

# ShootAR: ユーザ姿勢を考慮したモバイル AR のための操作スタイルの提案

島田 哲朗<sup>†</sup> 樋口 啓太<sup>†</sup> 暦本 純一<sup>†, ††</sup>

拡張現実感 (Augmented Reality: AR) をモバイル端末上で実現するモバイル AR は、カメラを備えた高機能なモバイル端末などの登場により一般的になりつつある。ほとんどのモバイル AR は、ユーザがモバイル端末のカメラを対象物体に向けてすることで、対象物体に関連する情報を合成してディスプレイ上に表示する。しかし、長い時間モバイル端末を対象物体に向けて掲げることは、ユーザに身体的疲労を与えることとなる。さらに、遠くにある物体や一部が遮蔽された物体をモバイル端末に搭載されたカメラで認識して、情報を合成することは困難である。本論文では、これらの問題を解決するために、“ShootAR” と呼ぶモバイル AR の操作スタイルを提案する。ShootAR では、モバイル AR におけるユーザの操作を 2 つのステップに分ける。まず、既存のモバイル AR システムと同じようにモバイル端末を掲げた姿勢で、ユーザは端末のカメラが捉える画像から情報を合成して欲しい対象物体 (あるいはその物体があるであろう空間) を探す。ここで、対象物体が見つかった場合、ユーザはその画像を撮影する (この動作を “shoot” という)。その動作後、システムは撮影した画像と、それに近いデータベース上の高精細パノラマ画像をすり替える。これにより、ユーザは端末をかざす必要はなくなり、実際の物体を映していたとき以上のインタラクションがどこでも楽な姿勢で可能になる。我々は、このスタイルに基づいたモバイル AR システムを開発し、既存のモバイル AR ではできないアプリケーションを示す。

## ShootAR: A Mobile Augmented Reality Interaction for Reconciling Comfortable Operation with Spatial Continuity

TETSURO SHIMADA,<sup>†</sup> KEITA HIGUCHI<sup>†</sup> and JUN REKIMOTO<sup>†, ††</sup>

Mobile augmented reality systems that use a smart-phone with a camera are becoming popular. These systems typically let users hold the device and see the real world through a live image overlaid by computer-augmented information. However, holding a mobile device toward the environment for a long time is not comfortable and suitable for precise operation. It may cause fatigue in users' arms. Moreover, the systems can not deal with distant and occluded objects because of hardware limitation including optical zooming and camera resolution. In this paper, we solve these problems by proposing a two-step interaction model called “ShootAR”. This model first let the user to target the real world scene like normal AR systems, and then the system suggests the user to “shoot” the scene if a matching between real and stored images can be done. Once this shoot operation succeeds, the user can interact with a stored image without holding the device toward the environment. This simple model reconciles AR's natural spatial continuity between real and virtual scene, with kinetically comfortable operation style. We developed a mobile AR system based on this model, and show applications that are not available by traditional AR systems.

### 1. はじめに

近年、現実の世界にコンピュータで生成された情報を合成して表示する、モバイル端末を利用したシステムが注目を集めている。それらのシステムは、モバイル端末において拡張現実感 (Augmented Reality: AR) を実現しているので、モバイル AR と呼ばれる。

モバイル AR というコンセプトは、1990 年代に提案された NaviCam というシステムで初めて示された<sup>1)</sup>。NaviCam は、持ち運び可能なテレビの背面に小型のカメラを備えたシステムで、ユーザがカメラを向けた方向の状況を映像から取得し、状況に合わせた情報を現実の映像と重ね合わせてテレビディスプレイ上に表示するものであった。今では、スマートフォンと呼ばれる高性能なモバイル端末をプラットフォームにしたモバイル AR システムが開発されている<sup>2)</sup>。すでに一般ユーザ向けに、iPhone や Android といったスマートフォン上で動作する、セカイカメラ<sup>3)</sup> や Laya<sup>4)</sup> と

<sup>†</sup> 東京大学

The University of Tokyo

<sup>††</sup> 株式会社ソニーコンピュータサイエンス研究所

Sony Computer Science Laboratories, Inc.



図 1 既存のモバイル AR インタラクション: ユーザはモバイル端末を操作中も対象物体に向けていなければならないため、長時間のモバイル AR の使用でユーザは腕に疲労を感じる。

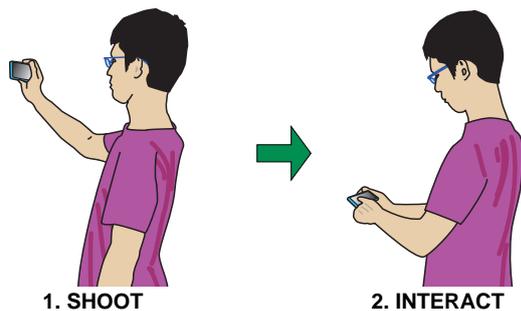


図 2 ShootAR インタラクションモデル: はじめにユーザはモバイル端末を対象物体に向けて掲げる。そして、システムは現在の画像を撮影する“shoot”可能かどうかをユーザに知らせる。この操作が行われた後は、ユーザは端末を対象に向けることなく楽な姿勢でモバイル AR を使用することができる。

いったモバイル AR システムが存在している。

一般的に、モバイル AR システムを利用するには、ユーザは常にモバイル端末を現実の物体に対して掲げている必要がある (図 1)。この姿勢において、カメラの中心軸とユーザのしている方向を合わせることは容易である。さらに、タッチスクリーンを備えたモバイル端末では、スクリーン上の物体を触るだけで離れた物体を操作できるという直感的なインタラクションが可能となる<sup>5)</sup>。しかし、顔の高さに手を持つてくるという姿勢を長時間続けると、ユーザの腕に疲労を生じることとなる。これは、垂直型タッチスクリーンにおいて議論されてきた、「ゴリラ腕」と呼ばれる症状と同じである<sup>6)</sup>。さらに、片手でモバイル端末を把持し、もう片方の手で操作をしようとする、ユーザはカメラのぶれに悩まされることとなる。

一般的に、画面を見ながらモバイル端末を操作する場合、ひじをゆったりと曲げて体に添わせた状態にする (図 2-2)。なぜなら、この姿勢がモバイル端末を操作する上で楽な姿勢だからである。モバイル AR で

ユーザにこの姿勢を実現させるためには、現実空間と合成された情報が同時に表示されている画面を一時的に静止させるという方法がある<sup>7)</sup>。しかしこのような方法であったとしても、対象物体が遠くにある場合や他の物体による遮蔽 (オクルージョン) が発生している場合などは、モバイル端末のカメラでは対象物体を捉えきれず、そもそも情報を合成して表示することができない。これを解決するためには、ユーザが対象物体の見える適切な位置まで移動して、カメラを適切な方向に向けて掲げることになる。

本論文では、モバイル端末を掲げてディスプレイを通して現実空間を見る姿勢、そして楽にモバイル端末を操作できる姿勢をその場でシームレスにつなぐモバイル AR のインタラクションモデルを提案する (図 2)。そして、ShootAR 操作スタイルを実現する具体的なステップを説明し、既存のモバイル AR よりも高機能なインタラクションが可能であることを指摘する。さらに、実際に ShootAR 操作スタイルを実装したプロトタイプを作成し、そこから議論を通じて ShootAR 操作スタイルの可能性を示す。

## 2. 関連研究

u-Photo は写真をユーザインタフェースとして利用するシステムである<sup>8)</sup>。ただし、我々のシステムと異なり、u-Photo ではユーザが撮影した写真を関連付けられた機器の操作画面を呼び出すブックマークのように扱っているため、後に紹介する“スーパーズーミング”の様なインタラクションはできない。また、u-Photo は環境の認識にビーコンを用いているが、我々のシステムは画像中の特徴点を用いているため、認識できる空間がビーコンをおける場所に限定されないという空間的柔軟性がある。

Object-oriented video は、工場内のノブやハンドルをビデオ画像とそれにオーバレイされた情報を通して遠隔操作可能にするシステムである<sup>9)</sup>。ただし、このシステムはモバイルではなく、固定式のカメラを用いている。

## 3. ShootAR

モバイル AR のメリットは、その直感的な操作性にある。つまり、ユーザがモバイル端末を掲げるという単純な操作をするだけで、モバイル端末のカメラの向けられた空間の情報を手に入れることができるということである。これにより、ユーザは空間と情報の対応付けを自然に理解することができる。しかし、実際の利用シーンを考えてみると、次のようなモバイル AR

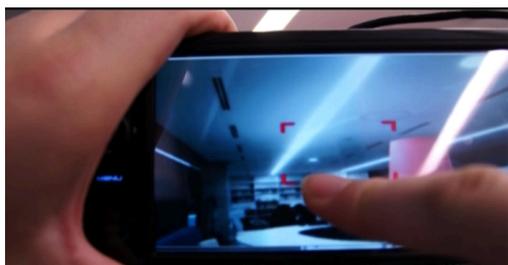


図 3 shoot をフィードバックする画面: ユーザは画面中にターゲットカーソルが出ている間であれば、撮影 (shoot) できる。この動作は、ユーザに明示的にその場所の情報をモバイル端末に取り込んだことを意識させる。

の問題点が浮き上がってくる。

- (1) モバイル AR を利用する場合、ユーザは常にモバイル端末を掲げる必要がある。
- (2) 遠くにある物体や一部が遮蔽されて見えない物体を認識して、情報を重ね合わせることが困難である。

現実空間に合成された情報を見るということと、カメラを掲げ続けるという行為は本質的には関係ない。むしろ、ユーザが合成された情報に集中したいと考えていても、カメラを安定させて持つという行為によって阻害される可能性がある。また、システムが現実の物体をうまく認識してくれない場合や表示された情報が現実の物体を覆い隠してしまう場合などは、合成された情報が適切に表示される位置までユーザはカメラを掲げながら移動しなければならない。

我々はこの 2 つの問題を、モバイル AR におけるユーザ動作を 2 つのステップに分けることで解決を図る。はじめに、従来のモバイル AR と同様に、情報を合成する対象物に対してユーザはモバイル端末を掲げる。そして、その対象を含めた空間がデータベースに含まれている場合、ユーザはその画像をシャッターを切ったり、画面をタッチすることで撮影する (この動作を “shoot” という)。これにより、ユーザにとって現実空間とこれから操作可能になるモバイル端末上の情報との自然な対応付けが可能になる (図 3)。その後、shoot された画像と同じ位置から見えるであろうパノラマ画像の一部をデータベースから引用し、すり替えることにより、これ以降ユーザはモバイル端末を掲げる必要はなくなる。ユーザは手元のモバイル端末で仮想的な現実空間を見ることができ、その中でモバイル端末のカメラの性能を超えたモバイル AR インタラクションを提供することが可能となる。

### 3.1 イメージベーストレンダリング

通常、仮想空間の記述と表示にはコンピュータグラフィックス技術が用いられている。しかし、広範囲に

わたる仮想空間を構成する正確な 3 次元形状モデルを作成することは容易ではない。特に複雑な形状を有する自然物を正確に記述することは困難である。このような問題を回避するため、イメージベーストレンダリングと呼ばれる手法が用いられる。イメージベーストレンダリングは、実際の写真を変形させて、任意の視点位置からの画像として提示するというものである。3 次元仮想空間を表示するのはモバイル端末上の 2 次元ディスプレイであるため、提示する 2 次元画像を適切に生成できさえすれば、3 次元形状モデルを用いなくとも 3 次元仮想空間を表現することができる。

最も単純な実装方法としては、必要な画像をすべて記録しておくという方法がある。ある視点位置における周囲の映像情報をパノラマ画像として記録しておき、観察者の視線方向に応じた部分を切り出して変形処理して提示する手法が提案されている<sup>10)11)</sup>。ShootAR では、後述の Shoot した画像からパノラマ画像に切り替わる場合、視点を切り替える場合に適切なトランジションエフェクトを表示することにより単純なイメージベーストレンダリングを実現する。パノラマ画像は、モバイル端末上もしくはネットワークのサーバ上にデータベースとして用意する。

### 3.2 特徴点辞書

ShootAR システムは最初の段階で、現在カメラが捉えている画像がデータベース上のパノラマ画像と置換可能かどうか判断する。その方法として、現在カメラが捉えている画像の特徴点とパノラマ画像の特徴点を比較する。この際に、パノラマ画像から特徴点を計算して比較判定に使用すると、比較範囲が広がるため計算時間がかかってしまう。そこで、あらかじめパノラマ画像の中から特徴点の一致判定に用いるランドマークを選択し、その特徴点を特徴点辞書として登録しておく (図 4)。パノラマ画像の代わりに特徴点辞書との特徴点比較を行うことにより、計算量を減らすことができる。

### 3.3 トラッキング

ShootAR システムの画像からの特徴点抽出には SURF アルゴリズム<sup>12)</sup>を採用した。特徴点の一致によって判断することにより、特殊なマーカや装置を環境に配置する必要がないため、マーカの大きさなどスケールの問題に依存せず、屋外や屋内を問わない空間に対して柔軟なモバイル端末インタフェースを提供することができる。この特徴点抽出処理はバックグラウンドで行っていて、この時点ではディスプレイ上に AR 表示は何もされていない。そのため、ユーザにとってはカメラに写った画像をそのまま見ているように感じ

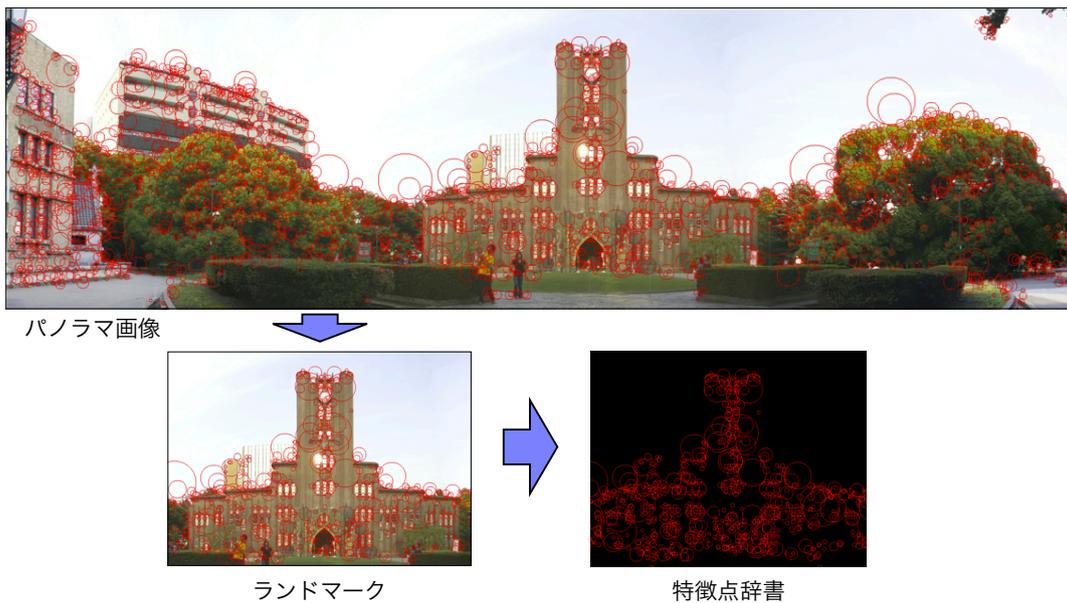


図4 特徴点辞書: マッチングを高速化するため, ランドマークを抽出し, その特徴点をパノラマ画像を参照するための辞書として用いる.

る (図5 - (1)). 計算された特徴点は, 特徴点辞書中の特徴点と比較される (図5 - (2)). カメラが捉えた画像の特徴点とランドマークの特徴点とが一定以上一致した場合, ShootAR システムはユーザに “shoot” 可能であることをディスプレイ上で知らせる (図5 - (3)).

### 3.4 Shoot

ユーザが shoot した場合, カメラが写していた画像を撮影し, 一時的にディスプレイ上に表示する (以降, 撮影された画像を “shoot 画像” と呼ぶ). ここで, ShootAR システムは, shoot 画像と特徴点一致したランドマークにリンクしているパノラマ画像をデータベースから取得する. パノラマ画像をそのままディスプレイに表示した場合, shoot 画像と大きく見え方が異なるので, ユーザに対していきなり操作画面が変わったという印象を与えかねない. そこで, パノラマ画像の特徴点と shoot 画像の特徴点と重なるように変換する特徴点同士のホモグラフィ (平面射影) 行列を計算する. そして, shoot 画像とホモグラフィ行列によって見え方が変更されたパノラマ画像の一部を, 対応する特徴点と重なるように滑らかに移行するトランジションエフェクトを表示する (図5 - (5)). これにより, 「操作画面にいきなり移行した」のではなく, 「現実の空間とモバイル端末の表示が連続している」という空間的連続性をユーザに意識させることができる. パノラマ画像そのままインタラクシオン可能な画面となるため, これ以降ユーザがモバイル端末を顔の高さ

まで掲げる必要はなくなる (図5 - (6)).

### 3.5 インタラクシオン

**スーパーズーミング:** ほとんどのモバイル端末に搭載されているカメラは, 解像度やズーム性能が劣るため, 遠くにある物体を正確に捉えることは困難である. そのため, 既存のモバイル AR では対象物体が小さすぎて認識できず情報を合成できなかつたり, 情報を合成できても合成して表示した情報で対象物体を覆い隠してしまうといった問題がある. この問題を解決するため, ShootAR システムは高精細なパノラマ画像をデータベース上に用意し, インタラクシオンの際にはパノラマ画像を操作可能にする. このことにより, モバイル端末に搭載されたカメラでは捉えきれないところまでズームし, 手ブレの影響なく適切な大ききで情報を表示することが可能となる. 我々は, この機能をスーパーズーミングと名付けた.

**スーパーパンニング:** モバイル端末のディスプレイサイズより十分大きなパノラマ画像をデータベースに用意しておく. インタラクシオン可能な画像がパノラマ画像に入れ替わっていれば, ユーザは現在見えている画像を簡単にパン (カメラの向きを振ること) した画像を得ることができる. モバイル端末がタッチスクリーンを備えていれば, ドラッグ操作をするだけで見えている向きを変えることができる (図6). この機能をスーパーパンニングと呼ぶ. 通常のモバイル AR では, ユーザ自身が向きを変えてカメラで捉えられる範囲を変更するが, ShootAR ではその場にいながら

ShootAR:ユーザ姿勢を考慮したモバイル AR のための操作スタイルの提案

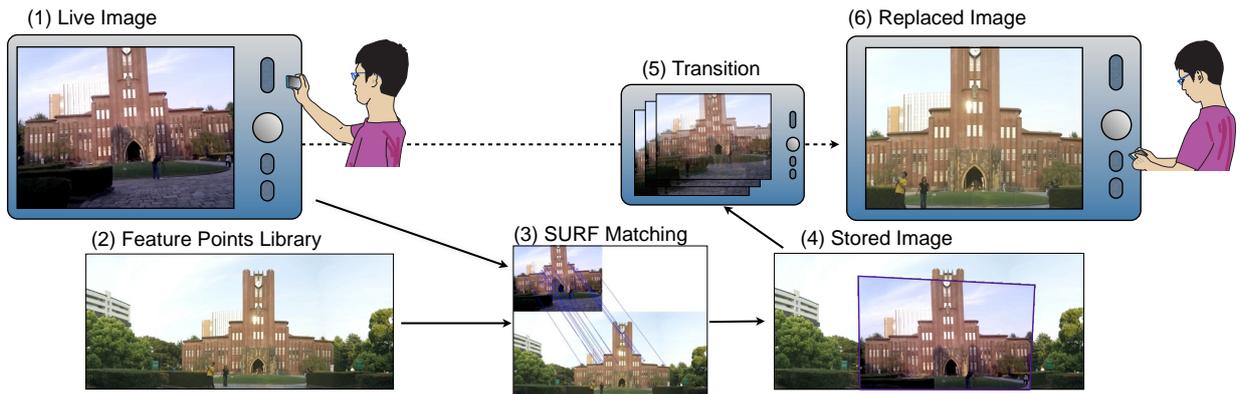


図 5 ShootAR システム構成: (1) ユーザには現在カメラが捉えている画像そのものが見えている。(2) 特徴点辞書にあるランドマークの特徴点を参照する。(3)SURF アルゴリズムを用いてカメラが捉えている画像の特徴点を計算し、特徴点辞書のランドマークと一致するものがないか計算する。(4) 一致した場合、ランドマークが含まれるパノラマ画像をデータベースから引用する。(5) パノラマ画像の一部と現在の画像がスムーズに移行するように、トランジションをする。(6) 元の画像をパノラマ画像と置き換える。



図 6 スーパーパンニング: 画面をドラッグするだけで、shoot 画像の外側を見ることができる。



図 7 疑似視点移動: ユーザが移動せずに付近の視点を得たい場合は、現在表示されている画像中の他の視点アイコンを選択すれば、その視点から見た画像に移行する。

画面操作だけで同等のことが可能となっている。

**疑似視点移動:** モバイル端末のカメラが高性能なズーム機能を備えていたとしても、遠くにある物体を安定してカメラに収めるには手ブレの問題や、オクルージョンを考慮しなければならない。通常、モバイル AR システムが情報を合成する対象をうまく捉えられなかった場合は、ユーザ自身が適切な位置まで移動する必要がある。ShootAR システムでは、ユーザ自身が動かずとも、別の場所からの視点に移動できる疑似視点移動の機能を提供する (図 7)。ユーザは画面内に表示された視点アイコンを選択することで、実際にその視点の位置に移動して見える周囲の風景を、その場で取得することができる。上記の 3 つの機能を組

み合わせれば、ユーザはその場にいながらにして 3 次元空間を移動したような視点を得ると同時に、その空間に電子的な情報を重ね合わせて見ることができる。

**シンクロナイズドオーバーレイ:** 既存のモバイル AR は、ビデオシーズルであるため、現実の情報はカメラが捉えたものをそのままディスプレイに表示すれば良い。一方で ShootAR システムの場合、ユーザがインタラクションすることができるのは同じ場所の過去の画像である。そのため、ユーザがインタラクション可能な画面中に、リアルタイムな情報が反映されていない可能性がある。しかし、静止画に対して情報を合成するという事は、既存のモバイル AR のよう

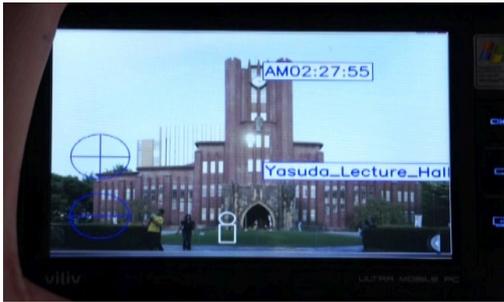


図 8 シンクロナイズドオーバーレイ: 時計や電光掲示板の表示を, リアルタイム情報を合成して表示することで, 視認性の高い形式や背景に溶け込む形式などで提示できる.



図 10 ユーザ操作画面

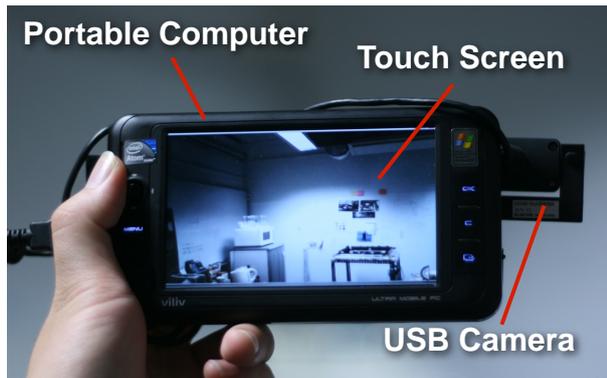


図 9 プロトタイプ: タッチスクリーンと背面に USB カメラを備えたポータブルコンピュータ.

にリアルタイムで実物体の位置を特定せずとも画像中の対象に合わせて AR を表示させることができるというメリットを ShootAR システムは持ち合わせている<sup>13)</sup>. そこで, ユーザに必要なリアルタイム情報を画像中に AR として表示させることにより, 通常のモバイル AR と同様にユーザがリアルタイム情報を得られるようにする (図 8). たとえば, 時計を実時間に合わせた情報でオーバーレイする, 電光掲示板の情報を最新のもので置き換えるといったことが考えられる. 情報の提示形式も, 実物と見分けがつかないように背景に溶け込む形式で表示させる, あるいは内容がはっきりと読み取れる視認性の高い形式やなどで表示するといったことが可能となる.

#### 4. ShootAR システム実装

##### 4.1 システム構成

我々は, ShootAR のコンセプトを実現するプロトタイプを製作した. 機器構成としては, タッチスクリーンを備えたポータブルコンピュータと, その背面に小型の USB カメラを装着している (図 9). タッチスクリーンの解像度は  $1024 \times 768$  pixel. USB カメラで取得する画像の解像度は  $320 \times 240$  pixel. データベー

ス上のパノラマ画像の解像度はおよそ  $3000 \times 1000$  pixel となっている. なお今回は, パノラマ画像 (計 5 視点) を保存しているデータベース, およびランドマークの特徴点辞書はポータブルコンピュータ内に用意した.

今回の実装の場合, ShootAR システムは USB カメラで取得した画像の特徴点抽出を常にバックグラウンドで行っているため, ユーザが shoot を行う前のディスプレイ表示は 1 秒当たり 2 - 3 フレームとなっている. そして, 取得画像中の特徴点と特徴点辞書中の特徴点が一定以上一致した場合, システムは画面上に “shoot” 可能であることを知らせるターゲットマークによるフィードバックを生成する.

##### 4.2 インタラクションの実例

“shoot” した後の画面におけるインタラクションとして, ズームイン/アウトをしたい場合は, 画面左下にある + (プラス) ボタンまたは - (マイナス) ボタンを押せば良い (図 10). また, スーパーパンニングについてはタッチスクリーンをドラッグすることで実装している. 画像中にある人間のシルエットを模したアイコンを選択すれば, その視点から見た画像に切り替えることができる.

## 5. 議 論

### 5.1 画像データベースの構築

今回のプロトタイプにおいては, 事前にカメラで撮影した画像から手動でパノラマ画像を作成した. しかし, すでに複数の写真からパノラマ画像を自動で生成する方法は存在している<sup>14)</sup>. これらの技術を応用して, 街中や駅といったパブリックスペースにおいては Flickr<sup>15)</sup> などのユーザが投稿した写真から, 美術館や博物館といった所有者が明確な空間では所有者が撮った写真から, 自動的に ShootAR で利用可能なデータ

ベースを作成することが可能になると考えている。そうであれば、現実空間にマーカやビーコンを設置するよりも、安易に低コストでモバイル AR システムが認識可能な空間を用意できるはずである。さらに、特定の対象を異なる場所から写した複数の写真から、3次元形状モデルを復元することも可能となっている<sup>16)17)</sup>。こうした技術と ShootAR を組み合わせれば、モバイル端末で1枚の写真を撮るだけで、それを手がかりとした周囲の空間の3次元情報を取得でき、手元で3次元空間を自由に動かして見るというようなインタラクションが可能になるであろう。

## 5.2 データ読み込み

今回のプロトタイプでは、画像データベースおよび特徴点辞書をモバイル端末内に用意したため、データを読み込む際のオーバーヘッドはほとんど存在しなかった。しかし、世界中で ShootAR システムが利用可能になったとすると、そのパノラマ画像の量は膨大なものとなり、とてもモバイル端末に置けるようなものではなくなる。そこで、モバイル端末の通信機能を使用し、ネットワーク上のサーバから画像データベースおよび特徴点辞書を適宜ダウンロードする方式が考えられる。ShootAR システムでは、データベースにあるパノラマ画像の容量は1枚当たり約30メガバイトになっている。1つの特徴点を記述するために必要な容量は512バイトである。特徴点辞書については、1つのランドマークにつき約500ヶ所の特徴点を保存しなければならぬとしてので、ランドマーク1つあたりの容量は約250キロバイトとなる。ここで、特徴点辞書全体の容量が問題となるが、モバイル端末の電子コンパスやGPSの情報を組み合わせることにより特徴点辞書の比較対象範囲を絞ることができれば、ユーザの利用場所で必要なランドマークを10ヶ所程度に限定することが可能であると考えられる。ランドマークが10ヶ所程度であれば、特徴点辞書の容量は2.5メガバイト程度になるので、モバイル端末の通信速度でもストレスの無い速さでダウンロードして、ShootAR システムのインタラクションを開始することが可能となるはずである。

## 5.3 実世界指向リモコン

ARにおける遠隔物体を操作することの最大の利点は、ユーザの見ているものと操作対象が1対1に対応しており、直感的に分かりやすいインターフェースとなっているということである<sup>18)</sup>。しかし、モバイル端末に搭載された性能の低い演算処理装置では、リアルタイムで現実空間を認識し正確な位置合わせを行った描画することは困難である。また、ユーザが片手で

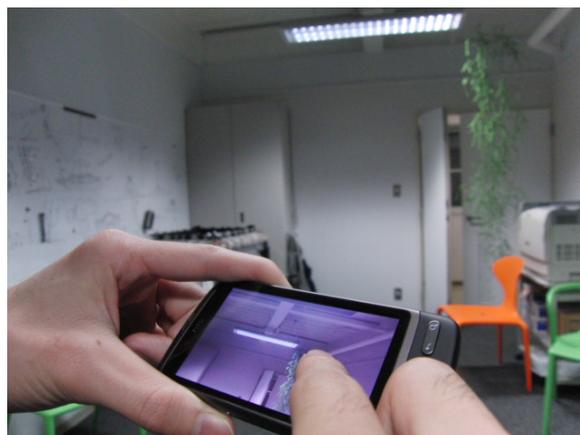


図 11 実世界指向リモコン: ShootAR を利用すれば明示的に操作対象を指定して、自由な姿勢で対象を操作できる。

モバイル端末を把持しもう片方の手で操作するとなれば、手ブレの影響を受けやすくなり、本来行いたい機器の操作ではなく、モバイル端末を安定させることに注力しなければならなくなる。ShootAR の場合、ユーザがインタラクションをする対象は現実空間を写した静止画であり、カメラの手ブレによる影響を受けることなく、ひとつひとつの物体を正確に分けて認識することができる(図 11)。物体が認識できれば、モバイル端末の通信機能を用いて個別に対象を操作することが可能である。

## 5.4 写真の拡張

ここまでで議論してきた ShootAR スタイルは写真の拡張によって実現していると見ることができる。つまり、通常の写真はデジタルであれアナログであれ、情報としては2次元の画素情報しか持たない。はじめにある写真を手がかりとして、はじめの写真には写っていなかった部分を見ることができるようになっていたり、その場所に関連付けられた情報を引き出せるようにすることは、技術的に可能である。これは、通常の写真ではなく、様々な情報にアクセスできる拡張された写真であると言える。この写真の拡張をモバイル端末上で実現することにより、モバイル AR におけるユーザが常にモバイル端末を掲げなければいけないという問題を解決することができる。

## 6. おわりに

本論文では、既存のモバイル AR における問題点を指摘した。ひとつは、モバイル AR 利用時のユーザの姿勢が負荷の高いものであるという点であった。もうひとつは、モバイル AR において物体が認識できない状況がしばしば発生するという点であった。これらの問題点を解決するため、我々は“ShootAR”というモ

バイル AR インタラクシオンのモデルを提案した。さらに、ShootAR モデルのプロトタイプを作成し、議論を通じて ShootAR モデルの可能性を明らかにした。ShootAR モデルでは、ユーザに直感的な AR 対象空間の指定をさせた後、データベース上の高精細パノラマ画像とディスプレイ上の画像をすり替えることにより、ユーザに空間的連続性を意識させたままインタラクシオン可能な状況に移行することができる。重要なのは、インタラクシオン時にユーザがモバイル端末を対象物体に向けて掲げる必要がないということであり、また従来のモバイル AR よりも高度なインタラクシオンが可能となっているということだ。今後の課題としては、実世界指向リモコンの実装することが挙げられる。モバイル端末上での 3 次元形状モデルのインタラクシオンについても研究したいと考えている。

### 参 考 文 献

- 1) Rekimoto, J. and Nagao, K.: The world through the computer: Computer augmented interaction with real world environments, *UIST '95: Proceedings of the 8th annual ACM Symposium on User Interface and Software Technology*, pp.29–36 (1995).
- 2) Mohring, M., Lessig, C. and Bimber, O.: Video see-through AR on consumer cell-phones, *ISMAR '04: Proceedings of the 3rd IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 252–253 (2004).
- 3) “セカイカメラ” : <http://sekaicamera.com/>.
- 4) “Layar” : <http://www.layar.com/>.
- 5) Boring, S., Baur, D., Butz, A., Gustafson, S. and Baudisch, P.: Touch projector: Mobile interaction through video, *CHI '10: Proceedings of the 28th International Conference on Human Factors in Computing Systems*, pp.2287–2296 (2010).
- 6) Wimmer, R., Schulz, F., Hennecke, F., Boring, S. and Hußmann, H.: Curve: Blending horizontal and vertical interactive surfaces, *Tabletop '09: Adjunct Proceedings of the 4th IEEE Workshop on Tabletops and Interactive Surfaces* (2009).
- 7) Lee, G.A., Yang, U., Kim, Y., Jo, D., Kim, K.-H., Kim, J.H. and Choi, J.S.: Freeze-Set-Go interaction method for handheld mobile augmented reality environments, *VRST '09: Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, pp.143–146 (2009).
- 8) Kohtake, N., Iwamoto, T., Suzuki, G., Aoki, S., Maruyama, D., Kouda, T., Takashio, K. and Tokuda, H.: u-photo: A snapshot-based interaction technique for ubiquitous embedded information, *Pervasive '04: Video Proceedings of the Second International Conference on Pervasive Computing*, pp.1–4 (2004).
- 9) Tani, M., Yamaashi, K., Tanikoshi, K., Futakawa, M. and Tanifuji, S.: Object-oriented video: Interaction with real-world objects through live video, *CHI '92: Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp.593–598 (1992).
- 10) Hirose, M., Hirota, K., Kijima, R., Kanno, M., Hayakawa, K. and Yokoyama, K.: A study on synthetic visual sensation through artificial reality, 計測自動制御学会ヒューマン・インタフェース研究論文集, Vol.1, No.1, pp.19–26 (1992).
- 11) Chen, S.E.: QuickTime VR: An image-based approach to virtual environment navigation, *SIGGRAPH '95: Proceedings of the 22nd annual Conference on Computer Graphics and Interactive Techniques*, pp.29–38 (1995).
- 12) Bay, H., Tuytelaars, T. and Van Gool, L.: SURF: Speeded up robust features, *Computer Vision ECCV 2006* (Leonardis, A., Bischof, H. and Pinz, A., eds.), Lecture Notes in Computer Science, Vol. 3951, Springer Berlin / Heidelberg, pp.404–417 (2006).
- 13) 内村圭一, 胡 振程, 三好正純: リアルタイム映像を用いたシームレス・ナビゲーションシステムの開発, 電気通信普及財団研究調査報告書 (CD-ROM), Vol.21, pp.04–01030 (2006).
- 14) Kopf, J., Chen, B., Szeliski, R. and Cohen, M.: Street slide : Browsing street level imagery, *ACM Transactions on Graphics*, Vol.29, No.4, pp.96:1–96:8 (2010).
- 15) “Flickr” : <http://www.flickr.com/>.
- 16) Snavely, N., Garg, R., Seitz, S.M. and Szeliski, R.: Finding paths through the world’s photos, *SIGGRAPH '08: Proceedings of the 35th annual Conference and Exhibition on Computer Graphics and Interactive Techniques*, pp.15:1–15:11 (2008).
- 17) Goesele, M., Ackermann, J., Fuhrmann, S., Haubold, C., Klowsky, R., Steedly, D. and Szeliski, R.: Ambient point clouds for view interpolation, *ACM Transactions on Graphics*, Vol.29, pp.95:1–95:6 (2010).
- 18) 綾塚祐二, 松下伸行, 暦本純一: 「見ているものに接続する」というメタファによる実世界指向ユーザインタフェース, インタラクシオン 2000, pp.181–188 (2000).