

6E-8

変動エントロピーによる 文字の手書き変動量の推定

長谷 博行、米田 政明、酒井 充、尾塩 浩
(富山大学工学部)

1. はじめに

手書き文字認識の研究が一応の成果を得た現在、文字変動に関する研究報告は久しくない。筆者らは文献(1)において相関法的な考え方で変動エントロピーなるものを提案したが、これは文字としてではなく文字画像としての変動評価量であった。即ち、評価量は文字線幅に依存したので本質的な(線幅に依存しない)手書き変動を表しているとはいいがたかった。これはETL8B, ETL9Bを使った実験で確かめられている⁽¹⁾。

しかし、筆者らが提案した変動エントロピーは、画像枠の大きさには依存せず、文字パターンの拡大・縮小にも不変であり、比較すべき標準パターンも必要とせずかつ一意に値が定まるという利点をもっていた。

本稿では、変動エントロピーを使って手書きによる文字の形や大きさのばらつきを反映する本質的な変動量を変動モデルと対応をとることによって推定を試みる。

2. 変動エントロピー

表記を簡単にするために1次元画像で考える。m枚の2値画像を重ね合わせたとき、第1画素に累積された黒画素の数を $P(i)$ と表す。このとき P を変動パターンと呼ぶが、そのエントロピー I_P は

$$I_P = - \sum_i \frac{P(i)}{N} \log \frac{P(i)}{N}$$

$$= - \sum_i \frac{P(i)}{N} \log \frac{P(i)}{m} + \log B$$

と表される。ここに $N = \sum P(i)$ であり、 B は1枚当りの平均黒画素数($= N/m$)である。上式の第1項は変動分を表し、第2項は大きさを表している。

そこで、変動エントロピー h_P を(1)式で定義する。

$$h_P = - \sum_i \frac{P(i)}{N} \log \frac{P(i)}{m} \quad (1)$$

この性質は上述の通りである。

3. 基本的考え方

1次元パターン g に変動を与える確率密度関数 $f(x; \sigma^2)$ を適用することにより g の変動パターン $\tilde{p}$ が作り出せる。但し、 σ^2 は分散で変動の大きさを表す。ここで、 σ を基準変動量ということにする。

$$\tilde{p}(x) = \int f(\tau; \sigma^2) g(x-\tau) d\tau \quad (2)$$

ここに、パターン g は(3)式とする。

$$g(\tau) = \begin{cases} 1 & -a \leq \tau \leq a \\ 0 & \text{else} \end{cases} \quad (3)$$

この $\tilde{p}$ からエントロピー $I_{\tilde{p}}$ をもとめることができ、変動エントロピー $h_{\tilde{p}}$ は(4)式で計算することができる。

$$h_{\tilde{p}} = I_{\tilde{p}} - \log 2a \quad (4)$$

文献(1)より変動エントロピーはパターンの伸縮変換に不変なので、 $h_{\tilde{p}}$ は a/σ に依存する。この $h_{\tilde{p}}(a/\sigma)$ を f を適用したときの変動エントロピー関数といい、変動の基準関数とする。もちろん f の形により基準関数は異なる。ここで $2a$ はパターンの大きさであるが、文字などの線図形では線幅に相当すると考えるのが妥当である。

さて、文字パターンは細線化や太め処理を施しても手書き変動には影響がないと思われる。実際に、ETL8B, ETL9Bは平均文字線幅が異なるのでそのまま変動エントロピーを計算しただけでは変動の比較はできない。そこで画像処理的に線幅を変えて変動エントロピーを計算し、補間により同一線幅のときの変動エントロピーを求めることで両データベースの変動を比較することができる。またデータベースから得られた変動エントロピー関数 h_P と基準関数 $h_{\tilde{p}}$ の対応を取ることが可能ならば基準変動量 σ を(5)式で計算することができる。

$$\sigma = a / \eta \quad | \quad a = d/2 \quad (5)$$

ここに、 η と d は $h_{\tilde{p}} = h_P$ となるとき値である。

図1にその様子を示す。対応の可否は両関数の形状が類似していることであり、 h に対する σ がほぼ一定値を取ることである。

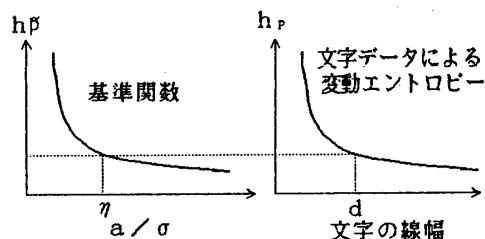


図1 モデルとの対応

4. 実験および考察

モデルは f として分散 σ^2 の正規分布関数と一様分布関数を用いた。

$$f(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

$$f(x) = \begin{cases} 1/2\sqrt{3}\sigma & -\sqrt{3}\sigma \leq x \leq \sqrt{3}\sigma \\ 0 & \text{else} \end{cases} \quad (6)$$

(2)~(4)、(6)式を用いて計算した結果を図2に示す。図では横軸が a/σ 、縦軸が変動エントロピーである。図より一様分布関数の方が大きな値をとることがわかる。

実際の文字データはETL8BとETL9Bからひらがな71文字種、各文字種につき100個用い、(1)式で計算した。線幅の制御は細線化した文字パターンから4段階の太め処理をすることにより行った。データ間の値は3次の準エルミート補間式を用いた。なお細線化(Hilitchのアルゴリズム)や太め処理はSPIDERを利用した。結果を図3に示す。各変動エントロピー値は71文字種の平均である。図よりETL9Bの方がETL8Bより大きな値をとる(変動が大きい)ことがわかる。原データの平均線幅はETL8Bが3.407、ETL9Bが3.877であり、図の曲線からはそれら点の変動エントロピー値が反転していることがわかる。文献(1)で得られたETL9Bの方が変動が小さいという結果はこのことが原因している。

次に、データから得られた変動エントロピー関数とモデルから計算した基準関数と対応付け、基準変動量を求めるを試みる。(5)式から計算した変動エントロピーに対する基準変動量のグラフを図4に示す。曲線はスプライン補間が施されている。図中4つの曲線が描かれてあるが「+」記号はETL8Bと正規分布関数との対応、「 Δ 」はETL9Bと正規分布関数、「*」はETL8Bと一様分布関数、そして「 $\diamond$ 」はETL9Bと一様分布関数との対応である。

図4において最も信頼のおける h_p の区間はおよそ1から2の区間である。それは図3から h_p が1以下または2以上では測定点が少なく補間計算の誤差が大きいためである。この1から2までの区間では、正規分布モデルの場合は σ の値がフラットではなく不適当な対応である。それに対して一様分布モデルでは両データベースともグラフが比較的フラットであり、実際の分布に近いモデルであろうと思われる。文献(1)においても同じ見解が得られている。結果的にETL8Bの基準変動量は約2.5、ETL9Bでは約2.9となった。これらの値は手書きの性質として2.5画素から2.9画素の標準偏差を持つ変動で文字が書かれると考えられ、文字データベースの品質を反映する値であろうと思われる。

5. まとめ

文献(1)では変動エントロピーを単純に用いたのでは異なったデータベース間の手書き変動の評価は困難であるという結論であった。それは相関法では線幅の影響を取り除いて変動を測定するのは困難であったからである。本稿ではそれを拡張し、基準変動量を求めることによって異なったデータベース間の品質評価を試みた。その結果ETL9BはETL8Bより手書き変動が大きいという、認識実験の結果を裏付けるような結果が得られた。本稿で提案した手法は線幅の制御が必要なため計算量が多い。しかし、データベースの客観的な品質評価ができ、

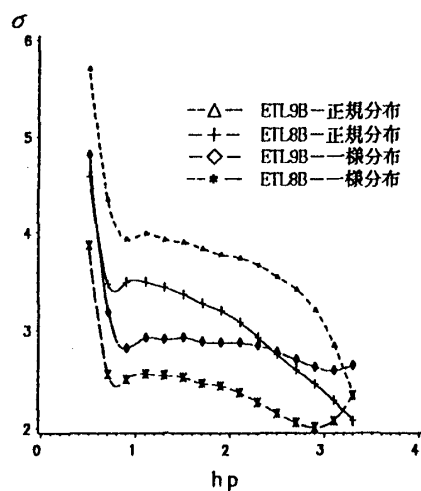
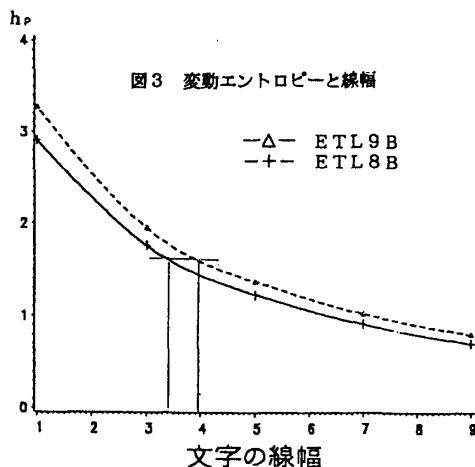
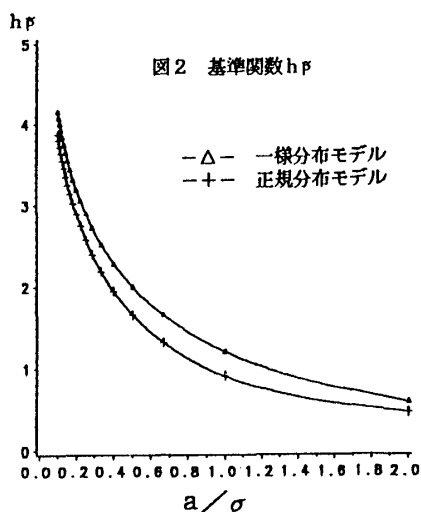


図4 変動エントロピーと基準変動量

認識実験のように多くの実験条件を設定しなくてもよいという利点がある。

[文献] (1)長谷、米田、酒井、吉田: "変動エントロピーによる文字変動の評価", 信学論D, Vol.71D, 6 (1988).