

Watershed 分割による Semantic Object の Ridgelet 表現

Ridgelet Representation of Semantic Objects Using Watershed Segmentation

正会員 長谷川 誠†

田島 慎一††

Makoto Hasegawa† and Shinichi Tajima††

Abstract We propose applying of the ridgelet transform to semantic objects segmented using watershed segmentation. The ridgelet transform effectively represents objects with singularities along lines; therefore, it is a powerful tool for coding. Moreover, it can easily rotate blocks in the ridgelet domain. Nevertheless, the targets to be rotated are not rectangular blocks but real objects. For that reason, we divide pictures into semantic objects using watershed segmentation, then convert each object into the ridgelet domain so they can be rotated easily. We experimented in coding and rotating objects.

キーワード: Ridgelet 変換, Watershed 領域分割, Semantic オブジェクト, 回転

1. ま え が き

画像符号化の基盤技術として、離散コサイン変換や Wavelet 変換が主流となっている現在、Candès によって提案された Ridgelet 変換が注目されつつある¹⁾。Ridgelet 変換は、濃淡画像を Radon 変換した後に Wavelet 変換することと等しく、輝度値を線積分する Radon 変換の特徴により、直線状のエッジを効果的に表現することが可能である。また、Wavelet 変換によって高符号化効率が実現される。Do と Vetterli とは、画像領域を四辺形ブロックで格子状に分割し、各ブロックの輝度値を Ridgelet 変換して変換係数を符号化している²⁾。また、画像領域を逐次十字形に可変サイズの四辺形ブロックで分割して変換する方法を示している (Quad-tree 符号化)³⁾。Donoho らは濃淡画像を階層的にサブバンド分解してから、各層のブロックを変換する方法を提案し、雑音除去への応用について述べている⁴⁾⁵⁾。

ところで、Ridgelet 変換の更に優れた特性として、ブロックの拡大縮小、移動、回転が容易であることがあげられる。

特に回転については、変換係数行列を左右にシフトすればよい。しかし、回転の対象はブロックではなくオブジェクトであり、個々のブロックを回転させてもオブジェクトは回転しない。そこで、Watershed 領域分割⁶⁾によって濃淡画像を意味的にまとまりをもった対象 (Semantic オブジェクト) に分割し、個々のオブジェクトを Ridgelet 変換する方法を提案する⁷⁾⁸⁾。これにより、Semantic オブジェクト単位の回転が可能となる。変換係数の操作によって物体を回転させることは、オブジェクトの同定に有効である。例えば、風景や人物、様々な動物や静物を被写体とするデジタル濃淡画像のデータベースを考えよう。被写体は様々な姿勢をとり、移動のみならず回転も含む。また、カメラの回転によって、映像全体が回転している場合もある。移動や拡大縮小は、離散コサイン変換や Wavelet 変換において、変換係数の操作によって実現できるが、柔軟な回転は難しい。ここでは、Ridgelet 変換係数を左右にシフトさせることによって、オブジェクトを回転させ、変換係数の比較によって検索する。画像を逐次復号化した後に、オブジェクトを切り出し、回転させて走査することは可能であるが、符号化データ上で実現することにより高速な処理が可能となる。また、動画像符号化の動き補償への応用にも有効である。なお、ここでは本方法の応用の可能性について述べたが、これらの実装については今後の検討とする。

ここでは、はじめに Ridgelet 変換と Watershed 領域分割とについて解説し、Semantic オブジェクトの Ridgelet 変換符号化について検討する。Watershed 領域分割によって濃淡画像から Semantic オブジェクトを抽出し、各オブジェクトについて重心を原点とするオブジェクト画像を生

2004 年 3 月 10 日受付, 2005 年 1 月 5 日最終受付, 2005 年 3 月 11 日採録

† 近畿大学 工学部 情報システム工学科

(〒739-2116 東広島市高屋うめの辺 1 番)

†† 新潟大学 工学部 情報工学科

(〒950-2181 新潟市五十嵐 2 の町 8050 番地)

† Dept. of Information and Systems Engineering, School of Engineering, Kinki University

(1, Takaya Umenobe, Higashi-Hiroshima City, Hiroshima 950-2181, Japan)

†† Dept. of Information Engineering, Niigata University

(8050, Ikarashi 2-no-cho, Niigata City 950-2181, Japan)

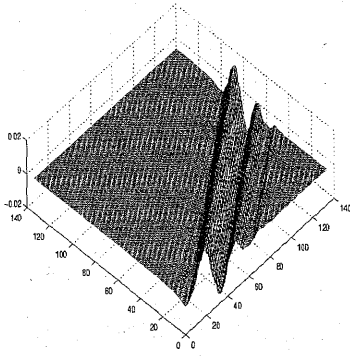


図 1 Ridgelet 基底関数 $\psi_{a,b,\theta}(\mathbf{x})$ の例
An example of ridgelet $\psi_{a,b,\theta}(\mathbf{x})$.

成する。各オブジェクト画像を Ridgelet 変換し、変換係数を符号化する。なお、オブジェクト画像においてオブジェクト外部の輝度値を零とすると、必ず零になる変換係数が存在する。これらを非有意係数として符号化対象から外す。Ridgelet 変換係数行列を左右にシフトすることによってオブジェクトを回転させる。標準動画「Claire」を用いてシミュレーションする。顔オブジェクトを符号化し、回転させた結果を示す。

2. Ridgelet 変換²⁾

濃淡画像が与えられている 2 次元領域 R^2 の座標を位置ベクトル \mathbf{x} で表し、輝度値を $f(\mathbf{x})$ とする。 $f(\mathbf{x})$ の Ridgelet 変換 D は、

$$Df = \hat{f}(a, b, \theta) = \int_{R^2} \psi_{a,b,\theta}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}, \quad (1)$$

と定義される。なお、ここでの $\psi_{a,b,\theta}(\mathbf{x})$ は Wavelet 変換関数 $\psi(\mathbf{x})$ を

$$\psi_{a,b,\theta}(\mathbf{x}) = a^{-1/2} \psi((\mathbf{x} \cdot \xi - b)/a) \quad (2)$$

と拡大縮小 a 、平行移動 b 、回転 θ したものである (図 1)。 $\xi = (\cos \theta, \sin \theta)$ とする。Ridgelet 変換 D は、Radon 変換 R

$$Rf = \tilde{f}(\theta, t) = \int_{R^2} f(\mathbf{x}) \delta(\mathbf{x} \cdot \xi - t) d\mathbf{x}. \quad (3)$$

を用いて、

$$Df = WRf = \int_R \psi_{a,b}(t) \tilde{f}(\theta, t) dt. \quad (4)$$

と書き換えることができる。すなわち、Ridgelet 変換は濃淡画像を Radon 変換した後に Wavelet 変換することと等しい。

3. Watershed 領域分割⁶⁾

濃淡画像を 3 次元空間内の地形図とし、ここに水を流し込むことを考える。水は低い領域 (輝度値の極小値) にたまり、いずれ水溜りの境界に分水嶺ができる。この水溜りを領域とする。図 2 に標準動画「Claire」を Watershed 領

研究速報 □ Watershed 分割による Semantic Object の Ridgelet 表現

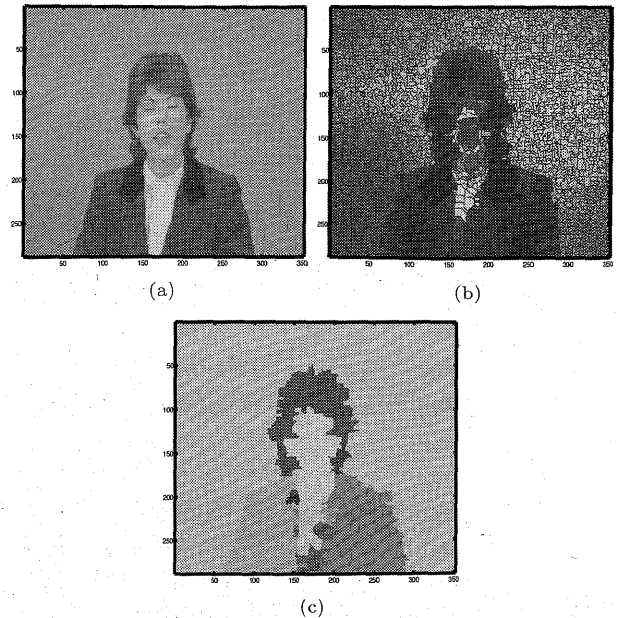


図 2 (a) 標準画像「Claire」、(b) Watershed 領域分割の例、
(c) 過分割統合結果

(a) An original image – a frame of Claire, (b) a result of watershed segmentation, (c) a result of merging over-segmented regions.

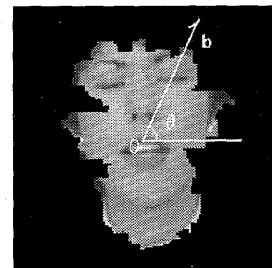


図 3 顔オブジェクト
A face object.

域分割した例を示す。一般的な Watershed 領域分割では、図 2(b) に示すように分水嶺が残るが、ここではモルフォロジー処理を進めて領域を接合させる。また、画像中の雑音などにより過分割状態になるため領域統合を行う。条件

$$(ave(R_a) - ave(R_b))^2 > T \quad (5)$$

を満たすまで再帰的に領域を統合する。なお、ここでの R_a と R_b とは互いに隣接している二つの領域であり、 $ave(R)$ は領域 R における輝度値の平均値を示す。 T は閾値である。図 2(c) は閾値 T を 32 に設定した場合の領域統合結果である。目黒は、Watershed 領域分割、および、過分割領域の統合によって、画像から意味的にまとまりをもった対象 (Semantic オブジェクト) が抽出できると述べている⁹⁾。

4. Semantic オブジェクトの Ridgelet 表現

Semantic オブジェクトごとにオブジェクト画像を生成する。オブジェクト o_i のオブジェクト画像 $f_i(\mathbf{x})$ を

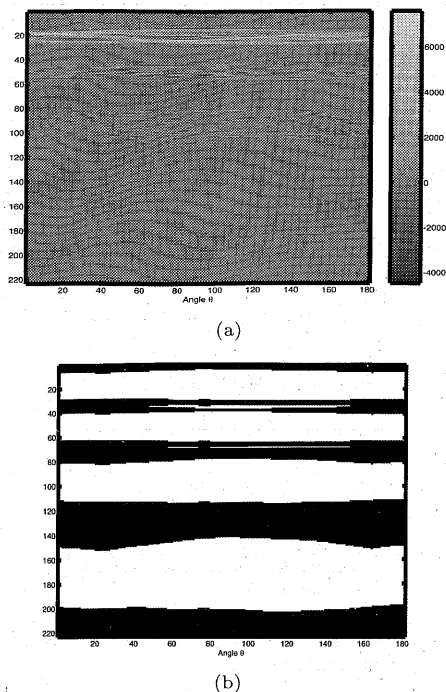


図 4 Ridgelet 変換係数行列 Df_i (a) とそのマスク g_i (b)
(a) A ridgelet matrix Df_i and (b) its mask g_i .

$$\hat{f}_i(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \mathbf{x} \in o_i, \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

と定義する。また、オブジェクト o_i の重心 \mathbf{w}_i を

$$\mathbf{w}_i = \int_{R^2} \mathbf{x} \hat{f}_i(\mathbf{x}) d\mathbf{x} / \int_{R^2} \hat{f}_i(\mathbf{x}) d\mathbf{x} \quad (7)$$

と定義し、重心が画像の原点に位置するように

$$\hat{f}_i(\mathbf{x}) = \hat{f}_i(\mathbf{x} - \mathbf{w}_i) \quad (8)$$

とする。図 3 にオブジェクト画像の例を示す。オブジェクト画像 $\hat{f}_i(\mathbf{x})$ を Radon 変換し、Radon 変換係数を偏角 θ 軸方向について 1 次元離散 Wavelet 変換する。例えば、各 Radon 変換係数を四つのサブバンドに分割すると、図 4(a) に示す Ridgelet 変換係数行列 Df_i が得られる。

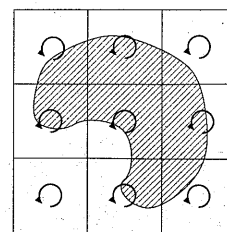
オブジェクト画像においてオブジェクト外部の輝度値は零であるので、必ず零になる Ridgelet 変換係数が存在する。この零係数を非有意係数とし、符号化対象から外すことを考える。まず、オブジェクト o_i について画像 $\hat{g}_i(\mathbf{x})$ を

$$\hat{g}_i(\mathbf{x}) = \begin{cases} 0 & \mathbf{x} \in o_i, \\ 1 & \text{otherwise} \end{cases} \quad (9)$$

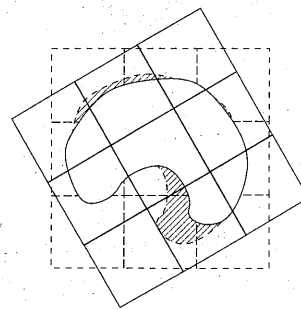
と定義する。画像 $\hat{g}_i(\mathbf{x})$ を Ridgelet 変換し、変換係数 $D\hat{g}_i$ を求める。マスク $g_i(a, b, \theta)$ を

$$g_i(a, b, \theta) = \begin{cases} 0 & (D\hat{g}_i \neq 0), \\ 1 & \text{otherwise} \end{cases} \quad (10)$$

と定義し、マスク値が 0 の場合、オブジェクト画像の Ridgelet 変換係数は非有意係数として符号化しない。マスク $g_i(a, b, \theta)$ の例を図 4(b) に示す。



(a)



(b)

図 5 (a) 格子状のブロックを回転、(b) ブロックの平行移動を含めた回転

(a) Rotation of lattice blocks and (b) a multi-block rotation with a shift operation on each block.

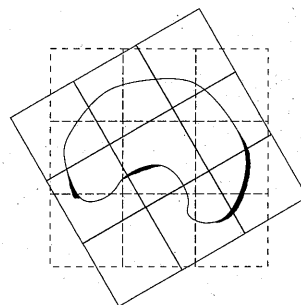


図 6 回転操作を含めた物体検索のためのオブジェクト比較
Comparison between two objects for searching with rotation.

5. Semantic オブジェクトの回転

Ridgelet 変換符号化は、一般的に画像領域を四辺形ブロックで格子状に分割し、各ブロックの輝度値を Ridgelet 変換する。比較的小さなオブジェクトが大きなブロックに含まれる場合、ブロックの回転によってオブジェクトは回転する。しかし、図 5(a) に示すように、大きなオブジェクトが複数のブロックに跨る場合、各ブロックは回転するがオブジェクトは回転しない。強制的に回転させるとオブジェクトは分裂する結果となる。各ブロックの回転軸を平行移動して回転させると、図 5(b) に示すように、複数のブロックをまとめて回転させることは可能である。しかし、回転操作を含めた物体検索などの応用を考えると、図 6 のように回転前後のブロックどうしの重なりは少なくなり、相関は小さくなる。特に、回転軸から離れたブロックについては著しい。そこで、ここでは画像を Semantic オブジェクトに分割し、各オブジェクトを Ridgelet 変換する。オブジェクトどうしの対応をとる必要があるが、図 7(a) に示すよう

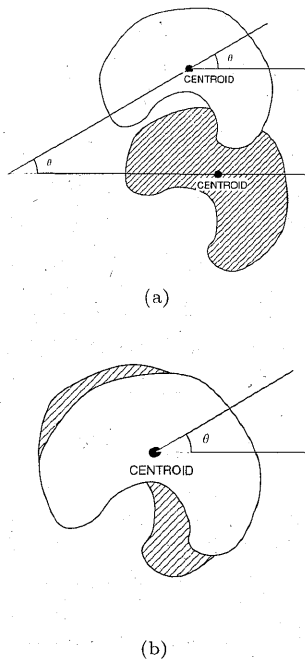


図 7 Semantic オブジェクトの回転
Rotation of semantic objects.

表 1 シミュレーション結果
A result of the simulation.

対象画像	Claire	Scissors
オブジェクト画像のサイズ	128 × 128 画素	256 × 256 画素
オブジェクト内部の画素数	4943 画素	4732 画素
オブジェクト内のビットレート	6.87 bpp	6.15 bpp
Ridgelet 変換係数行列のサイズ	223 × 180	368 × 180
有意係数の個数	24591	22650
Ridgelet 変換後のデータレート (256 ステップで均等量子化, 非有意係数を除く)	3.20 bits/coef.	5.21 bits/coef.
画質 (信号対雑音比 PSNR)	31.41 dB	27.80 dB

に回転軸がオブジェクトの外にあったとしても、図 7(b) に示すようにオブジェクトの重心を重ねて回転させるように捉えることが可能となる。

ところで、Ridgelet 変換係数 $\hat{f}(a, b, \theta)$ は

$$\hat{f}(a, b, \theta + \pi) = \hat{f}(a, -b, \theta) \quad (11)$$

の性質があるため、偏角 θ について 0 から π までの範囲を符号化する。偏角 θ が 0 以下、または、 π 以上の場合、Ridgelet 変換係数行列を上下反転すればよい。ただし、図 4(a) に示すように、Ridgelet 変換行列は層になっているため、各層で上下反転しなくてはならない。オブジェクト o_i を時計回りに θ_0 度回転させる場合、Ridgelet 変換係数 $\hat{f}(a, b, \theta)$ を

$$\hat{f}(a, b, \theta) = \hat{f}(a, b, \theta + \theta_0) \quad (12)$$

とする。すなわち、Ridgelet 変換係数行列 $\hat{f}(a, b, \theta)$ を θ_0 列左シフトさせればよい。

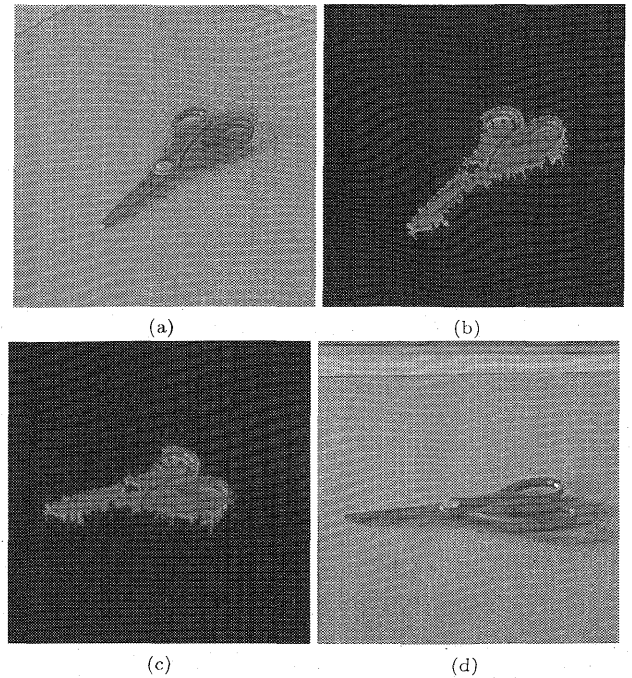


図 8 標準画像「Scissors」(a) とハサミを抽出して符号化した場合の再生画像 (b)、図 (b) を 30 度回転させた例 (c)、ハサミの角度が異なる標準画像 (d)

(a) Original test image “Scissors”, (b) a reconstructed image from encoding the extracted scissors, (c) an example of rotation for clockwise 30 degree, and (d) another scissor image.

6. シミュレーション

図 2(a) に示す標準動画像「Claire」と図 8(a) に示す標準画像「Scissors」を用いて、本方法をシミュレーションする。「Claire」はオハイオ州立大学で、「Scissors」は INRIA(フランス国立情報処理自動化研究所) で提供されている標準画像である。数値計算ソフトウェア MATLAB のライブラリ関数 (radon(), watershed() など) を用いてシステムを実装した。

まず、濃淡画像を Watershed 領域分割し、過分割領域を統合する。画像から顔とハサミを抽出し、オブジェクトの重心を中心とするオブジェクト画像を生成する。表 1 に示すようにオブジェクト画像は、「Claire」については縦 128 画素 × 横 128 画素、「Scissors」については縦 256 画素 × 横 256 画素とした。顔領域の画素数は 4943 画素であり、ハサミ領域の画素数 4732 画素である。顔領域のビットレートは 6.87bpp、ハサミ領域については 6.15bpp である。オブジェクト画像を Radon 変換する。なお、偏角 θ については 1 度刻みに、動径方向 b については 1 画素間隔刻みに標本化する。Radon 変換係数を偏角 θ 軸方向に Wavelet 変換する。Ridgelet 変換係数行列、および、そのマスクを図 4 に示す。Ridgelet 変換係数行列のサイズは、それぞれ縦 223 × 横 180、縦 368 × 横 180 となり、有意係数の個数は 24591、22650 である。有意係数のダイナミックレンジを 256 階調 (8bits) で均等量子化したところ、データレートはそれぞれ

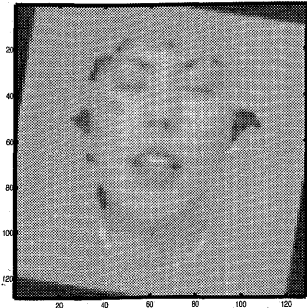


図9 図3のオブジェクトを回転させた例。回転の状況がわかるように画像の背景を顔領域の輝度値の平均値とした

An example of a rotating object in fig. 3 by ridgelet domain. In order to recognize rotation clearly, we fill in the background of image by the average of its luminances in the face.

れ3.20bits/coefficient, 5.21bits/coefficientとなり、画質は信号対雑音比 (PSNR) で31.41dB, 27.80dBとなった。符号化前のデータ量をビットレートとオブジェクト領域の画素数との積とすると、それぞれ約33958bits, 29102bitsである。一方、符号化後のデータ量をデータレートと有意係数の個数との積とすると、それぞれ約78691bits, 118007bitsである。今回のシミュレーション条件ではデータ圧縮効果は見られなかった。標本化、量子化を適切に調節する必要がある。

Ridgelet 変換係数行列を左右にシフトしてオブジェクトを回転させる。図9はRidgelet 変換係数行列を左シフトし、顔を時計回りに10度回転、図8(c)はハサミを時計回りに30度回転させた例である。図8(d)はハサミ姿勢を変えて撮影した標準画像である。図8(c)と図8(d)を比較し、ハサミを同定することを今後検討する。

ここでのシミュレーションによって、本提案方法の実現可能性を明らかにした。Watershed 領域分割によってオブジェクトを抽出、Ridgelet 変換し、PSNRで30dB程度の画質を有する再生画像の表現が可能であることを示した。なお、変換係数を256ステップで均等量子化する場合にはデータ圧縮の効果は確認されず、量子化のチューニングが必要であることを今後の課題として示した。また、Ridgelet 変換係数行列を左右にシフトすることによって、オブジェクトの回転が可能であることを示した。動画像符号化への応用を考えると、物体を回転させてイントラ符号化することによって、フレーム間の予測誤差を軽減させて圧縮効率を向上させること、物体を回転させて動き補償することによって画質を向上させる効果が期待される。これについては今後の課題とする。

7. む す び

濃淡画像をWatershed 領域分割してSemantic オブジェクトを抽出し、Ridgelet 変換符号化する方法を検討した。Ridgelet 変換は直線状のエッジを効果的に表現することが可能であり、データ圧縮効率も高く、符号化や雑音除去な

どに適用されている。また、変換係数行列を左右にシフトすることによって、変換対象ブロックを回転させることができる。しかし、回転の対象はブロックではなくオブジェクトである。そこで、ここではWatershed 領域分割を用いて濃淡画像からSemantic オブジェクトを抽出し、各オブジェクトをRidgelet 変換する方法を提案した。標準動画像を用いてシミュレーションし、本提案方法の実現可能性を示した。今後、本方法を画像データベース、動画像符号化に応用したアプリケーションの開発を検討する予定である。

本研究の一部は(財)山口育英奨学会と、新潟大学プロジェクト推進経費の助成による。

(文 献)

- 1) E. J. Candès : "Ridgelets: Theory and application", Ph.D. dissertation, Dept. Statistics, Stanford Univ., Stanford, CA (1998).
- 2) Minh N. Do and M. Vetterli : "The Finite Ridgelet Transform for Image Representation", *IEEE Trans. Image Processing*, **12**, pp.16-28 (Jan. 2003).
- 3) Minh N. Do, P. L. Dragotti, R. Shukla, and M. Vetterli : "On the Compression of Two-dimensional Piecewise Smooth Functions", *Proc. of IEEE ICIP '01*, Thessaloniki, Greece (Oct. 2001).
- 4) D. L. Donoho and M. R. Duncan : "Digital Curvelet Transform: Strategy, Implementation and Experiments", *Manuscript* (Nov. 1999).
- 5) J. Starck, E. J. Candès, and D. L. Donoho : "The Curvelet Transform for Image Denoising", *IEEE Trans. Image Processing*, **11**, pp.670-684 (June 2002).
- 6) S. Beucher and C. Lantuejoul : "Use of Watershed in Contour Detection", *International Workshop on Image Processing, Real-Time Edge and Motion Detection Estimation France*, pp.17-21 (Sep. 1979).
- 7) 長谷川誠, 田島慎一, 長井啓彦 : "Watershed 領域分割による Semantic Object の Ridgelet 表現", 信学総大, A-6-5 (Mar. 2004).
- 8) M. Hasegawa and S. Tajima : "A Ridgelet Representation of Semantic Object Using Watershed Segmentation", in *Proc. of IEEE ISCIT 2004*, Sapporo, Japan, Oct. 2004.
- 9) 目黒光彦 : "領域構造の解析に基づく動画像からのセマンティックオブジェクトの抽出", 電気通信普及財団, 研究調査報告書, **17** (2002).



は せ が わ ま こと
長谷川 誠 1991年、新潟大学工学部情報工学科卒業。2002年、同大学大学院自然科学研究科修了。1993年、富士写真フイルム(株)入社。新潟大学総合情報処理センター講師。現在、近畿大学工学部情報システム工学科助教授、博士(工学)。信号処理、画像認識、CG、コンピュータネットワークなどの研究に従事。正会員。



た じ ま し ん い ち
田島 慎一 1976年、東京工業大学理学部卒業。1981年、東京大学大学院理学研究科博士課程修了。理学博士。現在、新潟大学工学部情報工学科教授。偏微分方程式論と代数解析学を専攻。多変数複素解析、積分変換の超局所解析、計算代数解析アルゴリズムの研究に従事。