

Pick and Lap:カメラ付き携帯電話を用いた実空間コンテンツの重ね合わせによる効果的な情報閲覧

武田 十季[†] 鶴野 玲治^{††} 牛尼 剛聡^{††}

E-mail: [†]toki@gospel.aid.design.kyushu-u.ac.jp, ^{††}{tsuruno,ushiama}@design.kyushu-u.ac.jp

あらまし 近年、日常生活の様々な場面で、携帯電話でデジタル・コンテンツを閲覧する機会が増えている。本論文では、携帯電話を利用した情報活動を支援する Pick and Lap 手法を提案する。本手法では、実空間上のコンテンツを携帯電話のカメラを用いて拾い上げ (Pick)、他の実空間上のコンテンツへ重ね合わせる (Lap) により、複数の実空間上のコンテンツを意味的な関係性をふまえて仮想的に融合し、ユーザの目的に適した形式で情報を提示する。本論文では、旅行雑誌を例題として用い、重ね合わせによる実コンテンツの仮想的な融合の形式的なモデルを提案する。そして、実装中のプロトタイプシステムについて述べる。

キーワード 携帯電話、拡張現実、Peephole ディスプレイ、コンテンツ融合、実空間コンテンツ

Pick and Lap: An Information Browsing Technique by Overlapping Contents in the Real World with Mobile Camera Phone

Toki TAKEDA[†], Reiji TURUNO^{††}, and Taketoshi USHIAMA^{††}

E-mail: [†]toki@gospel.aid.design.kyushu-u.ac.jp, ^{††}{tsuruno,ushiama}@design.kyushu-u.ac.jp

Abstract Recently we have many opportunities to brows the digital contents on a mobile phone in variety of situations in our daily life. This paper describes the Pick-and-Lap technique, which supports a user to manage and utilize information with a mobile camera phone. In this technique, actual contents in the real world are picked up with a mobile phone using it's camera, and they are overlapped other actual contents, then they are re-composed virtually on the basis of their semantic relationships, and provide suitable information to the user. In this paper, we focus on a travel magazine as actual contents in the real world, and we introduce a formal model for recomposing contents by overlapping operation.

Key words Mobile phone, augmented reality, peephole display, merging contents, actual content

1. はじめに

現在、携帯電話は、日常的に帯同され、日常生活の様々な場面で必要なデジタル・コンテンツを閲覧するために利用されている。こうしたことから、近年は、実空間における情報閲覧のための汎用的なインフラストラクチャとして、重要な役割を担うようになった。

これまでに、携帯電話上で実空間上での情報利用を支援することを目的とした研究が活発に行われてきた。近年では、カメラ付き携帯電話を用いて、実世界上の紙面上に印刷された、二次元バーコードを読み取り、情報閲覧するサービスは広く普及した。また、ユーザのコンテキストに応じて現実空間へ仮想的に情報を重畳して提示する AR 技術を利用した研究も盛んに行われている。セカイカメラ [1] は、iPhone 内蔵のカメラによって目の前の景色が画面上に映し出された上に、その場所・対象

物 (建物・看板など) に関連する「エアタグ」と呼ばれる、文字や画像のような付加情報を重畳して提示する。セカイカメラをはじめとする従来型の AR の多くは、市街地における建物に対して名称を重畳して表示するように、ユーザを取り巻く風景を拡張するという、いわば、マクロな視点に基づくものであった。しかし、例えば、2 人のユーザが同じ位置に立って同じ方向を見ているからといって、同一の情報を必要としているとは限らない。ユーザの具体的な目標・意図に基づいた情報活動の支援のためには、ミクロな視点からのコンテキスト・アウェアネスの実現が必要である。本研究では、ミクロな視点からのコンテキスト・アウェアネスを実現する対象として、実空間上のコンテンツに注目し、効果的な情報閲覧の実現に向けた Pick and Lap 手法を提案する。

本手法では、実空間上のコンテンツを、携帯電話のカメラを用いて拾い上げ (Pick)、実空間上の他のコンテンツへ重ね合わ

せる (Lap)。このとき、Pick したコンテンツと Lap するコンテンツの意味的な関係に基づき、2 つのコンテンツを仮想的に融合し、ユーザの目的に合致した形式で情報を提示する。

我々の身の回りには、場所や時間に関連づけられるコンテンツが多く存在する。例えば、道端に提示されている広告には、映画上演や演劇公演、新製品等の様々な情報が含まれているが、それらは、公演日や上演時間、発売日、そして公開場所などの時間的な概念に関連づけることができる。また、日常空間では複数のコンテンツに含まれている情報を統合することが必要となることが多い。例えば、飲食店の紹介記事を掲載している地方情報誌では、店舗情報と地図が異なるページに掲載されており、それぞれのページを相互に参照することにより、店舗と地図という 2 種類のコンテンツ上の情報を融合して必要な情報を取得する。

本手法では、場所と時間という概念に基づいて、重ね合わせる複数のコンテンツを意味的に関連付け、Pick されたコンテンツが Lap された際の振る舞いを決定する。つまり、重ね合わせられる複数のコンテンツが持つ意味的な関連に応じて、コンテンツが再構築され、情報が整理されることによって、実空間上のコンテンツの効果的な閲覧を実現する。

Pick and Lap 手法を実現するためには、様々な技術的な課題を解決しなければならない。それらを大別すると次のようになる。

- (1) 携帯電話のカメラを利用した実コンテンツの認識手法の開発
- (2) 重ね合わせによる実コンテンツの仮想敵な融合の論理的なモデルの開発
- (3) サービス提供のためのインターネットを介したメタデータの管理機構の開発

本研究では、特に重ね合わせによる実コンテンツの仮想的な融合の形式的なモデルの開発に焦点を当てる。

2. 関連研究

これまでも、カメラ付き携帯電話を用いて、実空間上のコンテンツを効果的に閲覧する研究は多く行われている。HOT-PAPER [2] では、携帯電話を用いて紙媒体への情報を付加と、付加された情報の閲覧を実現している。情報の付加のためには、携帯電話のカメラで紙面を覗いた場所を特定することが必要であるが、従来のような二次元バーコード等の特別なマークを利用せず、文字列の空間的分布を利用する。具体的には、対象となる紙媒体を特定の大きさの矩形で分割しであらかじめ画像として登録しておき、カメラで覗いた領域とマッチングする。HOT-PAPER の応用例としては、不動産案内誌や会議のレジюме、学会のプログラムなど、事前に用意された紙媒体に対して、会場での現場の写真や、音声、手書きメモなどを、携帯電話を用いて仮想的に付加するといったものが考えられる。

Rohs ら [3] は、あらかじめ実空間上の地図にドットを格子状に配置し、各格子内の 32*32 ピクセルの領域を全て画像として登録しておくことによって、携帯電話のカメラから地図を覗いた際に、地図上のどの部分を覗いているかを認識する手法

を提案している。これを応用した研究に、Wikeye [4] や Real World [5] がある。Wikeye では、携帯電話のカメラを通して、地図上の任意のランドマークを覗くと、ランドマークの情報を記した Wikipedia のページが、ディスプレイ上に表示される。Real World では、地球型の 3 D オブジェクトを携帯電話のカメラを通して覗くと、覗いた場所の地名や山脈の情報を記した Wikipedia のページが表示される。

CaMus [6] の研究では、携帯電話とマーカーシートの位置関係から、携帯電話の移動や、回転、高さなどの動きの情報を取得し、その情報を MIDI データへ変換することによって、インタラクティブな音楽演奏を実現している。マーカーシートは、6×6 ドットから構成されるセルがシート全体に配置されている。各セル内には、白と黒のドットが規則的に記されており、これらのパターンの違いを携帯電話のカメラで認識して、携帯電話とマーカーシートとの位置関係を認識している。

これらの研究では、任意の実空間上のコンテンツに付加情報を対応づけ、それらの情報を携帯電話のカメラから取得している。このため情報を取得する閲覧者にとっては、あらかじめ対応づけられた、コンテンツと付加情報の関係のみの閲覧に限られている。本研究では、実空間上のコンテンツ自体を他の実空間上のコンテンツへ重ね合わせることで、閲覧者の選択の違いによって、取得される情報が変化する手法を提案する。

また上記の研究はいずれも、携帯電話を、実空間上のコンテンツに対する Peephole(覗き穴) ディスプレイ [7] として機能させており、ユーザにとって直感的な操作で閲覧することを可能としている。本研究では、任意のコンテンツを他のコンテンツへ重ねた (Lap) 際に、Peephole と同様のアプローチで、携帯電話のディスプレイ上に情報を提示する。

3. Pick and Lap 操作による情報取得

3.1 概要

本システムでは、携帯電話のカメラとディスプレイを利用して、複数の実コンテンツを意味内容に基づいて仮想的に融合し、ユーザに分かりやすい形式で必要な情報を提示する。重ね合わせによる仮想的な実コンテンツの融合モデルを構築し、有効性を検証するために、本研究では、対象とする実空間上のコンテンツ（以降、「実コンテンツ」と呼ぶ）の例として旅行情報誌を利用する。旅行情報誌には、ホテルや店舗、土産情報等を紹介するページと、それらの場所を確認するための地図のページ等、様々な種類のコンテンツが含まれている。本章では、これらのコンテンツに対して Pick and Lap の考え方に基づいた携帯電話上での効果的な情報の提示方法について説明する。

3.2 Pick 操作と Lap 操作

一般に、コンテンツは階層的な構造を有している。例えば、多くの旅行情報誌は対象とする地域をいくつかのサブ地域に分割し、サブ地域毎に章を構成する。一つの章は、その地域の観光スポット、レストラン、ホテル等、紹介する対象の分類に応じて節に分かれている。それぞれの節は複数のページから構成される。さらに、それぞれのページには複数の紹介記事が含まれている。それぞれの紹介記事は個々の観光スポット、レスト



図 1 店舗の情報を pick して地図に対して Lap する例



図 2 店舗の情報を Pick して地図へ Lap する操作の概念図

ラン、ホテル等に関する情報を記載している。

Pick 操作は、実空間上でユーザが興味のあるコンテンツを選択し、携帯電話に取り込む操作である。Pick するコンテンツの粒度としては様々なレベルが考えられるが、今回は、個々の観光スポット、レストラン、ホテル等に関する紹介記事を Pick 操作の単位とする。すなわち、旅行雑誌における pick 操作として、ユーザが興味を持った、観光スポット、レストラン、ホテル等を携帯電話に取り込むことを考える。一方、Lap 操作は、pick 操作により携帯電話に取り込んだコンテンツと他の実コンテンツとの関連を可視化するための操作である。

例えば、旅行雑誌の店舗一覧のページ上で、携帯電話のカメラを通して、興味のある店舗の写真を選択することによって、特定の店舗のコンテンツを Pick する。次に、別ページの地図に Lap すると図 1 に示すように携帯電話のディスプレイ上で、地図上の適切な位置に、Pick した店舗の写真が重畳して表示されるようにすることが考えられる。図 2 に、店舗の情報を pick して、地図上に lap する捜査の概念図を示す。

3.3 重ね合わせによる意図の推定

ユーザは、旅行雑誌を閲覧する際に、目的に応じて閲覧するコンテンツの順番が変わると考えられる。本研究では、Pick し Lap する過程がユーザの意図を表していると考えられる。

例えば、「飲食店」同士を重ね合わせたユーザは、「飲食店」を選択するために複数の飲食店を比較したいと推定できる。このように、同じカテゴリのコンテンツを重ね合わせた場合、ユーザは相互の比較を行いたいと考えられる。この時、携帯電話の画面上では図 3 のように、異なるコンテンツ同士を比較できる形で提示する。

一方、「飲食店」と「ホテル」を重ね合わせたユーザは、「飲

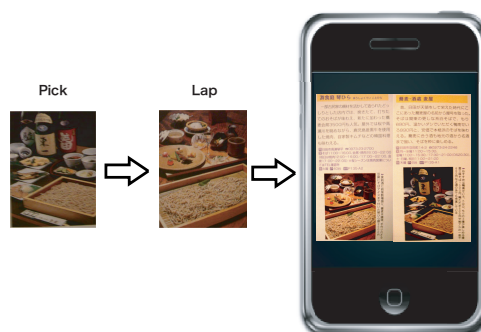


図 3 同一カテゴリのコンテンツを重ね合わせた場合の表示例



図 4 飲食店を Pick し、温泉へ Lap する



図 5 異なるカテゴリのコンテンツを重ね合わせた場合の表示例

食店」に近い「ホテル」、または、「ホテル」に近い「飲食店」を探していると推定できる。このように、異なるカテゴリ同士のコンテンツを重ね合わせた場合は、ユーザは互いのコンテンツの位置関係の比較を行いたいと考えられる。例えば、図 4 のように「飲食店」を Pick して、「温泉」へ Lap した場合、図 5 のように提示される。「温泉」へ Lap した際、画面上では、「温泉」を中心にした円周上に、「温泉」から見た飲食店の方向と距離を示す。この時、距離が近い飲食店ほど大きく表示され、遠い飲食店ほど小さく表示する。

3.4 コンテンツの重ね合わせ

コンテンツを重ね合わせた際に、重ね合わせられる複数のコンテンツが関連付け可能であるか認識することが必要である。また、関連付け可能な際に、それらのどの関連を利用して、どのようにユーザに提示するかを決定する必要がある。この問題に対して、旅行雑誌の例を用いて説明する。

本研究では、対象とする実コンテンツに関するメタデータがデジタルデータとして存在し、携帯電話内に保持されている、もしくはインターネットを介して携帯電話からアクセスできるようになっていることを前提とする。

実コンテンツに対して pick 操作が行われた場合、実コンテンツに対して、コンテンツ自身のメタデータが対応づけられる。また、lap 操作が行われる対象の実コンテンツに対しても同様



実空間上での二点の距離

$$\Delta X = \text{Long2} - \text{Long1}$$

$$\Delta Y = \text{Lat2} - \text{Lat1}$$

地図上における店舗の表示位置

$$X = \Delta X \times \text{地図の縮尺}$$

$$Y = \Delta Y \times \text{地図の縮尺}$$

図 6 地図上での店舗の表示位置の計算

にメタデータとの対応付けが行われる。そして、pick 操作が行われた実コンテンツのメタデータと lap 操作が行われた実コンテンツのメタデータを利用して、コンテンツの融合が行われる。

コンテンツの融合には様々な形式が考えられる。代表的な融合方法の一つとして、pick 下コンテンツをアイコン化し、lap したコンテンツ上に表示するものである。これは、lap したコンテンツが空間的な広がり表現しており、pick したコンテンツが lap したコンテンツにおける位置を利用者に提示する場合に適切な表現であると考えられる。旅行雑誌の例では、店舗情報が表示される実空間の地図上での座標を求めるために、店舗の緯度経度と、地図の基準位置の緯度経度から、二点の変位を求め、地図の縮尺によって、地図上での位置座標へ変換する必要がある。例えば、店舗の情報を地図上に表示する際の変換の計算方法を図 6 に示す。

ここで、Pick された実コンテンツが店舗情報であるとき、自身のメタデータとして店舗の「位置情報（緯度経度）」の値を持つ。この場合、店舗情報が重ね合わせることができる対象は、「基準値の緯度経度」と「縮尺」を自身のメタデータに持つコンテンツとなる。つまり、店舗情報を地図に重ねると、地図から「基準値の緯度経度」の値と、「縮尺」の値を得て、地図上での表示位置を計算する。この Lap の操作で、「店舗の位置情報」「基準値の緯度経度」「縮尺」の三つの値が得られ、それらの値を地図上での店舗の相対座標を求める式に代入し、店舗の表示位置を得ることができ、その位置に対応する店舗情報を提示する。

3.5 旅行雑誌における重ね合わせの例

地図と店舗情報の他にも、旅行雑誌内に存在する他のコンテンツの意味関係に着目して、様々な重ね合わせを考えることができる。図 7 は、旅行情報誌の一部を示している [8]。この情報誌では、観光スポットの紹介ページと、お土産やレストラン紹介



図 7 旅行雑誌の観光スポットの紹介ページ（左）とレストランの紹介ページ（右）

介のページが異なったカテゴリに分類され、異なったページに存在している。例えば、ユーザが P72 にある「マーライオン・パーク」（図 7 左）を訪問することを決め、その付近で食事をしたいと考えた場合、P150 以降のレストラン紹介ページを参照し、探さなくてはならない。この場合は、P197 (図 7 右) に「マーライオン・パーク」から近い店舗が存在する。このような場面で、本システムを応用すれば、「マーライオン・パーク」を Pick し、レストラン紹介ページに Lap したユーザは、「マーライオン・パーク」付近のレストランを探していると推測することができ、Pick した位置情報に近いレストランリストをユーザへ提示することが考えられる。このように、本システムを応用して、静的である実コンテンツの欠点を補い、実コンテンツを拡張することによって、ユーザの情報活動を支援することが可能となる。

3.6 スケジュール帳での例

道端にはイベントの時間や場所を記した広告等などの多くのコンテンツがあるが、通りすがりに興味を持った内容があっても、失念してしまうことがある。例えば、演劇の公演ポスターに興味を持った際に、ユーザはタイトル名などを覚えておき、自宅のパソコン等で、公演場所や、公演日、空席情報などの詳細な内容を確認する場合がある。このような通りすがりに目に入ってくる情報の量が多くなるにつれ、興味を持った情報までも忘れてしまう可能性が増加する。本システムを応用すれば、Pick したコンテンツが、開店時間や、新製品の発売日など、時間に関連した情報を含んでいた場合、スケジュール帳に Lap すると、スケジュール帳の当てはまる位置を表示領域として、コンテンツが提示される。またスケジュール帳には、異なったレイアウトが複数存在するため、そのレイアウト構造の特徴、またその用途によって、Lap した際のコンテンツの提示のされ方を変えることも可能である。携帯電話のパーソナル性と、携帯性から、日常の中で気になるコンテンツを Pick し、蓄積しておくことで、スケジュール帳を閲覧しながら、書き込まないような内容までも、振り返ることが可能となる。

4. プロトタイプシステム

4.1 システムの概要

提案手法の有効性を評価するためのプロトタイプシステムを実装した。本プロトタイプシステムでは、パーソナルコン

コンピュータ上で USB カメラを利用して動作する。

Pick の対象は、実コンテンツ上に存在する矩形の写真と図に限定する。店舗や地図の矩形画像は、あらかじめ取得して、関連するメタデータと共に管理サーバに登録しておく (図 8:(a)(b))。今回は、店舗のメタデータとして、店舗の位置情報 (緯度経度) を利用する。また、地図のメタデータとして地図の縮尺と、地図画像の左下を基準とした場合の左下の緯度経度を利用する。

Pick 時に、撮影領域内の矩形内の画像が抽出され、管理サーバに登録された画像との類似検索を行う。類似検索には画素地のヒストグラムを利用する。最も類似している画像に対して、登録されているメタデータを取得する (図 8:(c))。Lap 時にも同様に、Lap する対象のコンテンツに対応づけられたメタデータを取得する。Pick したコンテンツのメタデータと Lap するコンテンツのメタデータにより、提示のされ方を決定する。例えば Pick したコンテンツが店舗情報であり Lap するコンテンツが地図である場合には、Pick した店舗の画像のサムネイルを、Lap した地図上の適切な場所に重畳して表示する。

4.2 実コンテンツの認識手法

実コンテンツの認識手法に関する研究は画像処理、パターン認識の分野で活発に行われている。黄瀬ら [9] の研究では、撮影対象の局所特徴量を用いて、画像検索の処理時間の高速化を実現する手法を提案している。FPcode [10] は、人間の視覚特性に基づいて、人間の目に見えにくい形式で印刷画像にデータを埋め込むことによって、専用アプリケーションを搭載したカメラ付き携帯電話で、高速・確実に情報を読み取ることが可能である。これにより、印刷物のデザインを損なうことなく、フリーペーパーや雑誌、まちの案内ポスターなどの紙媒体から、携帯電話を介して動画や音声コンテンツを提供する。

本研究では、プロトタイプシステムにおいて評価実験を行うことができることを目標とし、認識機構の汎用性と認識性能よりも、実現のしやすさを考慮し、できるだけ単純な認識手法を採用した。

本プロトタイプシステムでは、データベース画像と撮影対象の画像との類似検索を、比較対象となる 2 画像の各ピクセル位置における RGB 値の差分から求める。差分距離のみを用いると、全く異なる内容の画像であっても、画像全体に渡った色の分布が等しい場合に同じ画像と認識されてしまうため、対応する 2 つの画像のピクセル位置ごとに重みづけを行い、色分布の空間的位置を考慮した図 9 の式で実験した。全画素数 N 個の、各画素ごとの重みを $ID(N)$ とする。各画素ごとの、撮影画像とテンプレート画像の画素値の差に重み $ID(N)$ を掛け合わせ、全画素 ($0 \sim N$) にわたって加算していく。その値を、 $-1.0 \sim 1.0$ の間におさまるように正規化する。最も二つの画像が似ている場合、この値はベクトル $T(1.0, 1.0, 1.0)$ であるので、このベクトル T と計算された値との距離が小さいほど二つの画像が類似していると考えられることができる。この手法は単純であるが、例えば図 10 のような特に色合いの似た地図画像においても正しく認識可能であった。

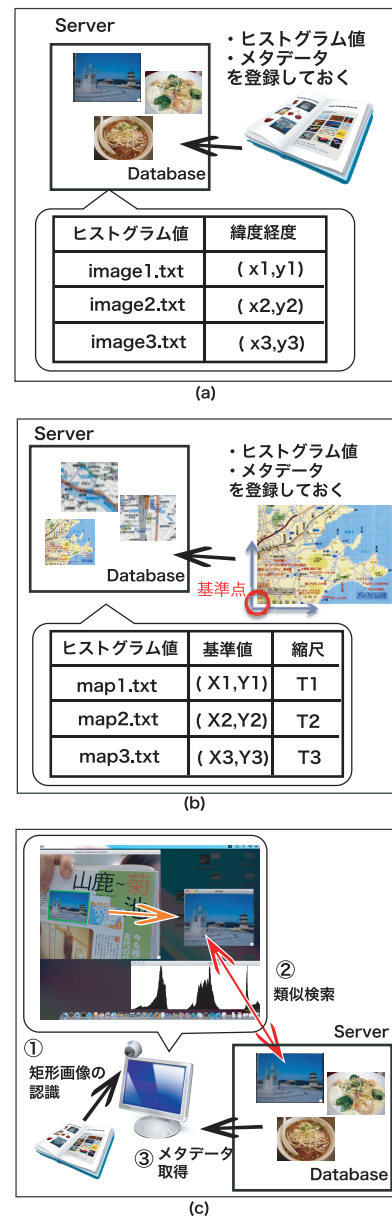


図 8 システムの概要

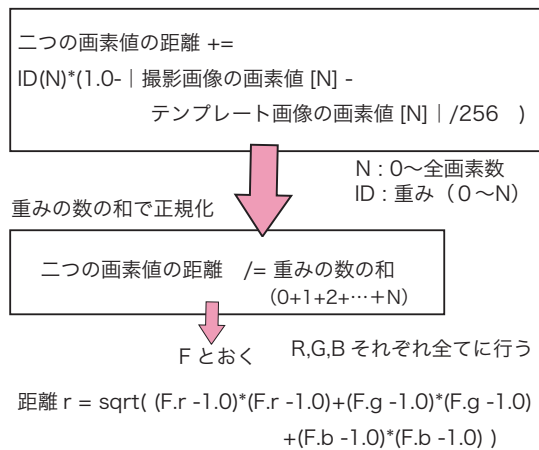
5. まとめと今後の課題

本論文では、携帯電話のカメラを通して、実コンテンツを矩形画像として認識し、他のコンテンツ上に重ね合わせると、相互の意味的關係に基づいてコンテンツの提示のされ方が決定される Pick and Lap 手法を提案した。また本手法を利用した、紙媒体を中心とした実空間上のコンテンツの効果的な閲覧について、旅行雑誌やスケジュール帳を例として議論を行った。

今後の課題として、今回のプロトタイプシステムを用いて、有効な評価実験の手法を考え、実験を行っていく。また、4章で述べた応用例について、よりユーザ視点に立った議論を行い、本システムの拡張を行っていく。

文 献

- [1] <http://support.sekaicamera.com/>
- [2] B.Erol, J.Graham, E.Antunez, J.J.Hull, "HOTPAPER demonstration: multimedia interaction with paper using mobile phones", M.M2008, pp.983-984 (2008).



r が小さいほど二つの画像は類似している

図 9 差分距離を用いた二つの画像の比較



図 10 実験に用いた地図画像

- [3] M.Rohs, J.Schoning, A. Kruger, and B.Hecht, "Towards real-time markerless tracking of magic lenses on paper maps", In Adjunct Proceedings of the 5th Intl.Conference on Pervasive Computing (Pervasive), pp.69-72(2007).
- [4] B.Hecht, M.Rohs, J.Schoning, A.Kruger, "Wikeye - using magic lenses to explore georeferenced Wikipedia content" In Proceedings of the 3rd International Workshop on PER-MID2007.
- [5] J.Schoning, M.Rohs, A.Kruger: "Mobile Interaction with the 'Real World' ", MobileHCI2009, no.106(2009).
- [6] M.Rohs, G.Essl, "CaMus2 - Collaborative Music Performance with Mobile Camera Phones ", ACE2007, pp.190-195(2007).
- [7] G.W.Fitzmaurice, "Situated information spaces and spatially aware palmtop computers ", Communications of the ACM, Vol.36, No.7, pp.39-49(1993).
- [8] "わがまま歩き シンガポール ジョホールバル ビンタン島". 実業之日本社 (2007).
- [9] 黄瀬浩一, 岩村雅一, 中居友弘, 野口和人, "局所特徴量のハッシングによる大規模が贈検索", 日本データベース学会論文誌, Vol.8, No.1, pp.119-124(2009).
- [10] <http://jp.fujitsu.com/solutions/fpcode/>