

6P-7

画像特徴を利用した文書画像の蓄積・検索方式

山田敬嗣 美濃導彦 坂井利之

京都大学 工学部

1. はじめに

近年、入出力装置や通信メディアの多様化が進み、論文や書類などの紙面を文書画像として計算機で処理したり蓄積する要求が高まっている。

文書画像を効率よく処理、蓄積、伝送し、しかも高品質に再生するためには、画像の局所的な特徴(局所的特徴)を利用することが重要である。この基本的な考えに基づき文書画像を二値化する方式1)、及び局所的特徴に適応して二値文書画像を圧縮符号化する方式2)については既に報告している。

画像の局所的な特徴を利用すれば、文書画像の領域分割およびその領域の特徴量を抽出できる。本報告では、これに基づき二値文書画像を領域ごとに蓄積管理し、領域の特徴を用いて検索する方式について述べる。

2. 文書画像の領域特徴

文書画像は、図1のような階層構造をもつ。信号レベル構成要素の持つ局所的特徴の観測に基づき、構造を把握することによって文書画像を領域分割できる。また、領域内の各階層の構造から人間の直感に合った特徴を抽出し記述すれば、それを検索キーとして文書画像の検索に利用できる。

二値文書画像の各領域の直感的な特徴を以下に列挙する。

◎テキスト領域

- (a) 短い線分を信号レベルの構成要素とする。
- (b) 線分に方向性は少ない。
- (c) 線分相互間の距離は小さい
- (d) 一定幅の文字行と行間を基本的にもつ。

◎図領域

- (e) 長い線分を信号レベル構成要素とする。
- (f) 線分の方向は一樣でない。
- (g) 線の交点、屈折点の様子がパターンレベル構成要素により異なる。

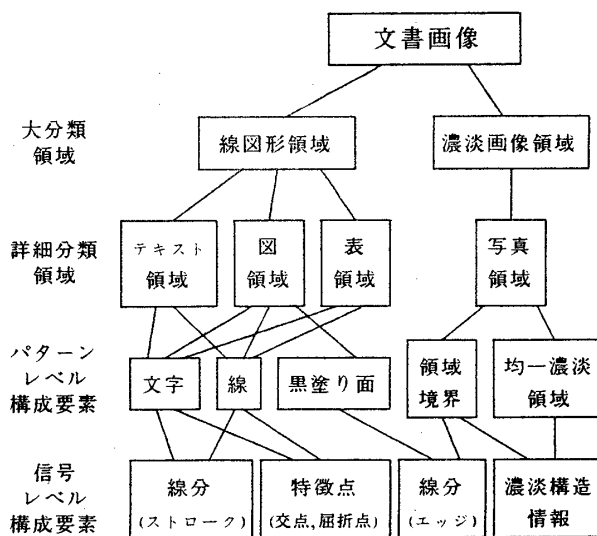


図1 文書画像の階層構造

◎表領域

- (h) 縦横の長い線分(線)が矩形の枠を構成する。
- (i) 枠の大きさや数は一樣でない。
- (j) 矩形内部には文字や記号を含む。

◎写真領域(擬似濃淡画像)

- (k) 画素間の相関が低く、ラン長が短い。
- (l) 濃淡が均一な部分が領域をなす。
- (m) 領域の境界によりその形状を表現できる。

これらの領域ごとの特徴のうち、(a), (c), (d), (e), (h), (k), (l)を用いれば領域分割が可能である。また、(b), (f), (g), (i), (j), (m)などを特徴量として記述すれば、個々の領域を区別でき、文書画像の検索キーとして利用できる。

3. システムの概要

文書画像の蓄積・検索システムの概要を図2に示

す。蓄積時の処理の流れは4つの段階からなる。

- 1) 原文書画像はイメージスキャナから多値画像として入力され、二値画像に変換される。ただし、濃淡画像領域は擬似濃淡表現される。
- 2) 二値文書画像を128×128画素の矩形ブロックに分割し、ブロックごとに符号化する。符号化の際に、ブロックが線図形領域、濃淡画像領域のいずれに属するかが判定される。
- 3) ブロックの領域判定の結果を利用して、さらにテキスト、図、表、写真へと詳細な領域に分割する。
- 4) 領域ごとの特徴量をもとめ記録する。

検索処理の流れは5段階からなる。

- 5) 検索対象となる領域の手書きのスケッチ画像(線図形)を入力する。
- 6) スケッチ画像から対象領域に適した特徴量を計算する。
- 7) スケッチ画像の特徴量を蓄積された特徴量とマッチングすることにより候補画像を選択する。
- 8) 候補画像の数が多過ぎる場合、同一文書画像内にある他の領域に対し5)、6)、7)の処理をおこなう。
- 9) 候補画像を適切な数にまで絞ったら、それを表示し、目的の文書画像を選択する。

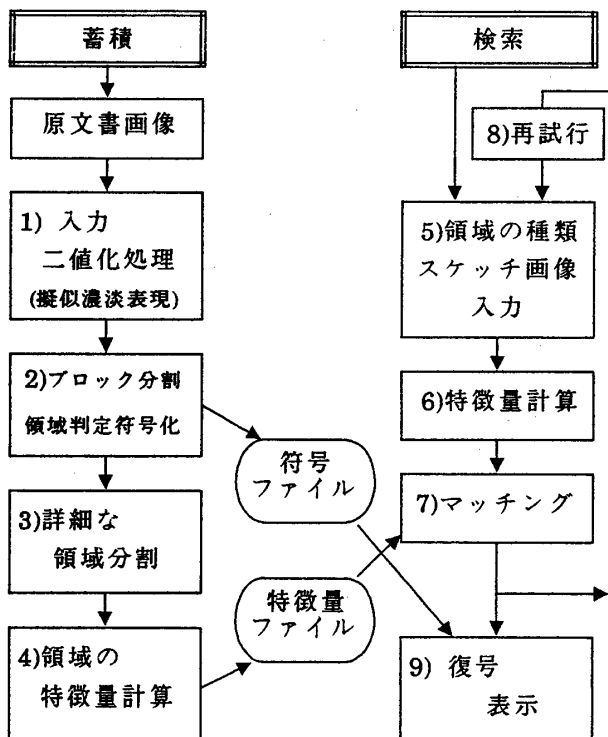


図2 システムの概要

4. 線図形領域のスケッチを用いた検索実験3)

線図形領域をラフスケッチした画像から特徴量を求め、それを用いて文書画像を検索する実験をおこなった。検索対象とする文書画像は、技術論文から無作為に選んだ66枚の画像で、各々は図、表、濃淡画像領域を含む。検索用のスケッチとしては手書きの概略図を用い、図表中の文字は省略する。

図領域の直感的特徴(2.の(f))を特徴量として記述するために、ア)2次水平方向差分、イ)2次垂直方向差分、ウ)縦横比、エ)傾斜方向、オ)曲線性について調べる。画像を符号化する際に、2×2画素に区切ったメッシュ毎に45°おきの8方向のうちの1つの方向をもつエッジが抽出される。そのエッジ方向の統計量としてア)~オ)の特徴量をもとめる。ア)、イ)は水平、垂直方向に隣接する2つのエッジ部メッシュのエッジ方向の差の平均を示す。ウ)は垂直エッジと水平エッジの比、エ)は左上がりエッジと右上がりエッジとの比であり、オ)は斜めエッジと水平垂直エッジとの比である。

表領域の直感的特徴(2.の(i))を特徴量として記述するために、カ)水平最頻ラン長比、キ)垂直最頻ラン長比をもちいる。これらは、水平垂直方向での出現頻度の高い白ラン長を捉える。

各特徴量の分散正規化距離によりスケッチ画像と蓄積文書画像との類似度を計算した。図領域を検索対象とした場合、66枚中で第10位までに目的の画像が得られたのは、86%であった。また、表領域の場合には、第10位までに目的の画像が得られたのは89%であった。

5. おわりに

本稿では、画像特徴を利用して文書画像を領域に分割し、その特徴とともに蓄積する方法について述べた。また、線図形領域を対象として、手書きスケッチを検索キーとして文書画像を検索する実験をおこない、抽出した特徴量の有効性を示した。今後は、線図形画像の特徴量についてさらに実験を進める。また、濃淡画像領域を対象とした手書きスケッチによる検索についても調査を続ける。

参考文献

- 1) 山田他、"メッシュ内の濃淡構造に適應する画像の二値表現方式", 信学会論文誌(D), vol. J69-D no.4 pp694~701(1985)
- 2) 山田他、"文書画像の領域に適應する符号化方式" 第29回情処全大6M-6, pp1143(1984)
- 3) 黒川他、"図表画像のパターン特徴の抽出", 第29回情処全大6M-8, pp1147(1984)