

Geometric Hashing を用いた帳票識別

6R-7

浅野三恵子 下辻成佳
(株)東芝 研究開発センター

1 はじめに

一般オフィスでは、各種伝票、名簿、一覧表、銀行の振込依頼書、企業の財務諸表など、多種多様の帳票文書が取り扱われている。近年、計算機を用いた画像ファイリングが一般オフィスで用いられるようになるにつれて、これらの帳票文書から、必要な情報のみをとりだし計算機に入力することが望まれるようになってきた。これらの情報を効率よくファイリングし、検索に利用するためには、帳票の種類ごとにデータを入力・管理することが必要であり、入力帳票の特徴を自動的に抽出し、帳票種別を自動識別することが要求される。

帳票識別を実用化するためには、

1. 文字の記入などによる画像の変動や入力画像のノイズやぼれなどによる情報の欠落の影響をうけない。
2. 画像の傾き・位置ずれ・伸縮などの変動を吸収する。
3. 多種多様の帳票に対応できるように、対象帳票の種類に依存しない。
4. 登録帳票が増えても検索時間が遅くならない。

の点を考慮しなければならない。即ち、入力で生じる画像の変動や、情報の欠落の影響をうけない柔軟かつロバスタなアルゴリズムを構築する必要がある。

我々は、これらの問題点を解決するために、帳票文書は枠領域（以下セルとよぶ）と、文字列から構成されていると考える。帳票上のセル領域は帳票上に文字が記載されても変わらない比較的普遍的な情報である。これに対し、文字情報は追記されることによって変動が大きい情報である。したがって、セルの位置情報を帳票の特徴として帳票識別を行うことを考える。

ここで、帳票文書をセルで表現し、セルの中心点をそのセルの代表点とした場合に、類似帳票識別は点マッチングの問題として考えることができる。図1に帳票文書からセルを抽出した例を示す。

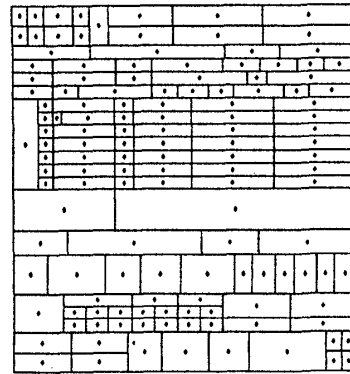


図1: セル抽出結果（・はセルの中心）

帳票の種別をセルの構造で識別することは、上記問題点を考慮した上での2枚の画像中の長方形領域の対応関係を検証することである。そこで、本論文では、領域対応を各領域の代表点対応問題と考え、点マッチングの方法として知られている Geometric Hashing^[1]を応用してこの問題を解くことにより帳票識別を行なう。Geometric Hashingとは、データの整理方法であるハッシュ法を画像認識に応用したものであり、モデルから2次元不変量を計算して Hash Table を作成した後、入力画像から計算された不変量に基づいて Hash Table 上を検索する手法である。以下では、本手法の詳細なアルゴリズムを述べると共に、実験によりその有効性を示す。

2 Geometric Hashing を用いた帳票識別

Geometric Hashing では Hash Table をモデル画像のデータベース表現に用いている。はじめに、登録したい全てのモデル画像から決められた特徴量を抽出し、モデル画像のアフィン変換に関する不変量を計算する。次にテーブル管理に必要なデータ（不変量、モデル名、不変量を計算するのに用いたパラメータ（以後、基底と呼ぶ））を構成する。最後に、この不変量を用いてデータを Hash Table に登録する。これらの処理を、登録したいモデル画像の枚数分、全ての基底について行う。

次に入力画像に対し同じく不変量を計算し、この不変量を用いて対応する Hash Table 上のエリアをアクセスする。

アクセスした位置に対して、条件を満たすモデル名（例えば、近傍位置のすべての位置に登録されたモデル名、またはモデル名が登録されている位置のうちその位置に最も近いもののモデル名）に投票する。これらの処理を入力画像内の特徴点数分繰り返し、最も多い投票数を得たモデルを認識結果とする。

Geometric Hashing は 3 点の点座標から算出した相対座標を用いているので、位置ずれや傾きの問題を同時に解決することが可能になるほか、Hash Table は登録時に作っておけばよいので、この方法によって観測画像の認識時間を大幅に短縮できる。このように、Geometric Hashing はモデルマッチングにおける画像認識の時間を短縮するのに有効な方法であるといわれている。

本処理ではモデル画像は全ての登録帳票、観測画像は入力帳票、利用する特徴量は各セルの中心座標とする。Geometric Hashing ではすべての 3 点の組合せについて Table 登録を行うが、帳票文書がセル構造であることから、Hash Table を作成するときの基底のとり方は、水平・垂直線分上に位置する 3 点の組合せのみでよいと考える。とくに、帳票の場合には連結する長方形枠が多く存在することから、基底には、連結長方形枠の外接図形の 2 辺を用い処理の簡略化を試みた。

はじめに、登録したい帳票の画像を入力する。次にセル抽出処理によってセル領域を抽出する。さらに、連結枠を抽出し、連結枠の外接長方形の 2 辺を基底とした相対座標で、各セルの中心座標を表現する。この座標値は、帳票画像のアフィン変換に関して不変な値となる。また、この座標値 (u, v) は全てのセルに対して $0 < u < 1.0, 0 < v < 1.0$ となる。次に、この点を 2 次元 Hash Table 上の (u, v) の位置に登録帳票名とともに登録する。

識別処理では、まず識別したい画像を入力する。次にセル領域抽出および連結枠抽出を行なう。その後、セルごとに Hash Table 上の座標値を計算し、この座標を中心とするウィンドウサイズ w の近傍（計算した座標値 (u, v) に対し、 $u-w < x < u+w, v-w < y < v+w$ の (x, y) 領域）に存在する登録帳票のセル中心点を検索する。もし、点が検出できた際はこの登録帳票名に投票を行う。これらの処理を入力画像内のセルすべてについて繰り返し、この投票数から得られる類似度が最大の登録帳票を認識結果とする。

ここで類似度は以下のように定義する。登録帳票 t に対する投票数を N_t 、登録帳票のセル数を R_t 、入力帳票のセル数を k 、類似度を S_t とすると、 $S_t = 2N_t / (k + R_t)$ で表す。

3 実験結果

482 種類の帳票を対象に以下の実験を行った。帳票データは 16 本/mm の解像度のスキャナによって取り込まれ、固定閾値により二値化されている。

はじめに、482 種の帳票を登録し、この中の 12 種の帳票 685 枚を用い、検索ウィンドウサイズ w を変動させながら識別率を測定した結果、 $w=0.07$ で最大識別率 0.987 をとり、 w が大きくなるにつれて、徐々に識別率が低下していった。

次に登録帳票の枚数を変動させながら Hash Table の作成時間と識別時間を測定した。図 2 は処理結果をグラフに表したものである。識別実験では、12 種の帳票 685 枚を用い、その識別に要する 1 枚あたりの平均時間を示している。テーブルの作成時間は登録帳票の枚数とほぼ比例関係にあるが、検索時間は登録枚数が増加してもほとんどかわらない。

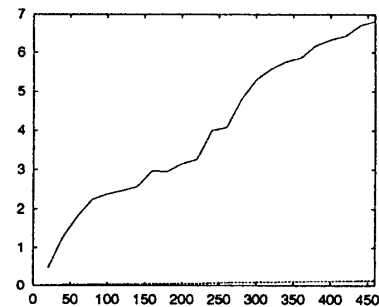


図 2: 登録帳票の枚数と処理時間の関係
(横軸: 登録帳票数, 縦軸: 処理時間 (sec))

実線: Hash Table の作成時間, 破線: 検索時間

4 おわりに

本報告では、帳票入力システムにおける Geometric Hashing を用いた帳票識別処理について述べた。今回の実験では、0.987 の識別率が得られた。また、登録帳票が増加しても一度 Hash Table を作成すれば検索時間にほとんど影響がでないことが確認でき、本手法は実用可能であることが検証できた。

参考文献

- [1] Y.Lamdan and H.J.Wolfson. "geometric hashing: A general and efficient model-based recognition scheme". *2nd ICCV*, pp. 238-249, 1988.