構造要素の選び方によって演算結果が大きく異なる特徴を持つ. モルフォロジーで定義される演算には, 膨張, 収縮, スケルトン化演算などがある.

モルフォロジーには 2 値画像を扱うバイナリーモルフォロジーと濃淡画像を扱うグレイスケールモルフォロジーがある. グレイスケールモルフォロジーは理論的にバイナリーモルフォロジーを含むため, ここではグレイスケールモルフォロジーを説明する.

整数全体の集合を  $\mathbb{Z}$ , 2 次元離散空間を  $\mathbb{Z}^2$  とし, 各  $(x, y) \in \mathbb{Z}^2, (u, v) \in \mathbb{Z}^2$  が濃淡値  $f(x, y) \in \mathbb{Z}, g(u, v) \in \mathbb{Z}$  を持つとする. ここで,  $f$  を処理対象画像,  $g$  を構造要素とすると,  $f(x, y)$  と  $g(u, v)$  の Minkowski 和及び Minkowski 差は式 2.12, 2.13 で与えられる. ただし,  $F, G$  は関数  $f, g$  のそれぞれの定義域とする.

$$\text{Minkowski 和} : (f \oplus g)(x, y) = \max_{\substack{(x-u, y-v) \in F \\ (u, v) \in G}} \{f(x-u, y-v) + g(u, v)\} \quad (2.12)$$

$$\text{Minkowski 差} : (f \ominus g)(x, y) = \min_{(u,v) \in \mathbf{G}} \{f(x-u, y-v) - g(u, v)\} \quad (2.13)$$

また,  $\check{g}(u, v) = g(-u, -v)$  とすると, opening, closing と呼ばれる演算は, それぞれ式 2.14, 2.15 で与えられる.

$$\text{opening} : f_g(x, y) = ((f \ominus \check{g}) \oplus g)(x, y) \quad (2.14)$$

$$\text{closing} : f^g(x, y) = ((f \oplus \check{g}) \ominus g)(x, y) \quad (2.15)$$

なお, バイナリーモルフォロジーを扱うには, 処理対象画像及び構造要素に式 2.16, 2.17 の制約条件を加える.

$$f(x, y) = \{0 \text{ or } 1 \mid f(x, y) \in \mathbf{F}\} \quad (2.16)$$

$$g(u, v) = \{0 \mid g(u, v) \in \mathbf{G}\} \quad (2.17)$$

処理対象を画像ではなく 1 次元関数とし, モルフォロジーの基本演算である Minkowski 和, Minkowski 差, opening, closing の演算例を図 2.3 に示す.

## 2.2 画像の多重解像度解析の研究例

画像を周波数スペクトルやパターンスペクトルにより多重解像度解析し, 各解像度の情報が有効な特徴量となることを数多くの研究が示している. 以下に周波数スペクトルとパターンスペクトルを特徴量に利用した画像の研究例を示す.

### 2.2.1 多重解像度解析に周波数スペクトルを用いた研究例

以下に, 離散フーリエ変換, 離散コサイン変換, 離散 Wavelet 変換を多重解像度解析に用いた研究例を示す.

- フーリエスペクトルを用いた顔画像識別  
赤松らは, 顔濃淡パターンをフーリエスペクトルに変換し, 主成分分析により次元圧縮を行って識別に用いる特徴ベクトルを求め, 識別を行った. [1]
- 離散コサイン変換による個人識別  
Stefan らは, 顔画像を離散コサイン変換し, 得られた DCT 係数を特徴情報とし確率モデルを用いて個人識別を行った. [13]
- Wavelet 変換を用いたテクスチャの特徴解析  
中田らは, テクスチャ画像を Wavelet 変換により多重解像度解析し, 多重空間で画像の持つ特徴を捉え, テクスチャ画像を分類した. [14]

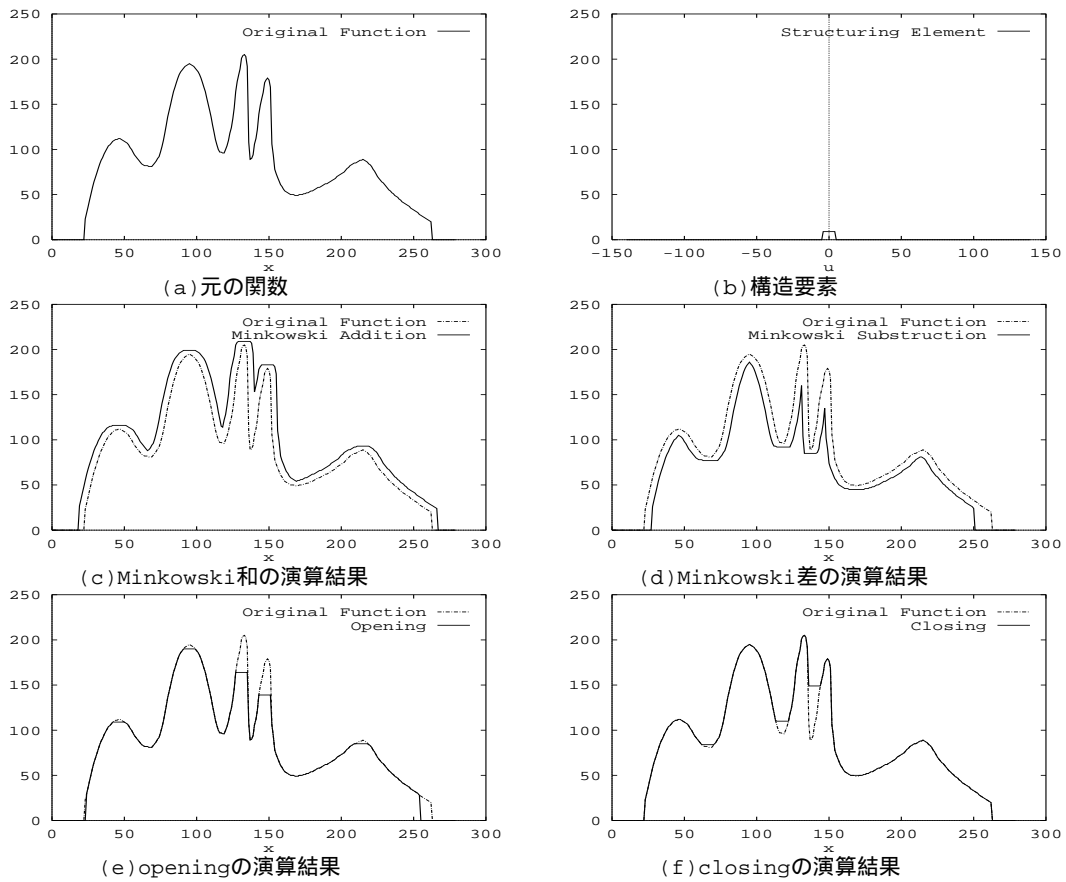


図 2.3: Minkowski 和、Minkowski 差、opening、closing の演算例 (図 (c)-(f) で実線は演算結果、点線は元の図形を示す。)

- 実ガボール変換による個人識別  
吉田らは, 顔の特徴点を自動抽出し, その特徴点のみを実ガボール変換し, その変換結果と特徴点の位置関係より個人識別を行った. [15]
- ウェーブレット変換を用いた南極域衛星画像のテクスチャ解析  
越中らは, ウェーブレット変換により得られた多重解像度の多次元ベクトル情報を用いて, 南極域の雲と地表の画像特徴を解析した. [16]

### 2.2.2 多重解像度解析にパターンスペクトルを用いた研究例

以下に, 多重解像度解析にパターンスペクトルを用いた研究例を示す.

- 3次元画像の構造解析  
入山らは, モルフォロジーを用いて3次元画像を多重解像度解析し, 3次元構造を解析した.[17]



## 第3章 顔画像の多重解像度解析

### 3.1 顔画像データ

顔画像の解析において、どのような画像を用いるのかは重要な問題である。顔画像の情報には、顔の照明変化を輝度の濃淡として捉える情報と、顔の3次元形状を顔と計測装置間の距離として捉える情報がある。ここでは、前者を2次元の濃淡パターンを用いるため顔濃淡画像とし、後者は顔の3次元形状を距離値として用いるため顔距離画像とする。

この2つの画像を比較すると、顔濃淡画像は、顔の濃淡変化などを捉えることができるが、正確な顔の形状情報を捉えることができない。一方、顔距離画像は、距離値により顔の3次元形状を捉えることができるが、顔の色などの色情報を捉えることができない。つまり、この2つの画像は、顔よりそれぞれ異なる情報を捉えていると考える。

そこで、本研究では顔の異なる情報を有している顔濃淡画像と顔距離画像の2つの画像を用いて顔の特徴解析を試みる。

#### 3.1.1 顔濃淡画像

顔濃淡画像は、顔の形状特徴を輝度の濃淡により表現したものである。顔濃淡画像の形状特徴は、照明光が顔にあたり、その反射特性が顔の形状特徴を与えると考える。よって、顔濃淡画像は、照明による顔の濃淡変化を、顔の形状特徴として捉えている。

顔濃淡画像の撮影条件を以下に示す。顔濃淡画像の撮影は、異なる時間帯に撮影した。

#### 顔濃淡画像の撮影条件

顔濃淡画像の撮影に用いた機材を表3.1に、モデルとカメラとの位置関係を図3.1に示す。顔濃淡画像は、顔の頭頂から顎までがフレームに入るように撮影した。(図3.2)。

実験に用いた顔濃淡画像は、目と鼻の下を基準に顔の位置、大きさ、傾きの正規化を行った。正規化方法は、3.1.3節に示す。

表 3.1: 顔濃淡画像の撮影に用いた使用機器

カメラ	NIKON COOLPIX950
撮像素子	1/2 インチ補色系インタレース CCD(総画素数 211 万画素)
レンズ	3 倍ズームニッコールレンズ, 焦点距離 : 7 ~ 21[mm], F2.6 ~ 4
記録画素数	1600 × 1200[pixels]

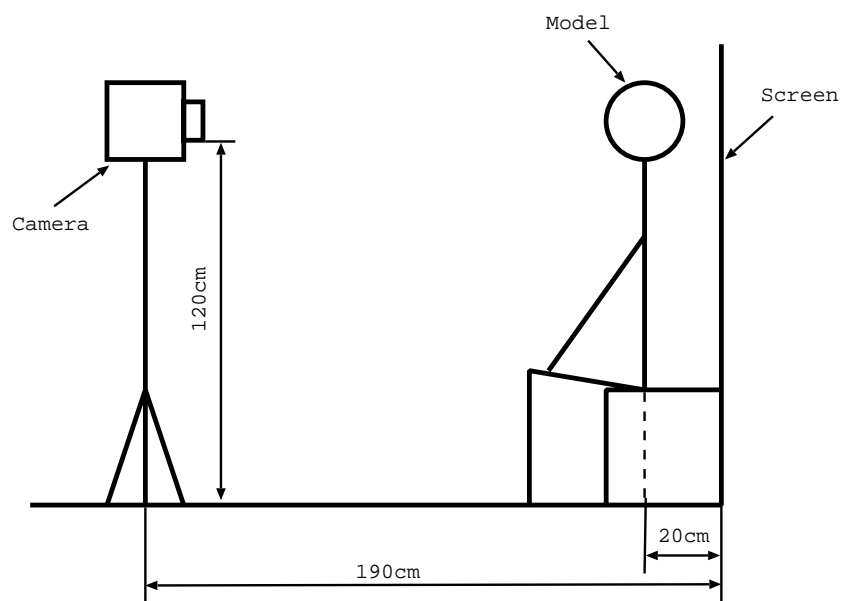
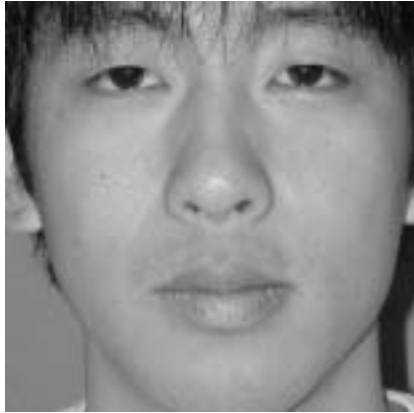
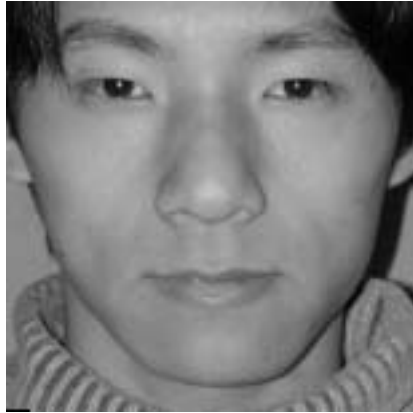


図 3.1: モデルとカメラの位置関係 (顔濃淡画像の場合)



(a) 顔濃淡画像 1



(b) 顔濃淡画像 2



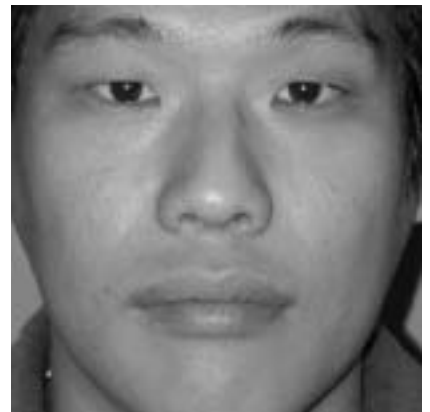
(c) 顔濃淡画像 3



(d) 顔濃淡画像 4



(e) 顔濃淡画像 5



(f) 顔濃淡画像 6

図 3.2: 顔濃淡画像の例

### 3.1.2 顔距離画像

顔距離画像とは、顔の直接的な形状特徴を捉えた画像である。顔距離画像は、3次元計測装置 VIVID700 を用い、モデルと VIVID700 の距離を計測した情報を用いた画像である。この距離値を2次元の濃淡画像の輝度値として捉え、2次元の画像に変換する。よって、顔距離画像の特徴は、照明変化の影響を受けにくい顔の直接的な形状特徴を捉えている。

顔距離画像の撮影条件を以下に示す。顔距離画像の撮影は、異なる時間帯に撮影した。

#### 顔距離画像の撮影条件

顔の距離値の撮影に用いた機材を表 3.2 に、モデルと3次元計測装置 VIVID700 との位置関係を図 3.3 に示す。ここでの顔の距離値とは、モデルと3次元計測装置の距離である。顔の距離値は、顔の頭頂から顎までがフレームに入るようた計測した。顔の3次元形状は、距離値の2次元配列を画像とみなして顔距離画像と呼ぶ。(図 3.4 に距離値によって可視化して表した顔距離画像を示す)。顔濃淡画像と同様に、目と鼻の下の点を基準に顔の位置、大きさ、傾きの正規化を行う。

表 3.2: 顔距離画像の撮影に用いた使用機器

カメラ	MINOLTA VIVID700 : NON-CONTACT 3D DIGITIZER
撮像素子	1/2 インチ 38 万画素フレームトランスファ CCD
レンズ	焦点距離 : 9 ~ 46[mm], F2.6 ~ 4
記録画素数	200 × 200[pixels]
対象物設置距離	0.6 ~ 2.5[m]
画像入力視野域 (視野域 $xy$ の一辺)	70 ~ 1100[mm]
有効画像入力奥行き範囲	画像入力視野域の 1 辺に相当

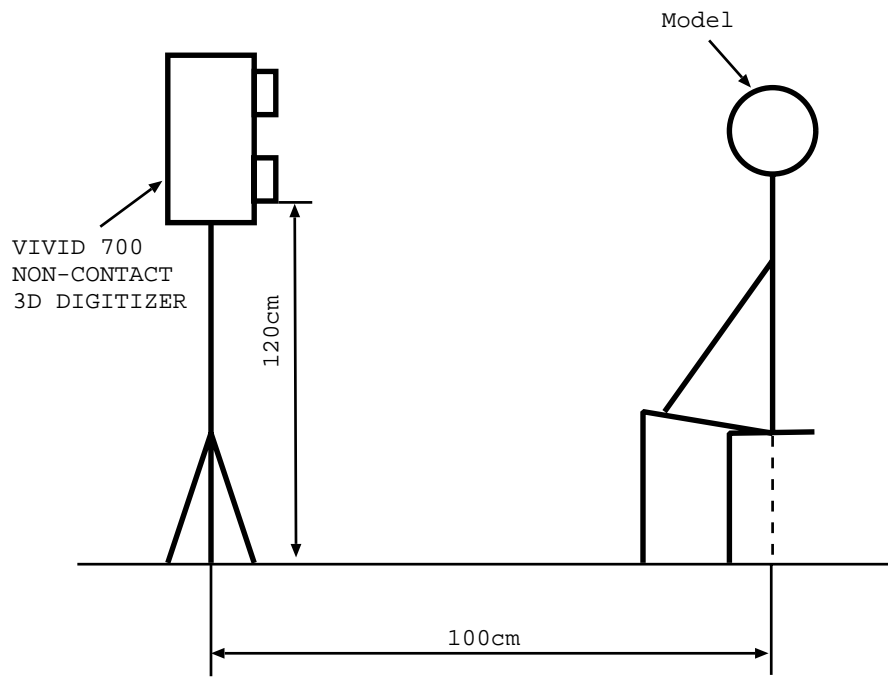


図 3.3: モデルとカメラの位置関係 (顔距離画像の場合)



(a) 顔距離画像 1



(b) 顔距離画像 2



(c) 顔距離画像 3



(d) 顔距離画像 4



(e) 顔距離画像 5



(f) 顔距離画像 6

図 3.4: 顔距離画像の例

### 3.1.3 顔濃淡画像と顔距離画像の正規化

顔濃淡・距離画像の撮影は、異なる時間帯に撮影しているため、照明条件や顔の大きさ・傾き・位置等が異なる。離散コサイン変換や離散 Wavelet 変換による解析では、位置情報を保有した解析手法であることから、解析対象の大きさ・傾き・位置は、特徴解析に影響を与える。そこで撮影した顔濃淡・距離画像を、実験の前処理として正規化を行った。顔濃淡・距離画像の正規化は、図 3.5 に示すように目と鼻の下を基準とし、顔の位置、大きさ、傾きの正規化を以下の手順で行う。また、図 3.5 のパラメータは、顔濃淡、距離画像ともに  $LE$  を 65[pixels]、 $LN$  を 63[pixels] と設定した。

1. 左右の目と鼻先に対応する領域の代表点を  $Er, El, N$  とする。
2.  $Er, El$  を結ぶ直線に鼻先  $N$  より垂線を下ろしたときに交わる点を  $O$  とする。ただし、 $Er_x, Er_y$  は各々  $Er$  の  $x$  座標,  $y$  座標を表し、 $El_x, El_y, N_x, N_y$  も同様に、それぞれ  $El, N$  の  $x$  座標,  $y$  座標を表す。

- $Er, El$  間の距離 :  $D(Er, El)$

$$D(Er, El) = \sqrt{(El_x - Er_x)^2 + (El_y - Er_y)^2} \quad (3.1)$$

- $O, N$  間の距離 :  $D(O, N)$

$$D(O, N) = \sqrt{\frac{\frac{El_y - Er_y}{El_x - Er_x}(N_x - Er_x) - (N_y + Er_y)}{1 + \left(\frac{El_y - Er_y}{El_x - Er_x}\right)^2}} \quad (3.2)$$

- $x$  軸に対する直線  $ON$  の傾き :  $\theta$

$$\theta = \tan^{-1} \frac{El_y - Er_y}{El_x - Er_x} \quad (3.3)$$

- $D(Er, El) = LE, D(O, N) = LN$  かつ直線  $ErEl$  に対して直線  $ON$  が垂直になるように顔画像全体を *affine* 変換する。  $[xy]^T$  を変換前の座標,  $[XY]^T$  を変換後の座標とすると以下のように書ける。

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \cdot \begin{bmatrix} LE/D(Er, El) & 0 \\ 0 & LN/D(O, N) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (3.4)$$

- 鼻を中心に、図 3.5 に示す矩形領域を切り出す。ここで  $X, Y = 256$  とする。

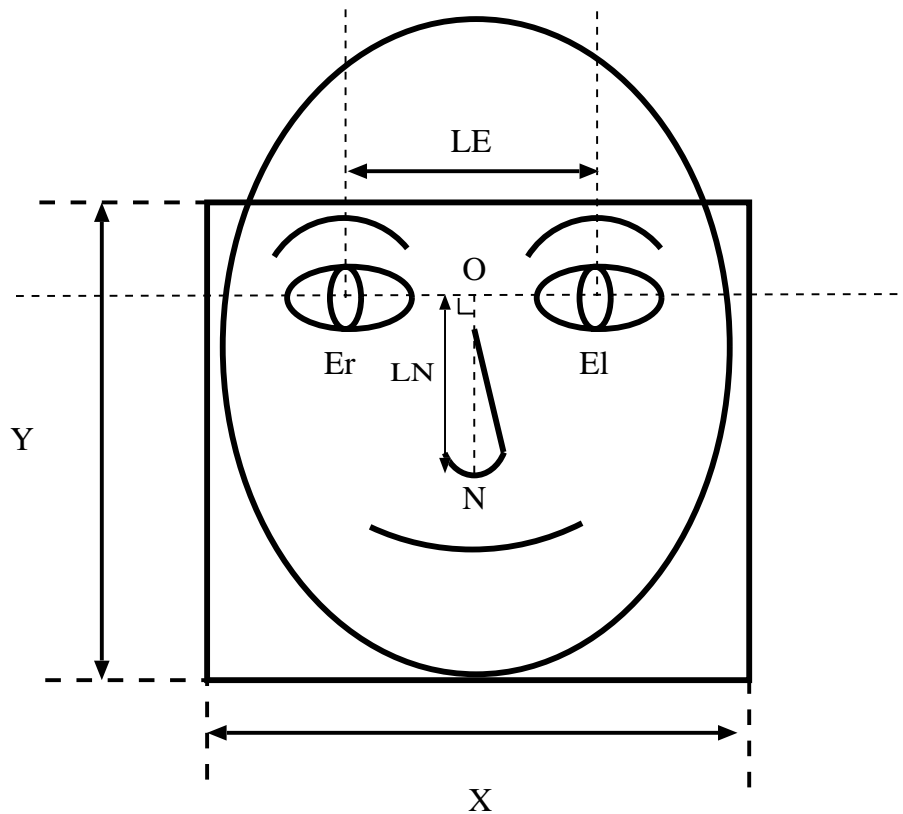


図 3.5: 顔の位置, 大きさ, 傾きの正規化

### 3.2 顔濃淡・距離画像の多重解像度解析によって得られた各係数の特徴

多重解像度解析は, 大局的・局所的な情報を異なる解像度から得る手法である. これを顔画像に適用することで, 顔の大局的・局所的解析が行える. 顔画像の特徴解析において, 多重解像度解析された顔濃淡・距離画像が, 各解像度ごとに顔のどのような特徴情報を捉えているかについて調べる. 実際に, 顔濃淡・距離画像を多重解像度解析し, 各解像度の特徴情報を可視化し, 視覚的に解析を試みる.

視覚から捉える顔の特徴として, 顔の部位やその部位の位置関係は特徴情報を有していると考え, 位置情報を保有した離散コサイン変換および離散 Wavelet 変換により多重解像度解析を行う.

式 3.5 により変換係数が 0 ~ 255 の輝度値の範囲の値に収まるように変換して, DCT 係数または Wavelet 係数の可視化し, 各係数の特徴を検討する.

$$\text{輝度値} = \frac{\text{DCT 係数または Wavelet 係数}}{\text{DCT 係数または Wavelet 係数の最大値と最小値の差}} \times 128 + 100 \quad (3.5)$$

顔濃淡画像と顔距離画像を離散コサイン変換し, DCT 係数を可視化した例をそれぞれ



図 3.7, 3.8 に示す. 図 3.7, 3.8 は, 顔濃淡画像と顔距離画像を離散コサイン変換により多重解像度解析し, 図 3.6 に示す基底画像の各基底成分別に可視化した顔画像である. 図 3.6 は,  $u = 0, v = 0$  の基底画像は直流成分を表し,  $u$  が大きくなるほど高い水平周波数を,  $v$  が大きくなるほど高い垂直周波数を表す. 図 3.7, 3.8 の各顔画像のラベルは, 可視化に用いる基底成分の番号を表す. 図 3.7, 3.8 より, 低域成分を表す基底  $(0, 0)$  による可視化の輝度値が高いことから, 顔の情報は, 低域成分に多いことが分かる.

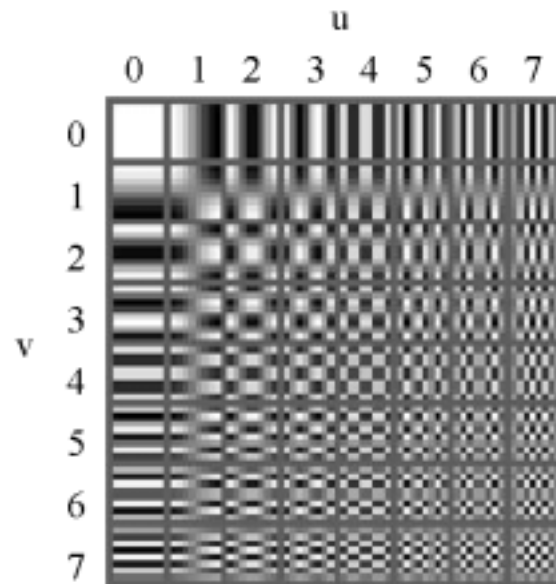


図 3.6: DCT の基底画像

また, 顔濃淡画像と顔距離画像を離散 Wavelet 変換し, Wavelet 係数を可視化した例をそれぞれ図 3.9, 3.10, 3.11, 3.12 に示す. 図 3.9, 3.11 は, 顔濃淡・距離画像に, 多重解像度解析を 1 回行って得られた, 各方向成分を可視化した画像である. 図 3.9, 3.11 の (c), (d), (e) には, 顔濃淡画像と顔距離画像の水平, 垂直, 対角方向のエッジ成分等の高周波成分が表れ, 図 3.9, 3.11 の (b) には, その他の低域成分が表れている. 図 3.9, 3.11 より, 離散 Wavelet 変換による顔画像の多重解像度解析では, 顔の部位や輪郭などのエッジ部分の情報を方向別に捉えていることが分かる. 図 3.10, 3.12 は, 低域成分に再度離散 Wavelet 変換を行い, 顔画像を多重解像度に分離した画像である.



(a) 基底 (0, 0)

(b) 基底 (0, 1)

(c) 基底 (0, 2)



(d) 基底 (1, 0)

(e) 基底 (1, 1)

(f) 基底 (1, 2)

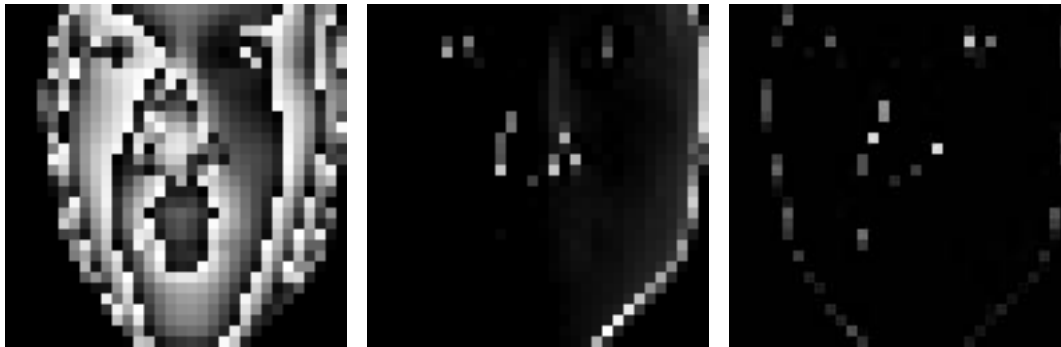


(g) 基底 (2, 0)

(h) 基底 (2, 1)

(i) 基底 (2, 2)

図 3.7: 顔濃淡画像を DCT した DCT 係数の可視化の例



(a) 基底 (0, 0)

(b) 基底 (0, 1)

(c) 基底 (0, 2)



(d) 基底 (1, 0)

(e) 基底 (1, 1)

(f) 基底 (1, 2)

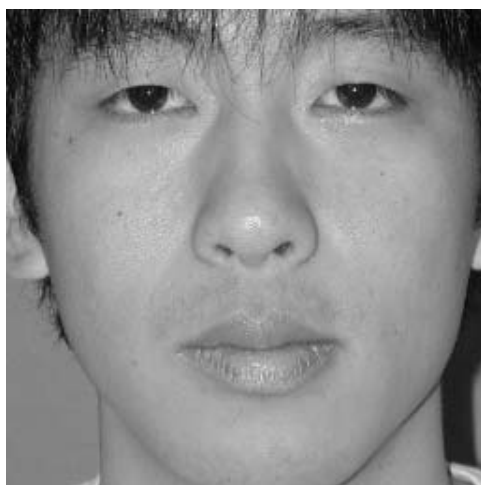


(g) 基底 (2, 0)

(h) 基底 (2, 1)

(i) 基底 (2, 2)

図 3.8: 顔距離画像を DCT した DCT 係数の可視化の例



(a) 顔濃淡画像  $f_j(x)$



(b) 低域成分  $f_{j-1}(x)$



(c) 水平方向成分  
 $g_{j-1}^h(x)$



(d) 垂直方向成分  
 $g_{j-1}^v(x)$



(e) 対角方向成分  
 $g_{j-1}^d(x)$

図 3.9: 顔濃淡画像を Wavelet 変換した Wavelet 係数の可視化 ( $j = 1$ , Level 1)

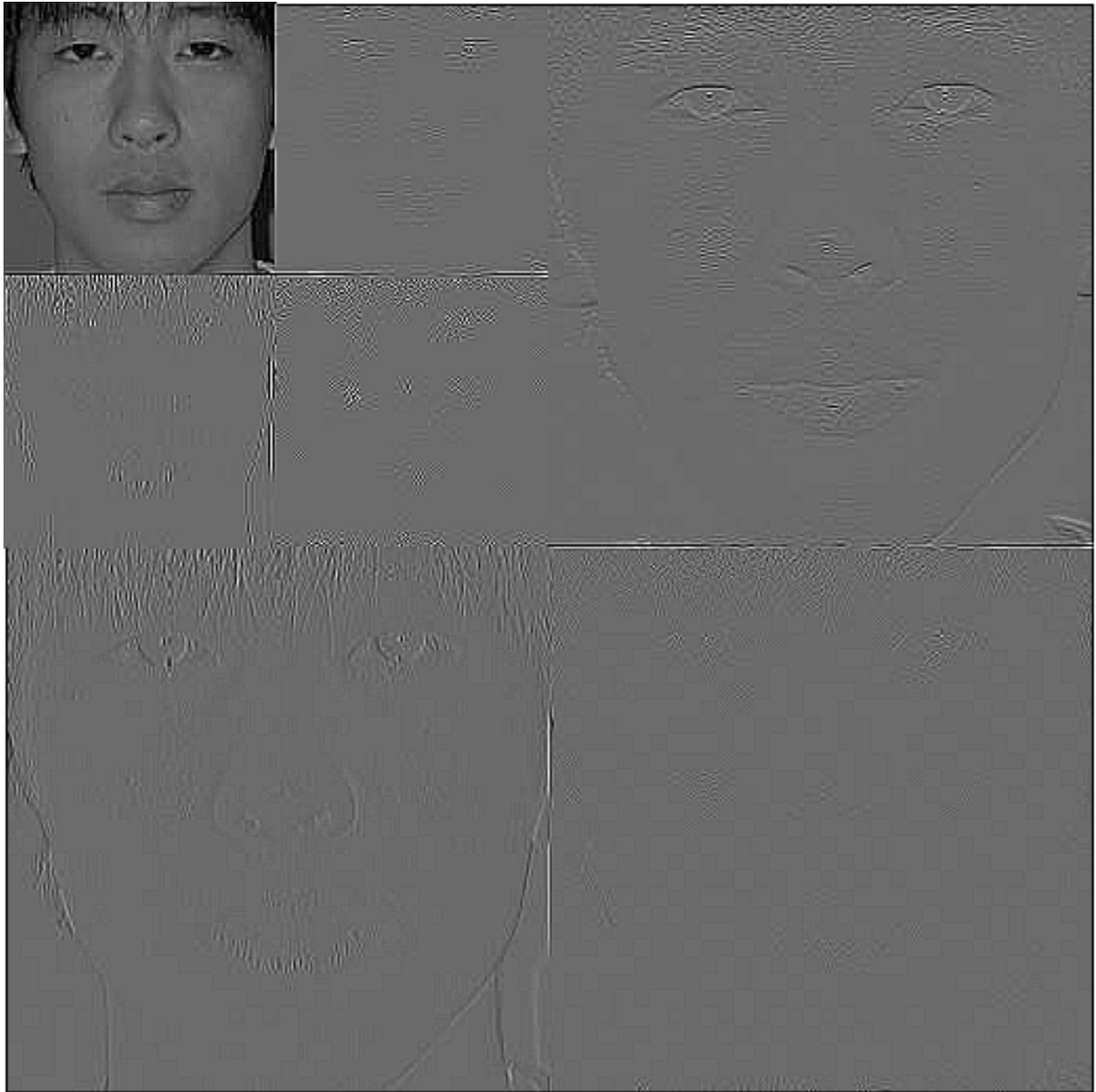
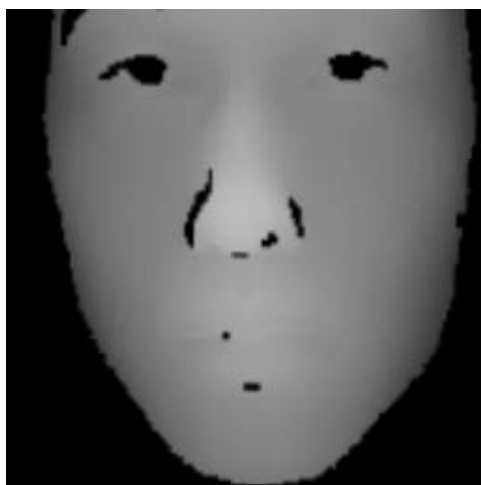


図 3.10: 顔濃淡画像を Wavelet 変換した Wavelet 係数の可視化の例



(a) 顔濃淡画像  $f_j(x)$



(b) 低域成分  $f_{j-1}(x)$



(c) 水平方向成分  
 $g_{j-1}^h(x)$



(d) 垂直方向成分  
 $g_{j-1}^v(x)$



(e) 対角方向成分  
 $g_{j-1}^d(x)$

図 3.11: 顔距離画像を Wavelet 変換した Wavelet 係数の可視化 ( $j = 1$ , Level 1)

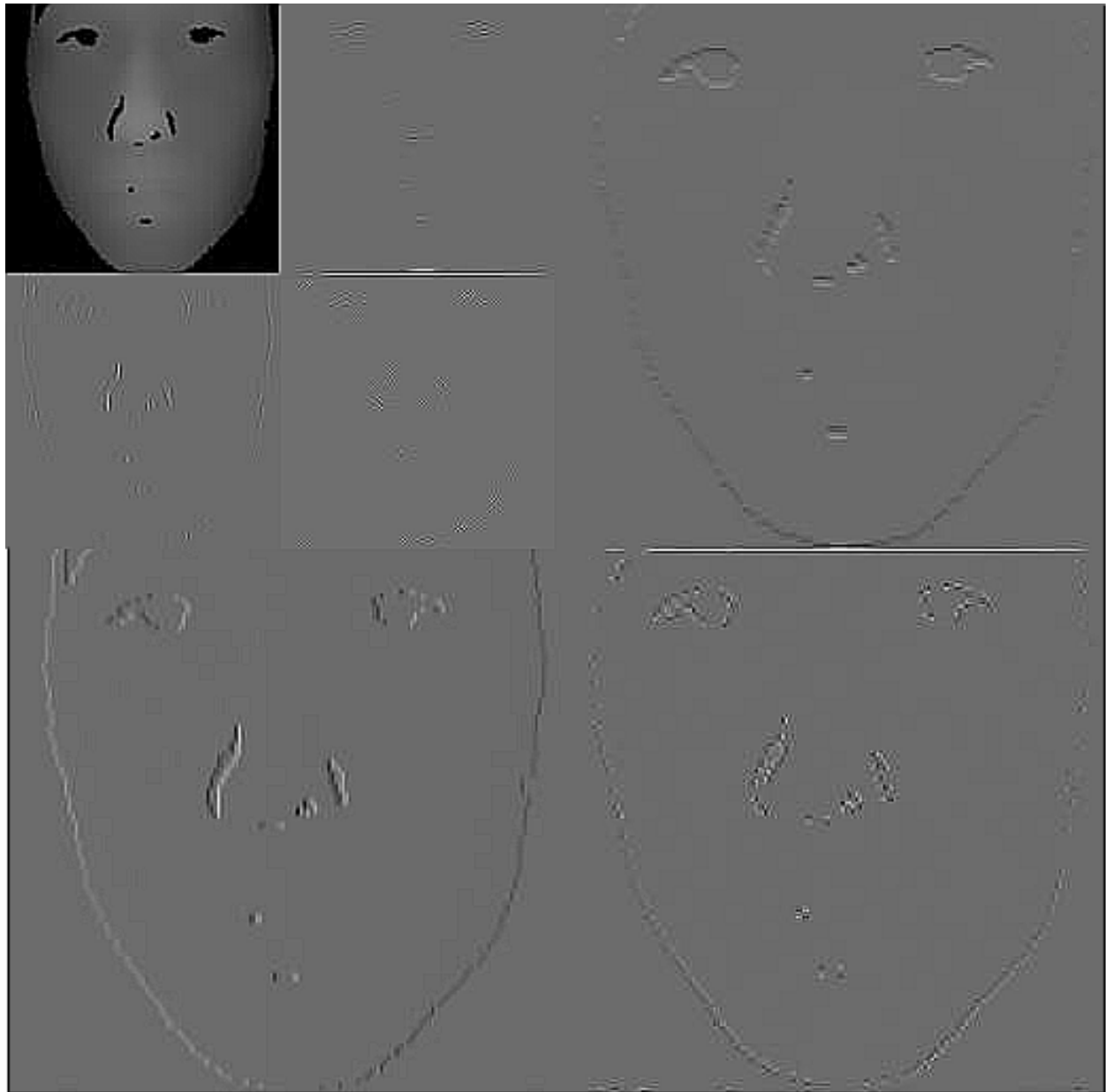


図 3.12: 顔距離画像を Wavelet 変換した Wavelet 係数の可視化の例

### 3.3 顔濃淡画像の多重解像度解析

顔の情報を捉える方法として、3.1 節で、照明による顔の濃淡の変換を顔の形状として捉える顔濃淡画像と顔の直接的な形状を捉える顔距離画像を示した。本節は、顔濃淡画像に着目し、顔濃淡画像を多重解像度解析し、顔画像の特徴解析を試みる。

顔の特徴は、顔の部位、輪郭、部位の位置関係などの位置情報を含んでいると考える。そこで、多重解像度解析の手法には、位置情報を保有した解析が行える離散コサイン変換と離散 Wavelet 変換を用いる。それぞれの手法において、個人を特定する情報（ここではこの情報を個人性とする）がどの解像度に含まれているかについて解析する。解析の評価は、多重解像度解析より得られた各解像度の変換係数を用いて個人識別を行う。また、離散コサイン変換と離散 Wavelet 変換の両手法は、異なる解析手法であるため、異なる個人性を捉えていると考えられる。よって、離散コサイン変換と離散 Wavelet 変換による個人識別特性の比較を行い、個人性の特性について検討した。

#### 3.3.1 離散コサイン変換による多重解像度解析と個人識別特性

画像を離散コサイン変換することにより求まる DCT 係数は、画像に基底画像がどの程度含まれるかの大きさを表す。顔濃淡画像を離散コサイン変換することは、顔濃淡画像にどのような基底画像を多く含むのかについて解析することになる。また、その基底のみの顔画像の情報が、個人性をどの程度含んでいるかについて、個人識別を行って解析する。顔濃淡画像を多重解像度解析し、各 DCT 係数に表れた個人性の特性について検討する。

以下に実験条件と識別方法、識別率の定義を示す。識別方法、識別率の算出方法は、以下の実験でも同様に用いる。

##### 実験条件

実験条件を表 3.3 に示す。識別に用いる画像は、各分割ブロックで求まる 64 個の基底より得られる  $32 \times 32$ [pixels] の 64 枚である。各基底より得られた 32 枚の画像を用いて識別実験を行った。識別方法は、入力画像と辞書画像を多重解像度解析し、各レベルの方向成分を得、入力画像と辞書画像の各レベルの方向成分との類似度を算出し、類似度が最大となる辞書のインデックスを識別された画像とする。識別の流れを図 3.13 に示す。識別の流れは、多重解像度解析に離散 Wavelet 変換を用いる場合、距離画像を用いる場合も同様な流れで識別を行う。



表 3.3: 顔濃淡画像の実験条件

画像サイズ	256 × 256[pixels]
画像の階調	256[bit/pixel]
識別に用いる画像の枚数	13人の無表情時の画像各8枚, 計104枚
辞書画像	各自の8枚から任意の1枚
入力画像	入力に用いなかった画像91枚
画像の分割ブロックサイズ	8[pixels]

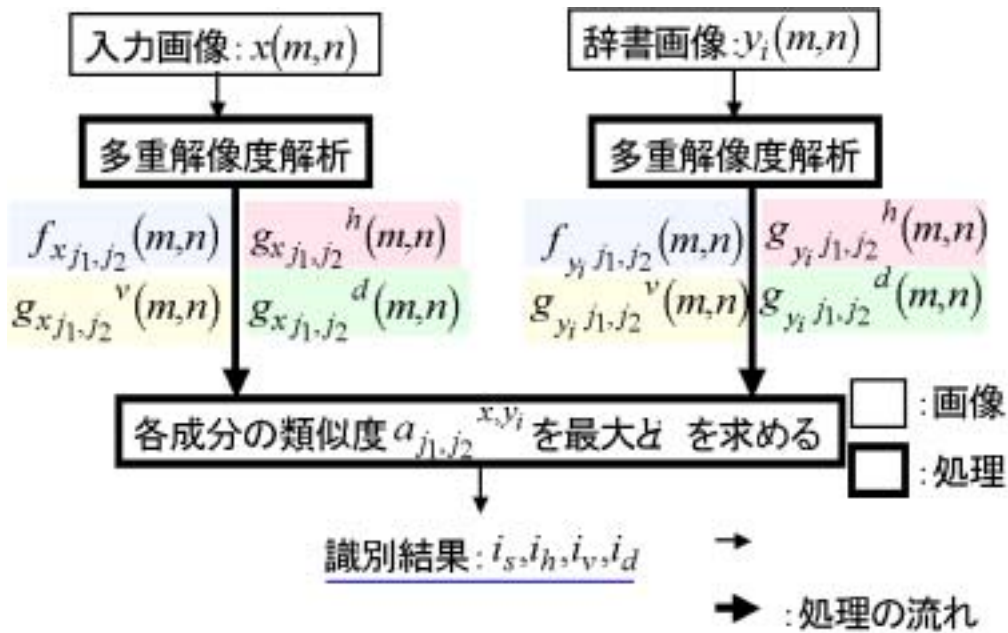


図 3.13: 識別の流れ図

## 類似度と識別率の定義

パターン間の類似性を求めるシンプルな方法として、単純類似度を用いる。単純類似度は、線形的な明るさの変化やコントラストの変化に強いという特長がある。単純類似度は以下の式で定義した [18]。

$$a_j^{vw} = \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_j^b(m, n) \cdot f_j^c(m, n)}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_j^b(m, n) \cdot \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_j^c(m, n)} \quad (3.6)$$

ここで、 $b, c$  は人物のインデックスを表す。 $m, n$  は画像の座標値を表し、 $f_j^b(m, n), f_j^c(m, n)$  は、レベル  $j$  における人物  $b$  および人物  $c$  の Wavelet 係数を表す。 $M, N$  は、それぞれ画像の縦、横の大きさを表す。識別率は、次式で算出した。

$$\text{識別率} [\%] = \frac{\text{識別に正解した数}}{\text{識別に用いた画像の枚数}} \times 100 \quad (3.7)$$

## 離散コサイン変換を用いた顔濃淡画像の多重解像度解析の識別実験結果および考察

離散コサイン変換を用いた顔濃淡画像の多重解像度解析の識別結果を表 3.4 に示す。基底 (5, 7), (6, 7) を用いた識別が、70.3% と最高の識別率となった。識別率の傾向としては、低周波成分による識別より、高周波成分を用いた識別の方が高い識別率が得られた。顔濃淡画像を離散コサイン変換した DCT 係数の可視化では、低域成分が多く情報を含まれていることが分かったが、個人性は高周波成分に含まれていると考えられる。図 3.14 に、基底 (5, 7), (6, 7) の DCT 係数を可視化した。図 3.14 から顔の輪郭や部位のエッジ部分を個人性として捉えていると考えられる。

また、水平方向と垂直方向の成分の比較では、垂直方向  $v$  の値が上がるにつれて識別率が高くなる傾向がある。一方、水平方向  $u$  の値が上がっても、識別率が上がることはほとんどない。よって、個人性は垂直方向成分に含まれると考えられる。垂直方向の識別率が高くなる傾向として、顔の構造において目・口等の水平方向の情報が多いことが考えられる。

垂直方向の成分の特性としては、基底  $v = 3$  で、識別率が低下する傾向があることが分かる。

表 3.4: 顔濃淡画像で DCT 係数を用いた識別率  $a_j^{bc}$

基底の位置 $v$	基底の位置 $u$							
	0	1	2	3	4	5	6	7
0	47.3	45.1	45.1	42.9	57.1	54.9	54.9	54.9
1	52.7	52.7	53.8	50.5	50.5	49.5	49.5	49.5
2	48.4	46.2	47.3	47.3	47.3	45.1	43.9	42.9
3	41.8	39.6	36.9	36.3	36.3	36.3	36.3	33.0
4	59.3	58.2	59.3	57.1	57.1	59.3	59.3	61.5
5	62.6	60.4	60.4	60.4	60.4	60.4	60.4	59.3
6	60.4	61.5	63.7	62.6	62.6	62.6	62.6	63.7
7	65.9	64.8	64.8	67.0	67.0	70.3	70.3	68.1



(a) 基底 (5,7)



(b) 基底 (6,7)

図 3.14: 識別率が高かった基底成分の可視化

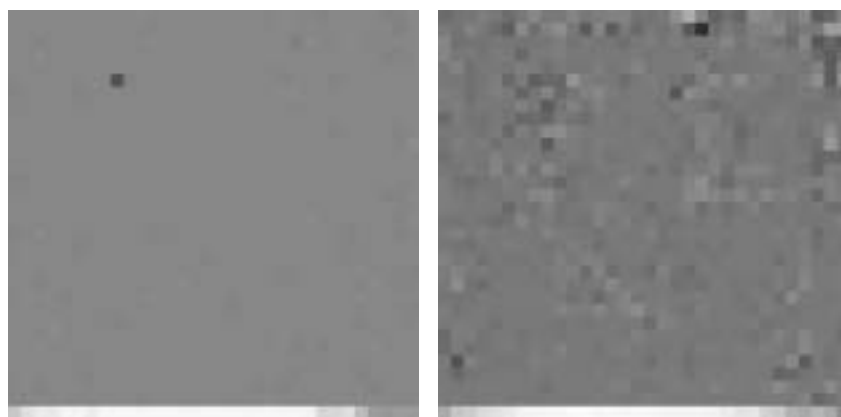
## 離散コサイン変換を用いた顔濃淡画像の多重解像度解析の識別実験において識別を誤った例

離散コサイン変換を用いた識別実験の特性を考察するために、識別に誤った場合に辞書画像と入力画像を比較する。離散コサイン変換を用いた顔濃淡画像の識別結果において、識別に誤った顔濃淡画像を図 3.15, 3.16, 3.17 に示す。図 3.15 は、離散コサイン変換する前の画像で、(a) は辞書画像であり、(b) は入力画像である。図 3.16, 3.17 は、識別率が最も高くなる入力画像と辞書画像の基底 (5, 7), (6, 7) の DCT 係数を可視化したものを示す。

基底 (5, 7), (6, 7) の DCT 係数を可視化した画像を比較すると、どのような特徴を捉えて識別を誤ったのか視覚による判断ができない。ただ、原画像と比較すると目・鼻・口等の部分に輝度の変化があることから目・鼻・口等のエッジ部分を捉えており、その部分において誤識別したと考えられる。



図 3.15: 誤って識別した時の辞書画像データと入力画像の例



(a) 辞書画像

(b) 入力画像

図 3.16: 誤って識別した時の辞書画像データと入力画像の DCT 係数 (基底 (5,7)) の可視化の例



(a) 辞書画像

(b) 入力画像

図 3.17: 誤って識別した時の辞書画像データと入力画像の DCT 係数 (基底 (6,7)) の可視化の例

### 3.3.2 離散 Wavelet 変換による多重解像度解析と個人識別特性

離散 Wavelet 変換は, Mother Wavelet を伸縮, 平行移動することによって, 解析する画像から Mother Wavelet と相似な様々なスケールの波形を画像の位置情報を保有したまま解析する手法である. 顔濃淡画像を離散 Wavelet 変換することは, 異なる解像度から大局的・局所の特徴を捉え, さらに Mother Wavelet により特徴情報を捉えている. また, 顔画像の離散 Wavelet 変換することにより, 顔濃淡画像の情報を低域, 水平, 垂直, 対角方向の成分に分離し, 各解像度ごとに解析が試みることができる. 顔濃淡画像では, 各成分がどのような個人性を抽出できるのかについて検討する. 個人性の抽出の評価は, 各解像度の 4 つの成分を用いて個人識別を行う. 実験条件を表 3.5 に示す.

#### 実験条件

表 3.5: 顔濃淡画像の実験条件

画像サイズ	256 × 256[pixels]
画像の階調	256[bit/pixel]
識別に用いる画像の枚数	13 人の無表情時の画像各 8 枚, 計 104 枚
辞書画像	各自の 8 枚から任意の 1 枚
入力画像	入力に用いなかった画像 91 枚

#### 離散 Wavelet 変換を用いた顔画像の多重解像度解析の識別実験結果および考察

表 3.6 に, 低域及び各方向成分の Wavelet 係数を用いた識別結果を示す. Level 0 は Wavelet 変換する前の画像輝度値そのものであり, Level 7 は最小解像度 ( $2 \times 2$ [pixels]) の Wavelet 係数を表す. 表 3.6 より得られる結果を以下にまとめる.

1. 低域 (s) 成分による特性では, 解像度が増加するにつれて識別率は増加し, Level 5 にて最大の識別率となり, その後わずかながら解像度の増加につれて識別率は低下する.
2. 水平方向成分 ( $w^{(h)}$ ) は低域成分と同様な特性を有し, Level 6 の解像度 ( $4 \times 4$ [pixels]) で識別率が最大となり, その後低下していく.
3. 垂直 ( $w^{(v)}$ ), 斜め ( $w^{(d)}$ ) 方向成分の識別への寄与はほとんどない.

一般的に, 画像データの解像度が高いほど識別に必要な情報が増えるため, 識別率も増加するはずであるが, Wavelet 変換を用いた実験では Level 5 あたりの解像度を超えると識別率が低下している. これは, ノイズや照明 (陰) が識別に影響を与え, 低解像度成分ではこれらの影響が小さくなったためと考えられる. 低い解像度成分でも高い識別精度が得

られていることから、低解像度成分が顔画像中の個人性についての特徴を充分有していることが分かる。

表 3.6: 顔濃淡画像で Wavelet 係数を用いた識別率  $a_j^{bc}$

Level j	識別率 (%)			
	s	$w^{(h)}$	$w^{(v)}$	$w^{(d)}$
0	72.5	-	-	-
1	71.4	31.9	6.6	9.9
2	71.4	31.9	4.4	3.3
3	71.4	37.4	8.8	6.6
4	73.6	41.8	5.5	7.7
5	79.1	64.8	6.6	8.8
6	73.6	69.2	1.1	13.2
7	57.1	62.6	1.1	1.1

各成分の識別特性について検討する。表 3.7, 3.8, 3.9 は低域成分と水平, 垂直, 対角の各成分の識別特性を比較し, 識別成功, 失敗の出現割合を計算したもので, " " は成功, " × " は失敗を表す。例えば, 表 3.7 の横方向について, s 成分が " " で  $w^{(h)}$  成分が " " の欄には, 低域成分, 水平方向成分共に識別に成功したことを表している。表 3.7, 3.8, 3.9, から識別には主に低域成分が寄与し, 低域成分以外による寄与は小さいことが分かる。従って, これらを組み合わせることで識別率が向上する可能性は小さいと考える。

表 3.7: 低域成分 s と水平方向成分  $w^{(h)}$  の識別特性 (濃淡情報)

識別結果		出現割合 [%]						
s	$w^{(h)}$	Level						
		1	2	3	4	5	6	7
		28.6	30.8	40.7	36.2	65.9	60.4	37.3
	×	38.5	40.7	40.7	37.4	12.0	14.3	11.0
×		5.5	2.2	3.3	5.5	2.2	7.7	20.8
×	×	27.4	26.3	15.3	20.9	19.9	17.6	30.9

表 3.8: 低域成分  $s$  と垂直方向成分  $w^{(v)}$  の識別特性 (濃淡情報)

識別結果		出現割合 [%]						
$s$	$w^{(v)}$	Level						
		1	2	3	4	5	6	7
		7.7	6.6	5.5	7.7	7.7	0.0	0.0
	×	60.4	64.8	65.9	64.8	69.2	75.8	47.3
×		0.0	1.1	2.2	1.1	0.0	1.1	0.0
×	×	31.9	27.5	26.4	26.4	23.1	23.1	52.7

表 3.9: 低域成分  $s$  と対角方向成分  $w^{(d)}$  の識別特性 (濃淡情報)

識別結果		出現割合 [%]						
$s$	$w^{(d)}$	Level						
		1	2	3	4	5	6	7
		8.8	5.5	13.2	14.3	12.1	2.2	0.0
	×	59.3	64.8	56.0	58.2	64.8	72.5	47.3
×		2.2	2.2	0.0	1.1	6.6	3.3	0.0
×	×	29.7	27.5	30.8	26.4	16.5	22.0	52.7

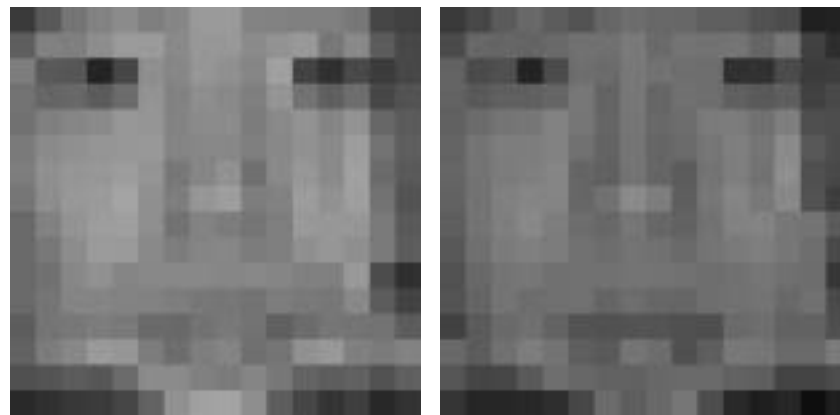


## 離散 Wavelet 変換を用いた顔濃淡画像の多重解像度解析の識別実験において識別を誤った例

離散 Wavelet 変換を用いた識別実験の特性を考察するために、識別に誤った場合に辞書画像と入力画像を比較する。識別に誤った場合の比較画像は、識別率が高い低域、水平方向の成分である。比較を行った画像を、図 3.18, 3.19, 3.20, 3.21 3.22 に示す。各図の (a) は辞書画像, (b) は入力画像である。

図 3.18, 3.19 は、それぞれ低域成分で識別率が高かった Level4, 5 のときの Wavelet 係数を可視化した画像である。図 3.18, 3.19 より、顔の肌色の違いによる輝度値の差があるが、顔の形状や部位の位置などが類似していることから誤識別したことが考えられる。

また、図 3.20, 3.21 3.22 は、それぞれ水平方向の Level4, 5, 6 のときの誤識別した画像である。識別に誤る原因として、Level を 6 以上上げるとパターンとしての分解能が低下するため、識別率が下がると考えられる。また、逆に Level を 3 程度に下げて、解像度を上げると、正規化しきれなかった大きさや位置のずれ、ノイズ等の影響が大きくなり、誤識別すると考えられる。



(a) 辞書画像

(b) 入力画像

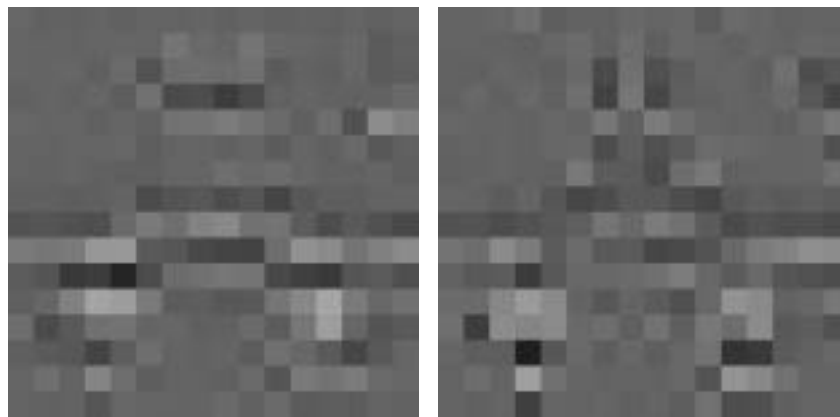
図 3.18: 誤って識別した時の辞書画像データと入力画像 (多重解像度解析した成分) の例 (低域成分 : Level 4)



(a) 辞書画像

(b) 入力画像

図 3.19: 誤って識別した時の辞書画像データと入力画像 (多重解像度解析した成分) の例 (低域成分 : Level5)



(a) 辞書画像

(b) 入力画像

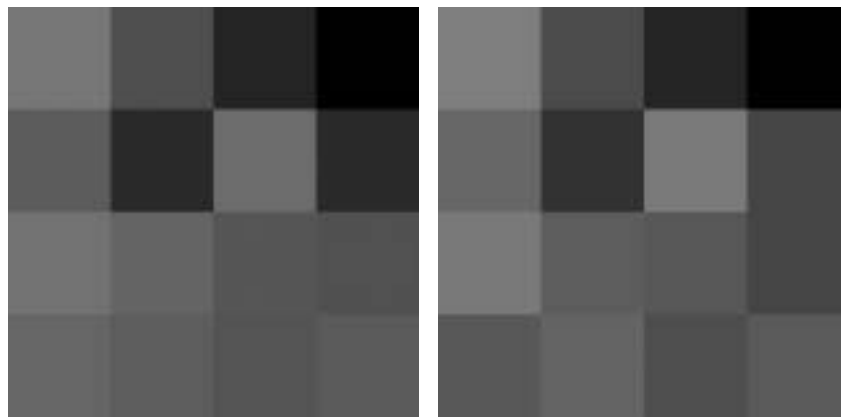
図 3.20: 誤って識別した時の辞書画像データと入力画像 (多重解像度解析した成分) の例 (水平成分 : Level4)



(a) 辞書画像

(b) 入力画像

図 3.21: 誤って識別した時の辞書画像データと入力画像 (多重解像度解析した成分) の例 (水平成分 : Level5)



(a) 辞書画像

(b) 入力画像

図 3.22: 誤って識別した時の辞書画像データと入力画像 (多重解像度解析した成分) の例 (水平成分 : Level6)

### 3.3.3 離散コサイン変換と離散 Wavelet 変換による多重解像度解析の個人識別特性の比較

2つの多重解像度解析手法の離散コサイン変換と離散 Wavelet 変換による個人性の抽出特性について比較した。2つの手法による識別結果の比較では、識別率が最も高くなるるときでは、約9%ほど離散 Wavelet 変換による識別率が高かった。よって、顔濃淡画像では、離散 Wavelet 変換の方が、より個人性と捉えていると考えられる。

識別特性の比較では、多重解像度解析に離散コサイン変換を用いた識別の傾向は、低域成分を用いた識別よりも高域の成分を用いた識別結果が高い傾向があった。一方の離散 Wavelet 変換では、低域成分が識別率が高くなる傾向があった。類似する特性としては、どちらの手法も目・口等の水平方向の成分をによる識別率が高いことから、顔濃淡画像では水平方向の情報に個人性を有していると考えられる。

## 3.4 顔距離画像の多重解像度解析

3.3節では、照明による顔の濃淡に変化による見掛けの形状による顔の特徴解析を試みた。本節では、顔の直接的な形状情報による特徴解析について解析を試みる。顔濃淡画像と同様に、顔の特徴は、顔の部位、輪郭、部位の位置関係などの位置情報を含んでいると考える。そこで、多重解像度解析の手法には、位置情報を保有した解析が行える離散コサイン変換と離散 Wavelet 変換を用いる。それぞれの手法において、個人性がどの解像度に含まれているかについて解析する。解析の評価は、多重解像度解析より得られた各解像度の変換係数を用いて個人識別を行う。また、離散コサイン変換と離散 Wavelet 変換の両手法は、異なる解析手法であるため、異なる個人性を捉えていると考えられる。よって、離散コサイン変換と離散 Wavelet 変換による個人識別特性の比較を行い、個人性の特性について検討した。

### 3.4.1 離散コサイン変換による多重解像度解析と個人識別特性

画像を離散コサイン変換することにより求まる DCT 係数は、画像に基底画像がどの程度含まれるかの大きさを表す。顔距離画像を離散コサイン変換することは、顔距離画像にどのような基底画像を多く含むのかについて解析することになる。顔距離画像においては、低域成分とは顔の形状に変化がない部分と考えられる。逆に、高域とは顔の形状に変化がある目・鼻・口などの部位である。また、その基底のみの顔画像の情報が、個人性をどの程度示しているかについて、個人識別を行って検討する。顔距離画像を多重解像度解析し、各 DCT 係数に表れた個人性の特性について検討する。実験条件を表 3.10 に示す。

## 実験条件

表 3.10: 顔濃淡画像の実験条件

画像サイズ	256 × 256[pixels]
画像の階調	256[bit/pixel]
識別に用いる画像の枚数	13 人の無表情時の画像各 8 枚, 計 104 枚
辞書画像	各自の 8 枚から任意の 1 枚
入力画像	入力に用いなかった画像 91 枚
画像の分割ブロックサイズ	8[pixels]
距離値の分解能	耳の付け根から鼻の頂点までの約 10[cm] の幅を 70[step] で分解

### 離散コサイン変換を用いた顔距離画像の多重解像度解析による識別実験結果および考察

離散コサイン変換を用いた顔距離画像の多重解像度解析の識別結果を表 3.11 に示す.

基底 (5, 7) を用いた識別が, 76.9%と最高の識別率となった. 識別率の傾向としては, 低周波成分による識別より, 高周波成分を用いた識別の方が高い識別率が得られた. 顔距離画像を離散コサイン変換した DCT 係数の可視化では, 低域成分が多く情報を含まれていることが分かったが, 個人性は高周波成分に含まれていると考えられる. 図 3.23 に, 基底 (5, 7) の DCT 係数を可視化した. 図 3.23 から顔の輪郭や部位のエッジ部分を個人性として捉えていることが分かる.

表 3.11: 顔距離画像で DCT 係数を用いた識別率  $a_j^{bc}$

基底の位置 $v$	基底の位置 $u$							
	0	1	2	3	4	5	6	7
0	69.2	67.0	67.0	60.4	70.3	68.1	68.1	68.1
1	68.1	63.7	63.7	63.7	59.3	54.9	54.9	51.6
2	52.7	48.4	46.1	44.0	44.0	37.4	42.9	41.8
3	41.8	41.8	41.8	40.0	40.0	36.3	36.3	37.4
4	56.0	56.0	56.0	56.0	65.9	67.0	67.0	65.9
5	65.9	63.7	63.7	63.7	63.7	64.8	64.8	64.8
6	64.8	67.0	67.0	59.2	59.2	71.4	71.4	71.4
7	71.4	72.5	71.4	75.8	75.8	76.9	75.8	75.8



図 3.23: 識別率が高かった基底成分 (5.7) の可視化

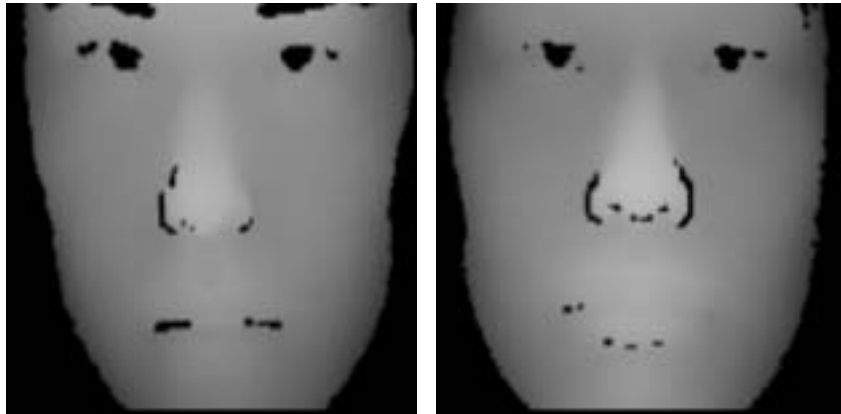
また, 水平方向と垂直方向の成分の比較では, 垂直方向  $v$  の値が上がるにつれて識別率が高くなる傾向がある. 一方, 水平方向  $u$  の値が上がっても, 識別率が上がることはほとんどない. よって, 個人性は垂直方向成分に含まれると考えられる. 垂直方向の識別率が高くなる傾向として, 顔の構造において目・口等の水平方向の情報が多いことが考えられる.

垂直方向の成分の特性としては, 基底  $v = 3$  で, 識別率が低下する傾向があることが分かる.

#### 離散コサイン変換を用いた顔距離画像の多重解像度解析の識別実験において識別を誤った例

離散コサイン変換を用いた識別実験の特性を考察するために, 識別に誤った場合に辞書画像と入力画像を比較する. 離散コサイン変換を用いた識別結果において, 識別に誤った顔画像をに図 3.24, 3.25 示す. 図 3.24 は, 離散コサイン変換する前の画像で, (a) は辞書画像であり, (b) は入力画像である. 図 3.25 は, 識別率が最も高くなる入力画像と辞書画像の基底 (5, 7) の DCT 係数を可視化したものを示す.

目・鼻・口等の部分に輝度の変化があることから目・鼻・口等のエッジ部分を捉えており, その部分において誤識別したと考えられる.



(a) 辞書画像

(b) 入力画像

図 3.24: 誤って識別した時の辞書画像データと入力画像の例



(a) 基底 (5,7)

(b) 基底 (5,7)

図 3.25: 誤って識別した時の辞書画像データと入力画像の DCT 係数 (5,7) を可視化した例

### 3.4.2 離散 Wavelet 変換による多重解像度解析と個人識別特性

離散 Wavelet 変換は, Mother Wavelet を伸縮, 平行移動することによって, 解析する画像から Mother Wavelet と相似な様々なスケールの波形を画像の位置情報を保有したまま解析する手法である. 顔距離画像を離散 Wavelet 変換することは, 異なる解像度から大局的・局所の特徴を捉え, さらに Mother Wavelet により特徴情報を捉えている.

顔距離画像を離散 Wavelet 変換することにより, 異なる解像度から大局的・局所の特徴を捉えることを試みる. また, 離散 Wavelet 変換では, 局所の特徴と大局の特徴は, 同じ階層構造により求まることから, 大局の特徴と局所の特徴の関係から顔距離画像の特徴の解析を試みる. また, 顔画像の離散 Wavelet 変換することにより, 顔距離画像の情報を低域, 水平, 垂直, 対角方向の成分に分離し, 各解像度ごとに解析が試みることができる. 顔距離画像では, どのような個人性を抽出できるのかについて検討する. 個人性の抽出の評価は, 各解像度の4つの成分を用いて個人識別を行う. 実験条件を表3.12に示す.

#### 実験条件

表 3.12: 顔濃淡画像の実験条件

画像サイズ	256 × 256[pixels]
画像の階調	256[bit/pixel]
識別に用いる画像の枚数	13人の無表情時の画像各8枚, 計104枚
辞書画像	各自の8枚から任意の1枚
入力画像	入力に用いなかった画像91枚
画像の分割ブロックサイズ	8[pixels]
距離値の分解能	耳の付け根から鼻の頂点までの約10[cm]の幅を70[step]で分解

#### Wavelet 変換用いた顔距離画像の多重解像度解析による識別実験結果および考察

表3.13に各 Wavelet 係数による識別結果を示す. 表より得られた結果を以下にまとめる.

1. 低域成分による識別特性では, Level 6( $2 \times 2$ [pixels])で識別率がほぼ飽和し, それ以上解像度を上げてても識別率はほとんど変わらない.
2. 水平方向成分による識別特性は Level 5( $8 \times 8$ [pixels])で識別率が最大となり, その後解像度の上昇につれて識別率は低下していく.
3. 垂直, 斜め方向成分は, 識別への寄与がほとんどない.



顔距離画像の低域成分とは顔表面の凹凸の概形を表すものであり、極めて低い解像度で識別率の特性が飽和していることから、個人性について顔の凹凸の特徴は余り高い解像度を有していなくてもよいことを示していると考えられる。

各成分の識別特性について検討する。表 3.14, 3.15, 3.16 は低域成分と水平, 垂直, 対角方向の各成分の識別特性を比較し、識別成功, 失敗の出現割合を計算したものである。表 3.14, 3.15, 3.16 から顔濃淡画像の場合と同様に識別は主に低域成分が寄与し、低域成分以外による寄与は小さいことが分かる。顔の凹凸の概形に個人性はかなり表わされており、凹凸の変化(水平, 垂直, 斜め方向成分)の寄与は少ないことが考察できる。

表 3.13: 顔距離画像で Wavelet 係数を用いた識別率  $a_j^{bc}$

Level j	識別率 (%)			
	s	$w^{(h)}$	$w^{(v)}$	$w^{(d)}$
0	73.6	-	-	-
1	72.5	41.8	7.7	12.1
2	72.5	42.9	5.5	11.0
3	74.7	44.0	6.6	9.9
4	73.6	58.2	12.1	7.7
5	72.5	68.1	7.7	4.4
6	73.6	65.9	0.0	3.3
7	51.6	44.0	0.0	0.0

表 3.14: 低域成分 s と水平方向成分  $w^{(h)}$  の識別特性 (距離情報)

識別結果		出現割合 [%]						
s	$w^{(h)}$	Level						
		1	2	3	4	5	6	7
		31.7	42.9	31.7	56.0	59.3	59.3	27.5
	×	37.4	31.7	33.0	20.9	12.0	15.4	26.4
×		5.5	4.4	4.4	6.6	7.7	6.6	19.9
×	×	25.4	21.0	30.9	16.5	21.0	18.7	26.2

表 3.15: 低域成分  $s$  と垂直方向成分  $w^{(v)}$  の識別特性 (距離情報)

識別結果		出現割合 [%]						
$s$	$w^{(v)}$	Level						
		1	2	3	4	5	6	7
		7.7	4.4	1.1	6.6	5.5	0.0	0.0
	×	58.2	71.4	74.7	69.2	69.2	75.8	53.8
×		0.0	0.0	5.5	2.2	1.1	0.0	0.0
×	×	34.1	24.2	18.7	22.0	24.2	24.2	46.2

表 3.16: 低域成分  $s$  と対角方向成分  $w^{(d)}$  の識別特性 (距離情報)

識別結果		出現割合 [%]						
$s$	$w^{(d)}$	Level						
		1	2	3	4	5	6	7
		4.4	7.7	4.4	5.5	2.2	1.1	0.0
	×	70.3	67.0	72.5	69.2	71.4	73.6	53.8
×		6.6	3.3	3.3	3.3	1.1	0.0	0.0
×	×	18.7	22.0	19.8	22.0	25.3	25.3	46.2

## 離散 Wavelet 変換を用いた顔距離画像の多重解像度解析の識別実験において識別を誤った例

離散 Wavelet 変換を用いた顔距離画像の多重解像度解析の識別結果において、識別に誤った例を示す。識別に誤った例は、識別率が高い傾向が得られた低域、水平方向の成分である。各図の (a) は辞書画像, (b) は入力画像である。図 3.26, 3.27 は、それぞれ低域成分で識別率が高かった Level5, 6 のときである。図 3.26, 3.27 より、Level が上がるに連れて顔の大局的な情報を捉えていることが分かる。この程度の解像度で 70 %の識別率が得られることから、顔距離画像において、ある程度の識別や分類が可能であると考えられる。

図 3.28, 3.29, 3.30, 3.31 は、それぞれ水平方向の Level3, 4, 5, 6 のときの識別が誤った画像である。識別に誤る原因として、識別に誤る原因として、Level を 6 以上上げるとパターンとしての分解能が低下するため、識別率が下がると考えられる。また、逆に Level を 3 程度に下げて、解像度を上げると、正規化しきれなかった大きさや位置のずれ、ノイズ等の影響が大きくなり、誤識別すると考えられる。

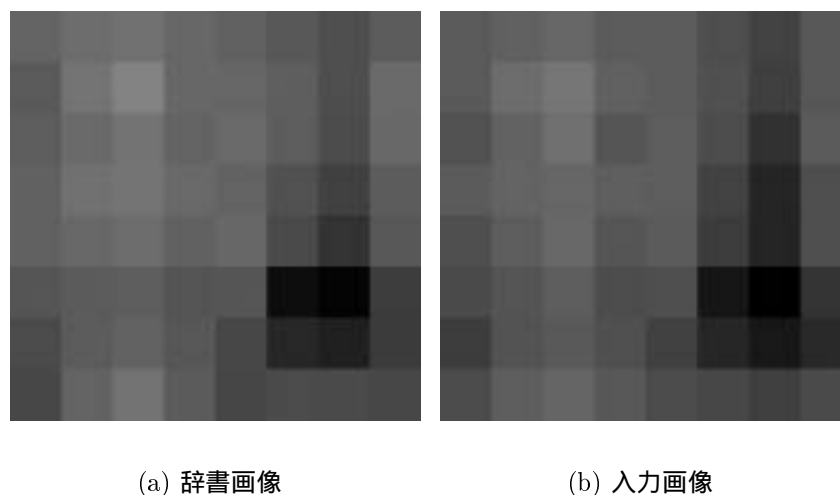
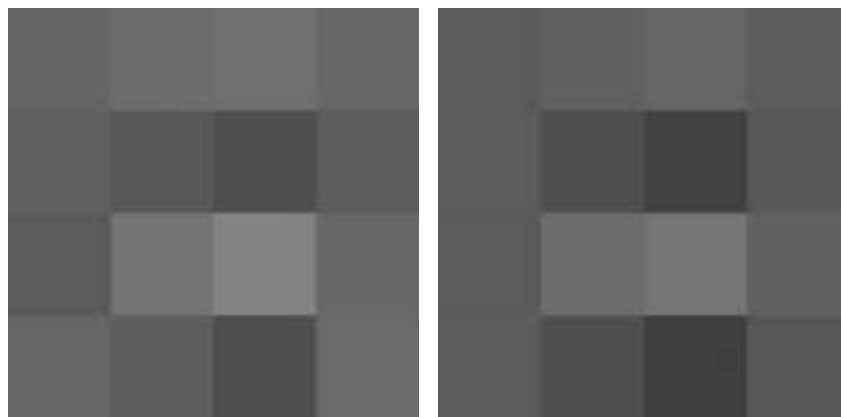


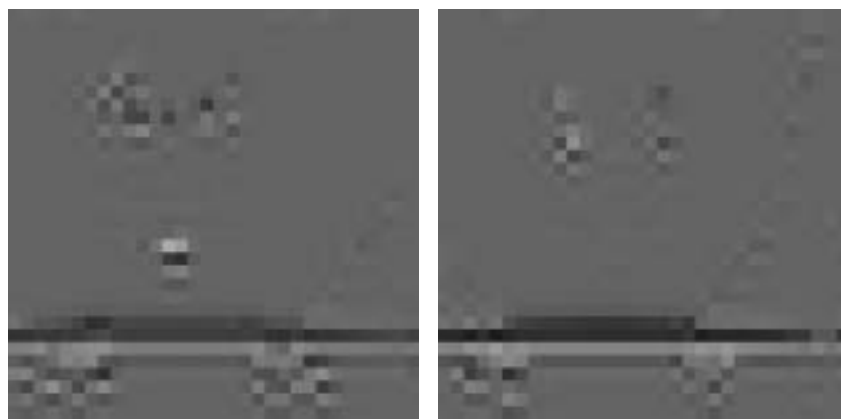
図 3.26: 誤って識別した時の辞書画像データと入力画像 (多重解像度解析した成分) の例 (低域成分 : Level5)



(a) 辞書画像

(b) 入力画像

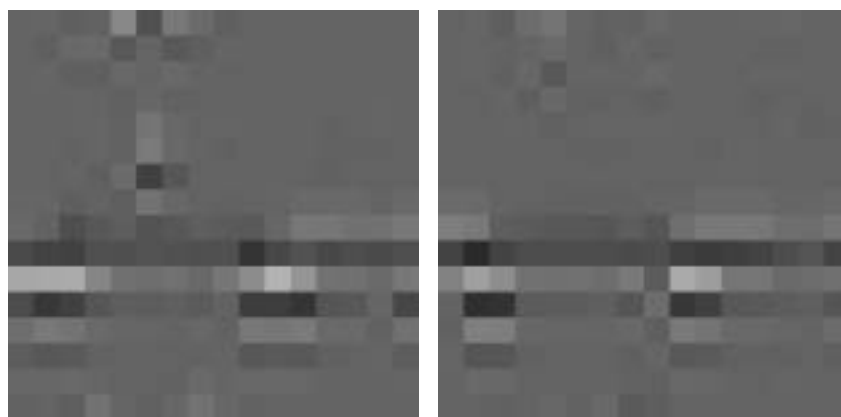
図 3.27: 誤って識別した時の辞書画像データと入力画像 (多重解像度解析した成分) の例 (低域成分 : Level6)



(a) 辞書画像

(b) 入力画像

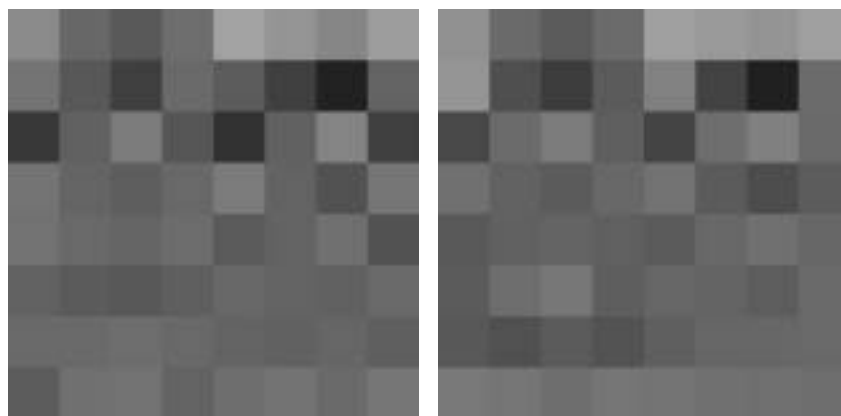
図 3.28: 誤って識別した時の辞書画像データと入力画像 (多重解像度解析した成分) の例 (水平成分 : Level3)



(a) 辞書画像

(b) 入力画像

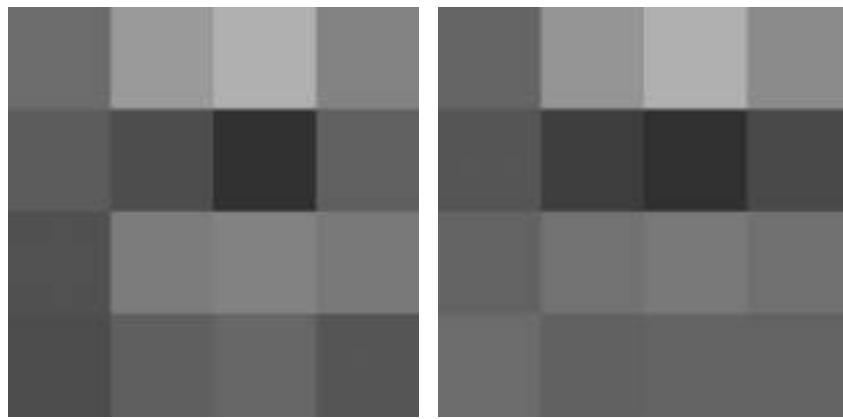
図 3.29: 誤って識別した時の辞書画像データと入力画像 (多重解像度解析した成分) の例 (水平成分 : Level4)



(a) 辞書画像

(b) 入力画像

図 3.30: 誤って識別した時の辞書画像データと入力画像 (多重解像度解析した成分) の例 (水平成分 : Level5)



(a) 辞書画像

(b) 入力画像

図 3.31: 誤って識別した時の辞書画像データと入力画像 (多重解像度解析した成分) の例 (水平成分 : Level6)

### 3.4.3 離散コサイン変換と離散 Wavelet 変換による多重解像度解析の個人識別特性の比較

2つの多重解像度解析手法の離散コサイン変換と離散 Wavelet 変換による個人性の抽出特性について比較した。2つの手法による識別結果の比較では、識別率が最も高くなる際には、約2%ほど離散コサイン変換による識別率が高かった。わずかな差であることから、顔距離画像では、どちらがより個人性と捉えているとは判断できない。

識別特性の比較では、多重解像度解析に離散コサイン変換を用いた識別の傾向は、低域成分を用いた識別よりも高域の成分を用いた識別結果が高い傾向があった。一方の離散 Wavelet 変換では、低域成分が識別率が高くなる傾向があった。類似する特性としては、どちらの手法も目・口等の水平方向の成分をによる識別率が高いことから、顔距離画像では水平方向の情報に個人性を有していると考えられる。

### 3.5 顔濃淡画像と顔距離画像の多重解像度解析結果による個人識別特性の比較

3.3, 3.4 節では、濃淡情報および距離情報を個別に用いた場合の識別結果を示した。特に、多重解像度解析において階層的な関係を保つ解析を行う離散 Wavelet 変換について、顔濃淡画像と顔距離画像について比較する。低域成分による濃淡情報と距離情報の識別特性について、表 3.17 に示す。表中で “ ” は正しく識別したもの、“×” は誤って識別したものを表している。

低域成分、水平方向成分、垂直方向成分、対角方向成分による識別特性をそれぞれ表 3.17, 3.18, 3.19, 3.20 に示す。これら 4 つの表から、各 Wavelet 係数の成分において、顔濃淡画像、顔距離画像のどちらかのみが識別に成功する場合が 40 % ほどあることが分かる。また、どちらも識別に成功した識別率が低いことから、人の顔の個人性について、顔濃淡画像と顔距離画像とは異なる情報を有しており、それぞれを統合することで識別率の向上を期待している。

表 3.17: 濃淡情報と距離情報の識別特性 (低域成分  $s$ )

識別結果		出現割合 [%]							
濃淡情報	距離情報	Level							
		0	1	2	3	4	5	6	7
		21.9	21.9	21.9	20.8	23.0	15.3	16.5	11.0
	×	30.0	28.6	30.0	30.8	30.8	35.2	28.6	23.0
×		20.9	18.7	18.7	19.8	18.7	22.0	18.7	15.4
×	×	27.2	30.8	29.4	28.6	27.5	27.5	36.2	50.6

表 3.18: 濃淡情報と距離情報の識別特性 (水平方向成分  $w^{(h)}$ )

識別結果		出現割合 [%]						
濃淡情報	距離情報	Level						
		1	2	3	4	5	6	7
		4.4	2.1	4.4	16.5	9.9	21.9	15.3
	×	6.6	19.8	19.8	18.7	22.0	24.2	30.8
×		24.2	17.6	18.7	22.0	31.9	29.7	16.5
×	×	64.8	60.5	57.1	42.8	36.2	24.2	37.4



表 3.19: 濃淡情報と距離情報の識別特性 (垂直方向成分  $w^{(v)}$ )

識別結果		出現割合 [%]						
濃淡 情報	距離 情報	Level						
		1	2	3	4	5	6	7
		0.0	0.0	1.0	0.0	0.0	0.0	0.0
	×	6.6	9.9	11.0	5.5	2.2	0.0	0.0
×		7.7	6.6	5.5	1.0	4.4	0.0	0.0
×	×	85.7	83.5	82.5	93.5	93.4	100.0	100.0

表 3.20: 濃淡情報と距離情報の識別特性 (対角方向成分  $w^{(d)}$ )

識別結果		出現割合 [%]						
濃淡 情報	距離 情報	Level						
		1	2	3	4	5	6	7
		1.0	1.0	0.0	0.0	1.0	0.0	0.0
	×	7.7	9.9	7.7	4.4	1.0	11.0	1.0
×		7.7	9.9	11.0	5.5	16.5	3.3	0.0
×	×	83.6	79.2	81.3	90.1	81.5	85.7	99.0

# 第4章 特徴量を用いた顔画像の多重解像度解析

## 4.1 顔画像の特徴量による解析

テクスチャ解析において、テクスチャ画像を Wavelet 変換し、得られた Wavelet 係数より特徴量を定義し画像を識別する研究 [14] がある。テクスチャ解析では、規則的なパターンを特徴として捉えたものである。そこで、顔画像をテクスチャとして捉え、テクスチャ解析に用いた特徴量を顔画像に適用し、顔画像の特徴情報を捉えることを試みる。

特徴量による解析は、多次元の画像情報を特徴量により次元をしていると見なすこともできる。削減した特徴量により、顔画像の特徴情報を捉えることができるか試みる。

## 4.2 特徴量の定義

顔画像の特徴解析に用いる 3 つの特徴量  $F_n^{l,p}$  を以下に再定義する。 $n$  は特徴量の番号、 $l$  はレベル、 $p$  は水平、垂直、対角方向成分を示す数である。図 4.1 に  $p$  の対応表を示す。3 つの特徴量より図 4.2 の特徴空間を定義し、さらに特徴ベクトル  $v_n^l$  を定義する。

0 (Low resolution component)	1 (Horizontal component)
2 (Vertical component)	3 (Diagonal component)

図 4.1: 各方向成分と  $p$  の対応図

特徴量 1 :  $F_1^{l,p}$

特徴量 1 は、各差分空間と全差分空間のエネルギーの比率であり、各差分空間でのエネルギー

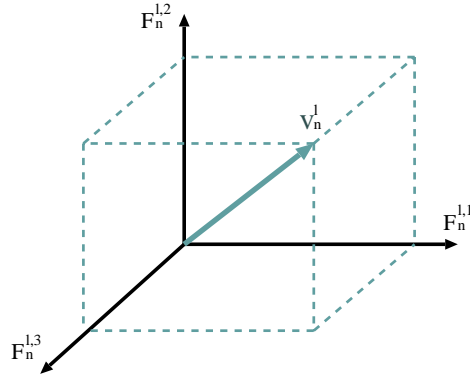


図 4.2: 特徴空間と特徴ベクトル

ギーを全エネルギーで正規化したものである。

また,  $W_{ij}^{l1}$ ,  $W_{ij}^{l2}$ ,  $W_{ij}^{l3}$  空間にはそれぞれ原画像の縦方向, 横方向, 縦と横が交差する部分の輪郭情報が含まれている. この性質により, 各レベル毎の顔画像の持つ方向性を, wavelet 変換で得られる差分空間から抽出することができる.

$$F_1^{l,p} = \frac{\sum_{i=0}^{m/2^l-1} \sum_{j=0}^{m/2^l-1} (W_{ij}^{lp})^2}{\sum_{l=1}^8 \sum_{p=1}^3 \sum_{i=0}^{m/2^l-1} \sum_{j=0}^{m/2^l-1} (W_{ij}^{lp})^2} \quad (4.1)$$

$m$  は画像サイズ,  $l$  はレベル,  $p$  は水平, 垂直, 対角方向成分を示す数である. 具体的には, 図 4.1 に示す.

### 特徴量 2 : $F_2^{l,p}$

特徴量 2 は, 特徴量 1 のように差分係数を 2 乗してエネルギーを見るのではなく, 差分係数の絶対値を用いる.

$$F_2^{l,p} = \frac{\sum_{i=0}^{m/2^l-1} \sum_{j=0}^{m/2^l-1} |W_{ij}^{lp}|}{\sum_{l=1}^8 \sum_{p=1}^3 \sum_{i=0}^{m/2^l-1} \sum_{j=0}^{m/2^l-1} |W_{ij}^{lp}|} \quad (4.2)$$

### 特徴量 3 : $F_3^{l,p}$

特徴量 3 は, 差分空間  $W^{lp}$  の係数の絶対値を, 各レベルの差分空間に画素数で割ることにより, 各差分空間毎に波の強さの平均を表している.

$$F_3^{l,p} = \sum_{i=0}^{m/2^l-1} \sum_{j=0}^{m/2^l-1} \frac{(|W_{ij}^{lp}|)}{(m/2^l \times m/2^l)} \quad (4.3)$$

以上の特徴量を用いて, 図 4.2 のような座標系を定義する. また, ベクトル  $v$  を表わす式は以下となる.

$$v_n^l = (F_n^{l,1}, F_n^{l,2}, F_n^{l,3}) \quad (4.4)$$

$n$  は, 特徴量の番号を表わす.

### 4.3 特徴量を用いた個人識別実験

顔画像をテクスチャとして捉え, テクスチャ解析に用いた特徴量が顔画像の解析に有効であるかについて検討する. また, 顔画像の多次元の情報を特徴量により 1 次元を縮退し, 次元縮退することの有効性についても検討した.

#### 4.3.1 実験条件と識別方法

##### 実験条件

表 4.1 に実験条件を示す.

表 4.1: 顔濃淡画像の実験条件

画像サイズ	256 × 256[pixels]
画像の階調	256[bit/pixel]
識別に用いる画像の枚数	16 人の無表情時の画像各 3 枚, 計 48 枚
辞書画像	各自の 8 枚から任意の 1 枚
入力画像	入力に用いなかった画像 32 枚
画像の分割ブロックサイズ	8[pixels]

##### 複合類似度 (complex similarity)

パターン間の類似性を求めるシンプルな方法として, 単純類似度を用いる. 単純類似度は, 線形的な明るさの変化やコントラストの変化に強いという特長がある.

単純類似度が 1 カテゴリ当たり 1 個だけの標準パターンを割り当てたの対し,  $K$  個の標準パターンを用意する. したがって, 複合類似度を  $s_n^l$  と表すことにすれば, その定義式は  $K$  個の各単純類似度の平均をとる.  $z_n^l, v_n^l$  は, 特徴ベクトルを表し,  $z_n^l$  は辞書ベクトル,  $v_n^l$

は入力ベクトルである.

$$s_n^l = \frac{1}{K} \sum_{k=1}^K \frac{z_n^l \cdot v_n^l}{\|z_n^l\| \cdot \|v_n^l\|} \quad (0 \leq s_n^l \leq 1) \quad (4.5)$$

識別率は, 次式で算出した.

$$\text{識別率} [\%] = \frac{\text{識別に正解した数}}{\text{識別に用いた画像の枚数}} \times 100 \quad (4.6)$$

### 4.3.2 各特徴量を単独に用いた識別実験

#### 特徴量を用いた識別実験結果と考察

特徴量を用いた識別結果を表 4.2, 表 4.3, 表 4.4 に示す. この3つ表は, 3つの辞書画像から2つの画像を選択するすべて組合せの結果である. 識別結果は, 62%が最もよい識別率であった. 各レベルにおける識別特性を比較すると, 各特徴量において Level 1, 8 が識別率が高くなる傾向がある. Level 1 は, 他の Level に比べ解像度が高いことから顔の情報が多く, 特徴量はその情報から個人性を捉えたことが分かる. しかし, Level 8 は 1[pixels]の最も小さい解像度である. 多重解像度解析の処理において, 顔の特徴ある情報が残ったと考えられる. 特徴量の比較では, 各 Level において識別特性は類似する傾向がある. 各特徴量は, 各 Level のにおいて類似した特徴情報を捉えていると考えられる.

識別結果から, 多次元の情報を特徴量によって1次元に縮退させて顔画像の特徴情報を捉えることを試みたが, 次元縮退によって個人識別に必要な分解能が得られなくなったと考えている. より高次元の特徴量を定義し, 分解能を上げる検討が必要と考えている.

表 4.2: 各特徴量を単独で用いた識別実験 1

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	44	56	56
レベル 2	37	44	44
レベル 3	44	50	50
レベル 4	37	31	31
レベル 5	50	50	50
レベル 6	37	37	37
レベル 7	37	44	44
レベル 8	56	56	56

表 4.3: 各特徴量を単独で用いた識別実験 2

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	50	37	37
レベル 2	31	37	37
レベル 3	44	50	50
レベル 4	31	44	44
レベル 5	19	31	31
レベル 6	25	19	19
レベル 7	50	50	50
レベル 8	37	44	44

表 4.4: 各特徴量を単独で用いた識別実験 3

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	44	62	62
レベル 2	37	37	37
レベル 3	37	37	37
レベル 4	12	37	37
レベル 5	25	37	37
レベル 6	31	25	25
レベル 7	44	37	37
レベル 8	50	56	56

### 4.3.3 各特徴量を組み合わせた個人識別実験結果

3つの特徴量は、多重解像度解析によって得られる各解像度の異なる情報を用いて定義されている。そのため、この3つの特徴量を組み合わせることで、特徴量の次元を上げる顔の特徴情報を捉えることを試みる。3つの特徴量の8つのレベルにおいて、その中から2つを組み合わせ、識別実験を行った。識別方法は、複合類似度法を用い、顔画像は無表情とした。

#### 特徴量の組合せによる識別実験結果と考察

識別結果は、最も高い識別結果は、62%となった。識別率としては、これまでの結果と変わらなかった。さらに、3つの組合せについても行ったが、識別率としては2つの場合と変わらなかった。

特徴量の組合せによる実験では、識別率が変わらなかったことから、各 Level の各特徴量は類似した特徴情報を捉えていると考えられる。よって、特徴量の組合せにより、顔画像の特徴情報を多く捉えることができないと考える。

### 4.3.4 笑い顔画像の識別実験

4.3.2, 4.3.3 節の識別実験では、無表情のみを用いた識別実験を行ってきた。しかし、人の顔には表情が含まれること。そこで、表情の変化が最も大きい笑い顔画像についての識別実験を行った。また、顔の表情に依存しない不変の特徴が、多重解像度解析により、捉えられるかについて試みる。

#### 笑い顔画像の識別実験結果と考察

結果を表 4.5, 4.6, 4.7 に示す。表 4.5, 4.6, 4.7 から、無表情画像の識別より 60% を越える識別率が多いことが分かる。これは笑い画像の方が、輪郭の成分が多くあるため、Wavelet 変換の性質である輪郭などの特徴がよくとらえられていることが分かる。

離散 Wavelet 変換による多重解像度解析では、エッジ部分を多く表出する顔画像の解析に有効であると考えられる。

識別特性は、無表情のときと同様の結果が得られた。よって、顔画像において、個人の不変な特徴を捉えていると考えられる。

表 4.5: 笑い濃淡画像を用いた特徴量による識別実験結果 1

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	31	56	56
レベル 2	44	44	44
レベル 3	19	31	31
レベル 4	12	31	31
レベル 5	12	25	25
レベル 6	31	19	19
レベル 7	37	50	51
レベル 8	69	62	62

表 4.6: 笑い濃淡画像を用いた特徴量による識別実験結果 2

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	50	50	50
レベル 2	37	31	31
レベル 3	31	25	25
レベル 4	19	25	25
レベル 5	12	25	25
レベル 6	19	12	12
レベル 7	44	37	37
レベル 8	56	62	62



表 4.7: 笑い濃淡画像を用いた特徴量による識別実験結果 3

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	37	50	50
レベル 2	44	31	31
レベル 3	19	37	37
レベル 4	19	19	19
レベル 5	25	25	25
レベル 6	25	19	19
レベル 7	50	44	44
レベル 8	62	50	50

### 4.3.5 辞書に無表情と笑い顔画像を組み合わせた場合の識別実験

4.4.3 節では, 笑いの表情を用いた顔画像の特徴解析では, 無表情の顔画像を用いた場合より特徴情報を捉えていた. 表情は, 同一の人間から得られた情報であることから, もし離散 Wavelet 変換による多重解像度解析が, 無表情と笑いにおいて異なる情報を与えるならば, 無表情時に加えて笑い時の情報も用いることにより識別精度を改善できるかもしれない.

そこで, 辞書画像に無表情と笑いの情報を用いた識別実験を試みる. 識別に用いる画像は, 無表情の顔画像とする.

#### 辞書に無表情と笑い顔画像を組み合わせた場合の識別実験結果と考察

識別結果を表 4.8 - 4.16 に示す. 識別結果より, 識別特性は無表情の場合と同様に, Level 1, 8 において識別率が高くなる傾向がある. また, 無表情のみを用いた実験との識別率の比較では, 識別率が変わらない傾向がある. 辞書画像に無表情と笑いの情報を用いても識別率に大きな差がないことから, 離散 Wavelet 変換による多重解像度解析では, 表情に依存しない顔画像の特徴を捉えていると考えられる.

表 4.8: 笑いと無表情の顔濃淡画像を用いた特徴量による識別実験結果 1

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	47	42	42
レベル 2	33	34	34
レベル 3	20	27	27
レベル 4	14	22	22
レベル 5	22	30	30
レベル 6	25	23	23
レベル 7	34	36	36
レベル 8	48	53	53

表 4.9: 笑いと無表情の顔濃淡画像を用いた特徴量による識別実験結果 2

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	34	42	42
レベル 2	34	28	28
レベル 3	20	28	28
レベル 4	8	30	30
レベル 5	17	37	37
レベル 6	20	20	20
レベル 7	34	37	37
レベル 8	44	52	52

表 4.10: 笑いと無表情の顔濃淡画像を用いた特徴量による識別実験結果 3

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	44	50	50
レベル 2	30	33	33
レベル 3	22	28	28
レベル 4	17	19	19
レベル 5	22	22	22
レベル 6	25	23	23
レベル 7	27	30	30
レベル 8	45	41	41

表 4.11: 笑いと無表情の顔濃淡画像を用いた特徴量による識別実験結果 4

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	33	41	41
レベル 2	28	25	25
レベル 3	16	27	27
レベル 4	17	27	27
レベル 5	20	23	23
レベル 6	19	14	14
レベル 7	41	18	18
レベル 8	34	42	42

表 4.12: 笑いと無表情の顔濃淡画像を用いた特徴量による識別実験結果 5

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	47	47	47
レベル 2	37	30	30
レベル 3	12	23	23
レベル 4	17	22	22
レベル 5	23	25	25
レベル 6	27	20	20
レベル 7	52	39	39
レベル 8	48	56	56

表 4.13: 笑いと無表情の顔濃淡画像を用いた特徴量による識別実験結果 6

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	39	52	52
レベル 2	25	33	33
レベル 3	23	33	33
レベル 4	22	20	20
レベル 5	28	30	30
レベル 6	22	27	27
レベル 7	36	28	28
レベル 8	42	50	50

表 4.14: 笑いと無表情の顔濃淡画像を用いた特徴量による識別実験結果 7

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	45	44	44
レベル 2	34	27	27
レベル 3	23	33	33
レベル 4	25	28	28
レベル 5	22	34	34
レベル 6	20	25	25
レベル 7	53	36	36
レベル 8	48	48	48

表 4.15: 笑いと無表情の顔濃淡画像を用いた特徴量による識別実験結果 8

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	33	45	45
レベル 2	28	28	28
レベル 3	19	27	27
レベル 4	17	28	28
レベル 5	27	30	30
レベル 6	30	22	22
レベル 7	45	44	44
レベル 8	48	53	53

表 4.16: 笑いと無表情の顔濃淡画像を用いた特徴量による識別実験結果 9

	識別率 (%)		
	特徴量 1	特徴量 2	特徴量 3
レベル 1	36	37	37
レベル 2	31	31	31
レベル 3	22	34	34
レベル 4	22	22	22
レベル 5	25	30	30
レベル 6	27	28	28
レベル 7	41	33	33
レベル 8	47	47	47

# 第5章 結言

## 5.1 まとめ

従来, 顔画像解析において, appearance ベース, model ベースの2つの手法により解析してきた. この2つの解析手法は, 大局的情報および局所的情報を捉える解析である. 画像解析において, 大局的, 局所的情報を異なる解像度から抽出, 解析する手法に, 多重解像度解析がある. 本研究では, 顔画像の特徴解析に多重解像度解析を用い, 多重解像度による大局的, 局所的情報を捉え, 各解像に顔のどのような特徴情報を含んでいるかを明らかにすることを試みた. 解析に用いる画像は, 照明による顔の濃淡変化を顔の形状特徴とし捉えている顔濃淡画像と顔の直接的な形状特徴を捉えている顔距離画像を用いた. その結果を以下に要約する.

### 顔濃淡画像の多重解像度解析

- 顔濃淡画像を離散コサイン変換, 離散 Wavelet 変換により多重解像度解析し, 顔画像の特徴情報を解析した. 顔濃淡画像の多重解像度解析より得られた特徴情報が, 個人を特定する情報を含んでいるかについて, 個人識別実験を行い評価した. 多重解像度解析に離散コサイン変換より, 離散 Wavelet 変換を用いた方が, 識別率が約9%と高い結果が得られた. 個人を識別する情報は, 離散 Wavelet 変換を用いて多重解像度解析した方が, 特徴情報を多く含んでいる.
  - 離散コサイン変換による多重解像度解析では, 各基底ごとに識別実験を試みた. 識別結果から, 高域成分を用いた識別率が高いことから, 顔濃淡画像は低域成分の基底を多く含むが, 個人性は高域成分に表れている.
  - 離散 Wavelet 変換による多重解像度解析では, 低域成分と水平成分による識別率がが高いことから, 低域成分と水平成分に特徴情報を含んでいる.

### 顔距離画像の多重解像度解析

- 顔距離画像を用いて, 離散コサイン変換, 離散 Wavelet 変換による多重解像度解析を行い, 顔画像の特徴情報を解析した. 顔濃淡画像と同様に, 顔距離画像の多重解像度解析より得られた特徴情報が, 個人を特定する情報を含んでいるかについて, 個人識別実験を行い評価した. 多重解像度解析に離散コサイン変換と離散 Wavelet 変換を用いた場合の識別率を比較した結果, 同じ識別率が得られた. 個人を識別する情報

は、離散コサイン変換と離散 Wavelet 変換の多重解像度解析では、同等の特徴情報が多く含んでいる。

- 離散コサイン変換による多重解像度解析では、各基底ごとに識別実験を試みた。識別結果から、高域成分を用いた識別率が高いことから、顔距離画像は低域成分の基底を多く含むが、個人性は高域成分に表れている。
- 離散 Wavelet 変換による多重解像度解析では、低域成分と水平成分による識別率がが高いことから、低域成分と水平成分に特徴情報を含んでいる。

#### 顔濃淡画像と顔距離画像の多重解像度解析結果の比較

- 顔濃淡・距離画像を離散 Wavelet 変換で多重解像度解析し、低域成分、水平方向成分、垂直方向成分、対角方向成分による識別特性をそれぞれ比較した。各 Wavelet 係数の成分において、顔濃淡画像で識別に成功するものと顔距離画像で成功するものが異なる場合が 40 %ほどあることが分かった。即ち、人の顔の個人性について、顔濃淡画像と顔距離画像とは異なる情報を有しており、それぞれを組み合わせることで識別率の向上が図れるのではないかと示唆される。

#### 特徴量を用いた顔濃淡画像の多重解像度解析

- 顔濃淡画像をテクスチャと捉え、テクスチャ解析に用いた特徴量を顔画像解析に適用し、顔濃淡画像の特徴解析を試みた。顔画像の多次元の情報を特徴量により 1 次元に縮退することは、個人性を表す情報が欠落する傾向がある。

本研究をまとめ、電子情報通信学会 通信方式 (CS) 研究会において研究成果 [20] を発表した。

## 5.2 今後の課題

- 多重解像度解析を用いた顔画像の表情解析  
本解析では、顔濃淡・距離画像を多重解像度解析した結果が、個人を特定する情報がどの解像度のどの様に含まれているかについて解析した。しかし、顔の情報には、表情の情報もあることから、顔の表情に関する解析を検討する。
- 動画の解析による識別 (時間軸方向の解析)  
本研究では、静止画のみを対象に解析を試みたが、多重解像度解析に用いた 2 つの手法は、実世界として動画技術に適用されているまたは期待されていることから、動画における個人識別、表情解析を検討する。
- JPEG2000 システムへの適応  
近年、新しい圧縮技術として期待される JPEG2000 システムでは、多重解像度解析



による圧縮を行っている。JPEG2000では、画像を Wavelet 変換し、多重解像度に分解して量子化および符号化を行っている。そのため、復号側では多重解像度に分解した Wavelet 係数をほとんど計算コストを掛けずに利用できる。この Wavelet 係数を用いて顔画像から個人識別を行うシステムについて検討する。

# 謝辞

本研究を進めるにあたり日頃から熱心に御指導頂きました本学 小谷 一孔 助教授に深く感謝致します。貴重な御意見、御鞭撻を頂きました本学 宮原 誠 教授に深く感謝致します。

また、日頃から御助言、激励頂きました 剣持 雪子 助手をはじめとする本学 像情報処理学講座の皆様に深く感謝致します。

# 付録A JPEG2000システムにおける顔画像認識システム

本研究において、顔画像の多重解像度解析に用いる Wavelet 変換は、次世代の圧縮技術となる JPEG2000[19] に用いられている。そこで、JPEG2000 システムにおける顔画像認識システムを提案する。図 A.1 に本研究が目指す JPEG2000 システムにおける顔画像認識のシステムを示す。JPEG2000 では、画像を Wavelet 変換し、多重解像度に分解して量子化および符号化を行っている。そのため、復号側では多重解像度に分解した Wavelet 係数をほとんど計算コストを掛けずに利用できる。本研究では、この Wavelet 係数を用いて顔画像から個人識別を行うシステムを提案する。特に、システム構築の基礎となる Wavelet 係数による個人識別の基本特性を明らかにし、パラメータの決定方法および適用範囲を示す。また、JPEG2000 システムで単に輝度や色情報を用いるだけでなく、画像として構成した距離画像（顔距離画像）による顔の3次元構造の多重解像度解析による個人識別の可能性も検討する。

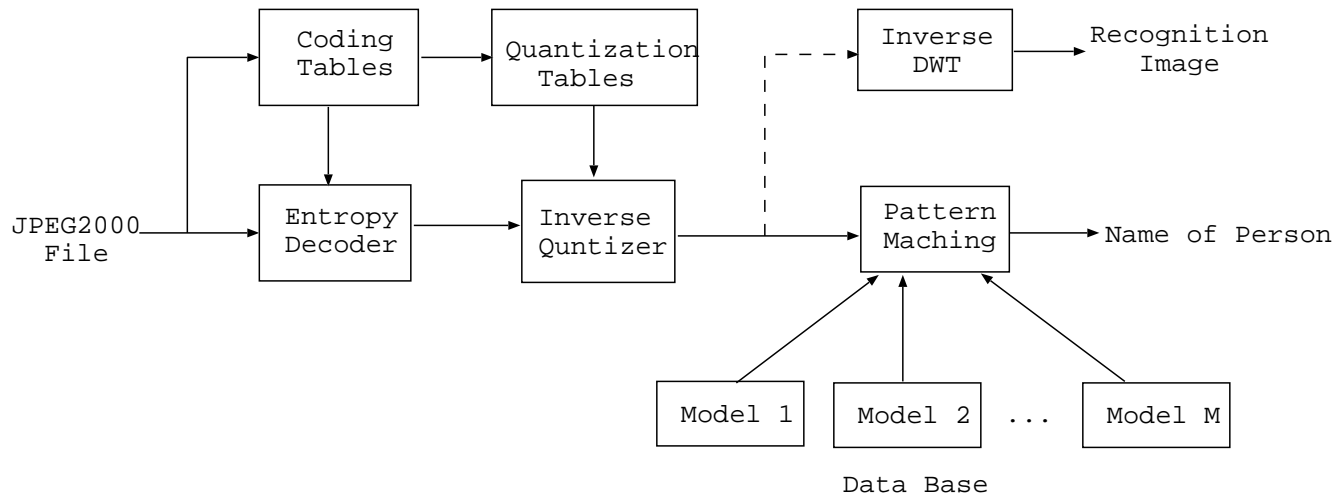


図 A.1: JPEG2000 システムにおける顔画像認識システム

## 関連図書

- [1] 赤松茂, 佐々木努, 深町映夫, 末永康仁, 濃淡マッチングによるロバストな正面顔の識別法—フーリエスペクトルのKL展開の応用—, 信学論 (D-II), pp1363-1373, 1993.
- [2] 赤松茂, コンピュータによる顔の認識—サーベイ—, 信学論 (D-II), J80-D-II, no.8, pp.2031-2046, 1997-08.
- [3] 長谷川修, 森島繁生, 金子正秀, 「顔」の情報処理, 信学論 (D-II), J80-D-II, no.8, pp.2047-2065, 1997-08.
- [4] 吉川敦則, 長坂保典, 鈴村宣夫, 直交変換を用いた顔画像の識別, 信学技法, PRU95-192, pp.15-22, 1996.
- [5] 山口修, 福井和広, 前田賢一, 動画像を用いた顔認識システム, 信学技法, PRMU97-50, pp.17-24, 1997.
- [6] 宋丙卓, 小沢慎治, 時系列顔画像処理による個人の認識, 信学技法, PRU93-66, pp.29-36, 1993.
- [7] 黒住隆行, 新座良和, 剣持雪子, 小谷一孔, 固有空間法と重判別分析による顔画像の個人性と表情解析, 信学技法, CS98-122, IE98-101, pp.57-64, 1993.
- [8] Harmon L. D., Khan M. K., Lasch R. and Ramig P. F., Machine identification of human faces, Pattern Recognition, 13, 2, pp.97-110, 1981.
- [9] 崔昌石, 原島博, 武部幹, 顔の3次元モデルを用いた顔面表情分析, 信学論 (D-II), J74-D-II, no.6, pp.766-777, 1991-06.
- [10] 阿部亨, 阿曾弘具, 木村正行, 3次元表面形状による人間の顔の自動識別—B-スプライン局面の制御点を利用して—, 信学論 (D-II), J73-D-2, 9, pp.1477-1484, 1990-09.
- [11] 新井康平, ウェーブレット解析の基礎理論, 森北出版, 2000.
- [12] ミノルタ, 非接触次元形状入力機 VIVID, ミノルタ, 1998.

- [13] Stefan Eickeler, Stefan Müller, Gerhard Rigoll, High Quality face Recognition in JPEG Compressed Images, Proc Int Conf Image Process, Vol.1999, No.Vol.1, pp.672-676, 1999.
- [14] 中田雄一郎, 小早川倫広, 星守, 大森堯, ウェーブレット変換を用いたテクスチャ特徴量, 信学技法, PRMU96-166, pp45-52, 1997.
- [15] 吉田幸生, 呉海元, 塩山忠義, 特徴点自動抽出を用いた実ガボール変換による個人識別, 信学技法, PRMU2000-163, pp23-30, 2001-1.
- [16] 越中洋, 久保守, 村本健一郎, ウェーブレットを用いた南極衛星画像のテクスチャ解析, 信学技法, CS2000-94, IE2000-104, pp43-48, 2000.
- [17] 入山彰夫, 剣持雪子, 小谷一孔, モフロフォジカルスケルトンによる3次元画像の構造解析, 信学技法, PRMU2001-216, pp.79-86, 2002.
- [18] 石井健一郎, 上田修功, 前田英作, 村瀬洋, わかりやすいパターン認識, オーム社, 1998.
- [19] 上野幾朗, 渥美栄司, 小野文孝, 静止画符号化の新国際標準方式 (JPEG2000) の概要, 映像情報メディア学会誌, Vol.54, No.2, pp.164-171, 2000.
- [20] 金森証, 小谷一孔, Wavelet 変換を用いた顔距離画像の特徴解析に関する研究 - 顔距離画像の Wavelet 係数による個人性の抽出 -, 信学技法, CS2001-118, pp.41-46, 2001.12.