

# 高性能携帯端末を用いたカメラ翻訳システムの提案

衛門 一樹<sup>†</sup>西山 裕之<sup>†</sup><sup>†</sup> 東京理科大学大学院理工学研究科

## 1 はじめに

近年、カメラを用いたインタラクションに関する研究は増えつつある。その理由として、携帯端末の高性能化により端末のカメラを用いてウェアラブルなアプリケーションを提供しやすくなった。その例として PACER[1] という携帯端末上で映した文書に対してデジタルな注釈表示やその文書内のコンテンツを操作するシステムなどが登場している。従来は見て調べる行為を人が行いコンピュータに入力していたが、カメラを用いることでコンピュータが直接画像から情報を抽出し利用することが可能となった。

またそのようなカメラを用いたインタラクションを翻訳に役立てる研究も存在している。Canedo[2] は、看板内の文字を英語からスペイン語へ翻訳するシステムを携帯端末上で実現している。また Jian[3] は携帯端末向けに行える文字抽出を研究し実用例で英語から中国語へ翻訳している。しかしこれらは、実際にユーザが使用する環境をあまり考慮していない。どちらの研究も画像全体から文字を抽出しようとしているが、実際翻訳する際は様々で、例えば一つの単語の意味が知りたい場合では画像全体をその単語だけで埋める必要がある。これはつまりユーザに対するインタラクションが考えられていないと言える。また翻訳結果の表示方法も考えられていない。視覚的な表示として現在ではカメラを用いた表現方法として拡張現実の手法や、他には音声出力など様々である。ユーザに対して効果的な表示を行うことでさらにカメラ翻訳システムの有効性を示すと考えられる。

そこで本論文では、カメラを用いた翻訳システムのインターフェイスの提案を目的とする。本論文の流れとしてまず設計方針において、入力及び出力インターフェイスについて関連研究および既存のソフトウェアを参考に設計を行い、次に全体の処理の流れを設計する。そして設計を基に本システムの実装について記述する。そして最後に考察と今後の展望について記述する。

## 2 設計方針

ここでは入出力インターフェイスとシステム全体の流れの設計方針を記述する。

### 2.1 入力インターフェース

入力インターフェイスは主にカメラ映像上をタッチパネルで範囲指定する方法を採用する。理由としては Chunyuan[1] によるとタッチパネルを使った指による入力方式は直感的で且つ細かな範囲指定を行えるからである。他には端末自身を動かすことで範囲指定する方法があるが、確認がし辛い事と指定範囲が細かい場合は操作し辛いため扱わなかった。

また端末をかざすことで抽出対象が暗くなったり、また抽出対象との距離が近すぎもしくは遠すぎたりといった場合による抽出失敗に対応することも考慮した。しかしながら端末自身の持つ補正機能では補正しきれないため、本研究では外部機器による補正を行う方法を採用した。そうすることでユーザに撮影環境を意識させる負担を和らげることが可能と言える。

### 2.2 出力インターフェース

出力インターフェイスは主に画面上に翻訳結果を表示することと音声による結果表示を採用した。

まず画面表示において単語意味は様々で複数の結果を表示してしまうと理解するまでに時間がかかる可能性がある。そのため必要な情報のある程度絞ることでユーザの理解度を高めるため画面上に表示する結果を少なくした。

また音声出力を同時に行うことで、画面表示だけでは読めない翻訳結果にも柔軟にユーザが対応できると考え採用した。さらに単語の意味をもっと知りたい場合はシステムから直接単語を Web 検索に掛けることが可能である。

現在は翻訳結果は文字表示と音声出力だけだが、今後は特定の単語が訳された場合などはアニメーションなどを付けて表現することを考えている。

Translation Camera for mobile phone

Kazuki Emon<sup>†</sup>, Hiroyuki Nishiyama<sup>†</sup><sup>†</sup> Faculty of Sci. and Tech, Tokyo University of Science

### 2.3 処理の流れ

処理の流れとしては、まずユーザがタッチパネルを使い指定範囲画像を生成する．そして生成した画像をサーバに送信し、サーバ側では画像の加工処理を行った後、OCR 処理をすることで画像内の文字のテキストデータを生成する．そして生成したテキストデータを端末側へ送信し、端末側で必要に応じて翻訳を行う．サーバを使う理由としては端末側での処理を少なくすることで他の端末でも変更点を少なくでき、遜色なく利用することが可能となるからである．外部センサに関しては照度センサを用いてライトの点灯を行い．距離センサを用いて撮影の有効範囲の表示を行っていく．処理の流れを表した図 2.3 を示す．

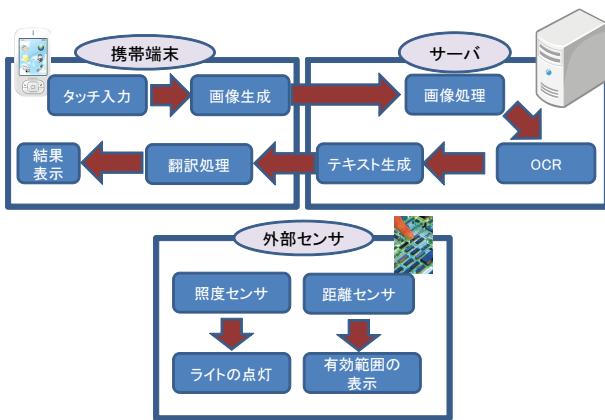


図 1: 処理の流れ

### 3 実装

実装について記述する．端末は HTC Desire を用いて開発を行った．端末のソフトウェアは AndroidSDK 利用し作成を行った．そのため端末に搭載されているタッチパネル及び通信、カメラなどの制御も AndroidSDK を利用して行っている．翻訳に関しては Google AJAX Language API[5] を用いて翻訳している．

サーバは Windows7 で開発し、通信部分は C++ を用いて行なっている．また画像処理部分は OpenCV2.1 を利用し処理を行なっている．そして OCR は Tesseract-OCR[4] を用いて行っている．

外部センサは Arduino を用いて制御を行い．照度センサは暗抵抗を用い、距離センサは SHARP:GP2Y0A21YK0F を利用している．端末とサーバの通信は Wifi を用いた無線通信でデータの送受信を行っている．実行画面を図 3 示す．まず起動画面上でカメラ映像とユーザが抽出した画像が表示されている．そしてユーザはタッチパネルを使ってカメラ映像上に直接抽出したい範囲を指定する．そして抽出した画像に触れると翻訳結果が表示される．

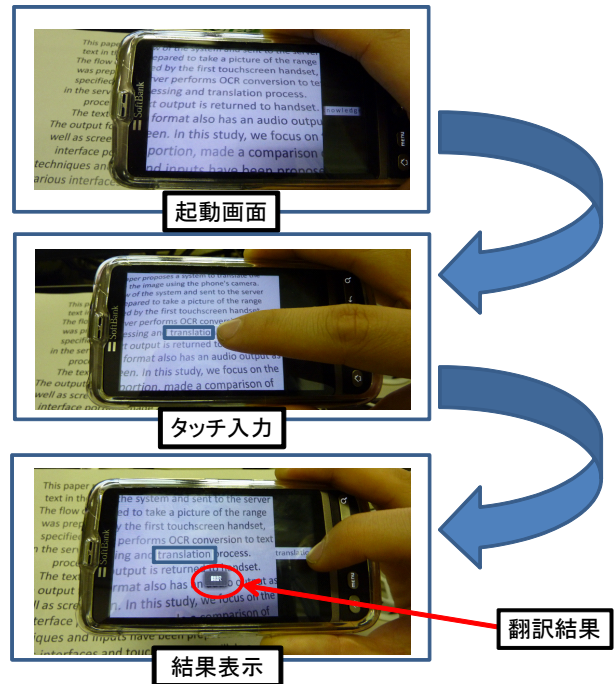


図 2: 実行画面 (青の矩形はタッチによる抽出範囲を示している．結果表示は音声による出力も行う)

### 4 おわりに

本研究では、カメラ翻訳システムのインターフェイスの提案を目的とした．そのために入力と出力のインターフェイスの設計を立て、入力としてタッチパネルを使った入力方法と外部機器による撮影環境の補正、出力として簡易的な結果表示及び音声出力を示し実装を行った．今後は出力結果の表現を文字だけではなく、アニメーションを使った描画等様々な表現を使うことで言葉の理解を深められると考えられる．

#### 参考文献

- [1] Chunyuan Liao, Qiong Liu, Bee Liew, Lynn Wilcox, "PACER: Fine-grained Interactive Paper via Camera-touch Hybrid Gestures on a Cell Phone", In Proc. of CHI 2010, April 10, 2010
- [2] Canedo-Rodriguez A., Soohyung Kim, Kim J. H., Blanco-Fernandez, Y. "English to Spanish translation of signboard images from mobile phone camera", Southeastcon, 2009. SOUTHEASTCON '09. IEEE
- [3] Jian Yuan, Yi Zhang, Kok Kiong Tan, Tong Heng Le, "Text Extraction from Images Captured via Mobile and Digital Devices", Advanced Intelligent Mechatronics, 2009. AIM 2009. IEEE/ASME International Conference on
- [4] "Tesseract-OCR", <http://code.google.com/p/tesseract-ocr/>
- [5] "Google AJAX Language API", <http://code.google.com/intl/ja/apis/ajaxlanguage/>