

ハンドマウスの応用： ユニバーサルリモコンと実環境中の文字情報取得*

蔵田武志[†]興梠正克[†]遠藤健[‡]坂上勝彦[†][†]産業技術総合研究所 知能システム研究部門[‡]メディアドライブ (株)

1 ハンドマウス

タッチパネルは、ウェアラブル/モバイル情報端末の中で比較的大型の表示部を備える PDA やタブレット PC に適した入力インターフェイスのひとつである。タッチパネルを用いると、表示部中の操作対象を手や指、ペンなどで直接触れて操作できる。このようなタッチパネルと同様の機能をもち、頭部着用型の小型ディスプレイ (HWD: Head Worn Display) に適したインターフェイスとして、筆者らは、“ハンドマウス”と名付けた入力インターフェイスを開発している。これは、ウェアラブルカメラ、主に頭部着用カメラ (HWC: Head Worn Camera) で得られた映像中の着用者の手や指を検出・追跡することで、その手や指にポインティングデバイスの役割を持たせようとするものである。

ハンドマウスは、キーボードのように長時間継続して多くの情報を入力する目的には向いていないが、とっさに何かを指示したい場合や簡単な指示で十分な場合にはとても適したインターフェイスといえる。このようなインターフェイスは文献 [5] ですでに述べられているが、具体的な実現方法については述べられていない。筆者らは、逐次型モンテカルロ追跡法 [2] の枠組みを用いて、動的な色モデル生成に基づく手領域と背景領域の大まかな分離処理及び手の輪郭モデルに基づく追跡処理を効果的に組み合わせることで、高速で安定な手や指の検出・追跡を実現している [4]。

本稿では、筆者らが開発中の視覚機能に優れたウェアラブルシステムである *VizWear* システム [1] 上に実装されているハンドマウスインターフェイスの応用例として、ユニバーサルリモコンの操作と実環境中の文字情報取得の 2 つについて述べる。

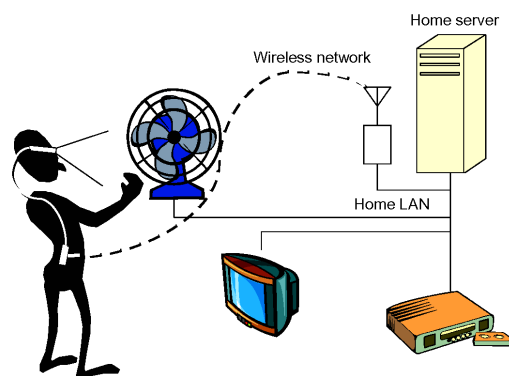


図 1: 見ている方向の家電を素手で操作できるユニバーサルリモコンシステム



図 2: ハンドマウスによるテレビ操作の例

2 ユニバーサルリモコン

筆者らは、事前知識のパノラマ画像群と HWC からの入力画像とを高速に位置合わせすることで着用者の位置と方向を推定し、同時に着用者視点に関する 2 次元の注釈情報を重畳提示するパノラマベースドアノテーション [3] と呼ばれる手法を開発している。Bluetooth などの無線デバイスのみで操作対象の方向を計測するのは一般に困難であるが、この手法を用いると、テレビやエアコンなどの操作対象の方向を見るだけでそれらの仮想リモコンを HWD に提示させることができる。この機能とハンドマウスを組み合わせると、図 1 に示すように、何も手に持たなくても見たものを操作できるユニバーサルリモコンを実現することができる。

VizWear システム上に実装されているユニバーサルリモコンの動作例を図 2 に示す。この例では、まず、パ

*Applications of The HandMouse: Universal Remote Control and OCR in Real Scenes

[†]Takeshi Kurata, Masakatsu Kourogi, Katsuhiko Sakaue, Intelligent Systems Institute, AIST, kurata@ieee.org

[‡]Ken Endo, Media Drive Corporation

ノラマベースドアノテーション機能により、着用者がテレビの方向を向いたことをシステムが認識し、HWDにテレビの仮想リモコンを表示している。着用者はそのリモコンをハンドマウスで仮想的ではあるが直接触れることでテレビを操作している。図左は電源を入れ、図右はボリュームを上げている様子を示している。ここでは、親指と人差し指が開いている状態をポインティング、閉じている状態をクリックとしている。

3 実環境からの文字情報取得

ウェアラブル/モバイル情報端末にフルキーボードを搭載することは一般に困難なため、如何に簡便に効率よく文字情報を入力するかが問題となる。本稿では、ハンドマウスを用いた実環境からの文字情報取得方式を提案する。これは、既にさまざまな方式が提案されている小型キーボードやペンタイプのインターフェイスとは競合するものではなく、補完し合うものと考えられる。本方式の概略図を図3に示す。まず、着用者はHWDに表示されている映像から、文字を含む範囲をハンドマウスで指定する。指定された範囲のみに文字認識手法を適用し文字情報を取得する。得られた文字情報は、ウェブ検索のキーとして利用できるのももちろん、図3のように、翻訳サービスや地名などに基づくナビゲーションサービス、また、施設予約サービスなどに適用できる。

VizWearシステム上に実装した文字情報収集機能の動作例を図4に示す。現在は、文字認識の範囲指定方法として、ハンドマウスでドラッグして得られた矩形領域を用いている。文字認識に使用する映像は、HWCで得られたライブ映像でも、すでに蓄積されたものでもよい。この図の例は、左手に持った本を見ながら右手で範囲を指定しているところである。

本システムの文字認識部には、メディアドライブ社の活字文字認識エンジンを使用している。この文字認識エンジンは、一般の手書き文字認識エンジンと同様、単語辞書を用いた文字認識が可能である。小型カメラで撮影された映像は一般に低画質であるため、このような単語情報の利用は認識率の向上に有効である。ただし、現状では撮影の仕方や矩形領域の位置によっては正しい認識結果が得られない場合もあり、今後の改良が必要である。

4 むすび

VizWearシステムは無線LANを介したクライアン

ト/サーバシステムであり、ハンドマウス機能は、ウェアラブルクライアントでの画像入力、結果出力処理、サーバでの手の追跡処理により構成されている。この構成はフレームレートの面では有利だが遅延の面では問題がある。現在の実装では、0.5~1秒の遅延があるため、快適に使えるようになるまでには慣れが必要である。今後は、クライアントとサーバでの処理の最適配置を決めるために、フレームレートと遅延がそれぞれ使用感にどう影響するのかを調査する必要がある。また、HWDとHWCの組み合わせ以外にも、ハンドマウスと同様の技術を適用できる応用について検討する予定である。

謝辞：本研究の一部は、経産省リアルワールドコンピューティング(RWC)プログラムの支援による。また、有益な議論を頂いた大隈隆史氏、加藤丈和氏に謝意を表す。

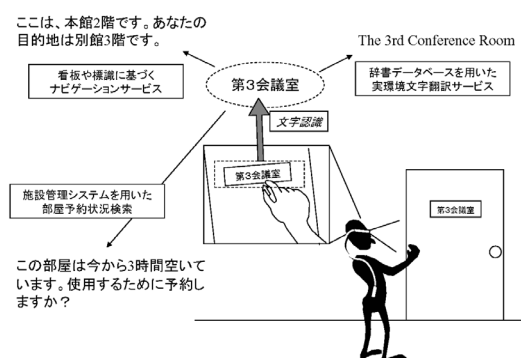


図3: ハンドマウスを用いた文字情報取得とその利用



図4: 実環境からの文字情報取得の例
(左)範囲指定中(右)文字認識結果の表示

参考文献

- [1] VizWear, <http://unit.aist.go.jp/is/hcv/vizwear/>.
- [2] Isard, M. and Blake, A.: ICONDENSATION: Unifying low-level and high-level tracking in a stochastic framework, *Proc. 5th ECCV*, pp. 893-908 (1998).
- [3] Kourogi, M., Kurata, T. and Sakaue, K.: A panorama-based method of personal positioning and orientation and its real-time applications for wearable computers, *Proc. 5th ISWC*, pp. 107-114 (2001).
- [4] 蔵田武志, 興梠正克, 加藤丈和, 大隈隆史, 坂上勝彦: ハンドマウスとその応用: 色情報と輪郭情報に基づく手の検出と追跡, *映像学技報*, VIS2001-103, Vol. 25, No. 85, pp. 47-52 (2001).
- [5] Starner, T., Mann, S., Rhodes, B., Levine, J., Healey, J., Kirsch, D., Picard, W. R. and Pentland, A.: Augmented Reality Through Wearable Computing, *Presence*, Vol. 6, No. 4, pp. 386-398 (1997).