

モバイルマジックハンド:2次元コードとオプティカルフローを用いたカメラベースのインタラクション

Mobile Magic Hand:Camera phone based interaction using a visual code and optical flow

吉田 悠一 宮奥 健人 佐藤 隆*

Summary. We propose the “Mobile Magic Hand” interface; it is an extension of the previous visual code-based interface system. Once the user acquires the visual tag of interest, the user can then manipulate the related virtual object/system without having to keep the camera centered on the visual code. Our new interface does this analyzing the optical flow as captured by the camera. For example, consider a visual code that represents a 3D object, such as a dial. After selecting the code, the user can freely rotate and/or move the virtual object without having to keep the camera pointed at the code. This interface is much more user friendly and is more intuitive the user’s hand gestures can be more relaxed, more natural, and more extensive.

In this paper, we describe “Mobile Magic Hand”, some applications, and a preliminary user study of a prototype system.

1 はじめに

モバイル端末においてボタンによる操作を実現するために、2次元コードに対する端末の相対的な位置・姿勢を計測し、ジェスチャによる操作を提供するインターフェースが提案されている [1, 2, 3]。この手法は、2次元コードから得られる情報をジェスチャによって操作できるため、入力の自由度が高く、操作が直感的であるという利点がある。

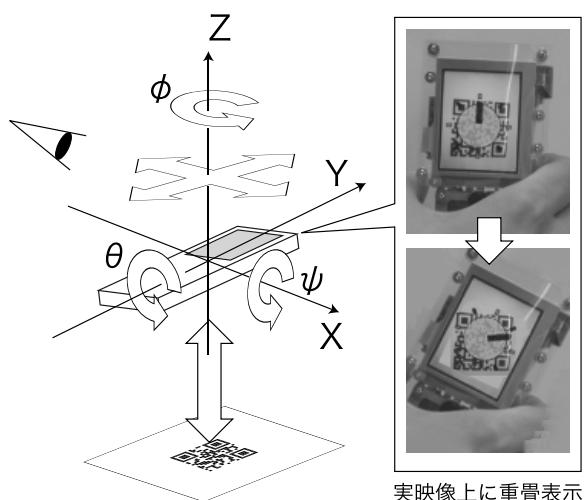


図 1. 2次元コードとモバイル端末の位置・姿勢関係を利用したインターフェース

Copyright is held by the author(s).

* Yuichi Yoshida, Kento Miyaoku, Takashi Satou, 日本電信電話株式会社 NTT サイバーソリューション研究所

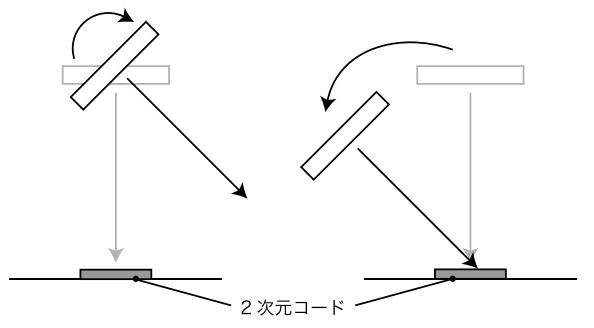


図 2. 空間的操作における制限の例

このインターフェースは、一般的に以下の手順で実現される(図1)。まず、カメラ付きモバイル端末で2次元コードを撮影し、IDを取得する。次に2次元コードの見かけの形状から、コードに対する端末の位置と姿勢を計算する。最後に、IDに対応するコンテンツに、この位置と姿勢に応じて変化を加えて、実映像上に重畠表示する。図1右の写真は、実映像に重畠されたダイアルを見ながら端末を回し、ボリュームを調整する例である。

しかしながら、この手法では、ジェスチャによる入力の間、ユーザは常に2次元コードをカメラの画角に収め続ける必要があり、端末を動かせる範囲や傾けられる範囲が限定される問題点がある。例えば「傾ける」という操作をする場合、図2-(a)に示すように端末の中心の軸周りに回転させることがジェスチャとして直感的である。しかし、そうするとすぐに2次元コードがカメラの画角から外れ、操作が

できなくなる。このため、2次元コードを用いた従来手法では、ユーザは図2-(b)のように2次元コードの位置を基準にモバイル端末を移動させるジェスチャを強いられることになる。

そこで、本研究では、2次元コードを検出する時に端末の初期位置・姿勢を計測し、その後のコードが検出されなくなったときの位置・姿勢を背景動画像のオプティカルフローによって算出し、ユーザのジェスチャによる操作可能な範囲を拡張する「モバイルマジックハンド」を提案する。モバイルマジックハンドでは、ユーザの動作が2次元コードを検出可能な画角に限定されないため、ユーザのジェスチャによる入力の操作性が向上することが期待される。

以下では、2次元コードとカメラを用いたジェスチャインターフェースの関連研究、そしてモバイルマジックハンドのシステムおよび特徴について述べる。さらに、プロトタイプシステムを用いたモバイルマジックハンドの実験結果、最後にそのアプリケーションの例について説明する。

2 関連研究

2次元コードに対するモバイル端末の位置・姿勢、および端末の動きを利用したインターフェースに関する関連研究について述べる。

2.1 2次元コードに対するモバイル端末の位置・姿勢を利用するインターフェース

このインターフェースでは、2次元コードに対するモバイル端末のX,Y,Z軸の移動量とこれらの3軸を中心とした回転量の合計6自由度のパラメータを用いる（図1の座標軸を参照）。X, Y軸方向への移動と、X, Y軸を中心とした回転をそれぞれまとめ、pointing, distance, tilting, rotationの4種類に分類する[2]。これら4種類の操作を本稿では空間的操作と呼ぶ（表1）。2次元コードと空間的操作を応用し、紙面上の印刷内容を活かした操作や実世界の動作をメタファとした操作を提供できる。例えば、ダイアルを回してボリューム調整をするような操作や、地図に虫眼鏡を近づけて、詳細を見るような操作を提供できる。

MIXISは、円形のマーカーおよび内部の図柄によって操作対象を識別、およびそれに対する端末の位置・姿勢を計算し、マーカーに関連づけられた情報を操作するインターフェースである[1]。地下鉄の路線図上の駅のマーカーを読み取り、その駅を経由する路線名をdistanceによってスクロールして閲覧するアプリケーションを構築している。このアプリケーションを使って、キーによる操作と比較した実験を行い、distanceの方が好まれるという主観的な評価を得たという。

Rohsは上記の4種類の空間的操作に加えて、キー

表 1. 空間的操作

名前	図	説明
pointing		(x, y) … X, Y 軸方向の移動
distance		(z) … Z 軸方向の移動
tilting		(ψ, θ) … X, Y 軸周りの回転
rotation		(φ) … Z 軸周りの回転

操作のkeystroke、静止状態のstayを提案し、これらを組み合わせたアプリケーションの操作性に関する評価を行っている[2]。

しかしながら、これらの従来研究では、2次元コードが検出可能な範囲に空間的操縦が限定される問題点に対するアプローチは提案されていない。

2.2 端末の位置・姿勢の測定手法

以下のような端末の位置・姿勢を測定する手法がある。

加速度センサ

加速度センサによって、姿勢や位置を測定することができる。加速度センサは、モバイル向けの入力手段の一つとしても提案されており、文字入力ためのインターフェースがすでに構築されている[4]。

ロータリーエンコーダ

マウスと同様にボールとロータリーエンコーダを用いることで、平面上の平行移動量を測定することができる。バーコード読み取り機能とこれを組み合わせたペン型デバイスFieldMouseが提案されている[5]。例えば、そのデバイスでバーコードを読み取った後、紙面をなぞり、任意の場所をクリックすると、そこに関連づけられた情報を取得することができる。しかし、FieldMouseは2次元平面上の操作を提供するものであり、奥行き方向の移動量や回転量による操作を実現することができない。

画像特徴量を用いた手法

カメラで撮影した実映像のフレーム毎の画像特徴量を利用し、端末の位置・姿勢を測定することができる。被写体を剛体と見なした動きベクトルを算出

するオプティカルフローや画像の領域毎の色ヒストグラムを利用するブロックマッチングなどいくつかの手法がある [6]。すでに、ブロックマッチングによって携帯電話の動きを検出し、上下左右方向の入力に利用するインターフェースが提案されている [7]。

加速度センサやロータリーエンコーダを利用する手法は、モバイル端末に何らかのデバイスを付加する必要があり、安価に実現できない。画像特徴量を用いる手法は、カメラ以外のデバイスが不要なため、一般的な携帯電話での実現性が高い。一方、この手法には、計算結果が実空間の移動量ではないという短所がある。ただし、2次元コードと組み合わせて利用する場合、コードのサイズとカメラパラメータを用いて、実空間での移動量を計算できる。

3 モバイルマジックハンド

モバイルマジックハンドは、2次元コードが検出可能な状態では、コードの見かけの形状から、端末の初期の位置・姿勢を測定する。一方、2次元コードが検出不可能な状態では、背景動画像のオプティカルフローによって端末の位置・姿勢を測定する(図3)。

モバイルマジックハンドは、QRコードを2次元コードとして利用する[8]。QRコードは、近年、新聞や雑誌、CD、ポケットティッシュなど様々な媒体上でWebへのダイレクトな移動手段として利用されており、モバイルマジックハンドの普及や実用化に向けた有用かつ安価な手段であると考える。

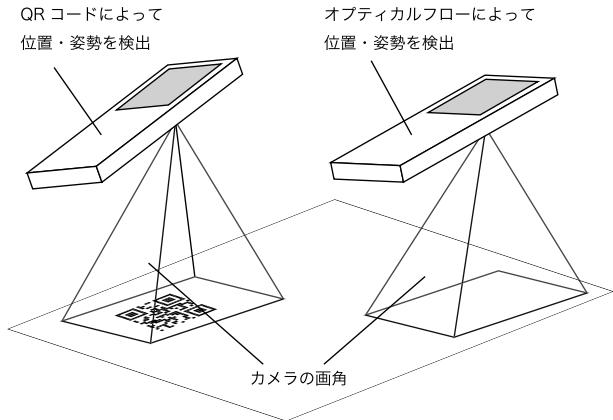


図 3. 2 次元コードとオプティカルフローの切り替え

3.1 実現手法

一連の処理の流れは、図4のようになる。まず、QRコードから取得されるIDを元に、コード自体の物理サイズとコンテンツ、コードに対して行う空間的操作の種類を取得する。

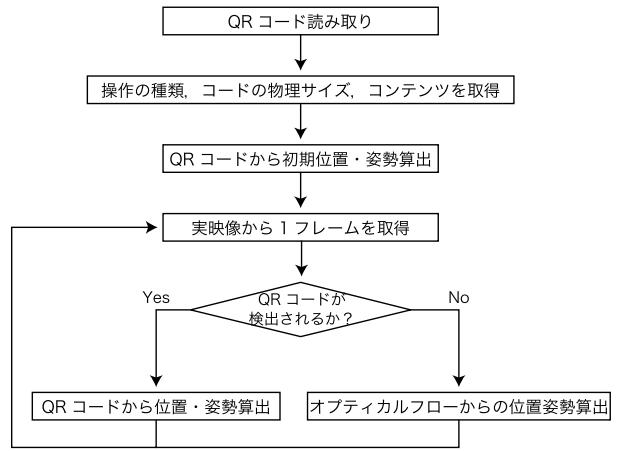


図 4. 処理の概要

QRコードからの位置・姿勢検出

QRコードが検出される場合は、画像中のQRコードの四隅の点から、QRコードが正方形であるという制限とコードの物理的なサイズを用いて、QRコードに対するモバイル端末の相対的な位置と姿勢を算出する[9]。

オプティカルフローからの位置・姿勢検出

まず、フレーム中の特徴点とオプティカルフローを算出する[6]。表1に分類された位置と姿勢を、各々以下のようなアルゴリズムで算出する。

pointing $\vec{v}_n = (v_{x n}, v_{y n})$ は、フレーム番号 n のときのオプティカルフローの平均ベクトルである。フレーム毎に、QRコード読み取り時の X, Y 軸方向の位置を初期値 $P_0 = (x_0, y_0)$ とし、それに \vec{v}_n に物理単位に変換するための尺度 k_P を乗算し、加算する。

$$P_n = P_0 + k_P \cdot \sum_{k=1}^n \vec{v}_k$$

distance D_n は、フレーム番号 n のときの特徴点の密度である。QRコード読み取り時の Z 軸方向の位置を初期値 z_0 とし、特徴点の密度のフレーム間差分 ($D_n - D_{n-1}$) を物理単位に変換するための尺度 k_z を乗算し、加算する。

$$z_n = z_0 + k_z \cdot \sum_{k=1}^n (D_k - D_{k-1})$$

tilting QRコード読み取り時の X 軸と Y 軸を中心とした回転角度をそれぞれ、初期値 ψ_0, θ_0 とし、フレーム番号 n のときのオプティカルフローの平均ベクトルに角度に変換するための尺度 k_ψ, k_θ を乗算

し，加算する．

$$\psi_n = \psi_0 + k_\psi \cdot \sum_{k=1}^n (v_y|_k - v_y|_{k-1})$$

$$\theta_n = \theta_0 + k_\theta \cdot \sum_{k=1}^n (v_x|_k - v_x|_{k-1})$$

rotation フレーム番号 n のときの特徴点全体の重心と各特徴点を結んだ直線が X 軸と成す角度の平均値を ω_n とする．QR コード読み取り時の Z 軸周りの回転角度を初期値 ϕ_0 とし，平均角度のフレーム間差分 ($\omega_n - \omega_{n-1}$) に誤差を減らすための尺度 k_ϕ を乗算し，加算する．

$$\phi_n = \phi_0 + k_\phi \cdot \sum_{k=1}^n (\omega_k - \omega_{k-1})$$

尺度の計算

QR コードが検出できる間も，同時にオプティカルフローからも位置・姿勢を検出する．その間，QR コードから算出される値 $(x, y, z, \psi, \theta, \phi)$ と，オプティカルフローから算出されるこれらの値との誤差を最小にする尺度 $(k_P, k_z, k_\psi, k_\theta, k_\phi)$ を 1 フレーム毎にそれぞれ計算する．

3.2 特徴

モバイルマジックハンドは，QR コードが検出不可能な空間でも空間的操作を実現することができる．このため，従来手法と比較し，以下のような効果を期待できる．

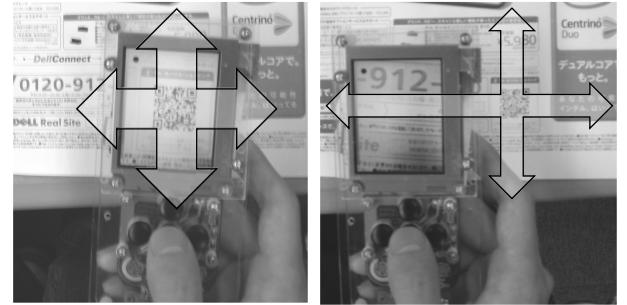
パラメータの値域の拡張

ユーザがモバイル端末を動かせる範囲や傾けられる範囲が拡張され，空間的操作の操作パラメータの値域が拡張される．例えば，pointing の場合，図 5 で示すように動かせる範囲が広がる．これを利用し，pointing によってメニュー項目を選択する場合，選択可能なメニューの項目数を増やすことができる．また，地図のズーム操作を distance に割り当てる場合，その縮尺の値域を広げられる．

空間的操作の操作性向上

ユーザは，常に 2 次元コードをカメラの画角に收め続ける必要がなくなる．例えば，図 6-(a) に示すように，ユーザは，2 次元コードを無視して傾けることができる．このため，ユーザは，実際に携帯電話を傾けることで，あたかも画面に表示されたオブジェクトを実際に傾けたような感覚で閲覧操作ができる．

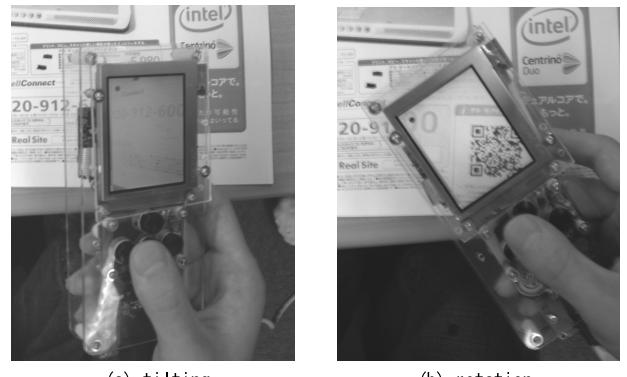
また，ユーザは，空間的操作の間，2 次元コードの検出状態を意識せずに操作できる．例えば，ユーザは図 6-(b) に示すような空間的操作中の一時的に



(a) 従来手法
検出範囲内に
pointing が制限される

(b) モバイルマジックハンド
pointing の範囲が拡張される

図 5. パラメータの値域の拡張の具体例



(a) tilting
「傾ける」ことで，操作できる．

(b) rotation
2 次元コードが画角から
外れても操作可能

図 6. 空間的operationの操作性向上の具体例

2 次元コードが検出されない状態でも操作ができる．このため，ユーザの動きに“遊び”ができる，空間的操作を滑らかにできる．

4 評価実験

プロトタイプシステムを利用して，空間的操作に対するモバイルマジックハンドの効果を評価するため，予備的な評価実験を行った．

4.1 実験条件

プロトタイプシステムのモバイル端末は，NTSC の液晶モニタ，USB の 4 つのボタン，USB の Web カメラ (320x240, 30fps) を備え，これらすべてのデバイスは，WindowsPC に有線で接続される．QR コード，オプティカルフローの検出，操作結果描画等の処理は，これらデバイスが接続された WindowsPC でおよそ 24fps で行われる(図 7)．

予備実験として，数名の被験者にこのシステムを利用してもらい，操作性を評価した．rotation, tilting, distance では，画面上に表示される距離や角度

Mobile Magic Hand: Camera phone based interaction using a visual code and optical flow

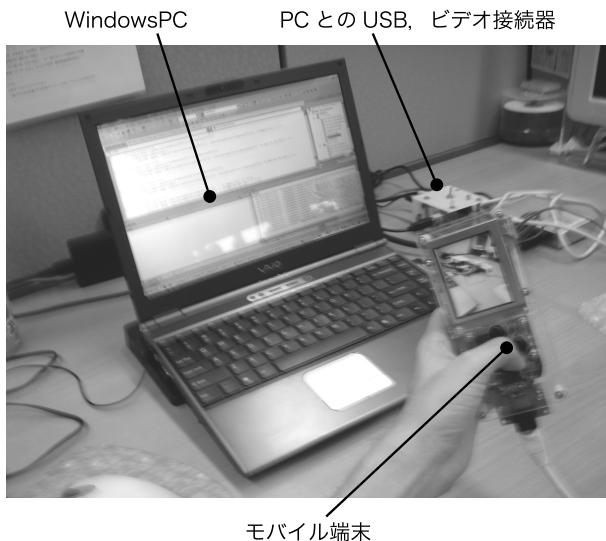


図 7. プロトタイプシステム

を，予め指定した数値に調節するというタスクを設定した。pointing では，QR コードの周囲に印刷されたオブジェクトのうち，指定された一つを選択するというタスクを設定した。各ユーザに各操作の方法に関する簡単な説明を行った後，各タスクを実施してもらった。

4.2 結果

rotation, tilting, distance に関しては，操作方法に関する簡単な説明を与えたのみで，全てのユーザが比較的スムーズにタスクをこなすことができた。

しかし，pointing については，指定した対象をうまく選択できないユーザがみられた。これは，ユーザが端末をすばやく移動させたとき，オプティカルフローによって算出される平行移動量の誤差が大きくなるシステムの制約に起因する。フレーム毎の特徴点を適切に抽出するために，端末の移動速度をある程度ゆっくり動かさなければならない。このユーザに対しては，ゆっくりと端末を動かすように促することで，うまくタスクが達成されるようになった。

4.3 考察

ユーザがモバイル端末を動かせる範囲が，2 次元コードを検出可能な空間に制約されないという提案手法の特徴は，rotation, tilting, distance を使いやすくすることに寄与すると考える。今後は，具体的な操作性への効果について既存手法と比較した評価を行う必要がある。

一方，pointing ではモバイル端末の移動速度が速い場合，平行移動量の計算に問題が生じ，操作性が悪くなることがわかった。したがって，ブロックマッチングなど他の手法を考慮し，移動速度にロバストな手法を検討する必要がある。

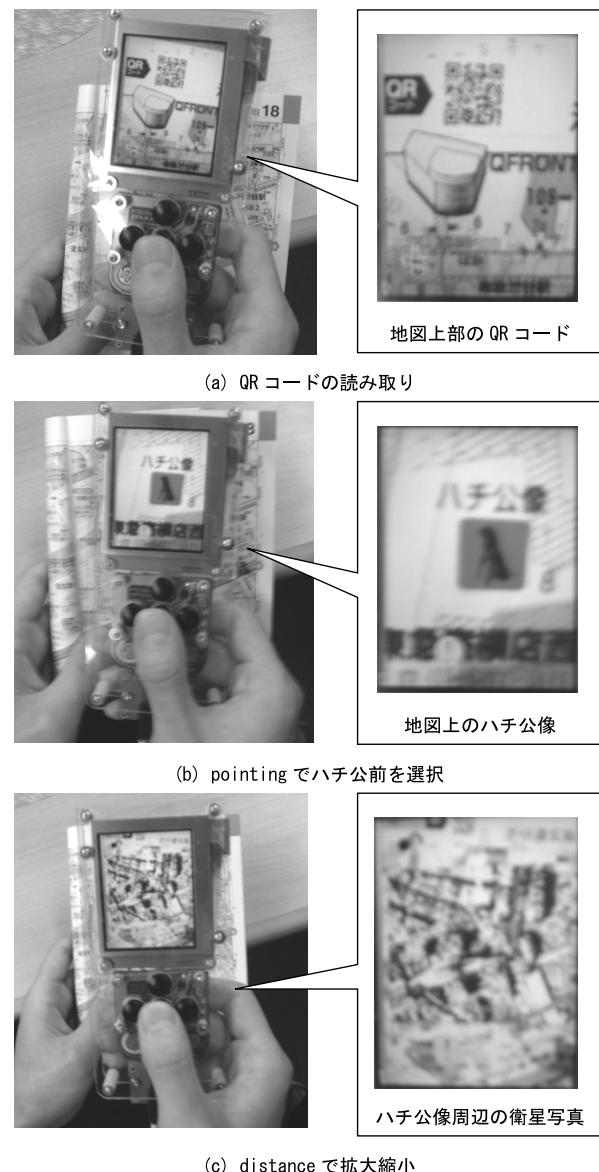


図 8. インタラクティブ地図

5 アプリケーション

モバイルマジックハンドの特徴を活かしたアプリケーションについて説明する。

5.1 インタラクティブ地図

モバイルマジックハンドのアプリケーションとして，pointing による媒体上の任意の領域を指定する操作と distance による操作を組み合わせて，インタラクティブ地図を構築できる(図 8)。QR コードには，地図上のランドマークの位置とランドマークに對して可能な空間的操作，操作時に利用する音，映像などのコンテンツが関連づけられている。

ユーザは，QR コード読み取り後(図 8-(a))，ランドマークのある場所まで，モバイル端末を平行移

動させる。pointing の間、ランドマークの位置と平行移動量を比較し、ユーザがどのランドマークを指定しているかを識別する(図 8-(b))。pointing の間、ユーザはランドマークが指定されたときにボタンを押すことで、画面上でランドマーク周辺の衛星写真を閲覧できる。さらに、ユーザは、その状態で distance をすることによって、その周辺の衛星写真を拡大縮小して閲覧できる(図 8-(c))。

5.2 アドベンチャーゲーム

モバイルマジックハンドを用いて、紙面上のオブジェクトにリアルな操作ができるペーパーアドベンチャーゲームを作成した(図 9)。図 9 に示す例では、ドアや金庫、タンス、テーブル上の書類のオブジェクトと、それに対応する QR コードが印刷されている。ドアは、tilting で開くことができる。また、ユーザは rotation によって、金庫のダイアルロックを空けることができ、下方向の pointing によってタンスの引き出しを空けることができる(図 9-(a)(b)(c))。他には、書類オブジェクトを pointing で動かして、机の上を調べられる。これら一連のアクションでシナリオを組み立て、紙のアドベンチャーゲームを実現できる。例えば、タンスを tilting で空けて金庫のダイアルロックを空ける暗証番号を見つけ、その暗証番号を rotation で金庫に入力し、金庫の扉を開けるといったシナリオが考えられる。

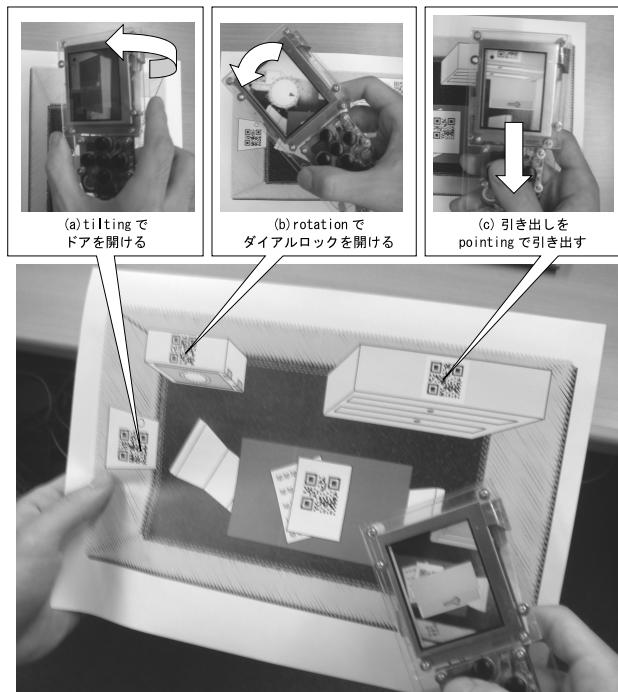


図 9. ペーパーアドベンチャーゲーム

6 まとめ

本稿では、2 次元コードが検出不可能な空間での位置・姿勢の変化をオプティカルフローを利用して算出するモバイルマジックハンドを提案した。プロトタイプシステムを用いた予備的な評価実験を行った。結果、rotation, tilting, distance の操作において、モバイルマジックハンドが有用である展望を得た。また、モバイルマジックハンドの特徴を活かしたアプリケーションについて述べた。

今後は、pointing の性能を向上させるために、より移動速度にロバストな算出方法を検討する。また、モバイルマジックハンドを用いた他のアプリケーションの可能性を検討し、モバイルマジックハンドの操作性の評価を行う予定である。

参考文献

- [1] T. R. Hansen, E. Eriksson, and A. Lykke-Olesen. Mixed Interaction Space - Expanding the Interaction Space with Mobile Devices. In People and Computers XIX - The Bigger Picture, Proceedings of British HCI, Springer. 2005.
- [2] M. Rohs and P. Zweifel. A Conceptual Framework for Camera Phone-based Interaction Techniques. Pervasive Computing: Third International Conference, PERVASIVE 2005, Lecture Notes in Computer Science (LNCS) No. 3468. 2005.
- [3] Y. Yoshida, T. Satou, K. Miyaoku, and S. Higashino. Mobile Reacher Interface for Intuitive Information Navigation. INTERACT2005. pp. 1091–1095. 2005.
- [4] D. Wigdor, and R. Balakrishnan. TiltText : using tilt for text input to mobile phones. Proceedings of the 16th annual ACM symposium on User interface software and technology. pp. 81–90. 2003.
- [5] 椎尾 一郎, 増井 俊之, 福地 健太郎. FieldMouse による実世界指向インターフェース. コンピュータソフトウェア Vol.18, No.1. pp. 28–38. 2001.
- [6] B. D. Lucas, and T. Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. International Joint Conference on Artificial Intelligence. pp. 674–679. 1981.
- [7] J. Wang, and J. Canny. TinyMotion : Camera Phone Based Interaction Methods. Conference on Human Factors in Computing Systems CHI2006 extended abstracts on Human factors in computing systems. pp. 339–344. 2006.
- [8] QR コードドットコム <http://www.qrcode.com/>
- [9] J. Rekimoto, and K. Nagao. The world through the computer:Computer augmented interaction with real world environments. Proceedings of the 8th annual ACM symposium on User interface software and technology. pp. 29–36. 1995.