

Morphology演算による高解像度古文書画像の Scale Space生成

藤井大輔* 村山健二* 岡田至弘*

*龍谷大学古典籍デジタルアーカイブ研究センター

要約

人間が文書画像を認知する場合、要求される詳細度に応じて多段階の解像度で対象画像をみている。これにより、行の方向検出、言語種の診断、段落、文章、単語等の認識など、1つの視野入力(画像)に対して解像度を变化させて認知的負荷を最適化している。本研究はこの特徴を応用しmorphology演算による多重解像度画像を生成し、その中から特徴抽出・解析過程に最適な解像度を対応させる。そして多種類の特徴抽出・解析処理を高効率化し、大量・高解像度画像データ処理システムのusabilityの向上を図る。提案手法の評価実験として、「大谷文書」紙片の行高および行間の計測を行う。パラメータを最適化したmorphology演算により生成した画像を処理することにより、計測精度が向上することを示す。

Generating Scale Space for High Resolution Historic Document Image by Morphology Processing

FUJII Daisuke*, MURAYAMA Kenji*, and OKADA Yoshihiro*

Abstract

In document image recognition, we observe an image in multiple resolutions according to required cognitive level of detail. This enables us to detect line direction, kind of language, paragraphs, sentences, or words from single view input under optimum cognitive load. Based upon this characteristics, in this research, we generate plural images in different resolutions by "Morphology Processing" from single source image, and adapt appropriate resolution to every kinds of image analysis methods. Then we aim to improve usability of large scale high resolution image data processing system. As a evaluation experiment of our method, we measure line height and line skip of fragmented paper in 'Otani Documents'. We show that using image generated by morphology processing with optimized parameter increases measurement accuracy.

*Ryukoku University Digital Archives Research Center

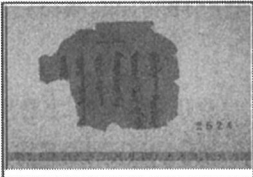
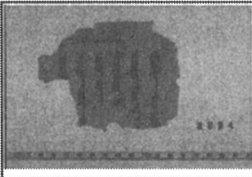
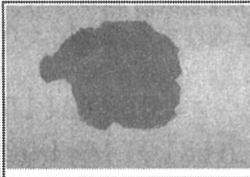
1. はじめに

人間はあるシーンを視覚的に認知する場合、その認知作業に要求される詳細度に応じてシーンの心理物理的解像度を多段階に変化させている。例えばジグソーパズルを組み上げる、移動する群集の中で特定の人物を発見する、広大な住宅地図から目的の家を特定するなどは、目的達成までの過程においてシーンや画像情報の解像度を変化させる典型例である。このことは人が文書画像を認知する場合にもあては

まる。人間は行の方向検出、言語種の診断、段落、文章、単語等の認識など、目的のタスクにあわせて1つの視野入力(画像)の心理物理的解像度を変化させ、認知的な負荷および認識精度を最適化している [1]。

本研究はこの働きを画像処理システムに応用し、システムの低負荷化および高精度化を目指す。すなわち、一般的に高負荷である高次画像解析・高次特徴抽出のさまざまな手法に対し、それぞれに低負荷かつ高精度の結果をもたらすような解像度を求める。その解像度を導き出す基盤として *Mathematical Morphology* 演算による *Scale Space (尺度空間) Filtering* のパラメータにより多重解像度画像を生成し、その中から当該高次画像解析・高次特徴抽出処理に最適な解像度を選び対応させ、それを *Scale Space* データベースとして保持する。

本研究の最終的目標は、こうして多くの種類の高次画像解析・高次特徴抽出を合理化し、大量・高解像度画像データ処理システムの低負荷化および高精度化——すなわち usability の向上をはかろうというものである。

(Resolution)	High	Middle	Low
(Image)			
(Applied Tasks)	Show Detail (Investigation by Human)	Line Detection	Shape Recognition, Color Categorization

[図1: 各解像度とそれに適した処理の関係]

以降、第2節では *mathematical morphology* 演算を用いた *scale space* の生成について述べている。本研究では *scale space* を生成する *Scale Space Filtering* の手法として *Mathematical Morphology* 演算に着目する。*Mathematical Morphology* の各種演算操作を組み合わせると1つの原画像から多次元かつ多段階の解像度を生成することができる。このことは、高次画像解析・高次特徴抽出を行うためのさまざまな要素的手法に対して、より柔軟な操作パラメータを設定することが可能であることを意味する。

第3節では第2節で提案した手法を龍谷大学が保有する古文書に対して適用し、行高行間計測処理を例として、より最適なパラメータの選択、選択された尺度における特徴抽出、そしてこれらのパラメータを含む画像アーカイブのための *scale space* データベースの設計を試みた。

第4節では本手法を適用した処理の精度評価を行い、選択された尺度において画像解析・特徴抽出処理を行うことにより計測精度が向上することを示す。

2. Mathematical MorphologyによるScale Space生成

2.1 Scale Space理論とMorphology演算

Scale Space 理論では通常、2次元Gaussian畳み込みを利用して画像解像度を変化させる [2] [3] [4]。この手法によれば、*scale* 値 σ をパラメータとする *scale space* を生成することができる。この *scale* による画像解像度の制御可能性を応用

し、処理の効率化・精度向上を目指す研究がなされており、本研究もこの流れの上に展開している。

Gaussian畳み込みによる演算は画像処理では2値画像を対象とする。このことはすなわち、処理対象や生成されたscale spaceの応用範囲が必然的に限定されるということである。原画像が有彩色である場合、2値化処理の精度(形状の保存、ノイズ除去)がGaussianの結果に大きく影響する。また色彩に関する処理を後続処理として直接適用できない。本研究では後述するように超高解像度のフルカラー画像を対象とするため、これらの画像の高次特徴抽出および解析には i) 解像度を十分低減できることおよび ii) カラー画像処理に直接対応できることが要求される。Gaussianの代わりにこれらの要求を満たす演算系として、本報告では *Mathematical Morphology* (以降 *morphology*) 演算に着目した。morphological systemをscale space理論に適用する例はすでに多く見られ、Gaussian scale spaceに対して *Morphological scale space* と呼ばれる [5]。本研究ではmorphology演算のパラメータと処理アルゴリズムを対応させ、さらに対象画像のメタデータとすることにより、コンテンツ管理の効率化を図る。

morphology演算は、少数の構成要素と少数の基本操作を組み合わせることにより、非常に幅広い変化を生成することができる [6]。まず、すべての操作の構成要素 G はある形状を持ち、この形状が操作結果に基本的な特徴を与える。これが第1のパラメータである。構成要素の典型的な形状は3種類挙げられる (*Ellipse (Circle)*, *Rectangle*, *Cross*)。

次にmorphology操作 O の種類である。これらはすべて *dilation* (拡張) と *erosion* (浸食) という2つの基本操作を組み合わせることによって実現される。代表的なものは、*opening*, *closing* である ([図2.1] 参照)。これらの操作の *dilation/erosion* による構成と、操作によって生じる画像の変化は概ね次の通りである [7]。またその他の応用的な操作も多くが *dilation/erosion* で置き換えることができる。

(I : Result Image, I : Source Image, G : Structuring Element)

- Dilation: 画像全体を太らせる。

$$I.Res = dilate(I.Src, G)$$

- Erosion: 画像全体を痩せさせる。

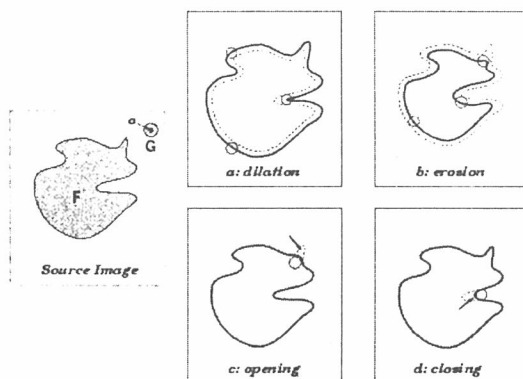
$$I.Res = erode(I.Src, G)$$

- Opening: 突出部分を削る

$$I.Res = open(I.Src, G) = dilate(erode(I.Src, G), G)$$

- Closing: へこんだ部分を埋める

$$I.Res = close(I.Src, G) = erode(dilate(I.Src, G), G)$$



[図2.1]: Geometric Operation in Morphology by Circle Element G.

これらの操作は回数に依存した不可逆的解像度低減変化を生じる。したがって操作の結果生成される低解像度画像の内容は構成要素の形状と操作の種類および回数の順列組合せとなり、多様な内容の低解像度画像を生成しうる。これらが残りのパラメータである。したがって、原画像 $I. Src$ に対する morphology 演算処理は以下のように表現することができる。

$$I. Res = Morph(I. Src) \{G. Form, O. Ope. 1, O. Num. 1, O. Ope. 2, \dots, O. Num. N\}$$

$I. Res$: 結果画像、 $I. Src$: 原画像、 $G. Form$: 構成要素の形状、 $O. Ope. N$: 操作の種類 (N番目)、 $O. Num. N$: $O. Ope. N$ の操作回数

このように morphology 演算は理論的には多次元特徴空間として無限の広さをもつ scale space を形成するが、本研究のように高次特徴解析への応用を目的とした場合、実用上有意な条件範囲を限定することができる。本報告ではエッジ抽出・色による分類・行特徴解析など、数種類の特徴抽出処理に対して比較的容易に判定可能な原画像を選び、それに morphology 操作の種類と回数を変化させる予備実験を行なった。その結果より、実用上多くの特徴解析処理に適用可能であると仮定的に判断し得る制約を以下のように与える。(当然全ての特徴抽出にこの仮定が有効である保証はない)

- 連続する単一操作は5回までとする。
- 論理的に入力画像と等価な出力になる処理は排除する。
- 1連の処理に利用される1種類の操作は3度までとする。

これらの制約により生成されるパラメータ空間の次元数はたかだか10次元であり、画像の種類はたかだか10の4乗オーダーである。これ以降の実験については上記制約を仮定としたパラメータ空間を用いる。

2.2 最適 scale space と scale space データベース

上記制約下における morphology 演算により生成される scale space を利用し、効率のよい処理を目指す。それには各々の処理に有効な scale 群もしくはパラメータ群を選別する必要がある。このため本報告では以下の手順を適用する。

1. あるパラメータによる down scale 画像すべてに処理を適用し、
2. その結果から評価上位となる scale を選びそのアルゴリズムにおける「最適

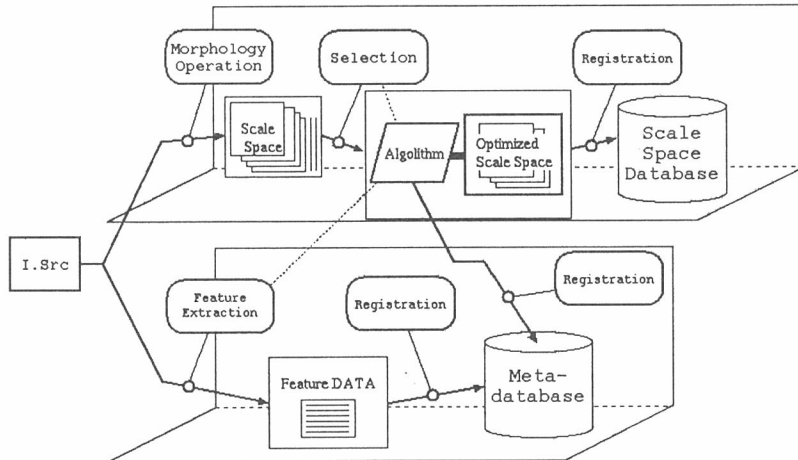
scale space) とする。

3. 最適scale spaceを利用して実際の特徴抽出処理を行なう。

こうして選択された最適scale spaceは処理アルゴリズムと対になるので、効率的な再利用のため、その処理アルゴリズムに関するメタデータベースとして構成する。これをここでは「Scale Spaceデータベース」と呼ぶ。

また、対象画像と処理目的によって利用する処理手法は当然異なる。処理目的が異なれば同じ対象画像に対しても異なったアルゴリズムが適用されるのは当然である。そこで、ある画像に適用されたアルゴリズムを、その画像に関するメタデータとして登録する。ここではこれを「画像メタデータベース」と呼ぶ。

これら2つのデータベースによって、特徴抽出処理および画像データのどちらからもアプローチ可能なコンテンツ管理システムを構成することができる。このシステムのメリットは他のメタデータと特徴抽出・解析処理およびScale Spaceパラメータ群を統合的に取り扱うことができる点である。以下にシステム概念図を示す。



[図2.2: コンテンツ管理システム概念図]
(本図はUser Interface部分を含んでいない)

3. 行特徴解析実験

3.1 行高行間計測

本節ではscale spaceを用いて、評価実験を行う。ここで使う処理は文書画像における行特徴解析の一種で、行の高さ(縦書きでは行の幅)・行間隔および行数を文書画像から計測・推定するものである[8]([図3.1参照])。ただし、本処理で算出されるのは実寸値であるが、複雑な手書き文字の行高・行間の厳密な計測は実際には困難であり、真値がわからないことが多いことを考慮すると、出力値は相対値として取り扱うのが正しいといえる。したがって本処理が適用されるのは行高・行間の近い紙片のグルーピングなど、絶対的実寸値が必ずしも必要とならない処理である。

実験の対象は龍谷大学所蔵の古文書をデジタイズ(PhotoCD形式)したもので、約3000枚の「大谷文書」データのうち、評価用サンプルとして300枚をピックアップした[9][10]。使用した画像データは画素数1536x1024、色数24bitである。

本処理のアルゴリズムは以下の手順で構成されている。

1. 画像からルーラを分離する。大谷文書の画像には実寸を目視するためのルーラが写されている。最初にこのルーラ部分と文書画像とを分離し、ルーラからは実寸とスケール比(実寸/ピクセル)を算出する
2. 対象画像から文字部分を抽出する。
3. 画像全体に対し文字と同色(または同輝度)のヒストグラムを算出する
4. ヒストグラムの集中している部分の幅と間隔を画素距離として計測する
5. ルーラより算出したスケール比により行高・行間を実寸値として算出する

これについてはすでに [8] において大谷文書紙片画像全データに直接適用した結果があり、本実験の結果との比較対象とする。

3.2 最適scale spaceの選択

次に最適scale spaceの選択を行なうが、制約条件により組合せの数が限られているために、特徴抽出処理の適用は一度行なえばよい。すなわち、

1. 上記手順 1. の直後にmorphology演算を行いscale spaceを生成し、
2. すべてのパラメータによる結果画像に上記アルゴリズムを適用し、
3. その結果を評価し最良のパラメータを選択する。

本実験では処理結果に対し、目視による「行数判定」を精度評価に用いた。大谷文書の紙片の状態を考慮した場合、行高・行間は絶対的な指標となりにくく、評価基準とはしなかった。これは紙片の断片化により行数のばらつきが大きいことや、行高・行間の平均値を手作業で測定するのが困難なためである。行数判定の基準は以下の通りとした。

- 目視で濃色の文字行が認識できなければ「解なし」
- 行数が真値と同値または±1の誤差であれば「正解」

行数に誤差を含める理由は、古文書の紙片が断片化しており、周端部に行がかかっていたり消えかけているものも多く存在し、その場合、行として認識しうるか否かは肉眼でも判定困難だからである。

また、評価としてprecision(正解率)とrecall(適合率)を算出するが、これは次の方式に従った。(解なし入り正解:「解なし」データを正解に含んで算出)

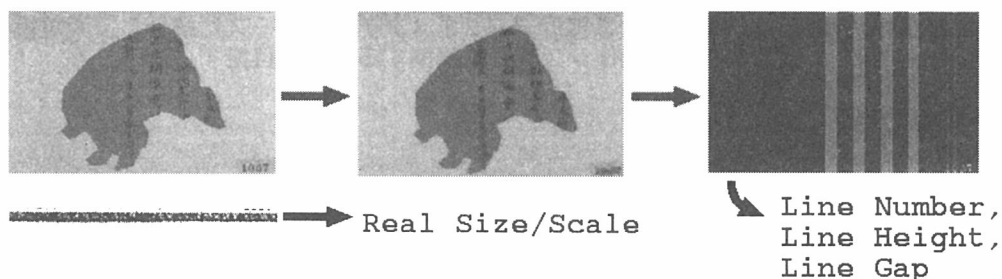
- Precision = 解なし入り正解/データ数
- Recall = 解ありのみ正解/解あり数

評価の結果、正解率の上位組合せは以下の通りであった。4位以下は正解率が低かった(50%以下)ので棄却した。

[表3.1: 最適パラメータの選択]

	条件名	構成要素	操作・回数	Precision
1	R. D2E2*3	Rectangle	(Dilate 2回 + Erode 2回) * 3回	72.0%
2	R. 05	Rectangle	Opening 5回	65.0%
3	R. E5	Rectangle	Erode 5回	58.0%

条件名: 構成要素名、操作名、操作回数により命名



[図3.1: 最適scale spaceを適用した行特徴解析実験]

4. 結果とまとめ

上記実験の結果、morphology処理をせず原画像より直接行抽出を行なった結果と比較すると、[表4.1]のようになる。(NoOpe:Morphologyなしの直接処理, ↑はNoRepとの差)

[表4.1: 行数の精度評価比較(単位:%)]

処理名	NoOpe	R. D2E2*3	R. O5	R. E5
Precision	52.0	72.0 (↑20.0)	65.0 (↑13.0)	58.0 (↑6.0)
Recall	30.4	56.5 (↑26.1)	49.3 (↑18.9)	39.1 (↑8.7)

[表4.1]に見るように絶対的な精度としてはまだ低いが、従来法と比較すると20%以上の精度向上となっている。なお、処理速度についてはここにはデータを掲載していないが、morphologyにかかる時間を別にすれば、本実験で採用した処理はピクセル単位での処理のためにほとんど変化はない。(心理解像度は低減するが、実際の画素数が低減するわけではない)以上より、morphology演算による最適scale spaceの有効性を検証することができた。

今回の実験では行数判定を目視により行なったが、真値の取得も含めたパラメータの評価基準を確立する必要がある。ただし、特徴抽出処理は2次元画像を対象とする基本的なものだけでも数十種類は存在するし、日々新たに研究開発されている。これらについて最適scale spaceをいかに効率良く選択・管理するかということは、システム全体の有効性にも関わる重大な課題である。

なお、取得された最適scale spaceはscale spaceデータベースとして構成してある。ただし登録されるアルゴリズムと最適scale spaceが多様になれば、それに対応した柔軟な構造の変更が要求される。この点を考慮にいれながら、現在XMLベースでの再構成を検討中である。

一方画像データはメタデータと共に、大規模なマルチメディアデータ向けのデータベースシステム“Cumulus”(Canto社[11])によって統合され、サーバ・クライアント方式で利用可能である。ただし、本報告ではこれらのデータを十分に活用しうる操作環境としてのユーザインタフェースについて、十分な検討が進んでいない。これについては実用上の評価を含めた形で検討する必要がある。

今後はこれらの課題を考慮しながら、本手法をさまざまな古典籍の解析処理に適用し、特徴抽出処理中心のscale spaceデータベースおよび画像中心のメタデータベースの両者によるコンテンツ管理・操作システムの構築を進める予定である。さら

には化学分析など他分野における解析データをメタデータベースに統合し、大規模なデータ群をよりスムーズに操作システムを構築することを目指す。

参考文献

- [1] 池田光男: '視覚の心理物理学', 森北出版.
- [2] Witkin, A. P.: 'Scale space filtering', Proc. of 8th International Joint Conference on Artificial Intelligence, pp.1019-1022, Karlsruhe, 1983.
- [3] Lindeberg, T.: 'Scale-Space Theory In Computer Vision', Monograph 1994, 1994.
- [4] Lindeberg, T.: 'Scale-space theory: A basic tool for analysing structures at different scales', J. of Applied Statistics, vol.21 (2), pp.224--270, 1994.
- [5] (Web Page), 'Morphology Digest Page', URL: <http://ams.jrc.it/mdigest/>
- [6] Giardina, C. R. and Dougherty E. R.: 'Morphological Methods in Image and Signal Processing', Printice Hall, 1988.
- [7] (Manual), 'OpenCV Reference Manual', Intel (R) Open Source Computer Vision Library, URL: <http://www.intel.com/research/mrl/research/opencv/index.htm>, 2003.
- [8] 藤井大輔, 岡田至弘: 'マルチエージェントによるメタデータ管理と画像アーカイブへの適用',
- [9] 小田義久, 大谷文書目録 No. 1001~NO. 3000, 龍谷大学仏教文化研究所, 1986.
- [10] 小田義久, 大谷文書目録 No. 3001~NO. 4500, 龍谷大学仏教文化研究所, 1998.
- [11] (Web Page), Canto's Web Page, URL: <http://canto.com/>