

仮説検証型認識知識処理における処理速度と精度の評価実験

6K-9

下村 秀樹

NEC ヒューマンメディア研究所

1. はじめに

文字認識技術の実用のためには、知識処理による認識候補の補正が不可欠である。また、より高い読み取り性能を追求するためには、パターン処理（個別文字切り出し／認識）や知識処理の要素技術だけでなく、両者の関係にまで踏み込んだ最適性を議論する必要がある。

従来の代表的な知識処理として、ボトムアップ型の方式がある。パターン処理の出力する候補群を単語辞書や単語接続関係といった知識と照合し、最もよく合致するものを出力する構成である。しかし、このような直列的な構成では、限られた処理時間での読み取り精度向上に限界がある。例えば、パターン処理出力の正解含有率を高くすることで知識処理での正読率を高めることはできる。しかし、そのためにパターン処理の出力する候補が多くなると、知識処理は膨大な処理時間を必要とし、しかも増加した処理時間の多くは正しい読み取りに直接関係しないという無駄が生じる。

これに対し、パターン処理と知識処理の関係という観点から、我々は仮説検証型の認識知識処理を提案している。概念図を図1に示す。この方式では、従来のボトムアップ型知識処理を読み取り候補を生成するための仮説生成処理と位置付ける。

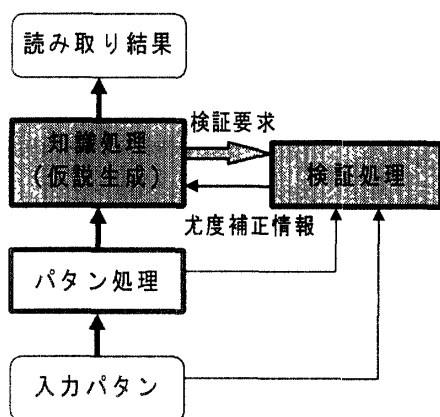


図1 仮説検証型認識知識処理

An evaluation of processing speed and precision of a knowledge-based processing method based on a hypothesizing and verification hybrid approach for string reading.

Hideki SHIMOMURA

Human Media Research Labs., NEC corp.

そして、そこで生成した仮説を文字認識／パタン的な情報から検証して尤度を補正し、候補選択や棄却を精密に行う処理を加えたことを特長としている。この構成において、知識処理はパターン処理に対して、仮説生成に必要な候補以上の過度な正解出力を要求しない。その結果、従来のボトムアップ型に比べ、パターン処理および知識処理での無駄な処理を削減することができる。そして、削減した時間の一部を出力判定に必要な不可欠な詳細処理（検証処理）に配分することで、全体として処理速度／精度の向上が期待できる。

すでに文献[1]でこの基本的なアイデアを報告し、検証処理の読み取り精度への寄与を評価した。本稿では、処理時間と読み取り精度の評価結果を述べる。

2. 住所読み取り実験システム

前記の方式を住所読み取りに適用した実験システムをすでに構築した[1]。入力した画像に対してパターン処理を行い、仮説検証型のアプローチで知識処理を行う。仮説生成の効率的アルゴリズムとして文字タグ法[2]を使用している。また、検証処理として、次の3つを実装している。

下位候補検証：仮説に含まれる推定文字が文字認識候補の下位に存在するかどうかを検証する

画数検証：認識不良領域の候補文字の平均画数と仮説に含まれる推定文字の画数が見合うかどうかを検証する

ブロックサイズ検証：認識不良領域の画像サイズが仮説に含まれる推定文字列の文字数に見合うかどうかを、前後の認識成功ブロックのサイズから検証する

3. 実験

実験は、仮説生成に使用する文字認識候補数を上位1,3,5,7,10候補と変化させ、地名判定の処理時間と正読率／誤読率を測定した。正読率と誤読率は次の定義である。

$$\text{正読率} = \text{正しく読めた画像数} / \text{全画像数}$$

$$\text{誤読率} = \text{誤って読んだ画像数} / \text{全画像数}$$

ほぼ同じ尤度を持つ地名候補の競合、あるいは全地名候補の尤度がしきい値を超えない場合はリジェクトとした。下位候補検証では、第12位候補までを確認範囲とした。また比較のため、検証

処理なしの場合の処理時間と精度も測定した。これは従来のボトムアップ型の処理に相当する。

実験に用いた画像は、東京のある区内向けの葉書上にかかれた住所 494 枚である。登録した地名の単語数は約 50 である。実験は、NEC 製ワークステーション EWS4800/360AD で行った。

実験結果をグラフで示す。図 2 は使用文字候補数と処理時間の関係である。図 3 と図 4 は、それぞれ処理時間と正読率／誤読率の関係で、文字候補数が 1,3,5,7,10 のときの値をプロットした。

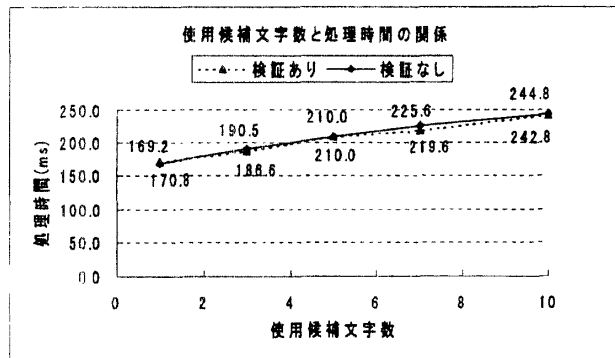


図2 候補文字数と処理時間の関係

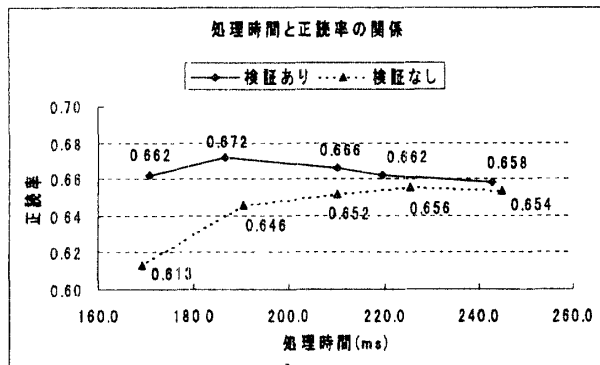


図3 処理時間と正読率の関係

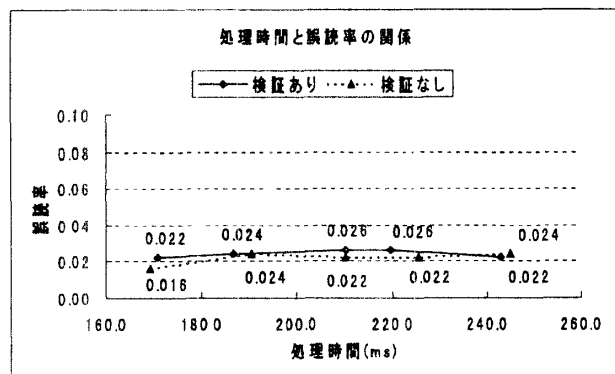


図4 処理時間と誤読率の関係

4. 議論

(1)使用候補文字数と処理時間の関係 (図 2)

仮説生成に使う候補文字数が増加すると、内部での仮説候補探索空間が広がるので、処理時間が

増加する。しかし、文字タグ法[2]の効果によりその増加割合はほぼ線形に収まった。また、検証の有無による処理時間の差はほとんど見られないことから、実装している検証処理は負担の軽い処理であることがわかる。

(2)処理時間と正読率の関係 (図 3)

処理時間が同じなら、検証ありの方が高い正読率となった。特に短い処理時間でも安定して高い性能となっている。例えば認識 1 位候補のみを使った場合、処理時間はほぼ同じ (170ms) で約 5% も高い正読率となった。これは、仮説検証型処理によって、処理時間が正しい読み取りのために適切に配分できた結果であると言える。

また検証ありの場合、仮説生成に使う候補数を増やしすぎても正読率は上がらず、逆に下がるという現象が見られた。実験では、文字認識 3 位までを使用したところをピークに、あとは正読率が減少している。これは、文字認識の尤度や順位を無視してすべての文字候補を同じ重みで処理したことにより、候補数の増加に伴って競合が増えたためである。認識上位候補文字を優先するように条件を変更すれば、正読率は上がると思われるが、誤読も増加するのでバランスの問題となる。いずれにせよ、パターン処理の候補を増やすことが知識処理の時間の増加とともに、処理精度にも悪影響を及ぼす可能性があるという重要な教示である。

(3)処理時間と誤読率の関係 (図 4)

処理時間による誤読率の変化は少なく、また検証の有無による差もほとんど現れなかった。処理時間の増加に関しても、検証の有無に関しても、改善と改悪がほぼ同数であった。検証によって起こる改悪に関しては、検証の適用条件を見直すことで性能向上が期待できるだろう。

5. おわりに

本稿では、仮説検証型認識知識処理を住所読み取りに適用した際の、処理時間と精度の評価を報告した。検証を導入することにより、限られた処理時間を判定に必要な箇所に適切に配分することが可能となり、短い処理時間内で高い読み取り精度が得られることを実証した。今回は認識多候補について検討したが、今後は文字切り出し多候補と仮説検証方式の関係なども検討したい。

参考文献

[1]下村他:ボトムアップ/トップダウン処理を融合した手書き文字列読み取り知識処理,情処 NL 研報告, 97-NL-118-20 (1997)
 [2]福島他:手書き文字列読み取りのための単語列探索アルゴリズム-文字タグ法-,情処論文誌, Vol.37, No.4, pp.500-510 (1996)