

視覚障害者のための環境内文字読み上げシステムのプロトタイプ開発

Development of Text Reading Prototype System from Natural Scene images for Blind Person

研究学生 プイ チュオン ミン 指導教員 江崎修央

1. まえがき

近年では福祉情報工学分野に関する研究が盛んに行われており、GPS利用した視覚障害者向けヒューマンナビゲーションシステム[1]やスクリーンリーダと併用して文章作成を行うワープロソフトウェアも開発されている。

ところで、我々が生活する空間上には、交通標識、看板や値札など、多くの文字情報が存在しているが、当然のことながら視覚障害者はこれらの情報を自身で得ることはできない。そこで我々は、先行研究[2]を参考にしつつ、視覚障害者向けの環境内文字読み上げシステムの開発を行う。本稿では、システムの基本設計デモンストレーション用プロトタイプシステムの開発について述べる。

2. システムの基本設計

本システムのハードウェアにおける構成要素は、ノートパソコンなどの携帯可能な情報端末とパン・チルト・ズームイン機能付きのカメラである。図1にシステムの流れを示す。画像のキャプチャ、文字候補領域の抽出、文字認識、文字情報の読み上げを順に行う。ここで初期キャプチャ画像に含まれる文字領域は文字が小さいことが予想されるため、文字認識するには十分な解像度が得られない可能性がある。そこで、パン・チルト・ズームイン機能を利用して文字領域を解像度の高い画像で再取得する。ここから文字認識を行って、認識した文字を音声で出力することにより視覚障害者に文字情報の提供を行う。

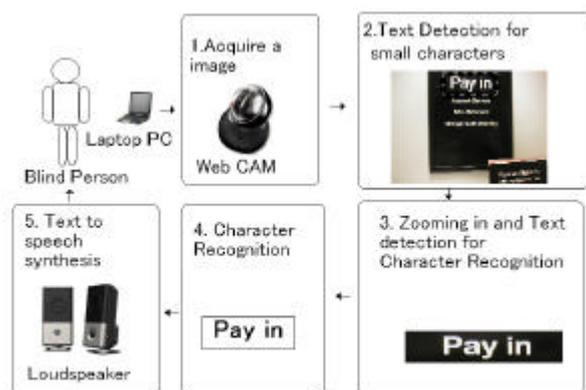


図1 システムの流れ

3. 文字領域の抽出

本システムでは初期画像に対してローカル領域への二値化適用する手法と Sobel エッジを利用する手法[3]の二つの手法を適用し、ズームイン後は Sobel 手法のみを利用して文字列の抽出を行う。

3.1 ローカル領域への二値化手法

画像を 32x32pixel 領域に分割し、各領域を判別分析手法により二値化する。この手法は全ての RGB チャンネルに適用され、それぞれの領域の二値化閾値と全体画像の閾値を比較し、その領域の閾値を決定する。二値化した後には

8色のみからなる画像ができ、画像のラベリングを行って、文字領域を抽出する。

3.2 Sobel フィルタによる手法

本手法は高さ 30pixel 以上の文字列を対象としており、キャプチャされる画像を Sobel フィルタによりエッジ抽出する。単純にグレイ画像をエッジ抽出するのは効果がよくないので、この Sobel フィルタは全てのチャンネル RGB に行われている。次に画像を二値化して、二値化した画像をラベリングし、各領域を比較する。この処理により認識用文字領域を抽出できる。

4. プロトタイプシステムの構築

これまでに述べてきた文字抽出手法の有効性を確認するためにノートパソコンおよびパン・チルト・ズームイン機能付きの Web カメラ (Logitech: QVR-1) を利用してデモンストレーションシステムを構築した。ユーザは文字情報を得たい任意の空間上で「キャプチャ」ボタンを押すことによりユーザの肩に置かれた Web カメラがその空間上にある文字列を自動的に走査するものである。キャプチャされる画像のサイズは 640x480 ピクセルであり、1 フレームあたり約 300 ミリ秒で文字列の走査が可能である。

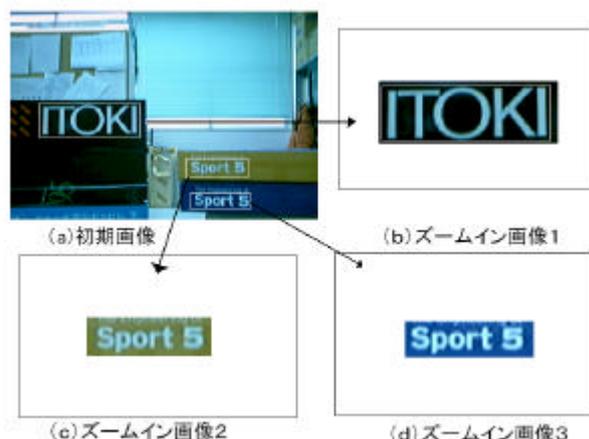


図2 文字列の走査

5. まとめと今後の課題

プロトタイプシステムの構築によりシステムの実現に対する見通しがえられた。今後はズームイン後の処理としてさまざまなフォントやサイズに頑強な文字認識処理手法の開発と詳細なシステムデザインを行っていく。

参考文献:

- [1] 音声を用いた歩行者ナビゲーションシステム東芝レビュー、Vol.59, No.4 2004
- [2] 視覚障害者のための環境内の文字情報抽出システム、鈴木、大西他、電気学会論文誌(C)、Vol.124-C, no.6, pp.1280-1287, 2004-6
- [3] Text Detection From Natural Scene Image: Towards a System for Visually Impaired Persons, Nobou Ezaki 他、International Conference on Pattern Recognition 2004