

# 携帯端末上で画像入りドキュメントの情報付加を行う実世界ブラウザ

鈴木 稔 人<sup>†</sup> 得 津 覚<sup>†</sup> 佐 藤 顕 治<sup>†</sup>  
矢 口 裕 明<sup>†</sup> 岡 田 慧<sup>†</sup> 稲 葉 雅 幸<sup>†</sup>

本研究では、携帯端末を用いて、実世界にあるチラシや雑誌、看板などのドキュメントに含まれる画像を認識し、それに関連する情報を Web 上から取り出してきて、スクリーン上に透過的に表示させることによって、実世界のドキュメントに、動的な情報を付加して表示するシステムを提案する。

## Real world browser displaying additional information to document including picture on mobile phone

NARITO SUZUKI,<sup>†</sup> SATORU TOKUTSU,<sup>†</sup> KENJI SATOU,<sup>†</sup>  
HIROAKI YAGUCHI,<sup>†</sup> KEI OKADA<sup>†</sup> and MASAYUKI INABA<sup>†</sup>

We propose the mobile-phone-based system that can display dynamic information on the real world documents, by recognition of the images in real world document such as flyers, magazines, and signboards, and displaying the related information from the Web transparently in the screen.

### 1. はじめに

画像の入ったパンフレットやレストランのメニューなどで、画像に携帯電話をかざすだけで、その画像項目に関するインターネット上の口コミ情報などを画面に重畳して表示し、かつ、その場で口コミ情報も入力できるシステムを携帯端末用実世界ブラウザと呼ぶ。

例えば、レストランのメニューで、料理の画像に端末をかざすと、それを食べた人の情報が画面に映る、また、食べ終わった後、それに対して感想を記入することで、お店のオンラインアンケート、口コミ情報をユーザーが共有できるようになる。

本研究では、画像の検索に SURF 特徴量<sup>1)</sup> を用い、XML の中にその特徴量と関連づけられた情報を埋め込み、シースルーブラウザを通してそのようなインタラクションをする実世界ブラウザを Android 携帯端末上で開発することを目的とする。

実世界にあるドキュメントを認識し情報を追加して表示する手法は、今までにもマーカー<sup>2)</sup> や文字<sup>3)</sup> を利用して実現する方法が提案されてきた。本研究では、レストランのメニューやパンフレットなど、人の目を引き付ける、多くの画像によって構成されるものを対

象とし、これらの画像を生かしたい。

また、携帯端末上で画像を取得し、サーバーに送信して Web 全体から検索し、関連情報を付加する Google Goggle<sup>4)</sup> も利用可能になっている。対して本研究では、パンフレット上などに探索範囲を狭めて、携帯端末上で認識できるようにすることで、軽快な動作で実世界のドキュメントに関連付けられた情報を閲覧できるようにするとともに、ユーザーや、パンフレットを発行する店舗などが情報を書き込んだりするなどの動的な情報生成を可能したい。

### 2. 実世界ブラウザの構成

画像を特徴としてコード化し、インターネットドキュメントとして登録する機能とその特徴量でインターネットドキュメントを検索してくる機能をもったアプリケーションとして実世界ブラウザを実現する。

図 1 がシステムの構成図である。ユーザーが実世界ブラウザを使い始めるとき、最初に画像特徴量と関連して表示する情報が記述された HTML コードがペアになって、画像の数だけ記述された XML ファイルを読み込む。ユーザーが携帯端末をパンフレットなどにかざすと、そのカメラ画像から特徴点・特徴量を計算する。その情報をもとに、事前に読み込んだ画像のデータベースと一致している画像があるかどうか計算する。もし一致した画像があるのであれば、関連す

<sup>†</sup> 東京大学大学院情報理工学系研究科  
Graduate School of Information Science and Technology,  
the University of Tokyo

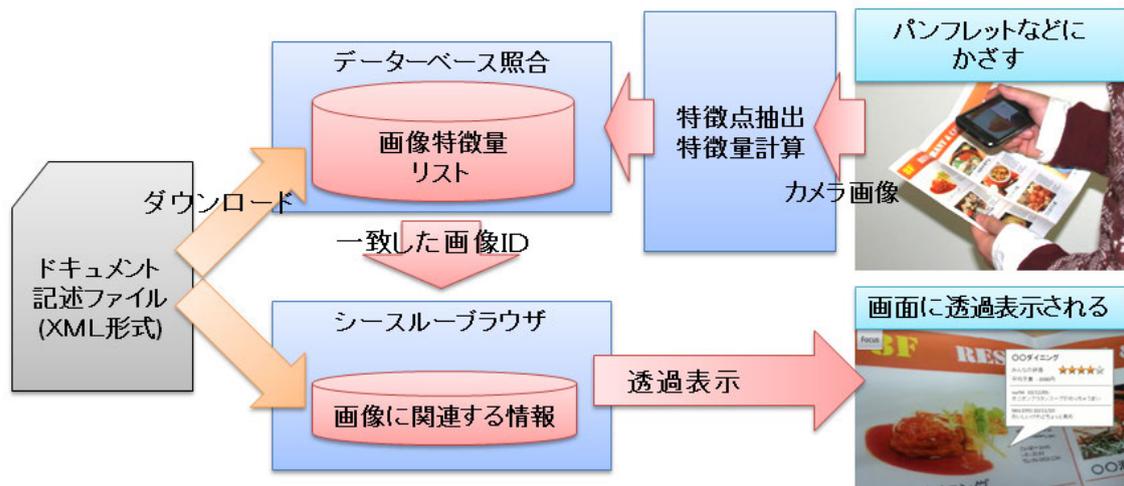


図 1 実世界ブラウザのシステム構成図

Fig. 1 system composition of realworld browser

る情報である HTML コードをカメラ画像の上にシー  
スルーで表示される Web ブラウザにレンダリングさ  
せる。以上の操作で、ユーザーはカメラの画面上に透  
過表示された情報を見ることができる。

画像の特徴点、特徴量計算については3章で、HTML  
コードのシースルーブラウザでの表示は4章で詳しく  
説明する。

### 3. 画像特徴量リストの特徴量と計算

画像どうしの一致を計算する手法は、テンプレ  
ートマッチング、カラーヒストグラム分布の距離計算、  
SIFT 特徴量の計算と比較など、様々な手法が提案さ  
れているが、ここでは、SURF の手法<sup>1)</sup>を用いて特徴  
量計算し、それを比較することで計算を行っている。

SURF では、まずヘッセ行列を用いて、いくつかの  
スケールを用いて、画像中の特徴点を検出する。図 2  
には、同一画像に対し、方向を変えて検出された特徴  
点の画像を示す。この際、積分画像を用いてヘッセ画  
像の値を近似することによって、処理の高速化を図っ  
ている。

次に、得られた特徴点のそれぞれに対し、Haar 型  
フィルタを用いて特徴的な方向を算出し、特徴的な方  
向を求める。そしてその特徴的な方向を軸にして画像  
を回転させ、特徴点の周囲を  $4 \times 4=16$  のブロックに  
分け、各ブロック内に存在するサンプル点で、Haar 型  
フィルタを用いて  $dx, dy$  を求める。最後にブロックご  
とに、 $\sum dx, \sum dy, \sum |dx|, \sum |dy|$  の 4 つの量を求め、  
計  $16 \times 4=64$  次元の特徴量を計算する。

この SURF 特徴量を、ドキュメント内のそれぞ  
れの画像に対して求め、4章で述べる XML ファイルに

記述しておく。ユーザーが携帯のカメラで画像をかざ  
したときにも、同様に SURF 特徴量を特徴点ごと  
に求める。ここで、この一つ一つの特徴点の特徴量ごと  
に、事前にダウンロードしてある XML ファイル中の  
特徴点の特徴量との距離を比較し、もっとも近い特徴  
点を持つ画像に投票する。尚、この比較は高速な近似  
近傍点探索手法の一つである LSH によって行った。  
最終的に、もっとも投票が集まった画像と、次点の画  
像を比べ、(次点の画像の得票数)/(もっとも投票が集  
まった画像の得票数)が、0.6 を超えた場合、もっとも  
投票が集まった画像を認識しているとみなした。

### 4. XML 表現とシースルーブラウザによるイ ンタラクシオン画面の構成

図 3 のような形式で、認識したい画像それぞれに対  
し、画像特徴量と関連する情報の HTML 表現をペア  
にして XML ファイルとして記述する。カメラが画像  
を認識すると、画面上に透過的に Web ブラウザを表  
示させ、関連する情報の HTML を図 4 のように表示  
する。

ここでの表示には、Android 携帯に標準に付属し、  
独自のアプリケーションの中で自由に組み込んで使う  
ことのできる WebKit ブラウザを利用した。

ここで HTML を利用して表示することの利点は、  
既存の Web ページ生成技術を利用して、コメント書  
き込みフォームなどを含んだ動的な情報生成が行える  
点である。

携帯端末上で画像入りドキュメントの情報付加を行う実世界ブラウザ



図 2 SURF による特徴点抽出結果  
Fig. 2 feature points extraction by SURF

```
<document>
<pair>
<features>
0.1,-0.00372637,0.00188171,0.00689026,0.00287667,0.0755842,-
0.014059,0.0790449,0.0353257,0.00134076,-0.00295633,0.00380834,0.00431284,-
0.000157443,0.000132124,0.000384727,0.000438843,-0.0157339,.....
</features>
<contents>
<h2>Kaiseki Ryori</h2>
<p>this sasaki is very good!!<br/><hr/>very resonable<br/><hr/></p>
<form method="POST" action="cgi-bin/xxx.cgi">
<div>Comment<input type="text" name="NAME"></div>
<input type="submit" value="Submit">
</form>
</contents>
</pair>
<pair>
<features>.....</features>
<contents>.....</contents>
</pair>
...
</document>
```

図 3 特徴量と関連情報が記述された XML ファイルの内容  
Fig. 3 marked up document of features and information



図 5 特徴点検出画像認識に対応してシースルーブラウザ上で関連情報の表示  
Fig. 5 Displaying related information corresponded to picture recognition



図 4 重畳されて表示された口コミ情報と書き込みフォーム  
Fig. 4 displayed comments and comment form

5. レストランフロアガイドでのお店紹介画像での認識実験

今回、図 6 の架空のレストランフロアガイドに含まれる店舗の紹介画像を対象に、画像を認識して関連する情報をシースルーブラウザを通して表示する実験をおこなった。図 7 の 5 枚の画像を認識対象としており、認識実験以前に SURF 特徴量、特徴点を計算して、携帯端末上にダウンロードした。また、対応する HTML で記述された関連情報もダウンロードした。今回の実験で用いた携帯端末は、Samsung GALAXY

S SC-02B (OS:Android2.2 メモリ:512MBRAM プロセッサ:S5PC110 1GHz) である。また、Android アプリケーションは、通常 Java で記述されるが、今回は、基本的なインターフェイス部分は Java で記述したほか、JNI(Java Native Interface) を通して、Android 上で動くようにコンパイルされた OpenCV2.1 ライブラリにアクセスし、処理を行った。

結果、図 7 に表示されているような特徴点を検出され、その特徴点の特徴量を手掛かりに、3 章の手法を用いて、どの画像かを認識することができた。さらに対応する情報がコードされている HTML 文章を、4 章で示した通り、Android 上の WebKit ブラウザプラグインを用いて、透過的に画面に表示することができた (図 5)

6. おわりに

携帯端末上で、ドキュメント上の画像を認識し、それに関連する情報を透過的に表示する実世界ブラウザを提案した。

画像認識には SURF の手法を用いて特徴点検出、特徴量取得を行い、カメラの画像の特徴量と比較して、携帯端末上で画像認識を行った。

# 8F RESTAURANT & CAFÉ

 <p> Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis              11:00~23:00              LO: 2230              TEL: 03-5555-1234</p> <p>〇〇ダイニング</p>	 <p> Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis              11:00~23:00              LO: 2230              TEL: 03-5555-1234</p> <p>〇〇海鮮井専門店</p>	 <p> Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis              11:00~23:00              LO: 2230              TEL: 03-5555-1234</p> <p>パスタ☆☆</p>
 <p> Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis              11:00~23:00              LO: 2230              TEL: 03-5555-1234</p> <p>☆☆懐石料理</p>	 <p> Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis              11:00~23:00              LO: 2230              TEL: 03-5555-1234</p> <p>□□そば屋</p>	 <p> Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis              11:00~23:00              LO: 2230              TEL: 03-5555-1234</p> <p>〇〇寿司</p>
 <p> Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis              11:00~23:00              LO: 2230              TEL: 03-5555-1234</p> <p>〇〇鍋</p>	 <p> Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis              11:00~23:00              LO: 2230              TEL: 03-5555-1234</p> <p>〇〇おでん屋</p>	 <p> Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis              11:00~23:00              LO: 2230              TEL: 03-5555-1234</p> <p>□□焼うどん屋</p>

図 6 実験に用いた架空のレストランフロアガイド  
 Fig. 6 the sample restaurant guide used in experiments

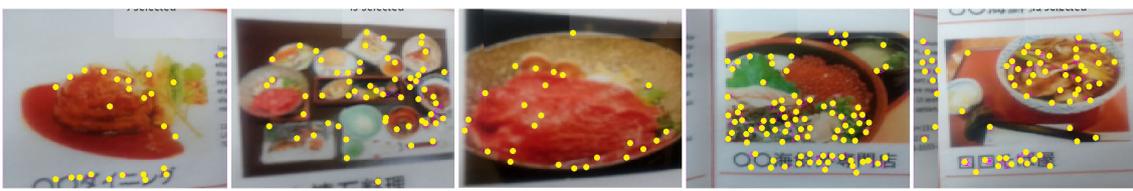


図 7 実験に用いた 5 つの画像の特徴点計算結果  
 Fig. 7 feature points calculation of five pictures in experiments

また、参照する画像の特徴量と関連する情報の HTML 表現をペアにして XML 内に記述し、HTML データをカメラの画面に重ね合わせて透過的に表示できる WebKit ブラウザを用いて表示した。

今後の課題は、より多くの画像に対して、ユーザーがストレスを感じない操作方法、処理時間でこのブラウザを使えるように画像認識を可能にすることである。

### 参 考 文 献

1) Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool, "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346-359,

2008  
 2) Billingham, Kato and Poupyrev: *The Magic-Book: a transitional AR interface*, Vol.25, No.5, pp.745-753 (2001)  
 3) 宮田, 塩原, 藤村: マーカーレス拡張書籍のための 2 次元文字ブロックインデクシング, インタラクシオン 2010 論文集 (2010).  
 4) Google Goggle: <http://www.google.com/mobile/goggles/>