

## 印刷文書認識システム AutoReco/2

2L-3

## -テキストプロセス-

山下晶夫<sup>†</sup>、加藤美治<sup>†</sup>、平山唯樹<sup>†</sup>日本アイ・ビー・エム株式会社<sup>†</sup>東京基礎研究所/<sup>†</sup>大和研究所

## 1 はじめに

AutoReco/2のテキストプロセスではOCRカード上で行なわれる文字認識とホストであるPC側で行なわれる日本語後処理、英語認識、ユーザー後処理、表示のユーザーインターフェースが同期を取りながら協調動作している<sup>(1)</sup>。本稿ではこれら制御の仕組みとアウトラインフォントを用いて認識結果をWYSIWYG表示しているユーザーインターフェースを中心に各処理を概説する。

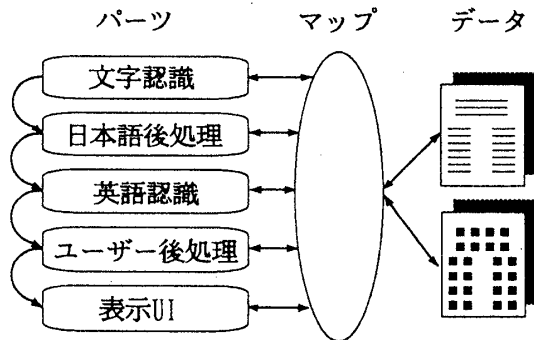


図1: パーツの協調動作

## 2 モジュールの協調動作

AutoReco/2では、各処理モジュールはパーツと呼ばれるオブジェクトとして定義されている。図1にパーツの協調動作の仕組みを示す。レイアウト解析が終了した時点でデータオブジェクトとしては、行および行の集合であるブロックが生成されている。各パーツはマップと呼ばれるデータオブジェクトの集合を管理しているオブジェクトにメッセージを送ってデータにアクセスし、新しいデータオブジェクトを生成したり、変更したりする。

文字認識パーツに起動のメッセージが送付されるとパーツは認識対象となるブロックをマップに問い合わせる。ブロックはブロックに対して行なわれるべき処理の

フラグ、処理が終了したか否かを示すフラグを持っていて、パーツはこのフラグによってデータを処理するかしないかを判断する。起動は1個のブロック毎に行なわれ、ブロックの位置座標から各文字の切り出し、認識を行なった後<sup>(2)</sup>、文字データオブジェクトを生成し、新たにマップに登録する。文字データオブジェクトには、文字の位置座標、最大5位までの候補文字が格納されている。データの登録と同時にブロックには文字認識終了のフラグが立てられ、終了メッセージが次のパーツ(日本語後処理)に送付される。

引き続きおこなわれるプロセスは、前のパーツの終了メッセージを受けて起動される。ブロックに処理フラグが立っていない場合には、パーツは処理を行わずに直ちに終了メッセージを発行する。日本語後処理は、1ブロックの単位で行なわれるわけではなく、文字行からストリーム(句読点等の区切り文字で囲まれた文)への変換が必要である。後処理パーツでは起動のメッセージがくる度にマップに認識結果を要求しデータを内部に蓄える。ストリームが生成できるだけのデータが蓄積されると1ストリームに対して日本語辞書(約10万語)と形態素レベルでの制約を用いた後処理を行なう<sup>(3)</sup>。後処理では最大5個の候補文字から単語列の最尤推定を行ない、認識誤りの可能性がある文字については下位候補との入れ換えや文字属性の変更(警告属性を付ける)を行なう。1個のブロックの全文字について後処理が完了した時点で、ブロックの終了フラグを立てて次のパーツに終了メッセージを渡す。日本語後処理では、文字種や未知語の情報から英語部分の推定も同時に行なっている。また単語の切り出し及び辞書引きを行なう時に、単語列に品詞コードがふられるので、自立語を出現頻度と共に検出し、キーワード候補として記録している。

英語認識、ユーザー後処理、表示のパーツは1個のブロックを単位とした処理である。パーツが起動されると、処理可能なブロックを1個マップに要求し、処理した後に次のパーツに終了メッセージを送付する。このプロセスを処理可能なブロックが無くなるまで繰り返す。英語の認識では、英語用のテンプレートを用いて、日本語後処理により検出された英語部分について再度文字カラムの切り出し、認識を行なう。さらに英語の単語辞

書を用いてスペリングチェックを行ない、日本語と同様に認識エラーの自動訂正を試みる。

ユーザー後処理はユーザーに解放されたパーツで、通常は何も処理しないでメッセージをやりとりしているだけだが、必要に応じてプログラムを加えることでユーザー独自の処理を行なうことができる。

最後のパーツである結果表示では後処理の結果を反映している。後処理が認識結果を修正した場合と修正してはいないが誤りの可能性が高い場合にはそれぞれ黄色、赤色で認識結果を表示し、オペレータに注意を促すようにしている。結果表示のパーツからの終了メッセージはメインプログラムに返される。メインから要求を出した全てのブロックの処理が終了した時点でテキストプロセスの終了とする。

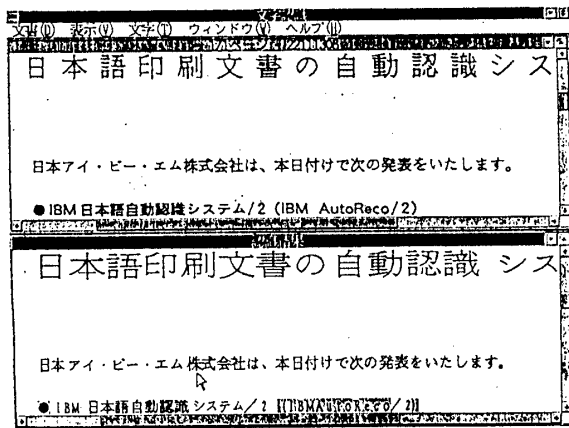


図 2: WYSIWYG 表示の例

### 3 WYSIWYG ユーザーインターフェイス

AutoReco/2では、認識結果の表示の際にアウトラインフォントを用いて、位置と大きさをオリジナルイメージと同じように表示している。OCRでは文字情報だけでなく、位置や大きさについての情報も検出できるので、図2に示したように、上段ウィンドウのイメージと下段のテキストの比較がしやすいだけでなく、レイアウト情報を出力してDTPシステムの入力として対応することができる。実際に検出された位置にアウトラインフォントをそのままの大きさに当てはめて出力するとベースラインが一致せず、大きさもまちまちになるので必ずしも美しい表示とはならない。そこで次に示すような処理を施し、出力結果の位置座標や大きさを整形している。

1. OCRで検出された文字の外接矩形はいわゆるフォントのボディに相当する。フォントのボディとフォントボックスの比率は文字種により固定されているので、各文字の外接矩形から、フォントボックスの大きさを計算する。
2. 各フォントボックスのベースライン、高さを揃え、文字種(平仮名、漢字、カナ、英数字)別にグルーピングする。
3. グループ毎に文字数で文字幅を均等割する。
4. 整形したフォントボックスの大きさを、アウトラインフォントを表示する。

イメージのウィンドウとテキストのウィンドウは同じサイズで表示され、一方をスクロールさせると他方も連動する。マウスクリックで文字または文字列を指定することにより、認識結果の修正を行なうこともできる。

### 4 おわりに

テキストプロセスでのパーツは行を単位としたデータによる事象駆動型の制御を受けており、認識とその他の処理は並行に処理されるため、後処理を含めた認識スピードは文字認識単独のスピードとはほぼ同等である。また、表示のユーザーインターフェースも1個のパーツとしてインプリメントされているために、認識その他の処理が行なわれている最中でも結果の修正を同時に行なうことができる。認識結果はWYSIWYG表示され、確認修正時に元のイメージとの比較が容易である。本稿のような協調動作の仕組みでは、パーツ間のデータの授受や同期のインターフェースが共通化されているため、ユーザー後処理の例のように、パーツを追加したり差し替えることが容易に実現できる。今後はこのような仕組みの下で、パーツの追加や改良を行なっていく予定である。

### 参考文献

- [1] 天野ほか: マルチメディア文書入力のための文書画像システム:DRS, 情報処理学会マルチメディア通信と分散処理研究会, 48-6, pp.41-48 (1991).
- [2] 高橋: 細線連結素方向による簡易手書漢字認識, 信学技報PRL82-8, pp.57-62 (1982).
- [3] 伊東: OCR入力された日本語文の誤り検出と自動訂正, 情報処理学会論文誌, Vol. 33, No. 5, pp.664-670 (1992).