JAIST Repository
https://dspace.jaist.ac.jp/

Title	離散Wavelet変換によるボケ成分を用いた画像認識手法の特性解析
Author(s)	鈴木,康彦; 小谷,一孔
Citation	情報処理学会研究報告 : オーディオビジュアル複合情報処理, 2004(126): 119-124
Issue Date	2004-12
Type	Journal Article
Text version	publisher
URL	http://hdl.handle.net/10119/3279
Rights	社団法人 情報処理学会, 鈴木康彦／小谷一孔, 情報処理学会研究報告 : オーディオビジュアル複合情報処理, 2004(126), 2004, 119-124. ここに掲載した著作物の利用に関する注意: 本著作物の著作権は（社）情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。 The copyright of this material is retained by the Information Processing Society of Japan (IPSJ). This material is published on this web site with the agreement of the author (s) and the IPSJ. Please be complied with Copyright Law of Japan and the Code of Ethics of the IPSJ if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. All Rights Reserved, Copyright (C) Information Processing Society of Japan.
Description	

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

社団法人　情報処理学会　研究報告
IPSJ SIG Technical Reports

2004－AVM－47（20）
2004／12／10

離散Wavelet変換によるボケ成分を用いた画像認識手法の特性解析

鈴木　康彦†　　小谷　一孔†

† 北陸先端科学技術大学院大学
〒923-1292 石川県能美郡辰口町旭台 1-1
E-mail: †{yasu-s,ikko}@jaist.ac.jp

あらまし　飯島は視覚パターンが満たすべき基礎方程式を導出し，同方程式を満足するガウス型の空間フィルタ特性によるボケを基礎にしたパターン認識理論を構築した[1]．一方，我々は顔画像を離散ウェーブレット変換（DWT）し，多重解像度成分に分解した上で，個人認識を行うことにより良好な識別精度を得ている．DWTによる分解は上記理論のボケ成分に対応するものと予想されるが，DWTに用いられるMother Wavelet関数の複雑さ等の理由から数理的証明はなされていない．本報告では，本手法とガウス型フィルタを基にした飯島理論とを対応させて検討・解析し，特性解析を試みる．加えて，実画像に対する画像認識実験を行って，解析結果との対応を検討する．

キーワード　離散ウェーブレット変換，多重解像度解析，パターン認識，ボケの変換

Characteristic Analysis of the Image Recognition Method Using the Blur Component by Discrete Wavelet Transform

Yasuhiko SUZUKI† and Kazunori KOTANI†

† Japan Advanced Institute of Science and Technology
Asahidai 1-1, Tatsunokuchi-machi, Nomi-gun, Ishikawa, 923-1292 Japan
E-mail: †{yasu-s,ikko}@jaist.ac.jp

Abstract　Iijima had built up his own pattern recognition theory by deriving a fundamental equation of visual pattern based on blur for Gaussian space filter [1]. We had proposed a new facial recognition method using multi resolution component of image by discrete wavelet transform (DWT), and it had fairly good accuracy of classification. Although decomposition of DWT correspond to blur of above theory, a mathematical proof is not yet carry out because mother wavelet function of DWT is complicate. This report discusses a mathematical analysis of correspondence between our method and Iijima's theory. We also discuss the comparison by experiments of pattern classification.

Key words　discrete wavelet transform, multi resolution analysis, pattern recognition, blur transformation

1. はじめに

顔画像を用いた個人認識において，顔画像を離散Wavelet変換（Discrete Wavelet Transform:DWT）して得られる多重解像度成分を用いた手法を検討したところ，個人認識では低域成分が識別に大きく寄与し，また解像度が低下しても識別率が殆ど低下しないという結果を得た[2]．一方，高域成分を用いると低い識別率しか得られなかった．このことから，識別に必要な顔の個人性を与える特徴の大部分が低域成分に含まれ，高域成分は顔の大きさ，位置ズレやノイズ成分など，外乱や不要変動成分が含まれているものと考えられる．これらの結果を基に，DWTを符号化アルゴリズムに採用しているJPEG 2000プロセスに顔画像識別を組み入れた結果，Eickelerらが提案したJPEGプロセス中のDCT係数を用い

た識別 [3] に比べて高圧縮時においても高い識別率が得られた [4].

これらの実験結果から顔画像を用いた認識においては，細かい情報である高域成分よりも大域的な情報を含む低域成分が重要であることが推測される．しかし，DWT の低域成分が顔画像の個人認識に有効であることについて，理論的裏付はできていない．

一方，飯島は視覚パターンが満たすべき基礎方程式を導出し，同方程式を満足するガウス型の空間フィルタ特性によるボケを基礎にしたパターン認識理論を構築した [1]．この理論を我々は飯島理論と呼んでおり，パターン認識におけるフィルタ処理や正規化処理などの重要な基礎理論となっている．この理論によると，パターンにボケを加えることによりパターン認識上で数々の利点が得られる．一方，ボケによってパターン認識に重要な情報が失われる可能性があるが，この問題を解決することで特に文字画像の認識においてボケは非常に重要な役割を果たした．

DWT による認識手法は，その実験結果から飯島理論のボケの変換に通じる特徴があると考えられる．そこで，DWT による多重解像度成分を用いた画像認識を飯島理論を基に解析・検討し，特性解析を試みる．この際，ガウス型のフィルタ特性と DWT に用いられるスケーリング関数と厳密には対応しない可能性があり，その場合，その違いによる歪が実際の画像認識にどのような影響を与えるのか実験によって検討する．DWT による我々の手法が飯島理論に十分対応するのであれば，それを基にして高効率・高精度な認識手法を構築することへつなげたい．

2. DWT のボケ成分のパターン認識への応用

本章では，まずガウス型のボケ変換の基礎理論となる視覚パターンの基礎方程式について説明する．その後，DWT による多重解像度解析が満足するエネルギー保存則を導出し，その保存則と視覚パターンの基礎方程式との対応関係を検討する．

2.1 DWT の処理機構を含んだ画像認識過程

ガウス型のフィルタ特性によるボケが持つ利点と同様に，DWT による処理はパターン認識過程における前処理過程の一部を担うものと考えられる．識別過程に入力する前に DWT によってノイズの除去，誤認識につながる微細変形の吸収，正規化誤差の影響の減少等が図られる．また，画像の解像度が低下することによって後の過程の効率化も達成できる．また，DWT による処理は低周波成分の抽出，すなわち特徴抽出過程の役割も果たしているとみなすことができる．本報告で考える DWT の多重解像度成分を用いた画像認識過程を図 1 に示す．

2.2 視覚パターンの基礎方程式 (飯島理論) [1]

画像は 2 次元平面上の光情報であり，光の伝播は物理現象であるからエネルギー保存則を満足する．ちなみに，飯島は鮮明な映像だけでなくその前後に現れるボケを含んだ映像群を一括して視覚パターンと呼び，特に鮮明な映像を画像として再定義している．視覚パターンにボケの変化が生じる時の光エネルギー保存則を微分形式で書くと，

$$\frac{\partial}{\partial \tau} f(\boldsymbol{r},\tau) + \nabla \cdot \boldsymbol{I}(\boldsymbol{r},\tau) = 0 \tag{1}$$

ただし，$\boldsymbol{I}(\boldsymbol{r},\tau)$: 光エネルギーの強さ (図流)

$\boldsymbol{r} = \boldsymbol{i}x + \boldsymbol{j}y$

τ : ボケのパラメータ

となる．式 (1) を変形すると，

$$\begin{aligned} -\frac{\partial}{\partial \tau}\left\{\frac{1}{2} f^2(\boldsymbol{r},\tau)\right\} &= -f(\boldsymbol{r},\tau)\frac{\partial}{\partial \tau} f(\boldsymbol{r},\tau) \\ &= f(\boldsymbol{r},\tau)\nabla \cdot \boldsymbol{I}(\boldsymbol{r},\tau) \\ &= \nabla \cdot \{f(\boldsymbol{r},\tau)\boldsymbol{I}(\boldsymbol{r},\tau)\} - \nabla f(\boldsymbol{r},\tau) \cdot \boldsymbol{I}(\boldsymbol{r},\tau) \end{aligned}$$

となる関係式が得られる．ここで

$$\begin{cases} w(\boldsymbol{r},\tau) = \dfrac{1}{2} f^2(\boldsymbol{r},\tau) \\ \boldsymbol{F}(\boldsymbol{r},\tau) = \nabla f(\boldsymbol{r},\tau) \\ q(\boldsymbol{r},\tau) = \boldsymbol{F}(\boldsymbol{r},\tau) \cdot \boldsymbol{I}(\boldsymbol{r},\tau) \\ \boldsymbol{P}(\boldsymbol{r},\tau) = w(\boldsymbol{f},\tau)\boldsymbol{I}(\boldsymbol{r},\tau) \end{cases} \tag{2}$$

なる 4 種類の量を定義すると，それらの間には次の関係が成り立つ．

$$\frac{\partial}{\partial \tau} w(\boldsymbol{r},\tau) + \nabla \cdot \boldsymbol{P}(\boldsymbol{r},\tau) + q(\boldsymbol{r},\tau) = 0 \tag{3}$$

ただし，$q(\boldsymbol{r},\tau) \geqq 0$ である．この $q(\boldsymbol{r},\tau)$ は視覚パターンがボケていくにつれて消費される量であり，図象損失と呼ばれている．

さて，式 (1) を満足するような $\boldsymbol{I}(\boldsymbol{r},\tau)$ は一意に決まらな

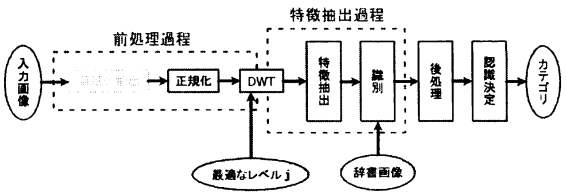


図 1 DWT による前処理を行なうパターン認識過程

い．ボケは本来理想位置にない映像の与える印象を，できるだけ弱めることを目的として備わった機能であった．そこで，図象損失の図流に対する比の 2 乗として

$$J[\boldsymbol{I}] = \frac{\{q(\boldsymbol{r},\tau)\}^2}{\|\boldsymbol{I}\|^2} \tag{4}$$

なる量を与え，この値が最大となるように $\boldsymbol{I}(\boldsymbol{r},\tau)$ が定まると仮定する (図象損失最大の原則)．先の光エネルギー保存則 (式 (1)) にこの原則を適用して次式が導かれる．

$$\nabla \cdot \nabla f(\boldsymbol{r},\tau) = \frac{\partial}{\partial \tau} f(\boldsymbol{r},\tau) \tag{5}$$

この方程式は視覚パターンの基礎方程式と呼ばれ，視覚パターンのボケはこの方程式を満たさなければいけないとされる．この方程式を満たすようなボケの特性はガウス型のみである事が示されている．

2.3 DWT による多重解像度解析が満足する光エネルギー保存則

DWT によって多重解像度成分に分解する場合も光エネルギー保存則は満足するものと考えられる．そこで，この場合に成り立つエネルギー保存則について考える．ここで，2 次元の DWT を水平垂直方向に分離して 1 次元 DWT により 4 つのサブバンド分割するモデルと仮定する．よって，今後の説明においては記述を簡単にするために 1 次元表記により説明する．

入力信号を

$$f(x), \left(x = 0, 1, 2, \ldots, N-1 \left(\text{ただし,} N = 2^M\right)\right)$$

とする．DWT によって多重解像度成分に分解する場合に成り立つ保存則は，

$$\Delta_{-j} f_j(x) = f_{j-1}(x) - f_j(x)$$

となる．上式の右辺は多重解像度解析より次のように変形できる．

$$\begin{aligned}\Delta_{-j} f_j(x) &= -g_{j-1}(x) \\ &= -\sum_{k=0}^{2^{j-1}N-1} d_k^{(j-1)} \psi\left(2^{j-1}x - k\right)\end{aligned} \tag{6}$$

式 (6) は DWT による多重解像度解析において成り立つ光エネルギー保存則とみなせる．ただし，

$$g_j(x) = \sum_{k=0}^{2^jN-1} d_k^{(j)} \psi\left(2^j x - k\right)$$

$$f_j(x) = \sum_{k=0}^{2^jN-1} c_k^{(j)} \phi\left(2^j x - k\right)$$

$\psi(x)$ ：ウェーブレット関数

$\phi(x)$ ：スケーリング関数

$d_k^{(j)}$ ：ウェーブレット係数

$c_k^{(j)}$ ：スケーリング係数

j ：レベル $(j = -1, -2, \ldots, -M)$

である．

2.4 DWT と飯島の基礎方程式との関係

飯島の視覚パターンの基礎方程式 (式 (5)) の左辺を差分形で書くと，

$$\begin{aligned}\Delta_x^2 f_j(x) &= f_j(x+1) - 2f_j(x) + f_j(x-1) \\ &= \sum_{k=0}^{2^jN-1} C_k^{(j)} \Big\{ \phi\left(2^j(x-1) - k\right) - \\ &\qquad -2\phi\left(2^j x - k\right) + \phi\left(2^j(x+1) - k\right) \Big\} \\ &= \sum_{k=0}^{2^jN-1} C_k^{(j)} \Delta_x^2 \phi\left(2^j x - k\right)\end{aligned} \tag{7}$$

一方，レベル $j-1$ のウェーブレット係数は分解アルゴリズムによってレベル j のスケーリング係数から次式より求めることができる．

$$d_k^{(j-1)} = \frac{1}{2} \sum_l h_{2k-l} c_l^{(j)} \tag{8}$$

ただし，h_k は分解数列であり，直交ウェーブレットの場合 h_k はツースケール数列 q_k から次式で求められる．

$$h_k = \bar{q}_k$$

よって，式 (6) の右辺は，

$$\begin{aligned}&\sum_{k=0}^{2^{j-1}N-1} d_k^{(j-1)} \psi\left(2^{j-1}x - k\right) \\ &= \sum_{k=0}^{2^{j-1}N-1} \left\{ \frac{1}{2} \sum_l h_{2l-k} C_l^{(j)} \sum_t q_t \left(2^j x - 2k - t\right) \right\} \\ &= \sum_{k=0}^{2^{j-1}N-1} \left\{ \frac{1}{2} \sum_l \bar{q}_{k-2l} C_l^{(j)} \sum_t q_t \left(2^j x - 2k - t\right) \right\}\end{aligned} \tag{9}$$

例えば，N=2 の Daubechies の場合，

$$\Delta_{-j} f_j(x)$$

$$= \frac{1}{2}\sum_{k=0}^{2^{j-1}N-1}\Big\{0.18C^{(j)}_{2k-2}\phi(x-2-2k) +0.32C^{(j)}_{2k-1}\phi(x-1-2k) -1.18C^{(j)}_{2k}\phi(x-2k) +0.68C^{(j)}_{2k+1}\phi(x+1-2k)\Big\} \quad (10)$$

上式において，式 (7) の基礎方程式と一致するのは，

$$\begin{cases} C_{2k-2}=0 \\ C_{2k-1}=0.27C_{2k} \\ C_{2k+1}=0.58C_{2k} \end{cases}$$

の場合であり，この条件は一般的とは考えられないため，DWT のエネルギー保存則は飯島の基礎方程式とは一致しない．しかし，式 (7)，(9)(或いは (10)) を見比べると，ある程度の対応関係は見えるので，DWT とガウス型のフィルタ特性によるボケの対応関係はあると考えられる．そこで，このことを次章以降の実画像を用いた認識実験で確める．

3. 実画像による認識実験

前章において，DWT によるボケのエネルギー保存則はガウス型のフィルタ特性によるボケが満足する基礎方程式と完全に一致しないことが分かった．この章では，この影響を解析するために DWT 及びガウスフィルタのボケを用いた認識実験を行う．

3.1 識別アルゴリズム

本報告で用いた識別アルゴリズムについて説明する．DWT によるボケ成分は式 (11) の 2-D 分解アルゴリズムによって求めることができる．

$$c^{(j-1)}_{k_1,k_2} = \sum_{l_1,l_2} g_{2k_1-l_1}g_{2k_2-l_2}c^{(j)}_{l_1,l_2} \quad (11)$$

ただし，g_k : 分解数列

$c^{(j)}_{k_1,k_2}$: レベル j のスケーリング係数

$c^{(0)}_{k_1,k_2} = f(m,n)$: 入力画像

今回は Wavelet 関数として N=2 の Daubechies を用いた．

一方，ガウス型の空間フィルタは式 (12) で表わすことができる．

$$G(m,n) = \frac{1}{4\pi\sigma}\exp\left(-\frac{m^2+n^2}{4\sigma^2}\right) \quad (12)$$

ガウス関数によるボケの変換は $G(m,n)$ と画像 $f(m,n)$ との畳み込み演算で計算する．

入力画像，及び辞書画像に前述した 2 種類のボケ変換によって前処理されたものを用いる．ただし，今回は両者のボケの違いを比較するために，正規化等の他の前処理や空間変換等の特徴抽出は行わない．また，画像間の類似性を求める識別関数は式 (13) で求められる複合類似度を用いた．複合類似度を用いた識別法は，注目するグループに属する類似画像集合全体の分布の模様に照らして識別を達成しようとするものである．

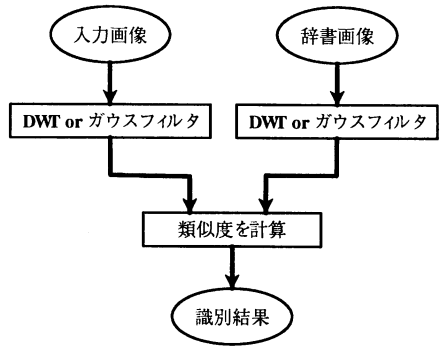


図 2 本実験の識別アルゴリズム

$$d_\alpha = \frac{1}{a}\sum_{\alpha\in A}\frac{\sum_{m=0}^{M}\sum_{n=0}^{N} f_{\mathrm{in}}(m,n)\cdot f_\alpha(m,n)}{\sum_{m=0}^{M}\sum_{n=0}^{N} f^2_{\mathrm{in}}(m,n)\cdot\sum_{m=0}^{M}\sum_{n=0}^{N} f^2_\alpha(m,n)} \quad (13)$$

ただし，$f_{\mathrm{in}}(m,n)$:入力画像

$(m=0,1,\ldots,M-1, n=0,1,\ldots,N-1)$

$f_\alpha(m,n)$:クラス A(a[枚]) に属する辞書画像

以上から本実験の識別アルゴリズムを図 2 に示す．

3.2 顔画像による比較実験

これまでの検討結果から DWT によるボケのエネルギー保存則はガウスフィルタによるボケが満足する基礎方程式と一致しないものの近い特性があるものと期待するため，違いが実画像による認識にどれほど影響するのか検討する．ここではこれまで我々の手法が有効であることが確認されている顔画像認識において DWT とガウスフィルタの比較実験を行なった．実験には ORL Face Database [5] の画像を使用した．

3.2.1 実験条件

DWT とガウスフィルタによるボケの違いが実画像認識においてどの様に影響するのか比較するために Open Data(入力画像が辞書画像に含まれない場合) に対して実験を行った．辞書画像に 1 人につき各 5 枚 ×20 人の計 100 枚，入力画像に同様に残りの 100 枚を用いた．図 3(a) に今回用いた顔画像の例を，(b) には DWT によって得られるボケ成分，(c) に

(a) 原画像　(b) DWT　(c) ガウスフィルタ

図 3　原画像とボケ処理の画像

はガウス型の空間フィルタ特性によるボケ画像を示す.

3.2.2 実験結果

図 4(a), (c) に DWT によるボケ成分を用いた場合の認識結果を，図 4(b), (d) にガウス関数によるボケ画像を用いた場合の認識結果を示す．識別率に関しては両者の違いはほとんど表れなかった．また，各クラスのクラス内分散はボケを加えるにつれて小さくなっていくのが分かる．これは，各クラスのばらつきがボケを加えるにつれて小さくなっていくことを示している．両者の値は若干違いが認められるものの，傾向は一致している．すなわち，DWT によるボケもガウスフィルタと同様の役割を果しているものと考えられる.

3.3 文字画像による認識実験

飯島は文字画像にボケを加えることで高い認識性能を実現している．そこで，DWT のボケ成分を利用した数字画像の認識実験を行い，その結果を検討した．実験には，The MNIST database of handwritten digits の数字画像を使用した [6]．図 5 にその一例を示す.

3.3.1 実験条件

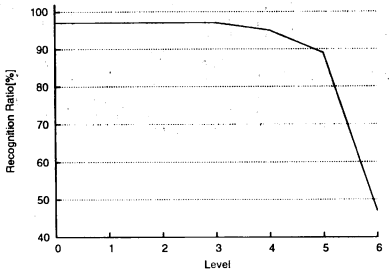
本実験も先の実験同様 Open Data に対して行った．辞書画像として各数字につき 1,000 枚 ×10(0〜9) の計 10,000 枚，入力画像には各 100 枚 ×10 の計 1,000 枚を用いた．その他の手法に関しては図 2 に示したアルゴリズムを用いる.

3.3.2 実験結果

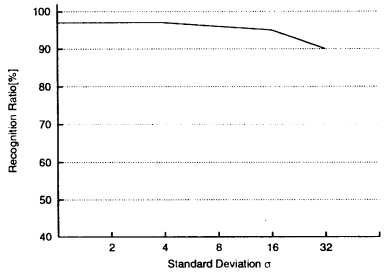
図 6 に認識結果を示す．文字画像による認識においても顔画像の場合と同様にレベルを上げて情報が失われても認識率に大きな低下は無い．これは文字認識においても識別に必要な情報が低域領域に含まれていることを示している．また，DWT とガウスフィルタの認識率を比較してみると，顔画像認識の場合と同様に近似した特性を示している.

4. おわりに

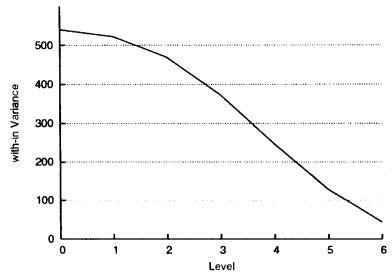
DWT によって多重解像度成分に分解することで得られる低域成分をボケ成分と見なし，飯島の基礎方程式を満足する



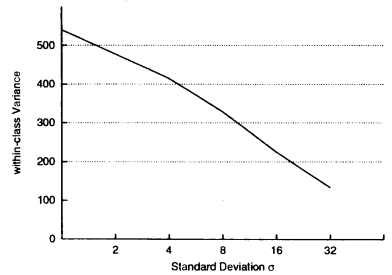
(a) DWT によるボケ成分の認識特性



(b) ガウスフィルタによるボケ画像の認識特性



(c)DWT によるボケ成分のクラス内分散の特性



(b) ガウスフィルタによるボケ画像のクラス内分散の特性

図 4　顔画像を用いた認識実験

ガウス型の空間フィルタ特性によるボケとの対応関係を解析した．その結果，両者は完全には一致しないものの DWT によるボケ成分の特性はガウスフィルタによるボケ画像に近い特性を有しているののと考えられる．そこで，完全に一致しないことが実画像の認識にどの程度影響するのか検討するた

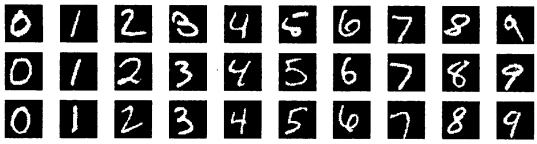
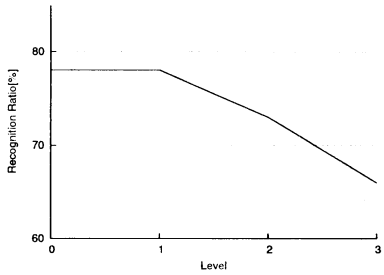
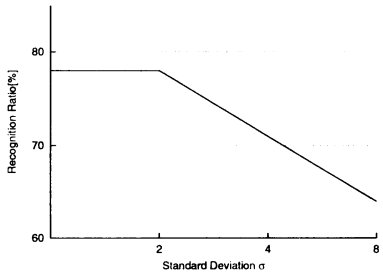


図 5　標準数字画像パターンデータベース MNIST の例



(a) DWT によるボケ成分の認識特性



(b) ガウスフィルタによるボケ画像の認識特性

図 6　数字画像を用いた認識実験

めに顔画像認識で比較実験を行った．その結果，両者の違いが実際の認識に与える影響はわずかであることが分かった．

加えて飯島によってボケが有効であることが実証されている文字画像の認識に DWT によるボケ成分を利用する実験を行った．その結果，顔画像認識と同様に文字認識においても，DWT によるボケ成分とガウスフィルタのボケの特性は近い認識特性を持っていることが分かった．

DWT により導出した光エネルギー保存則が飯島の基礎方程式と一致しないことの影響は実用的には問題無いレベルと考えている．今後，更に数理的な解析を行い定量的に違いを示したい．また，飯島の理論との対応関係が見られたことが何を意味するのか，どのような新たな知見が得られるのか，より深い解析を行う．その結果を活かし，各パターンに適用した認識手法を構築することを目指す．

文　献

[1] 飯島 泰蔵 “視覚情報の基礎理論 –パターン認識問題の源流–” コロナ社，1999.

[2] 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–” 電子情報通信学会技術研究報告，CS2001-118，pp.41-46，2001.12.

[3] S.Eickeler, S.Müller, G.Rigoll “High Quality face Recognition in JPEG Compressed Image” Proc International Conference on Image Processing, Vol.1999, No. Vol. 1, pp.672-676, 1999.

[4] 中村 一也，剣持 雪子，小谷 一孔 “JPEG2000 プロセスに適応させた Wavelet 変換による顔画像認識” 電子情報通信学会技術研究報告，CS2002-146，pp.73-78，2002.12.

[5] AT&T Laboratories Cambridge, http://www.uk.research.att.com/

[6] Yan LeCun, http://yann.lecun.com/

[7] 榊原 進 “数理科学 ウェーブレットビギナーズガイド” 東京電機大学出版局，1995.5.

[8] 中田 雄一郎，小早川 倫広，星 守，大森 尭 “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，CS2000-94，IE2000-104，pp43-48，2000.