

卒業論文

マルチスケール解析による  
飾り文字認識に関する研究

東北大学工学部通信工学科阿曾研究室

井上政樹

# 目次

<b>第1章 序論</b>	<b>1</b>
1.1 本研究の背景と目的	1
1.2 本論文の構成	2
<b>第2章 スケールスペース解析</b>	<b>3</b>
2.1 はじめに	3
2.2 スケールスペース	3
2.3 主曲率による曲面形状の表現	6
2.4 スケールの変化と濃淡曲面の変化	7
2.5 特徴点の抽出	8
<b>第3章 飾り文字の構造抽出</b>	<b>9</b>
3.1 はじめに	9
3.2 飾り文字の構造抽出	9
3.2.1 飾り文字認識に必要な構造	9
3.2.2 大局構造の抽出	10
3.3 実験	13
3.3.1 実験方法	13
3.3.2 実験結果・考察	13
3.4 問題点	17
<b>第4章 再帰的な谷点抽出による構造補完</b>	<b>18</b>
4.1 はじめに	18
4.2 構造の補完	18
4.2.1 ぼかしによる構造の補完	18
4.2.2 谷点による構造の補完	19

---

4.2.3	補完を用いた飾り文字の構造抽出 . . . . .	21
4.3	飾り文字の構造抽出の実験 . . . . .	22
4.3.1	実験方法 . . . . .	22
4.3.2	考察 . . . . .	22
4.4	飾り文字の認識実験 . . . . .	26
4.4.1	実験方法 . . . . .	26
4.4.2	考察 . . . . .	26
<b>第5章</b>	<b>結論</b>	<b>30</b>
5.1	まとめ . . . . .	30
5.2	今後の課題 . . . . .	30
	<b>謝辞</b>	<b>32</b>
	<b>参考文献</b>	<b>33</b>

## 目 次

2.1	スケールの変化による画像の変化 . . . . .	5
2.2	主曲率平面 . . . . .	6
2.3	スケールの変化による主曲率の変化 . . . . .	7
2.4	スケールスペース解析による特徴の抽出 . . . . .	8
3.1	ぼかしによる飾り文字「I」の濃淡曲面の変化 . . . . .	11
3.2	飾り文字「I」のスケールの変化による主曲率の変化 . . . . .	12
3.3	飾り文字「I」の大局構造 . . . . .	12
3.4	実験対象とした飾り文字の一覧 . . . . .	14
3.5	大局構造の抽出 (1) . . . . .	15
3.6	大局構造の抽出 (2) . . . . .	16
4.1	ぼかしによる構造の補完 . . . . .	19
4.2	谷点抽出の例 . . . . .	20
4.3	再帰的な谷点の抽出と構造の補完 . . . . .	20
4.4	飾り文字の構造抽出の流れ . . . . .	21
4.5	飾り文字の構造抽出結果 (1) . . . . .	23
4.6	飾り文字の構造抽出結果 (2) . . . . .	24
4.7	飾り文字の構造抽出結果 (3) . . . . .	25
4.8	認識実験結果 (1) . . . . .	27
4.9	認識実験結果 (2) . . . . .	27
4.10	認識実験結果 (3) . . . . .	27
4.11	実験対象とした飾り文字の大局構造 . . . . .	28
4.12	実験対象とした飾り文字の補完後の構造 . . . . .	29

# 第 1 章

## 序論

### 1.1 本研究の背景と目的

近年、計算機の性能は飛躍的に向上し、扱う情報量も膨大なものとなっている。このような高度情報化社会において、計算機が我々の生活の中で性能を発揮し共存する上で重要な課題の一つがマン・マシンインターフェースのさらなる発展である。その中でも文字認識は重要な位置をしめる。我々が情報交換に用いている文字を機械によって識別させ、コード情報に変換し計算機の入力情報として処理させる。この文字認識の手法により、多量の文書の保存、検索が計算機によって容易に行なうことが可能となる。

しかし、我々が普段目にする文書体は一様なものではない。大きく分けて手書き文字、印刷文字の2つにわかれ、後者の印刷文字にはゴシック体や明朝体といったフォントが多数存在する。中でも新聞・雑誌等の見出しに用いられる飾り文字は読者の目を引き付ける、レイアウトのデザイン性を高める等の特殊な目的を持つため、既成のフォントにとらわれない様々な形状を形作っている。それら飾り文字の中には文字をスリット状にする、射影のみで文字を表現する等、人間の認知特性を用いた書体が用いられることがある。

人間が飾り文字を読みとる際には、視覚の誘導場等の視覚的効果やあらかじめ自身が持っている文字についての情報を駆使し、その飾り文字がどの文字を表しているかを認識する。対して文字認識における構造解析法等の手法では、多くの場合文字を黒画素の連結性に着目することにより抽出・認識を行なっている。しかし飾り文字は本文書体とは異なり、黒画素の連結性に着目しただけではその文字の形状を表すとは限らない。そのため従来の文字認識手法を直接飾り文字に用いただけでは、正確に認識されることは難しい。また飾り文字用の辞書を作り対応しようとしても、見出し等の飾り文字にはその場でしか使用され

ない、特別に製作された飾り文字が使われることが多く、全ての飾り文字の対応できる辞書を作るのは困難であると言える。このため飾り文字を認識するためには、人間が視覚的効果によってその構造を得るように、飾り文字からその飾り文字が表す文字の構造を抽出することが前提になると考える。

一般に飾り文字は元の書体のある程度変形させ、飾りを加えたり削ったりといった処理を与えたものが大半を占める。飾り文字を認識するためには加えられた飾りを消去し、削られた部分を補うといった過程を必要とする。そこで画像にぼかしを与え、大局的な構造を得ることによりこれらの問題を解決する。

しかし、画像を一様にぼかしただけではさまざまな形状をした飾り文字に対応させることは困難である。そこで本研究ではスケールスペース解析を用いる。Lindeberg [1] は各位置ごとに適切なぼかしのスケールを与える手法を提案し、エッジ抽出等に応用した。また本谷ら [2][3] は画像の濃淡曲面から曲面の主曲率に着目することにより、ぼかしによる濃淡曲面の形状の変化を観測し、画像の特徴的な構造の抽出を行なった。

本研究では曲面の主曲率に着目したスケールスペース解析を応用し、様々な飾り文字の文字構造を表す画像を得ることを目的とする。

## 1.2 本論文の構成

本論文の構成は以下の通りである。

第1章 序論であり、本研究の背景と目的を述べた。

第2章 スケールスペース解析について述べる。

第3章 飾り文字の構造抽出法を提案する。また、提案手法を評価するため実験を行ない、問題点を提起する。

第4章 谷点を利用した構造の補完について述べる。また、改善後の提案手法を評価するための実験として、飾り文字の構造抽出と認識実験を行なう。

第5章 結論であり、まとめと今後の課題について述べる。

## 第2章

# スケールスペース解析

### 2.1 はじめに

本章では主曲率に着目したスケールスペース解析 [1][2][3] について説明する。

スケールスペース解析は画像処理において画像の濃淡特徴を記述する手法である。ガウスイフィルタ等によって画像にぼかしを与えていくと、スケールの変化によって濃淡曲面が変化する。変化の仕方は一様ではなく、構造の大きさを反映して変化する。そこでこの手法ではスケールの変化に対する濃淡曲面の主曲率を用い、曲面の形状の変化を追うことによって、尾根状の曲面等の特徴的な構造を抽出する。各スケールにおいて主曲率の変化を用い適切な特徴を抽出し、最後に全スケールの特徴を重ね合わせることによって、各位置ごとに適切なスケールでの特徴抽出を可能とする。

### 2.2 スケールスペース

画像  $f(x, y)$  をガウス関数でぼかすことを考える。このときスケールスペース  $L(x, y; t)$  とは、ぼかしのスケール  $t$  を 0 から次第に大きくするときに得られる画像のスタックであり、以下のように定義される。

$$\begin{aligned} L(x, y; t) &= g(x, y; t) * f(x, y) \\ &= \int g(x', y'; t) f(x - x', y - y') dx' dy' \end{aligned}$$

ただし  $g(x, y; t)$  は分散  $t$  のガウス関数である。この  $t$  をスケールパラメータと呼ぶ。

しかし実際には画像は離散であるから、以下に示すフィルタを  $t/\Delta t$  回繰り返して作用さ

せることにより、ぼかしを与える。 $\Delta t$ は微量なぼかし量であり、 $\Delta t < 0.5$ である。

$$\begin{pmatrix} \frac{1}{12}\Delta t & \frac{1}{3}\Delta t & \frac{1}{12}\Delta t \\ \frac{1}{3}\Delta t & 1 - \frac{5}{3}\Delta t & \frac{1}{3}\Delta t \\ \frac{1}{12}\Delta t & \frac{1}{3}\Delta t & \frac{1}{12}\Delta t \end{pmatrix}$$

図 2.1 にスケールパラメータ  $t$  による画像の変化の例を示す。左が画像の変化、右が濃淡曲面の変化である。スケールの値が小さいときには放射状の線分に対応する位置に尾根状の構造が見られるが、スケールの値が大きくなるにつれ円環状の尾根に形状を変化させていく様子が見られる。



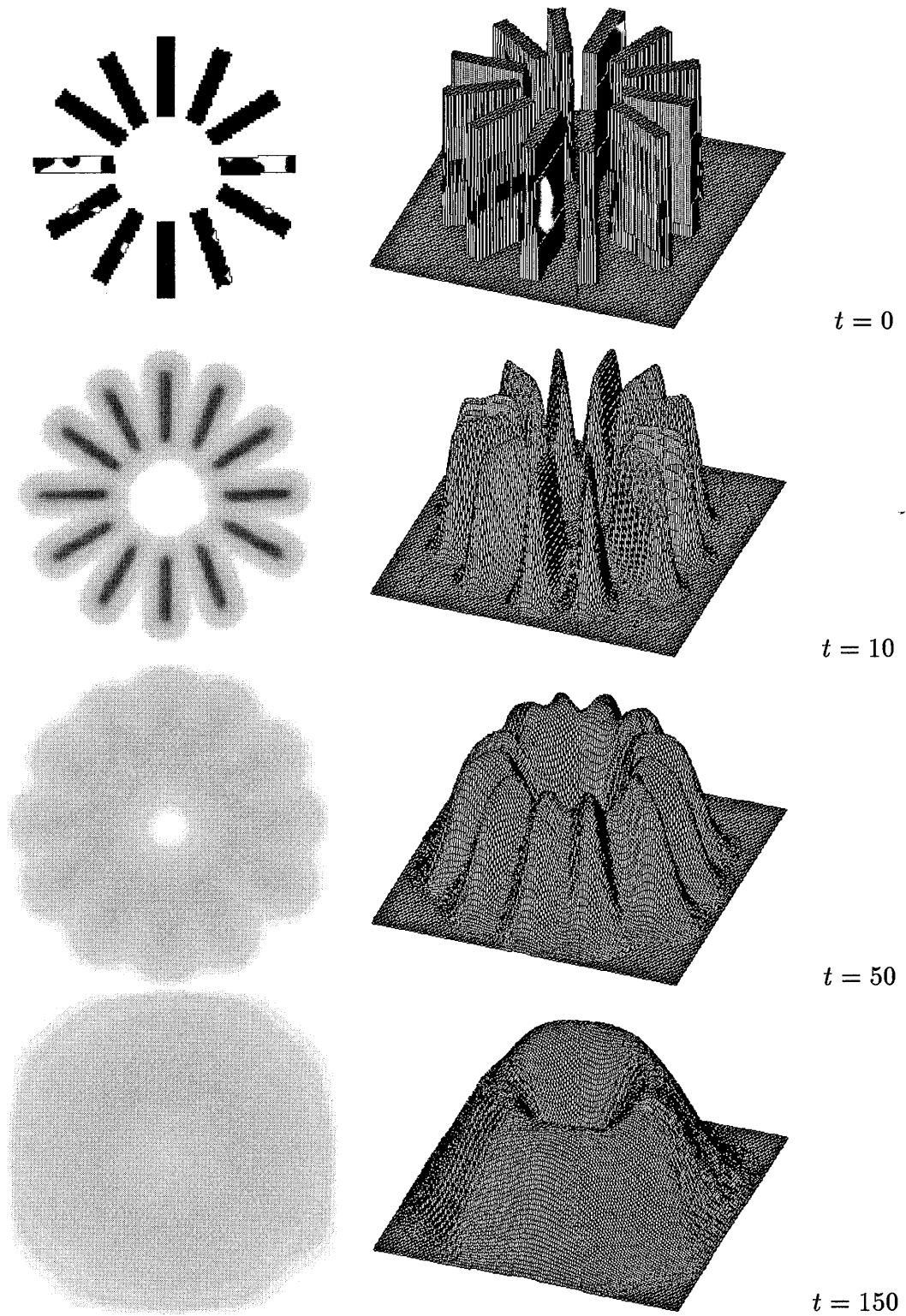


図 2.1: スケールの変化による画像および濃淡曲面の変化

### 2.3 主曲率による曲面形状の表現

スケールスペース解析では主曲率の変化を用いて、特徴を抽出する。各位置において求まる二つの主曲率の比は、その位置における曲面の形状を表す。スケールスペース解析ではこの曲面の主曲率および主曲率方向を用いて曲面の変化を観測し、画像の特徴を抽出する。

主曲率および主曲率方向は第二基本形式の固有値および固有ベクトルによって求める事ができる。 $\partial_x L(x, y; t)$  を  $L_x$  などと表すと、主曲率  $K_1, K_2$ 、主曲率方向が画像  $x$  軸となす角  $\theta_1, \theta_2$  は以下の式で与えられる。

$$\begin{cases} K_1 = \frac{(L_{xx}+L_{yy})+\sqrt{(L_{xx}-L_{yy})^2+4L_{xy}^2}}{2\sqrt{1+L_x^2+L_y^2}} \\ K_2 = \frac{(L_{xx}+L_{yy})-\sqrt{(L_{xx}-L_{yy})^2+4L_{xy}^2}}{2\sqrt{1+L_x^2+L_y^2}} \\ \theta_1 = \arctan\left(\frac{(L_{yy}-L_{xx})+\sqrt{(L_{xx}-L_{yy})^2+4L_{xy}^2}}{L_{xy}}\right) \\ \theta_2 = \arctan\left(\frac{(L_{yy}-L_{xx})-\sqrt{(L_{xx}-L_{yy})^2+4L_{xy}^2}}{L_{xy}}\right) \end{cases}$$

主曲率  $K_1, K_2$  を直交する二軸にとった平面を主曲率平面と呼び、この平面のどの象現に求めた主曲率が位置するかによって曲面の形状を知ることが出来る。図 2.2 に主曲率平面と曲面の形状との関係を示す。

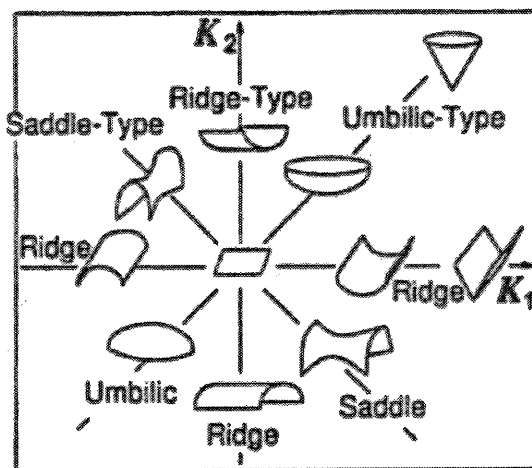


図 2.2: 主曲率による曲面の記述

主曲率平面内の位置を極座標で表現すると、動径成分  $R = \sqrt{K_1^2 + K_2^2}$  は曲がり具合を、角度成分  $\alpha = \arctan(\frac{K_2}{K_1})$  は曲面形状を表している。例えば主曲率が  $K_1 = K_2 = 0$  ならば

主曲率平面では原点に対応し、その点は平面である。同様に  $K_1 < 0, K_2 = 0$  となれば尾根点、 $K_1 > 0, K_2 < 0$  ならば鞍点、 $K_1 > 0, K_2 > 0$  ならば臍点に対応する、また  $R$  の値が大きいほど曲がり具合が鋭くなり、濃淡曲面状でははっきりとした特徴を示すことになる。

## 2.4 スケールの変化と濃淡曲面の変化

2.2節ではスケールの変化による画像の濃淡曲面の変化を図2.1を用いて述べた。本節では主曲率平面を用い、具体的に濃淡曲面がどのように変化するか、また特徴抽出に必要なスケールをどのようにとるか、といったことについて説明する。

図2.3の右のグラフは左の元画像「H」の中心の点における、スケールの変化による主曲率の変化である。スケールを  $t = 0$  から徐々にぼかしていく過程で主曲率が変化し、尾根点、鞍点、臍点と曲面の形状が変化する様子を主曲率平面によって知ることができる。

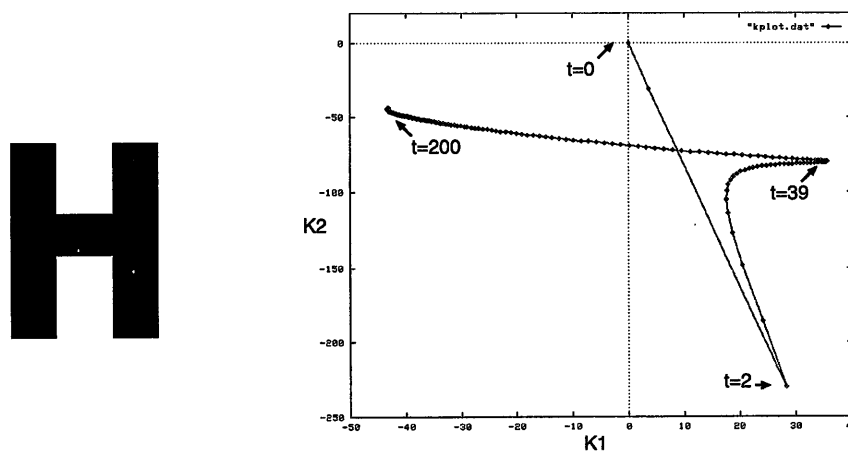


図 2.3: 左：元画像 右：スケールの変化による主曲率の変化

ここで主曲率の軌跡から原点までの距離  $R$  に着目する。主曲率平面において動径成分  $R$  は曲がり具合を表し、値が大きいほど曲がり具合が鋭くなる。つまり同じ角度成分  $\alpha$  に対して  $R$  の値が大きいほど、その曲面の形状をはっきり表していることになる。そこで特徴抽出に適したスケールとして、 $R$  が極大点となるスケールをとる。図2.3の右のグラフでは2つの極大点が存在し、 $t = 2$  では局所的な尾根点に、 $t = 39$  では大局的な鞍点にそれぞれ対応する。これよりこの位置では尾根点と鞍点の2つの特徴が存在することがわかる。

## 2.5 特徴点の抽出

濃淡曲面の各位置における主曲率を用い、画像の特徴点を抽出する。対象となる画像が文字であることを考慮し、ここでは画像の尾根点にあたる構造を抽出することを考える。

各点において図 2.3 のようにスケールの変化による主曲率の変化をプロットする。このうち  $R$  が極大点をとる点が画像の特徴を表すが、尾根点を対象としていることを考え、 $R$  が尾根点に相当する極大点を持つ位置を特徴として抽出する。実際の抽出では、尾根点だけでなく鞍点や頂点も考慮に入れるため次のような条件を定める。

主曲率  $K_1, K_2$  の絶対値を比べ、値の大きい方を  $K$  とする。主曲率  $K$  に対応する主曲率方向を  $p$  とし、 $\partial_p L$  を  $L_p$  などと表すと、抽出の条件は以下で示される。

$$\begin{cases} L_p = 0 \\ K < 0 \end{cases}$$

上記条件によって、主曲率方向の一方が濃淡の勾配方向と一致する尾根点や鞍点が抽出される。各スケールにおいて上記条件を満たす点を抽出し、全特徴点を重ね合わせることで、各位置ごとに適切なスケールで構造をとることができる。

図 2.4 は飾り文字「I」についてスケールスペース解析を行なった画像である。左の元画像から局所構造、大局構造を抽出できたことがわかる。

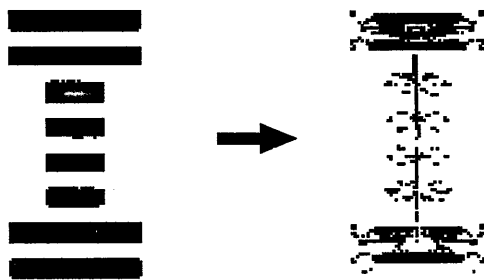


図 2.4: スケールスペース解析による特徴の抽出

## 第 3 章

# 飾り文字の構造抽出

### 3.1 はじめに

前章ではスケールスペース解析によって画像の局所的、大局的構造を抽出できることが確かめられた。しかしこれを直接飾り文字に用いれば、飾り部分の局所構造まで抽出されてしまい、飾り文字が表す文字の構造を得ることは難しい。

本章ではこれを飾り文字の構造解析という本研究の目的のためにどう応用するかを提案する。また提案手法を用いて飾り文字の大局構造の尾根点を抽出する実験を行ない、そこで生じた問題点を解決するために提案手法を改善する。

### 3.2 飾り文字の構造抽出

#### 3.2.1 飾り文字認識に必要な構造

局所的、大局的な構造を全て抽出するスケールスペース解析を飾り文字に適応させた場合、大局的な構造を抽出することによって飾り文字の削れた部分を補うことはできるが、局所的な構造も抽出してしまうため飾りを加えた部分の構造も抽出してしまう。

そこで飾り文字の構造解析においては、大きなスケールでぼかすことによって局所構造を消去し、大局構造を表す尾根点を抽出することが重要であると考えられる。

### 3.2.2 大局構造の抽出

スケールスペース解析ではスケールの変化による主曲率の変化を主曲率平面にプロットし、 $R$ が極大となる尾根点を抽出した。しかし、 $R$ が極大となる時点でのスケールは考慮していない。そこで本手法では尾根点を抽出する際の  $R$  に適当なしきい値をもうけ、一定スケール以上で抽出された尾根点を特徴点として飾り文字の構造とする。以下、図を用い具体的に説明する。

図 3.1 はスケールに対する飾り文字「I」の濃淡曲面の変化を表す。また、図 3.2 のグラフは飾り文字「I」の画像において  $(x, y) = (59, 69)$  の点でのスケールの変化による主曲率の変化である。図 3.1 において  $(x, y) = (59, 69)$  に相当する点に着目すれば、スケールが大きくなるにつれて曲面の形状が平面、尾根点、鞍点と変化する様子を知ることができ、図 3.2 では形状の変化に対応して主曲率が変化していく様子を見ることができる。ここで図 3.2 において 2 つの極大点が現われているが、両者の表す構造の性質は全く異なっている。 $t = 1$  で現われる極大点は  $x$  軸方向の尾根点を示しており、これは局所構造を表す。また、 $t = 32$  で現われる極大点は  $y$  軸方向の尾根点を示しており、これは大局構造を表す。もし、ここで大局構造を得るためのスケールのしきい値を  $t = 30$  と決めたなら  $(x, y) = (59, 69)$  は  $t = 32$  で極大点を持つため、大局構造を表す尾根点になりうる。このようにして各点において特徴点を抽出する際にしきい値をもうけることにより、飾り文字の大局構造を得る。図 3.3 に上記手法によって抽出した飾り文字「I」の大局構造を示す。スケールのしきい値は  $t = 30$  である。

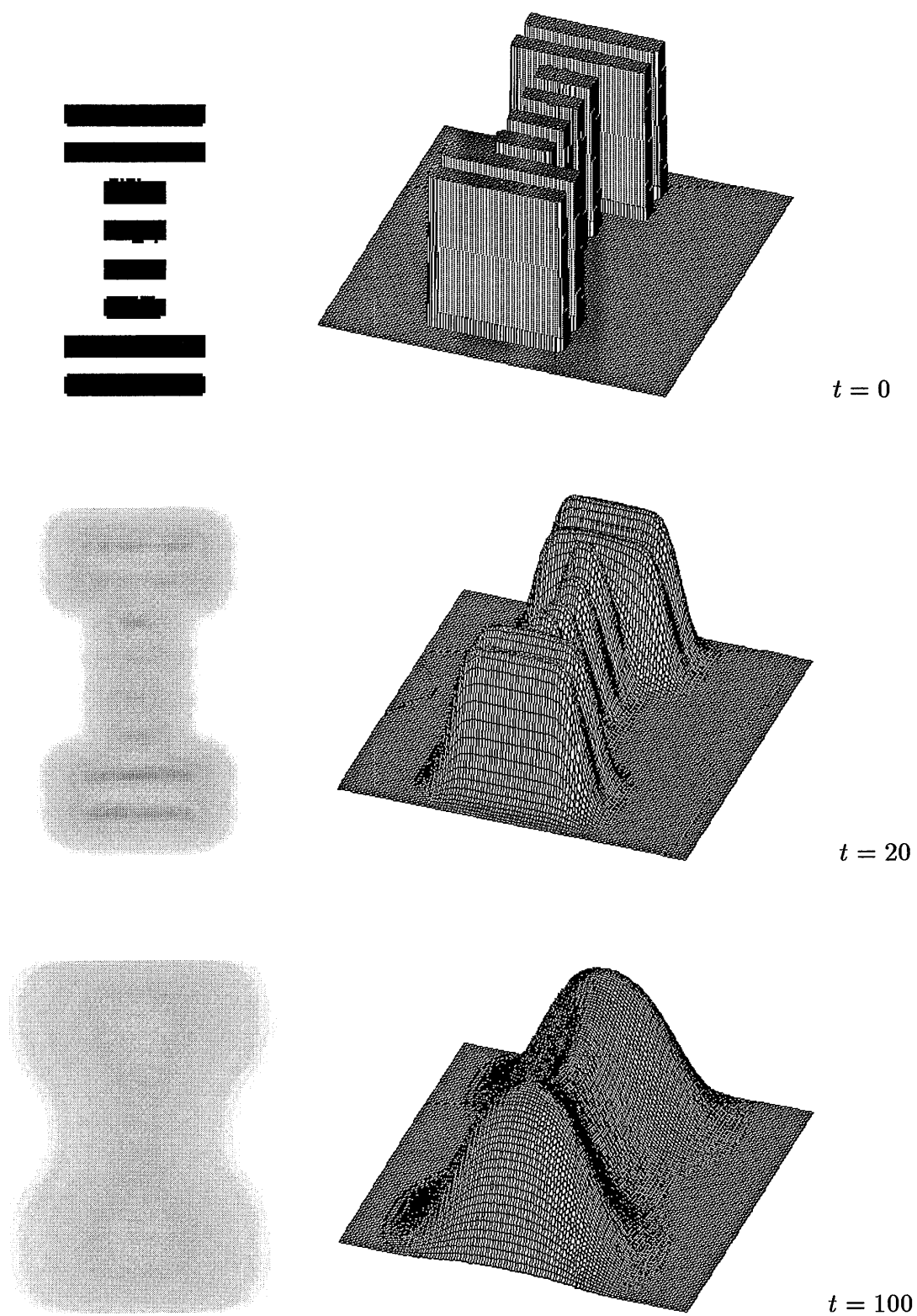


図 3.1: ほかしによる飾り文字「I」の濃淡曲面の変化

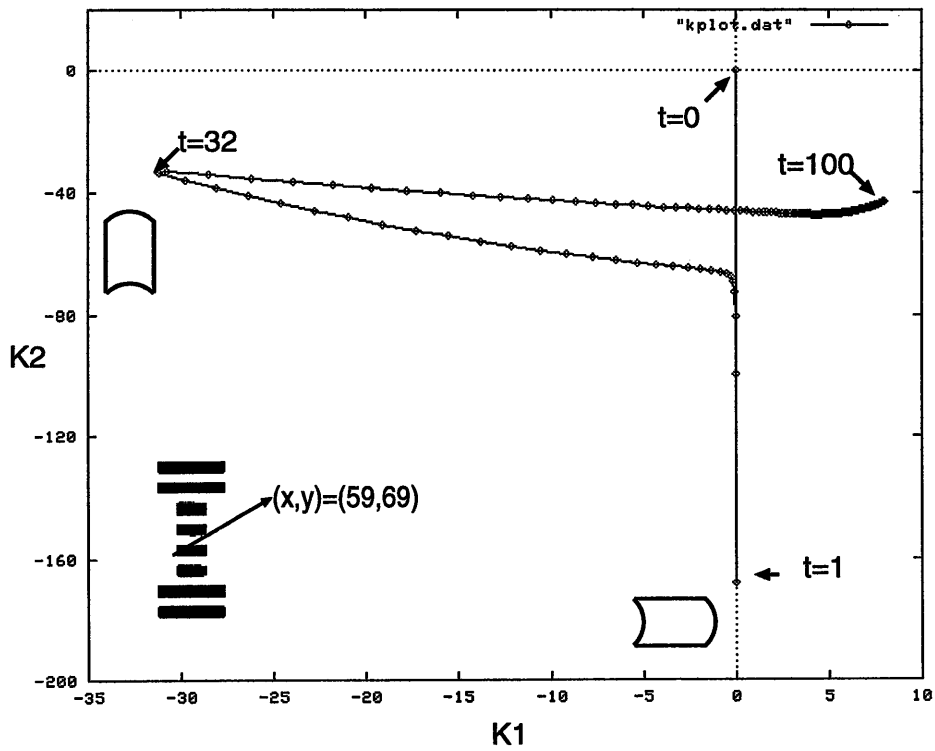


図 3.2:  $(x,y)=(59,69)$  での主曲率の変化 ( $0 \leq t \leq 100$ )

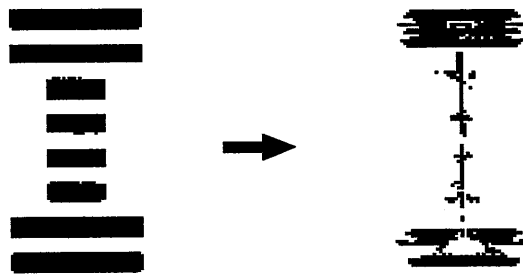


図 3.3: 飾り文字「I」の大局構造の抽出 左：元画像 右：抽出後の尾根点



### 3.3 実験

提案した大局構造の抽出法に対して実験を行う。対象とした飾り文字 26 種を図 3.4 に示す。画像のサイズは各種 128 ピクセル四方である。なお、「IBM」「NEW」においては今後特に断りのない限り一文字ずつ切り出して用いたものとする。

#### 3.3.1 実験方法

対象とする画像をスケール  $t = 0$  から  $t = 100$  までぼかしていき、大局構造となる尾根点を抽出した。なお、しきい値は予備実験の結果  $t = 30$  と設定した。

#### 3.3.2 実験結果・考察

図 3.5 に「IBM」の文字列の画像を用いた結果を示す。(b) は本谷らの手法 [2][3] を用い、局所構造から大局構造まで特徴点となりうる全ての尾根点を抽出したものである。また、(c) は提案手法によって尾根点を抽出したものであり、局所構造が消去され、大局構造のみが抽出されていることがわかる。

図 3.6 では様々な飾り文字「A」において大局構造を表す尾根点を抽出した。元画像の形状によって大きく成果が左右されており、大局構造は抽出できたが「A」の構造が抽出できたとはいえない、という結果になっている。

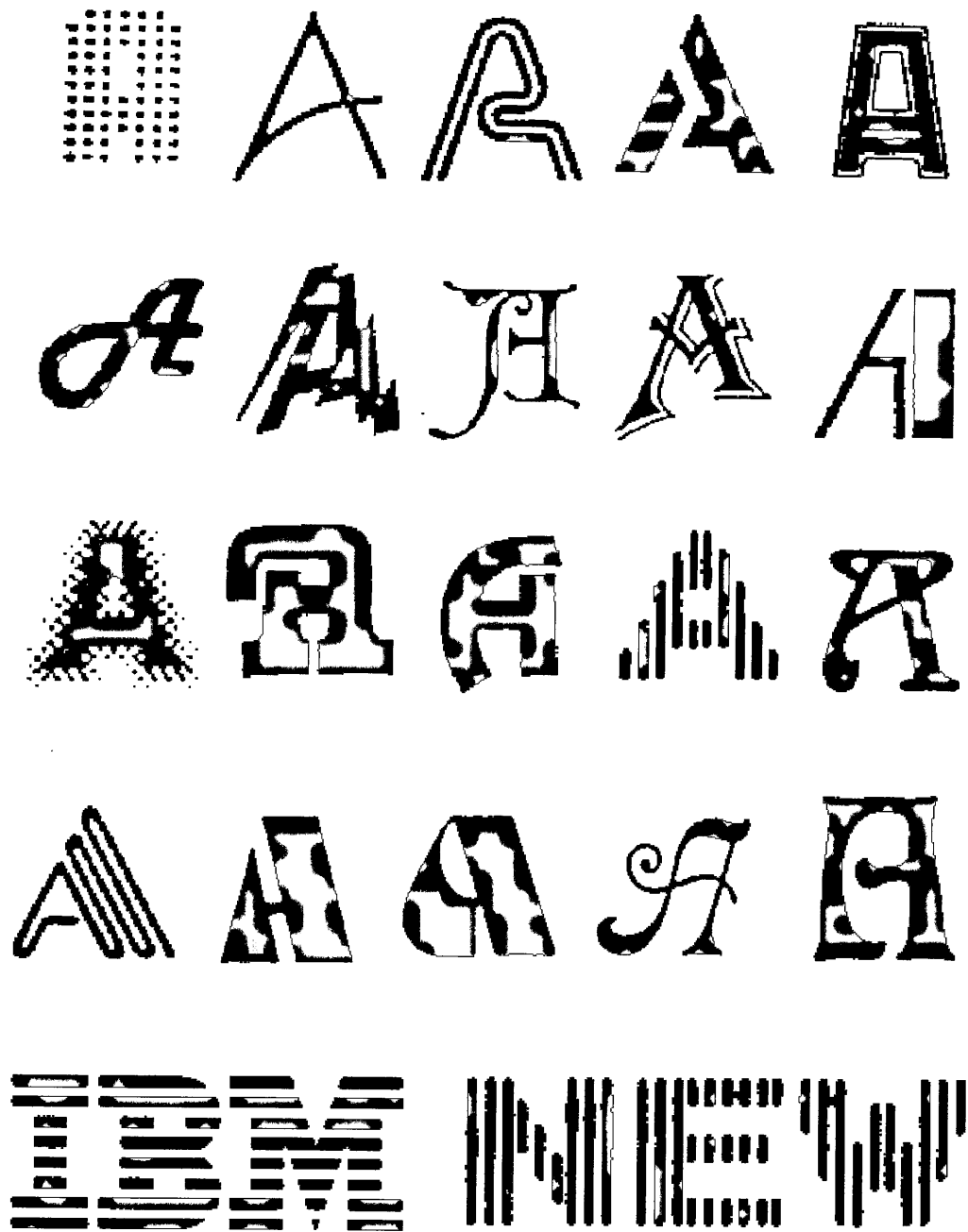
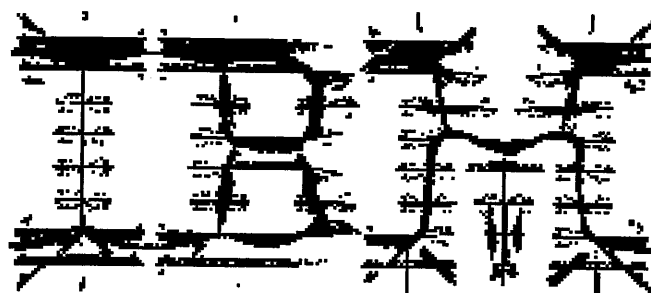


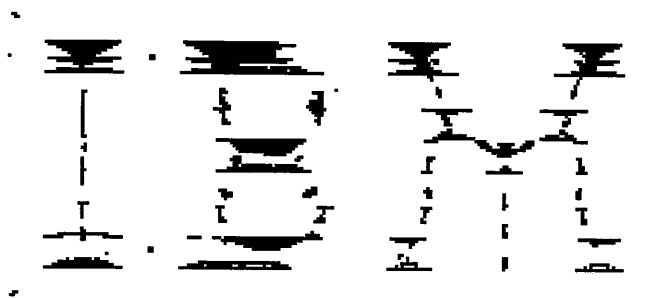
図 3.4: 実験対象とした飾り文字



(a) 元画像



(b) 本谷らの手法



(c) 本手法

図 3.5: 大局構造の抽出 (1)

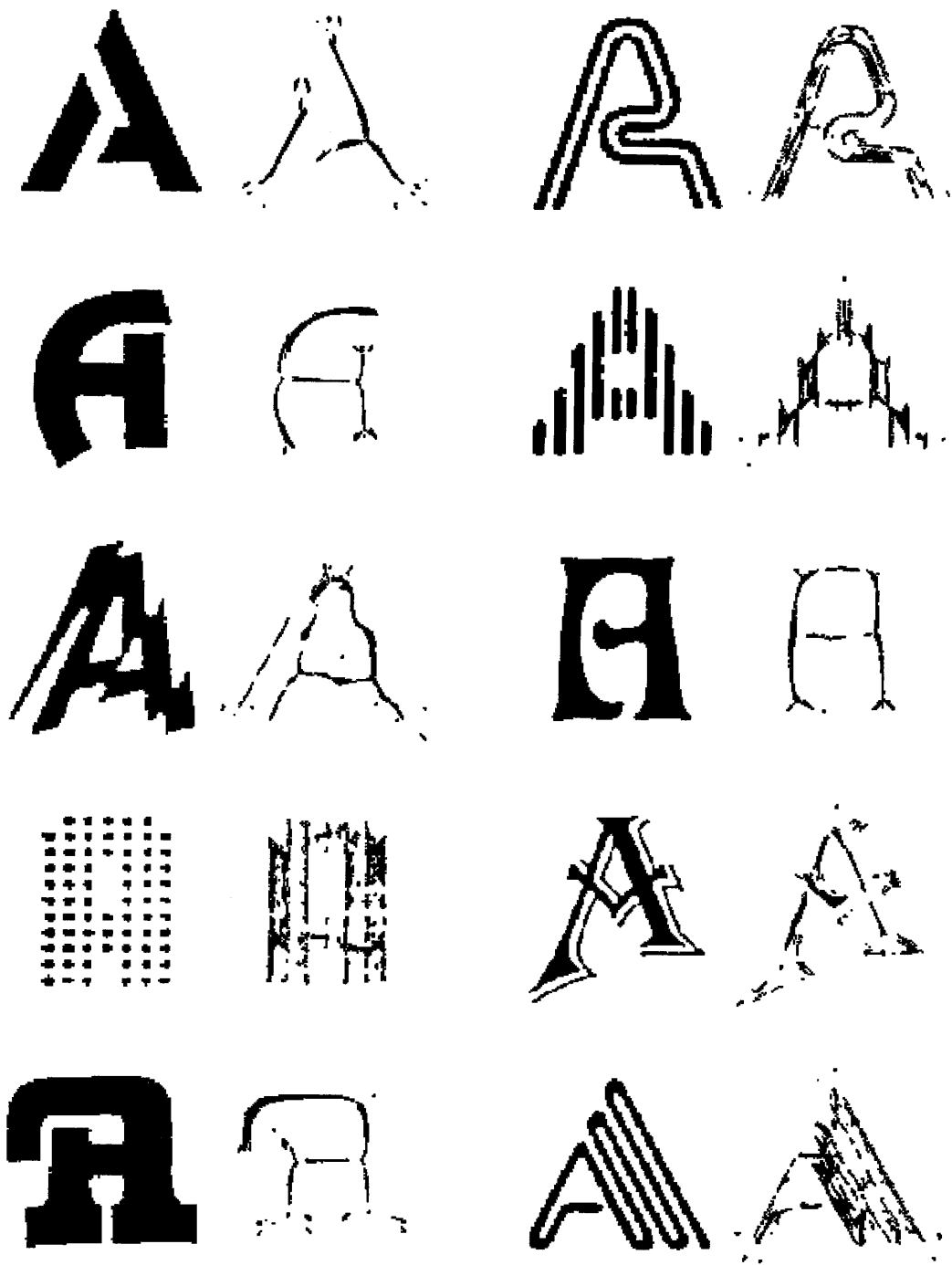


図 3.6: 大局構造の抽出 (2) 左: 元画像 右: 抽出後の尾根点

### 3.4 問題点

実験によって画像の大局構造を抽出できることを確認した。大局構造を抽出後の画像を見る限り、線分が連続した構造を抽出することができておらず、切れ切れになっているのが確認できる。飾り文字の認識を可能にするには、文字の持つ文字情報としての特徴を表現する構造を得ることが重要であると考えられるため、線分の不連続な構造では不十分であると考えられる。

このような結果となった原因として、全ての画像において  $R$  に一定のしきい値を設けて尾根点を抽出したため、様々な形状に対応しきれなかったことが考えられる。しかし、飾り文字の形状、線分の太さ、画像の大きさ等の要因によって大局構造が抽出されるスケールが異なるため、一文字ごとに適当なしきい値を与えることは難しく、また連続した構造が得られるしきい値を設定したとしても消去せねばならない局所構造までも抽出されてしまう可能性が高くなる。このことより、この手法で尾根状の点に着目しただけでは連続した構造を抽出することは困難であると考えられる。

そこで次章ではしきい値を変動させることはせずに、不連続な構造の線分間を補完することによって、連続した文字構造を得る手法を提案する。

## 第4章

# 再帰的な谷点抽出による構造補完

### 4.1 はじめに

スケールスペース解析の応用から飾り文字の大局構造を得られることが、実験によって確かめられた。しかし尾根点だけに着目しては線分が連続した構造を得ることは難しい。文字情報としての構造を抽出し、飾り文字の認識を可能にするためには前章の実験で得られた構造では不十分であると考えられる。求める構造を得るためには尾根状の点のみならず、他の形状の特徴点にも着目する必要があると推測できる。

本章ではまず画像をぼかし、二値化することで不連続な線分の埋めるべき箇所を補完することを提案し、それによって生じる問題点を提起する。そこで画像から抽出した谷点を利用することによって問題点を解決し、不連続な構造を補完する手法を提案する。また、実験として提案手法を用いた様々な飾り文字の構造抽出と認識実験を行なう。

### 4.2 構造の補完

#### 4.2.1 ぼかしによる構造の補完

画像を適当なスケールでぼかし、二値化することで構造を補完することができる。図 4.1 にぼかしによる構造の補完の例を示す。左が元画像であり、元画像を適当なスケールでぼかしたのが中の画像である。これを二値化することによって右のように線分間を埋めた画像を得ることができる。

図 4.1 上段の文字列「IBM」では、ぼかしによる補完によって文字の構造が得られている。しかし、ぼかしのスケールと二値化のしきい値は人間が調整し、設定してやらねばな

らない。下段の例では線分間が広いため、補完させるためにはぼかしのスケールを大きくせねばならず、結果として期待される補完後の画像とは形状が大きく異なってしまう。

このような形状の大きな変化が生じてしまうため、ぼかしによる補完では埋めてはいけない文字の特徴まで埋めてしまう可能性がある。そこで補完が必要な線分間の谷点を再帰的に抽出し、線分間を狭めることにより、ぼかしのスケールを最小限に抑える方法を提案する。

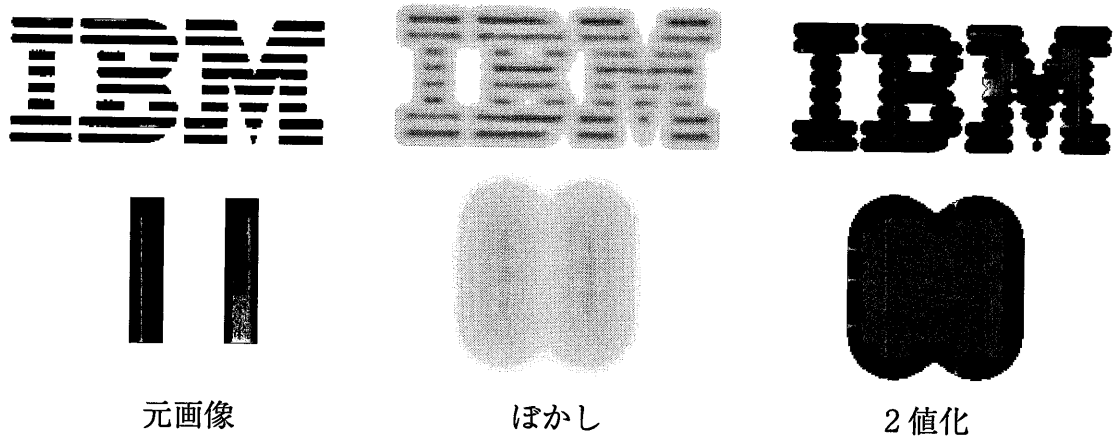


図 4.1: ぼかしによる構造の補完

#### 4.2.2 谷点による構造の補完

谷点の抽出手法は尾根点の場合とほぼ同様である。主曲率平面（図 2.2 参照）において、尾根点抽出のための主曲率の条件は  $K_1 < 0, K_2 = 0$  もしくは  $K_1 = 0, K_2 < 0$  であった。この条件を  $K_1 > 0, K_2 = 0$  もしくは  $K_1 = 0, K_2 > 0$  とすることによって、谷点に相当する点を抽出することができる。スケールを大きくとるほど、より幅が広い線分間の谷点を抽出することになる。

図 4.2 に谷点抽出の例を示す。右は元画像の「H」であり、左はスケール  $t = 0$  から  $t = 100$  までぼかす間に抽出された全谷点と、元画像を重ね合わせたものである。ぼかしを与えた際に谷状になる位置が、谷点となって抽出されていることがわかる。この谷点によって構造の線分間を埋める為に、谷点を再帰的に抽出する。すなわち谷点を抽出した画像から、さらに谷点を抽出していく。

まずスケールの初期値を設定し、このスケールで谷点を抽出できる線分間を補完の最大幅とする。最大幅の谷点を抽出した後の画像からもう一度谷点を抽出するが、このとき初

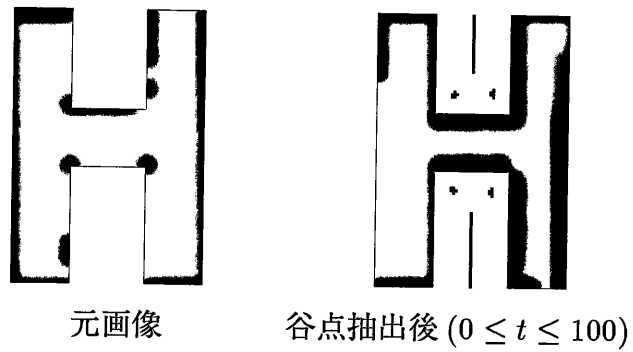


図 4.2: 谷点抽出の例

期値で定めたスケールの $\frac{1}{2}$ の値で谷点を抽出する。つまり谷点を抽出できる線分間の幅が半分になったことになる。このように谷点の抽出とスケールの半減を繰り返すことにより、文字に必要な特徴を埋めてしまう等の過度の補完を避けることが可能となる。上記の手順によってある程度谷点を抽出し終えた画像に対し、ごく小さなスケールでぼかしを与え二値化することにより、大幅な形状の変化なしに補完を行なうことが可能になる。

図 4.3 に構造の補完の例を示す。元画像から 1 回目、2 回目と谷点を再帰的に抽出する。このとき 1 回ごとに谷点抽出のスケールを半減させ、谷点を抽出できる幅を半分になっている。

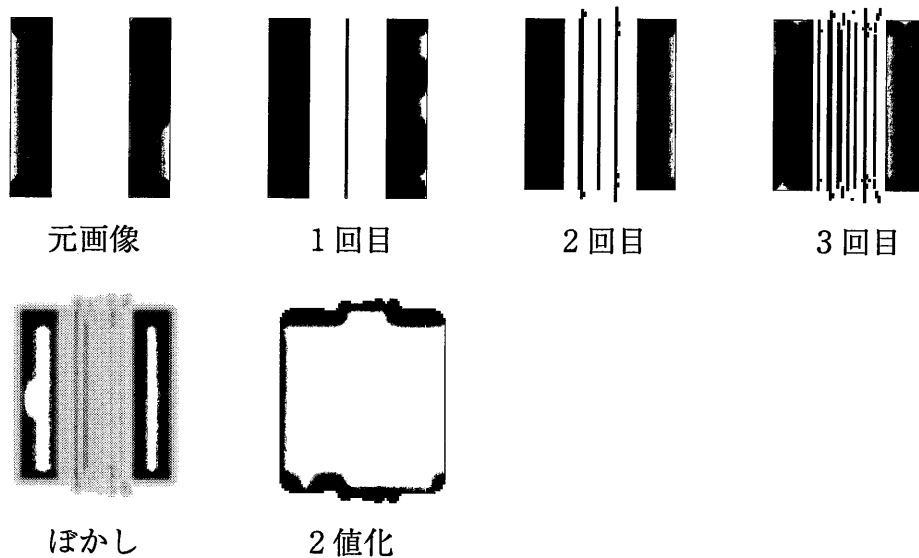


図 4.3: 再帰的な谷点の抽出と構造の補完



### 4.2.3 補完を用いた飾り文字の構造抽出

提案手法を飾り文字の構造抽出に適用する。第3章にて抽出した線分の不連続な大局構造を再帰的な谷点によって補完し、連続な構造を得る。全体の手順は以下の通りである。

1. 飾り文字画像の大局構造を抽出する。
2. 得られた大局構造から再帰的に谷点を抽出する。
3. 谷点抽出後の画像全体をごく小さいスケールでぼかす。
4. 二値化によって補完し、飾り文字の構造を得る。

図4.4は各手順での飾り文字「B」の画像である。スケールスペース解析の応用によって得られた大局構造を、再帰的な谷点の抽出によって補完し「B」の文字情報としての構造が得られた様子が見える。

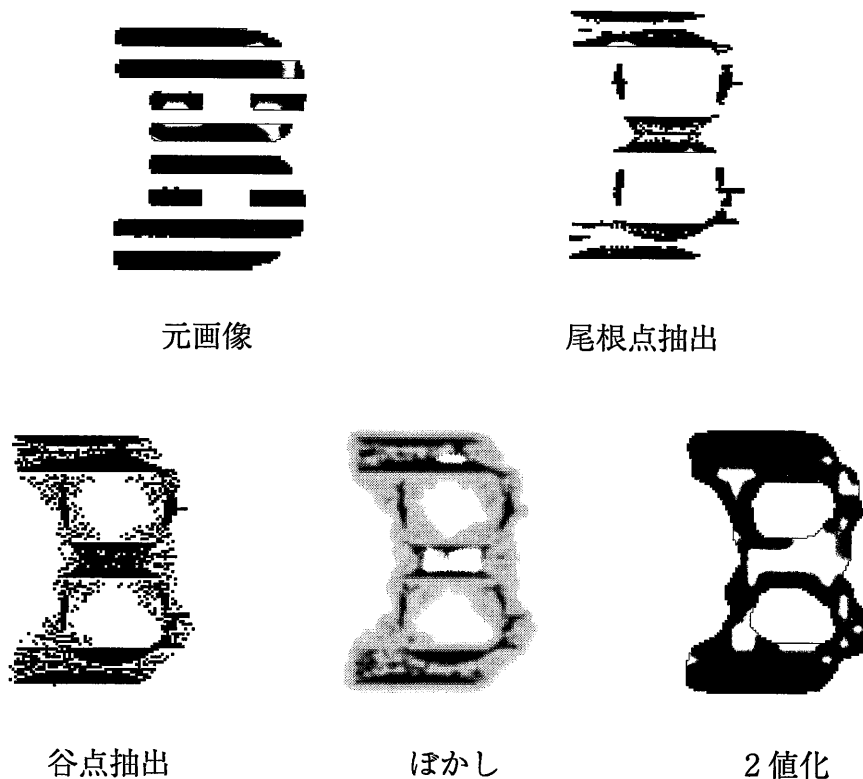


図 4.4: 飾り文字の構造抽出の流れ

## 4.3 飾り文字の構造抽出の実験

### 4.3.1 実験方法

提案手法を用いて様々な飾り文字に対して構造を抽出する実験を行なった。対象とした画像は図 3.4 に示した 26 種類である。

対象とする画像をスケール  $t=0$  から  $t=100$  までぼかしていき、大局構造となる尾根点を抽出した後、再帰的な谷点抽出により構造を補完した。なお、大局構造の抽出に用いる  $R$  を採用するしきい値は  $t=30$  とし、谷点抽出におけるスケールの初期値を  $t=20$  とし、再帰的な谷点抽出を 5 回繰り返した。結果を図 4.5～図 4.7 に示す。

### 4.3.2 考察

図 4.5 は主にスリット状に変形された飾り文字について実験を行なった結果である。補完することにより、飾り文字への変形の元となった文字の構造が得られたことがわかり、再帰的な谷点による構造の補完が正常に作用したことが結果からうかがえる。

図 4.6、図 4.7 は様々な飾り文字「A」に対して実験を行なった結果である。図 4.6(a)～(e) は補完によって「A」の特徴を保存することができた例である。対して、図 4.7(f) のように文字以外の新たなノイズが発生したものや、図 4.7(g)(h) のように大局構造を補完しきれなかったもの、図 4.7(i)(j) のように大局構造と構造抽出時に発生したノイズの間を補完してしまったために形状が変化したものなどの失敗例も存在する。

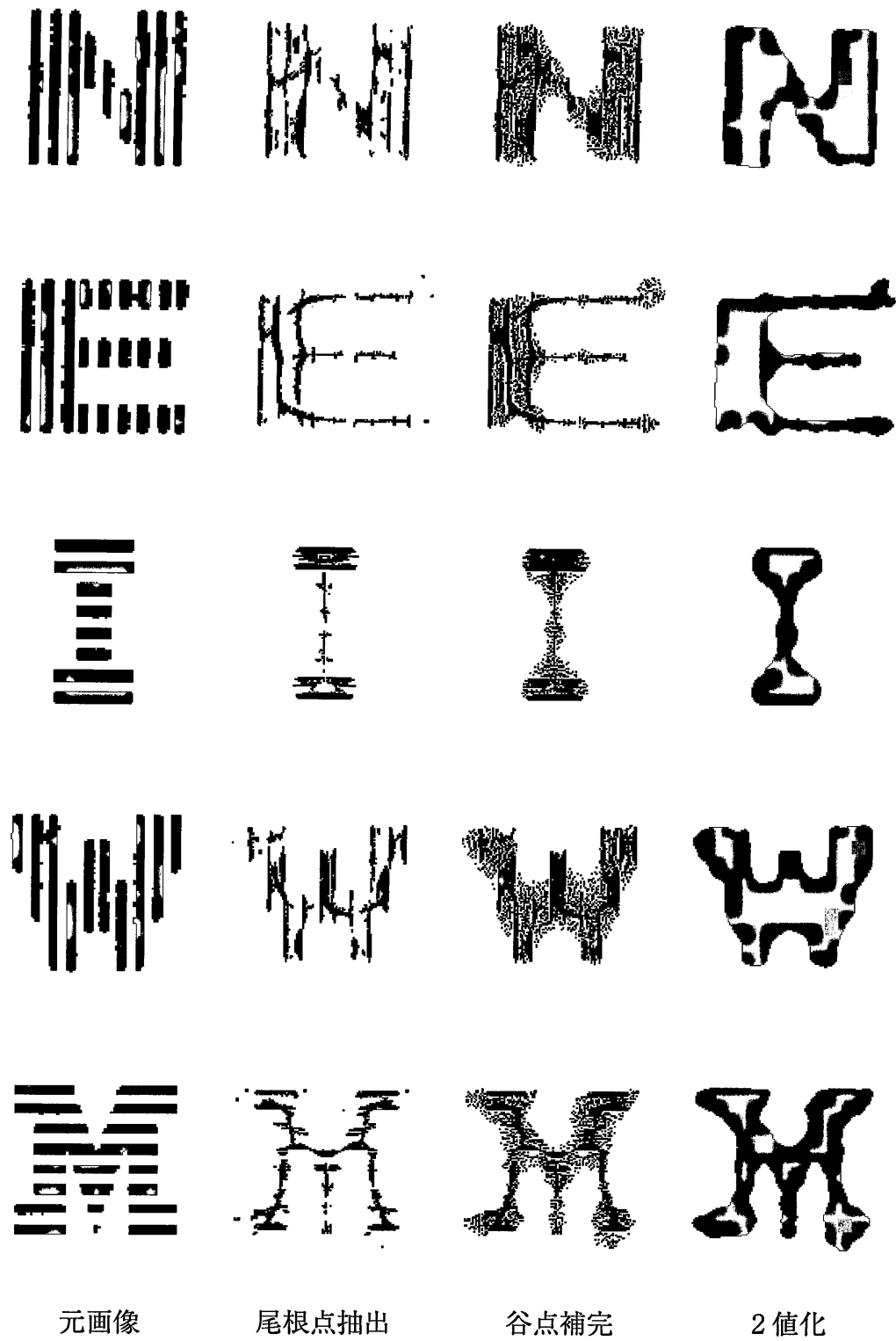


図 4.5: 飾り文字の構造抽出結果 (1)

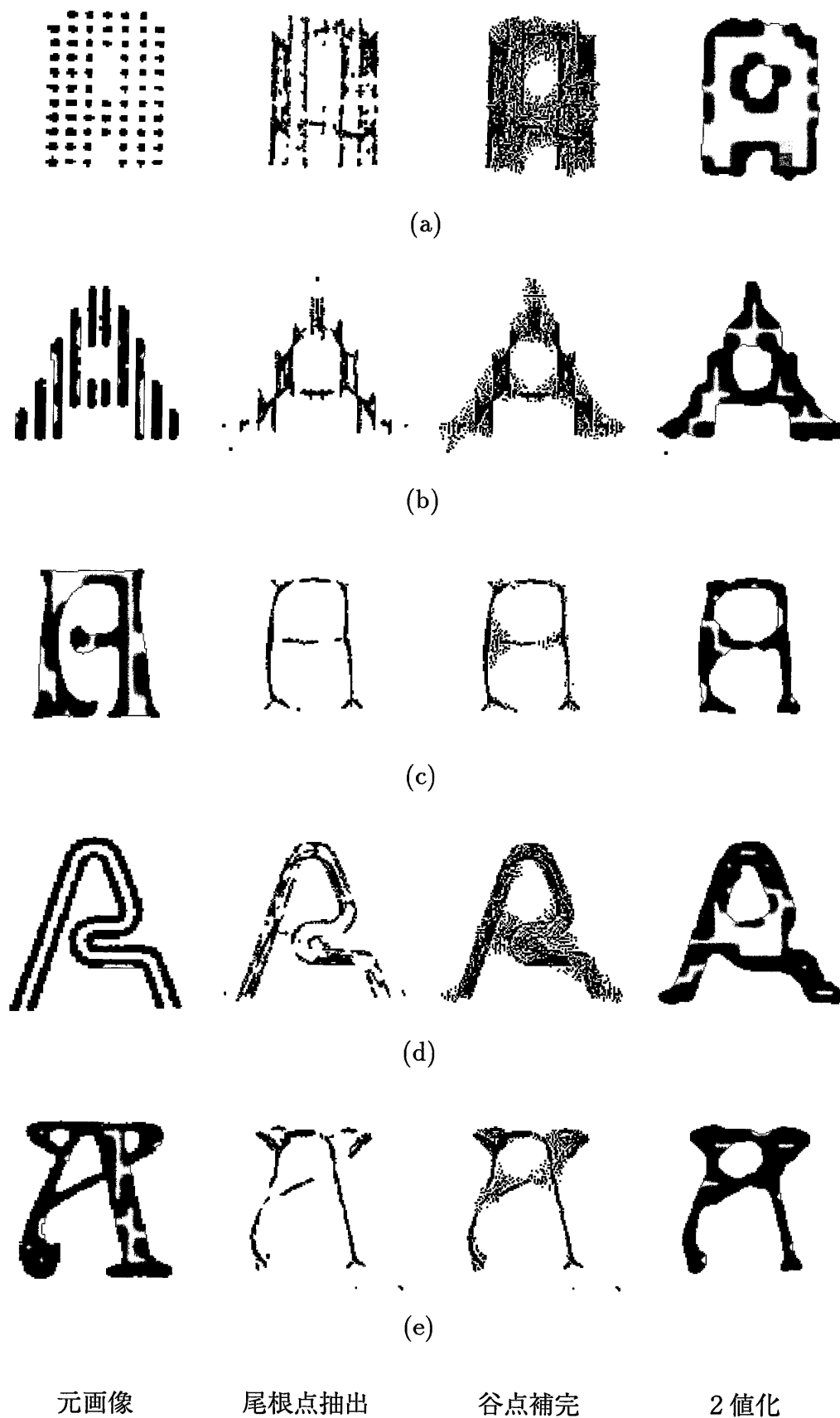
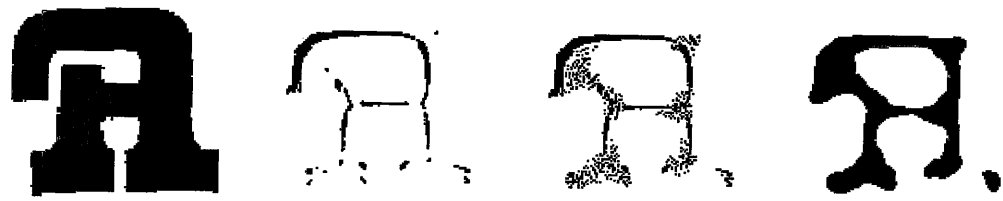


図 4.6: 飾り文字の構造抽出結果 (2)



(f)



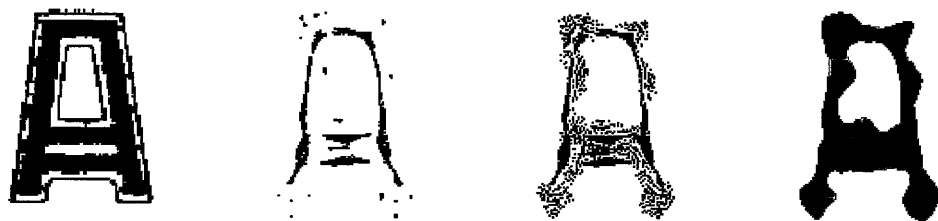
(g)



(h)



(i)



(j)

元画像

尾根点抽出

谷点補完

2値化

図 4.7: 飾り文字の構造抽出結果 (3)

## 4.4 飾り文字の認識実験

### 4.4.1 実験方法

アルファベット 26 文字の明朝体 14 種 38 セットより辞書を作成し、方向線素特徴量を用いて図 3.4 の 26 種類の飾り文字に対して認識実験を行なった。今回は飾り文字の構造を得ることを主な目的としたため、通常の本文書体に用いる認識手法を用いている。提案手法を評価するために以下の 3 つの実験を行なった。

1. 飾り文字を直接認識対象とし、認識実験を行なう。
2. 線分の不連続な大局構造を認識対象とし、認識実験を行なう。
3. 補完後の構造を認識対象とし、認識実験を行なう。

それぞれの手法で正しく認識された文字画像を図 4.8～図 4.10 に示す。

### 4.4.2 考察

実験 1 においては 11 種の飾り文字が正しく認識された。本文書体で適用される方向線素特徴量を用いた認識でも、ある程度の飾り文字の認識が可能であることが分かった。

実験 2 においては 11 種の飾り文字が正しく認識されたが、実験 1 の結果で得られた 11 種とは一部異なっている。大局構造が得られたとしても、構造の線分が不連続な状態のままでは認識率の向上が期待できないことが確かめられた。

実験 3 においては 15 種の飾り文字が正しく認識された。実験 1 と比較して、新たに 4 種の飾り文字が正しく認識されている。提案手法を用いる事により正しく認識された文字数が増え、飾り文字認識において提案手法が有効であるということが確かめられた。

4.3 節の実験結果において「M」や「E」等の文字構造を得ることができた飾り文字が、認識実験では誤認識されるという結果が出ている。これより、本文書体で使用される認識手法ではなく、本手法によって得られた構造を認識するために適した認識手法を考慮する必要があると思われる。

図 4.11、図 4.12 に対象とした飾り文字すべての大局構造と、補完後の構造を示す。

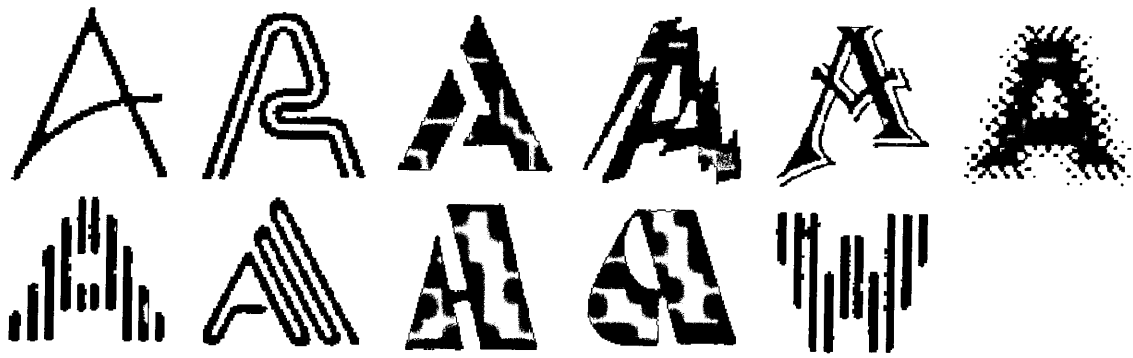


図 4.8: 認識実験1において正しく認識された飾り文字

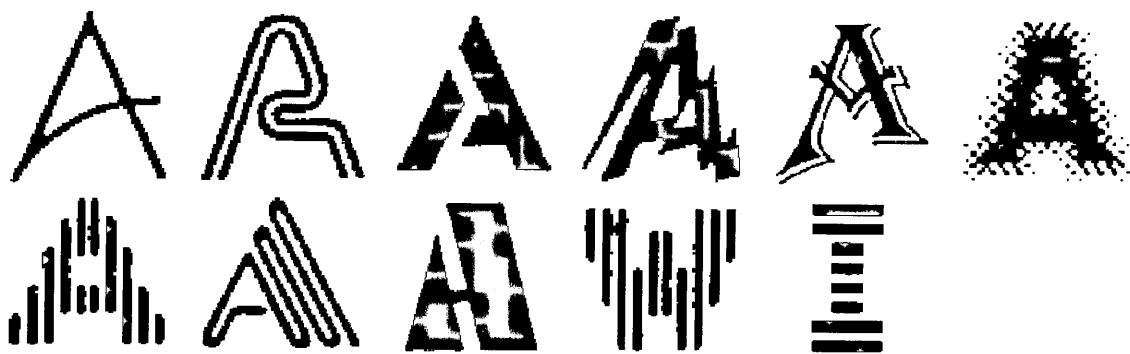


図 4.9: 認識実験2において正しく認識された飾り文字

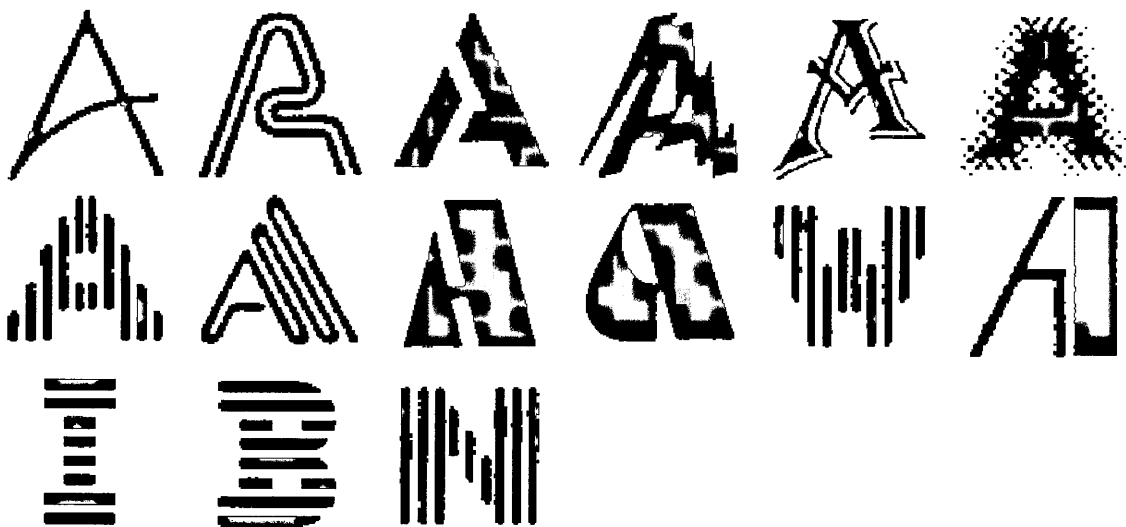
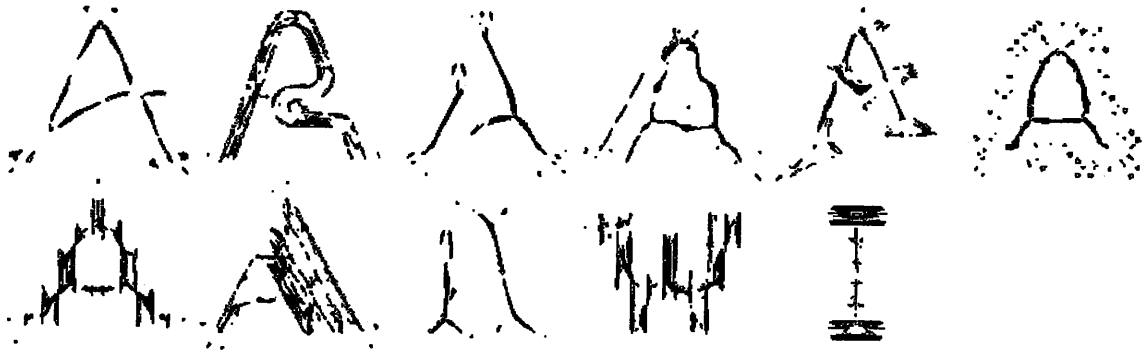
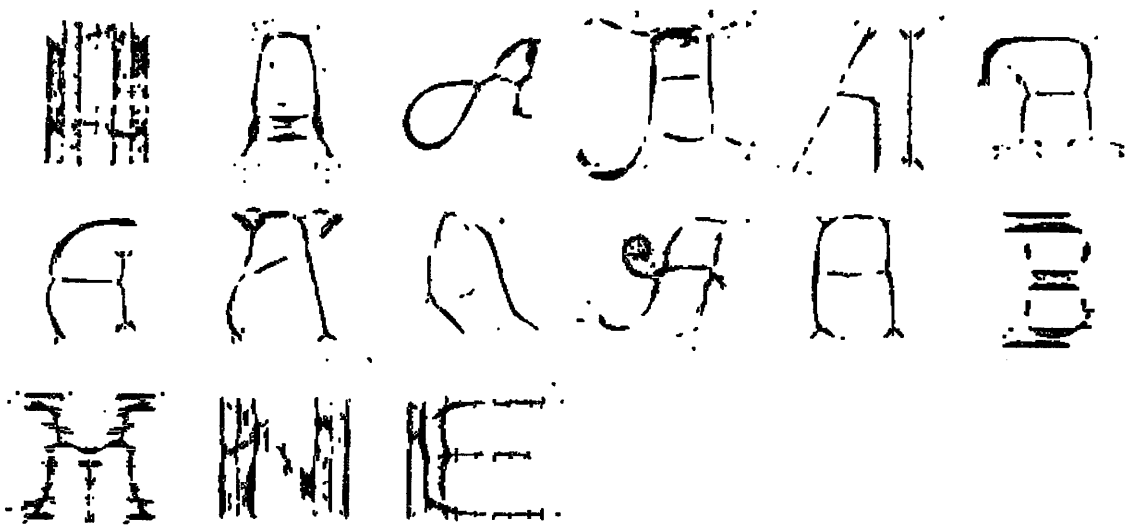


図 4.10: 認識実験3において正しく認識された飾り文字



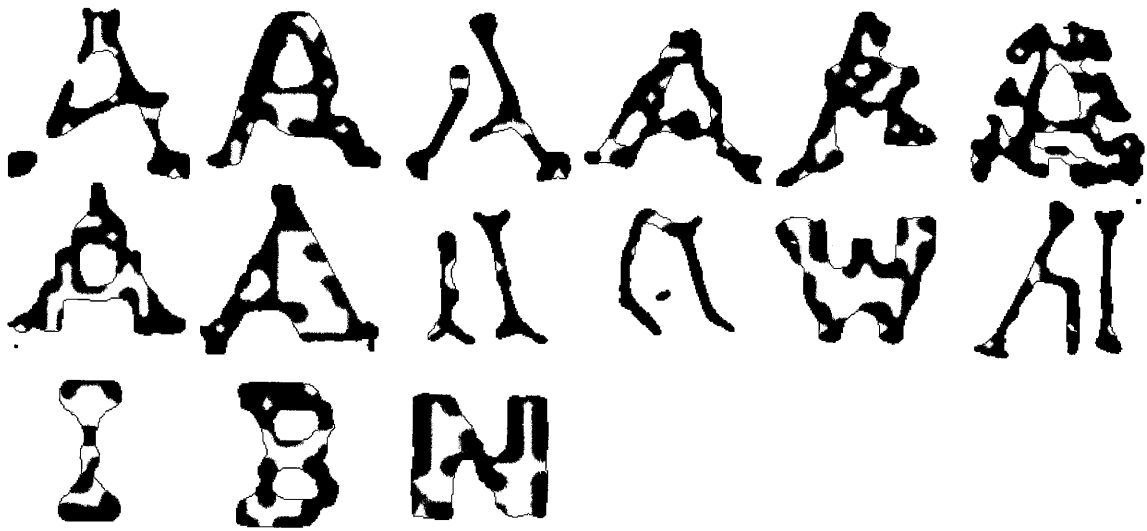
実験2にて正しく認識された大局構造



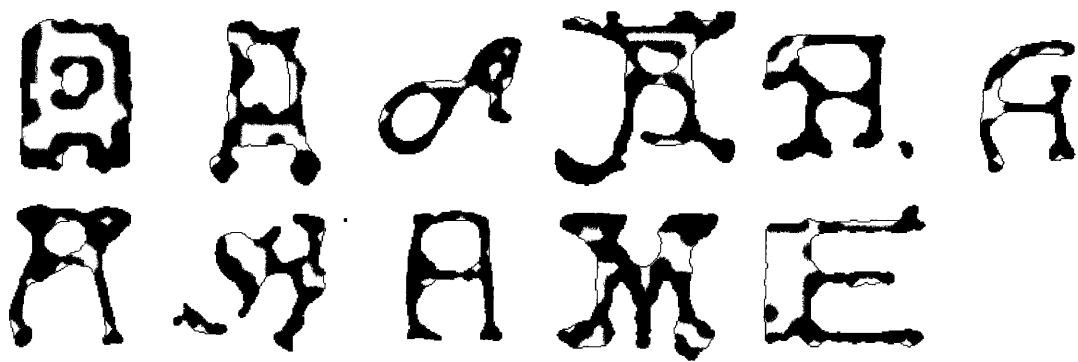
誤認識された大局構造

図 4.11: 実験対象とした飾り文字の大局構造





実験3にて正しく認識された構造



誤認識された構造

図 4.12: 実験対象とした飾り文字の補完後の構造

## 第5章

### 結論

#### 5.1 まとめ

本論文では様々な飾り文字に柔軟に対応でき、飾り文字の表す文字構造を得ることを目的とした。その研究の成果は以下の通りである。

第3章ではスケールスペース解析を応用することによって、飾り文字の表す文字構造を得る手法を提案した。飾り文字の大局構造を表す尾根点を抽出する実験を行なったが、結果として大局構造の一部を抽出するに留まり、尾根状の点に着目しただけでは線分の連続した構造を得ることが困難であることがわかった。

第4章では再帰的な谷点の抽出を用いた、線分の不連続な構造の補完を提案した。第3章で提案した大局構造の抽出と併せ、飾り文字の表す文字構造を得る実験を行なった。また抽出した文字構造を認識対象とし、文字認識実験を行ない、本手法が飾り文字認識において有効であることが確かめられた。

#### 5.2 今後の課題

提案手法によって飾り文字の文字構造を抽出し、認識率の向上を確かめることができたが、実用化にはまだ十分とは言えない。本手法では抽出対象の飾り文字の形状を考慮にせず、一様なパラメータで構造抽出を行なっている。そのため様々な形状に対応しきれず、文字構造が抽出できない例も少なからずみられた。対象とする飾り文字それぞれに対してパラメータを設定することができれば、様々な形状の飾り文字にも柔軟に対応できるはずである。

本研究では認識実験を行なう際に本文書体で用いられる手法と同様の手法を採用した。その結果として認識率の向上が実現できたが、本文書体ではなく提案手法によって得られた構造を認識するために適した手法を用いることができれば、飾り文字認識の高精度化を実現することが可能となるはずである。

また、飾り文字によって形成された文字列を一文字単位に切り出す場合において、本文書体に用いるような、例えば黒画素の密度に着目したりといった手法では正しく切り出すことができない。文字がスリット状になっていたり、文字同士が繋がっていたり、という例が存在するからである。認識の前段階として、このような飾り文字の文字列を切り出す手法の考察が必要であると思われる。

## 謝辞

本研究を進めるに当たり、全般的な御指導・御助言を戴くと共にこの研究の機会を与えて下さいました東北大学工学部 阿曾弘具教授に心より感謝致します。

本研究全般に渡り親身に御指導、御助言を賜りました東北大学大学院工学研究科助手 大町真一郎博士に心より感謝致します。

貴重な御意見、御助言を頂いた東北大学大学院工学研究科 森大毅氏、佐藤俊治氏に感謝致します。

最後に研究のみならず、日頃の生活においてお世話になりました阿曾研究室の御先輩方、共に研究活動に励んだ同研究室の皆様に感謝致します。

## 参考文献

- [1] Tony Lindeberg :”Edge Detection and Ridge Detection with Automatic Scale Serection”,In Computer Vision and Pattern Recognition(San Francisco),pages 465-470,1996
- [2] 本谷秀堅, 出口光一郎:「画像濃淡を表す曲面の主曲率に着目したマルチスケール形状記述手法」, 情報処理学会研究報告,96-CVIM-102,p17-24,1996
- [3] 本谷秀堅, 出口光一郎:「画像のスケールスペース解析にもとづく線図形の抽出」, 情報処理学会研究報告, 97-CVIM-105,p9-16,1997
- [4] Li wang,Theo Pavilidis:”Direct Gray-Scale Extraction of Features for Character Recognition”,In Transactions on Pattern Analysis and Machine Intelligence,vol 15,NO 10,1993
- [5] 電子情報通信学会編,小川 英光編著:「パターン認識・理解の新たな展開 -挑戦すべき課題-」 電子情報通信学会,1994

# マルチスケール解析による 飾り文字パターンへの構造解析

阿曾研究室 4年 井上政樹

1. 背景
2. スケールスペース解析
3. 飾り文字の構造解析
4. 谷点による補完
5. 認識実験
6. まとめ

## 1. 背景

• 本文書体 ↔ 飾り文字

… 人間の認知特性を利用

• 本文書体

黒画素の連続 ———→ 文字 … 認識対象

• 飾り文字

黒画素の連続 —✕—→ 文字 … 認識が困難

例.



認識困難な飾り文字

### ◎目的

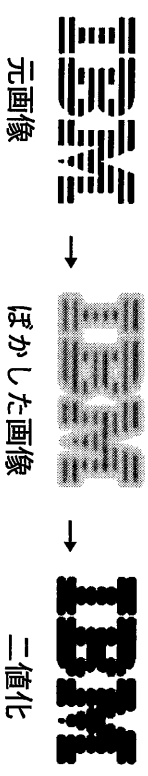
- ・様々な飾り文字パターンの構造抽出

### ◎手法

- ・スケールスペース解析  
画像にぼかしを与える → 濃淡局面の形状解析  
→ 構造の抽出

### ◎ぼかしの効果

- ・利点 … 細かな構造の消去  
大局的な構造の抽出



- ・欠点 … 単一スケールでは対応できない

(スケール：ガウスマイラタをかける回数)



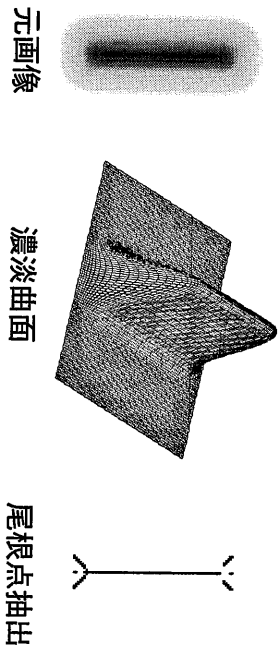
### スケールスペース解析

複数のスケールを用いて構造を得る

## 2. 曲面の主曲率を用いた スケールスペース解析

- 濃淡曲面の尾根点に着目

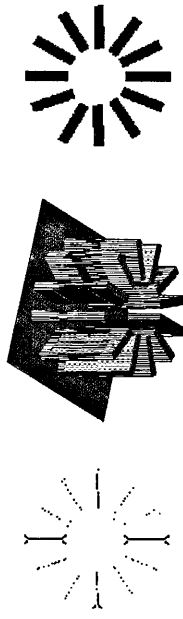
例.



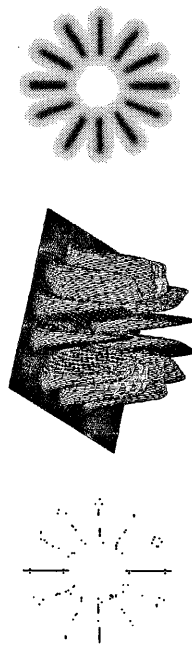
- 曲面の尾根点を各スケールにおいて抽出  
→ 各点の主曲率の変化を用いる

- 全尾根点を重ね合わせる

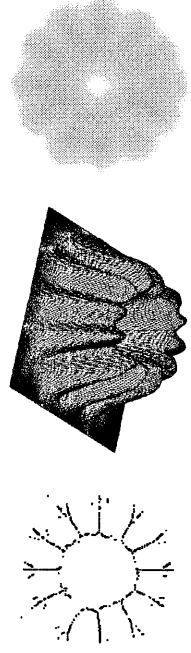
スケール:  $t = 0$



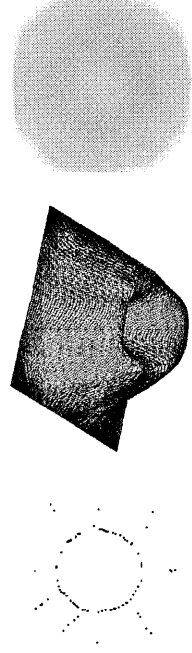
$t = 10$



$t = 50$



$t = 150$



全尾根点を重ね合わせる →





### ◎主曲率

• 各位置において求まる主曲率  $K_1, K_2$  の比は曲面の形状を示す。

$$\begin{cases} K_1 = \frac{(L_{xx} + L_{yy}) + \sqrt{(L_{xx} - L_{yy})^2 + 4L_{xy}^2}}{2\sqrt{1 + L_x^2 + L_y^2}} \\ K_2 = \frac{(L_{xx} + L_{yy}) - \sqrt{(L_{xx} - L_{yy})^2 + 4L_{xy}^2}}{2\sqrt{1 + L_x^2 + L_y^2}} \end{cases}$$

$$\begin{cases} \theta_1 = \arctan \left( \frac{(L_{yy} - L_{xx}) + \sqrt{(L_{xx} - L_{yy})^2 + 4L_{xy}^2}}{L_{xy}} \right) \\ \theta_2 = \arctan \left( \frac{(L_{yy} - L_{xx}) - \sqrt{(L_{xx} - L_{yy})^2 + 4L_{xy}^2}}{L_{xy}} \right) \end{cases}$$

$$R = \sqrt{K_1^2 + K_2^2}, \alpha = \arctan \left( \frac{K_2}{K_1} \right)$$

( $L_{xx}, L_{yy}, L_{xy}$ ...画像の偏微分)

$K_{1,2}$  : 主曲率

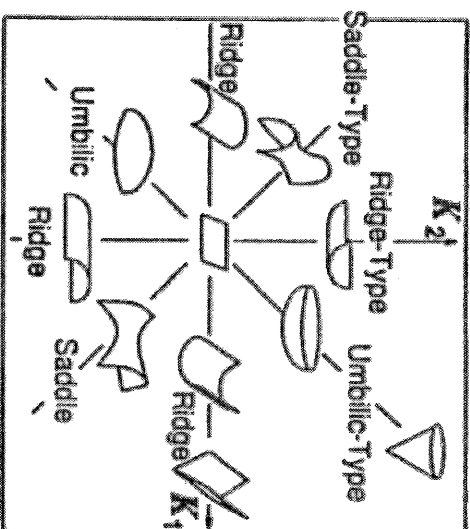
$\theta_{1,2}$  : 主曲率方向が画像の X 軸となす角

$R$  : 曲面の曲がり具合

$\alpha$  : 曲面の形状

### ◎主曲率

• 主曲率  $K_1, K_2$   
...各点における曲面の形状を表す



主曲率による曲面の記述

例.

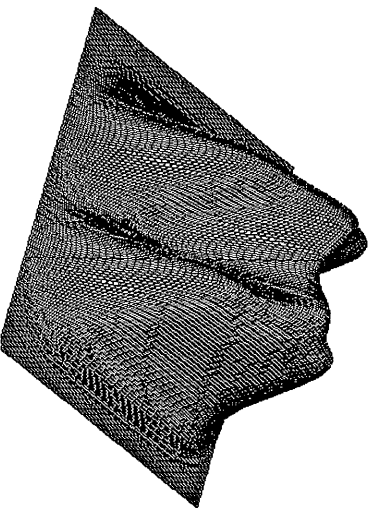
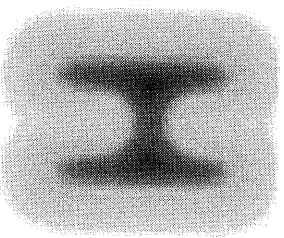
$K_1 = 0, K_2 = 0$  ... 平面

$K_1 < 0, K_2 = 0$  ... 尾根点

$K_1 > 0, K_2 < 0$  ... 鞍点

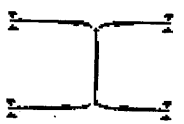
•  $R = \sqrt{K_1^2 + K_2^2}$  が大  $\rightarrow$  曲がり具合が鋭い  
 $\rightarrow$  重要な特徴

### ◎主曲率による特徴抽出



スケール:  $t = 100$  での「H」の形状

•  $t = 100$  までに抽出された特徴点



$K_1 < 0, K_2 = 0$   
尾根点



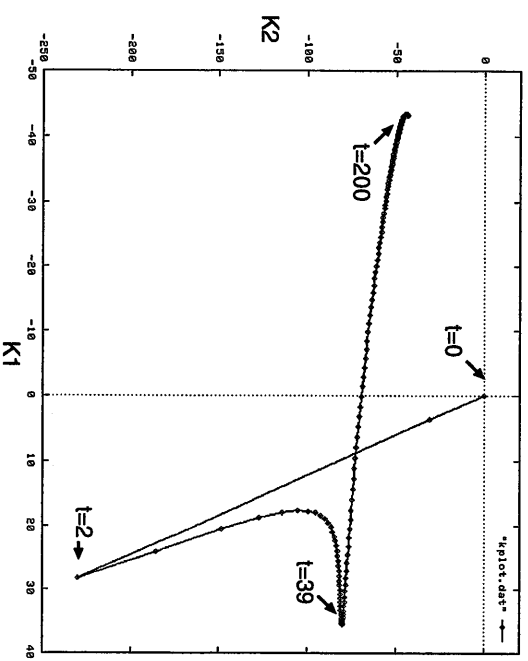
$K_1 > 0, K_2 = 0$   
 $K_1 = 0, K_2 > 0$   
谷点



$K_1 > 0, K_2 < 0$   
 $K_1 < 0, K_2 > 0$   
鞍点

### ◎スケールスペース解析による特徴点抽出

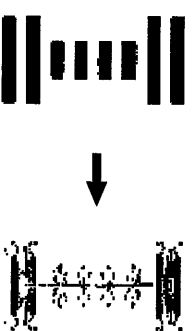
• スケールの変化に対する主曲率をプロットする



主曲率平面

• 抽出すべき特徴点 ... Rが極大をとる点

抽出例.

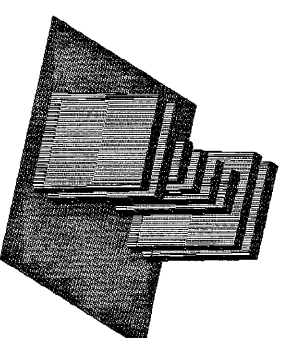


### 3. 飾り文字の構造抽出

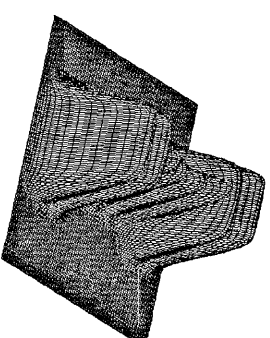
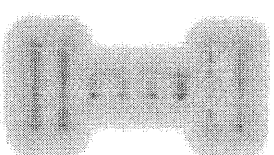
- ・スケールスペース解析を飾り文字に応用
- ・飾り文字の構造… 大局的な構造が必要
- ・大局構造の特徴点
  - 大きなスケールで現れる特徴点
  - あるスケール以上でRが極大をとる点を抽出

・ぼかしによる「罫」の形状の変化

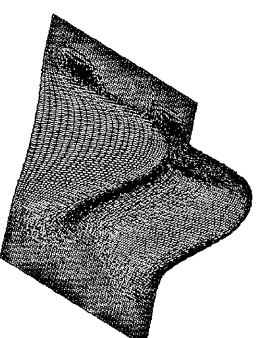
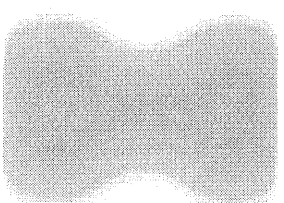
スケール： $t = 0$

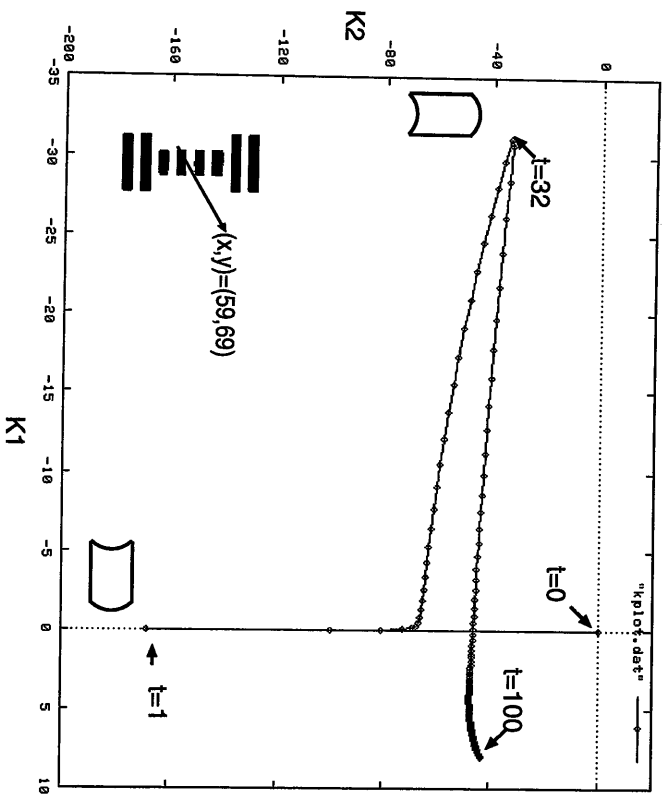


$t = 20$

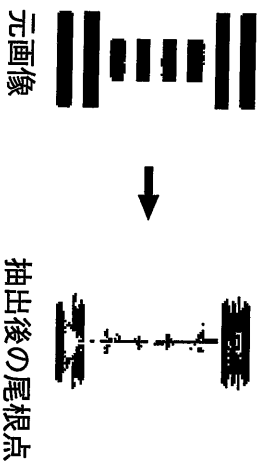


$t = 100$





「理」(x,y)=(59,69)での主曲率の変化(0 ≤ t ≤ 100)



### ◎実験

各種の画像において、主曲率を用いたスケールスペース解析により尾根点を抽出する。

### ◎実験結果

• 元画像



• 従来のスケールスペース解析



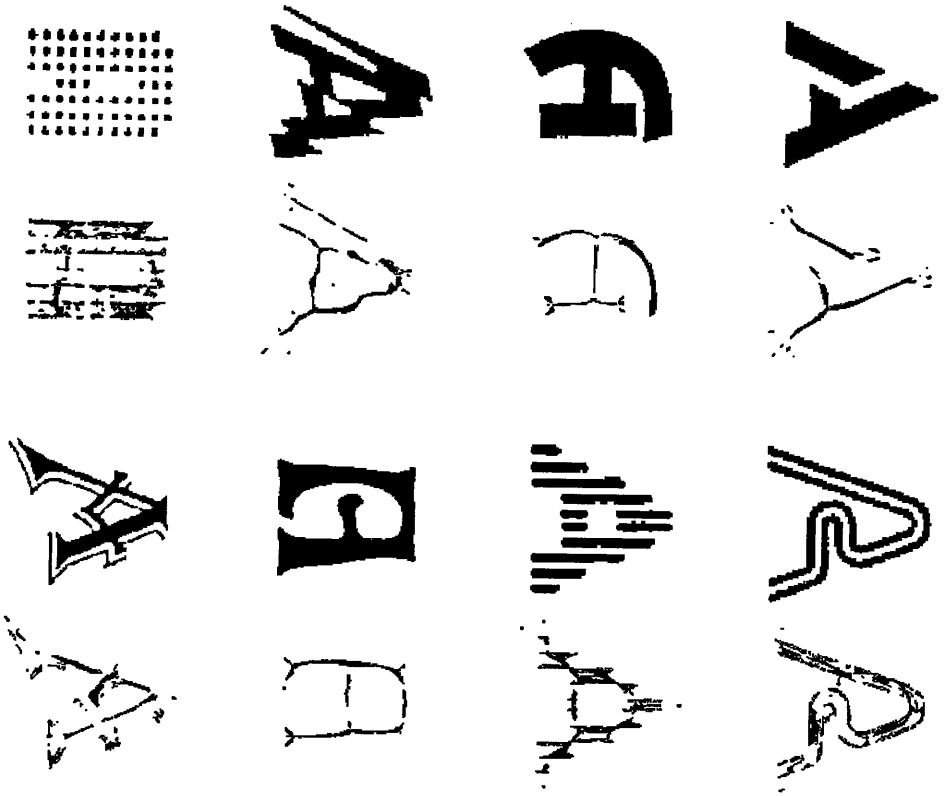
(t=100 までの全尾根点)  
微細構造まで抽出されている

• 本手法による大局構造の抽出



大局構造のみの抽出が可能

## 尾根点の抽出



左：元画像      右：抽出後の尾根点

## ◎問題点

・尾根点のみでは連続した構造が抽出できない例。



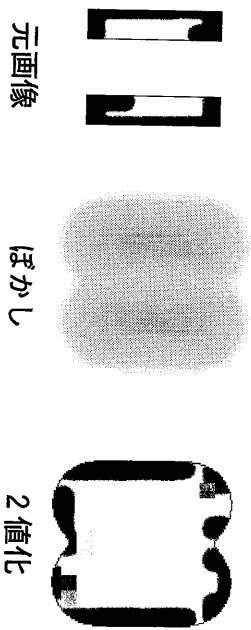
左：元画像      右：抽出後の尾根点

→ 尾根点抽出後の画像を補完する

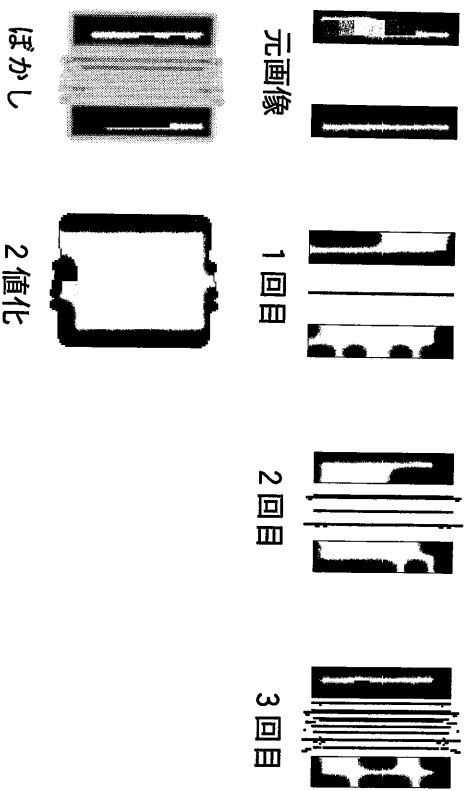
谷点による構造の補完

#### 4. 谷点による補間

- ぼかしのみで補完 → 形状の変化



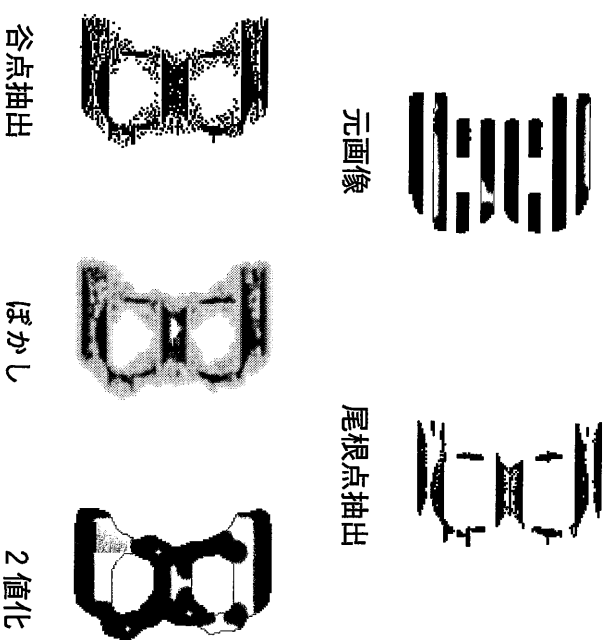
- 谷点を再帰的に抽出 → 隙間を補完



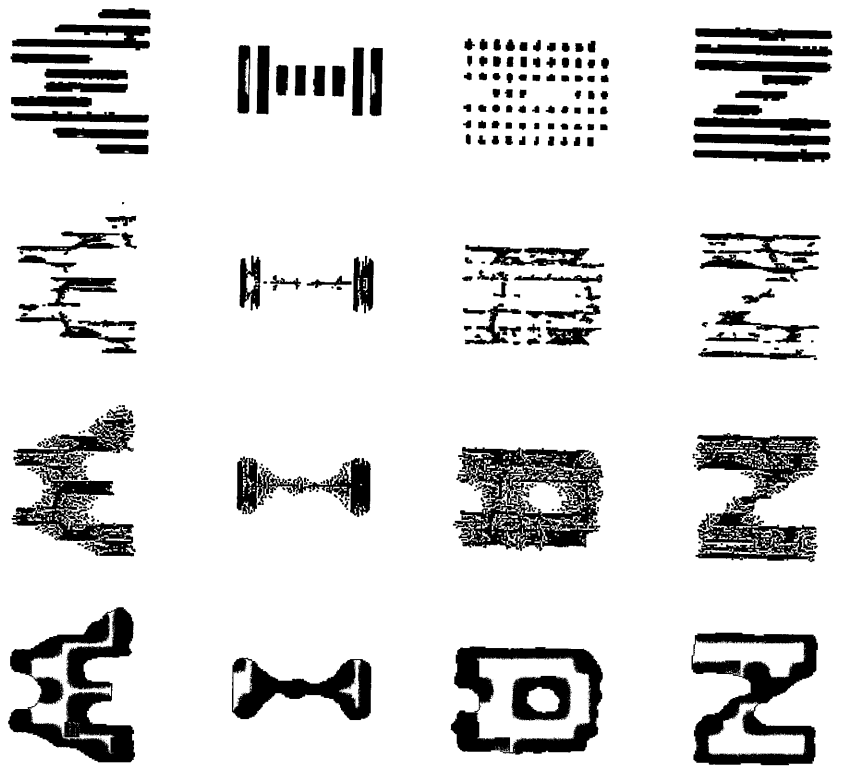
#### ◎適切な谷点による補完

- 一定幅以下で谷点を抽出 → 文字の特徴を保存
- 谷点補完のアルゴリズム

  1. 画像の尾根点を抽出
  2. 再帰的に谷点を抽出
  3. 画像全体を少しだけぼかす
  4. 2 値化

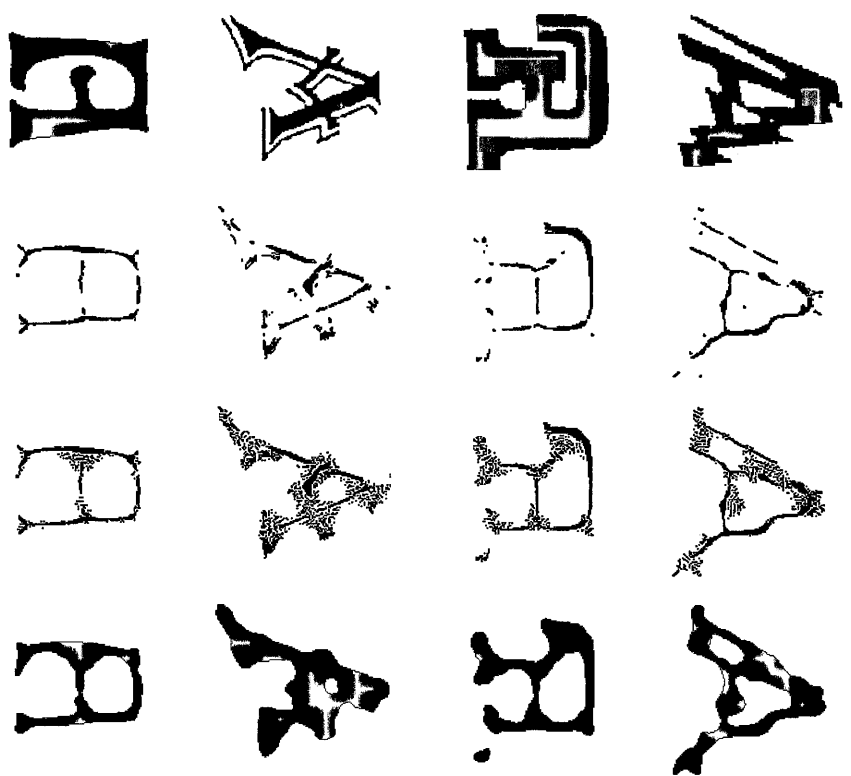


谷点の補間 (1)



元画像      尾根点抽出      谷点補完      2値化

谷点の補間 (2)



元画像      尾根点抽出      谷点補完      2値化

## 5. 認識実験

飾り文字 26 種に対し認識実験を行なった。

・方向線素特徴量

・辞書：アルファベット 26 文字の明朝体

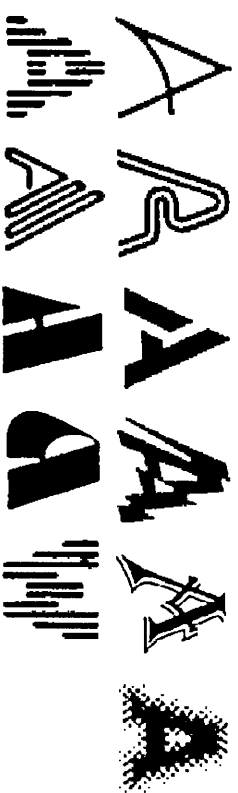
14 字種 38 セット

### ・実験方法

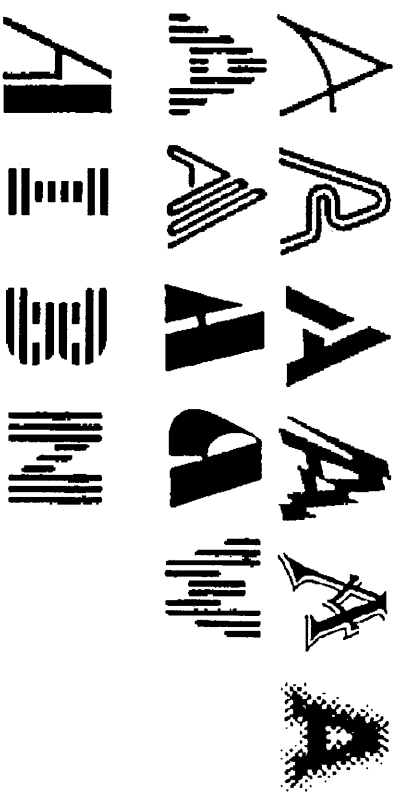
1. 飾り文字を直接認識させる。
2. 谷点補間後の画像を認識させる。

### ◎認識結果

結果 1. 認識可…26 種中 11 種



結果 2. 認識可…26 種中 15 種



…新たに認識された飾り文字



## 6. まとめ

- ・スケールスペース解析によって画像の尾根点を抽出し、飾り文字の構造を得ることができた。
- ・谷点による構造の補完を行ない、より認識に適切な構造を得る事ができた。
- ・認識実験の結果、本手法による認識率の向上を確認できた。

## ・今後の課題

- ・連結している飾り文字の切り出し
- ・スケールスペース解析によって抽出した構造を有効に利用する認識方法の検討

