

文字写真混在画像からの文字エッジ点候補抽出法

5M-7

八木啓介

美濃導彦

池田克夫

京都大学工学部
情報工学教室京都大学工学部
高度情報開発実験施設京都大学工学部
情報工学教室

1 はじめに

文書画像の多くは、新聞に代表されるように、文字領域と写真領域が互いに重ならない矩形で配置されている。また、用いられる文字フォントは、その大きさ、種類などが限られている。しかし、カタログなどこれらの条件を満たさない文書画像も多く存在する。これらの文書はカラー刷りであることが多いので、処理対象となる画像(文字写真混在画像)は、カラー画像である。これらの画像から文字列を抽出する時には文字領域と写真領域が重ならないことを仮定している従来の方法は適用できない。

文字写真混在画像は、人間が情報を伝えようとする意図を持って文字を配置しているため、文字と背景は、明度、色相、彩度において明確に区別されている。そこで、文字領域内と写真領域内のエッジ強度の統計的性質に着目し、文字と背景の境界となる、文字エッジ点の候補を抽出する手法を提案する。

2 文字エッジ点のモデル

カラーイメージスキャナより入力した文字写真混在画像において、互いに隣接している画素を、RGBの各輝度を座標軸とする、RGB直交座標空間上の点に写像する。この2点間の、RGB直交座標空間におけるユークリッド距離を求め、その値が、あるしきい値(T)よりも大きい場合、その画素間にはエッジ点が存在するものとする。この時、しきい値(T)と画像中から抽出されるエッジ点数(E)の関係は、図1の様にモデル化できる。

しきい値が十分小さければ、写真領域中に存在する弱いエッジ点が多く抽出され、エッジ点数は多くなる。しきい値を大きくしていくと、写真領域内のエッジ点が増えるため、エッジ点数は減少する。しきい値がある値(X)を越えると文字領域内のエッジ点(文字エッジ点)のみが抽出されるため、エッジ点数は低い値で安定する。このような性質を持つしきい値-エッジ点数グラフ(Threshold - Edge Point Number Graph: 以下TEグラフと略す)において、 X よりもしきい値の小さい区間を混在区間、大きい区間を文字区間と呼ぶ。本研究の主眼は文字領域を抽出することにあるので、文字にかすれがなく、写真領域内のエッジ点を可能な限り除去するように文字区間を決定し、これを用いて文字エッジ点候補を抽出する。

3 文字区間の決定法

TEグラフにおける文字区間は、対象画像によって変動する。しかし文字区間は、直線で近似できると考えられるので、2次以上の微係数が0となる。そこで、TE 2次微分グラフを最小2乗法によって多項式近似し、そのグラフが横軸に接する値をもってしきい値 X を決定する。

3.1 最小2乗法によるTE 2次微分グラフの多項式近似

TE 2次微分グラフの多項式近似は、横軸に接することを保証するため次の $E_i(A, B, C)$ を用いる。

$$E_i(A, B, C) = A^{-1}(t_i + B)^3(t_i + C)$$

いま求めるA、B、Cの初期近似値を \tilde{A} 、 \tilde{B} 、 \tilde{C} として、

$$A = \tilde{A} + \Delta A, \quad B = \tilde{B} + \Delta B, \quad C = \tilde{C} + \Delta C$$

と表す。これらを用いて E_i を Taylor展開し、 E_i の測定値を q_i 、さらに簡単のため、

$$q_i - E_i(\tilde{A}, \tilde{B}, \tilde{C}) = q'_i \\ \left(\frac{\partial E_i}{\partial A}\right)_1 = A_i, \quad \left(\frac{\partial E_i}{\partial B}\right)_1 = B_i, \quad \left(\frac{\partial E_i}{\partial C}\right)_1 = C_i,$$

とする。ただし、 $\left(\frac{\partial E_i}{\partial A}\right)_1$ などは、 E_i をAで偏微分し、 $A = \tilde{A}$ 、 $B = \tilde{B}$ 、 $C = \tilde{C}$ を代入したものである。上の文字を用いて書き換えると、

$$q'_i = A_i \Delta A + B_i \Delta B + C_i \Delta C$$

より、連立1次方程式

$$\begin{pmatrix} \sum(A_i^2) & \sum(A_i B_i) & \sum(C_i A_i) \\ \sum(A_i B_i) & \sum(B_i^2) & \sum(B_i C_i) \\ \sum(C_i A_i) & \sum(B_i C_i) & \sum(C_i^2) \end{pmatrix} \begin{pmatrix} \Delta A \\ \Delta B \\ \Delta C \end{pmatrix} \\ = \begin{pmatrix} \sum q'_i A_i \\ \sum q'_i B_i \\ \sum q'_i C_i \end{pmatrix}$$

を解いて ΔA 、 ΔB 、 ΔC を求め、初期近似値を修正する。

さらに、修正されたA、B、Cを新たな初期近似値として ΔA 、 ΔB 、 ΔC を求める。これを反復することにより近似値は改善され、 $-B$ がしきい値となる。

3.2 初期近似値の推定と近似区間

写真領域中に、同じ強さを持ったエッジ点がある程度存在すると、TEグラフには小さな棚が現れる。これは、TE 2次微分グラフに、高周波成分として現れる。

これを考慮してTE 2次微分グラフが正の値をとる区間のうち最大幅のものに着目し、その左端のしきい値を $-C$ の初期近似値、右端のしきい値を $-B$ の初期近似値とする。さらに A は、TE 2次微分グラフと横軸で囲まれる領域の面積を用いて求める。

すなわち、

$$\int_{-C}^{-B} E_i(A, B, C) dt = \frac{1}{20} A^{-1} (B - C)^5 = \sum_{i=-C}^{-B} q_i$$

より、

$$A = \frac{(B - C)^5}{20 \sum q_i}$$

これより求まる値を A の初期近似値とした。

また、 E_i が近似式として有効なのは初期近似値によって挟まれた区間内であるから、この区間についてのみ近似を行なう。

4 結果と考察

図2に、図3の画像のTE 2次微分グラフと、その近似グラフを示す。高周波成分の乗ったグラフがTE 2次微分グラフである。近似グラフの右端をしきい値とし、これを用いて処理を行った結果を図4に示す。本手法によって、文字エッジ点候補が的確に抽出されていることが分る。

また、TEグラフから直接文字区間に相当する直線を抽出し、その左端をしきい値として用いることもできる。そこでRamerの手法[1]によってTEグラフを切片線形近似し、得られた切片群の中で横軸への写像が最大となる直線を文字区間に相当する直線とし、その左端をしきい値とする手法も実現し、前述の方法と対照実験を行った。その結果は、どちらの手法で決められたしきい値を用いても概ね良好であった。

ただし切片線形近似では、直線とみなすためのパラメータを与えなくてはならない。このパラメータを大きくとれば、しきい値が過度に小さくなる。逆にパラメータを小さくとると各切片が短くなり、適切な文字区間が得られない

ことがある。実験では、このパラメータは試行錯誤的に決定した適当な値を用いた。

しかし多項式近似による方法では、この種のパラメータが不要であるため、画像に対して安定に、適切なしきい値が得られる。

また、文字領域を含まない画像を処理した場合、TEグラフには混在区間のみが存在する。切片線形近似を用いると、混在区間にしきい値をとってしまう。しかし多項式近似を用いると、混在区間の端に適切なしきい値が得られる。

5 おわりに

文字写真混在画像に対して文字エッジ点のモデルを作成し、文字エッジ点候補を抽出する手法を提案した。本手法はパラメータを一切与える必要がないため、画像に対して安定に処理が行えるという特徴があるが、近似に用いる多項式の自由度が限られているため、しきい値が最適にならない場合があった。今後はこれを前処理として、さらに文字パターンを抽出する手法を検討していく予定である。

参考文献

- [1] U.E.Ramer: "An interactive procedure for the polygonal approximation of plane curve," CGIP1, pp.244-256, 1972.
- [2] 谷口, 河口: "二値画像の複雑さと多値画像のしきい値処理に関する考察," 信学論, Vol. J70-D, No. 1, pp.164-173, 1 1987.
- [3] 塩: "情景中文字の検出のための動的二値化処理法," 信学論, Vol. J71-D, No. 5, pp863-873, 5 1988.

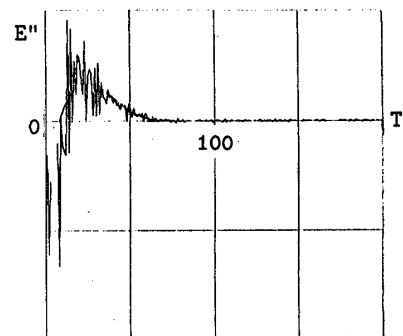


図2: TE 2次微分グラフとその近似グラフ

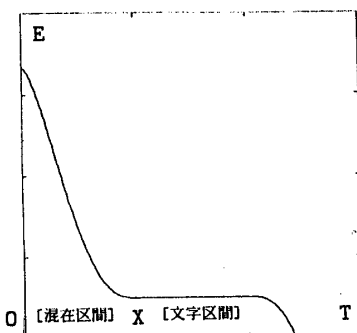


図1: TEグラフのモデル

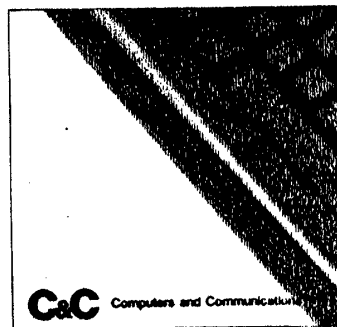


図3: 原画像

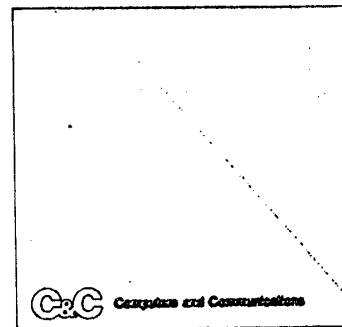


図4: 処理結果