

スマートフォンカメラとピンホールを用いた 3次元入力の実装と評価

田中 智泰^{1,a)} 山田 渉^{2,b)} 真鍋 宏幸^{1,c)}

概要: 本稿では、ピンホールを貼付したスマートフォンカメラの近傍に、マーカを設置し読み取ることで、リアルタイムな多次元入力を可能とする入力手法を提案する。ピンホールによって被写界深度が深まるため、カメラから10mm程度の位置に設置したマーカであっても高精細な画像が取得できる。そのため、高い精度でのマーカ位置推定、追跡が可能になる。取得したマーカ画像から、3次元座標 x , y , z および座標軸に対する角度 $roll$, $pitch$, yaw を推定することで、最大6次元の入力が可能である。今回は、アナログパッドを用いた x , y , yaw の3次元入力が可能なソフトウェアを実装し、アナログパッドを実際に動かした値とピクセル上の値から精度の評価を行った。精度評価の結果、高精度な入力が可能であることが分かった。また、アプリケーション活用例として絵画鑑賞アプリの実装を行った。

1. はじめに

近年スマートフォンは様々な事例で活用されており、非常に汎用性の高いデバイスである。そして、今後も活用できる事例は増えていくと考えられる。新たな事例での活用において、既存の入力が常に有効であるとは限らない。また、既存の事例においても、現在使われている手法よりも効率的な入力が存在する可能性もある。例えばスマートフォン HMD において現在では、単純なボタン入力のみとなっていることが多い。入力語彙を増やすために、タッチパネルでの入力やヘッドトラッキングを用いることも可能であるが、ゴーグルの形状に制限を受けるためタッチパネルは操作しにくく、ヘッドトラッキングはユーザの負担が大きい。別体のコントローラを使う方法もあるが、スマートフォンとのペアリングが必要であったり、コントローラ自体にも充電が必要など手間が多い。HMDに限らず、ブラウザや地図など大きい幅や高さを持つアプリケーションや、3Dゲームなど奥行きを持つアプリケーションは、何度もスワイプを繰り返す、複数の指を使った操作を使い分けるなどの入力が要求される。

このような課題を解決するため、また新たな用途に適した入力を行うために、既存の入力を拡張する様々な試みが

なされてきた。これまであまり入力として用いられてこなかった、スマートフォン背面を入力面とする試みや、スマートフォン内蔵センサの中でも特に多くの情報を得られる、カメラを用いたアプローチが数多く提案されている。

我々は、スマートフォンカメラとピンホールを用いた、超接近型の多次元入力が可能な手法を提案し、アナログパッドに貼り付けたマーカ座標を算出することで提案手法が機能することを示してきた [1]。本稿では、提案手法の精度評価結果と、提案手法を活用したアプリケーションについて報告する。

2. 関連研究

一般にカメラは写真や動画を撮影するために用いられるセンサであるが、コードの読み取りやジェスチャー認識など入力として用いることもできる。例えば、スマートフォンに搭載した小型プロジェクターでキーボードを壁に投影する Dotenco らの手法がある [2]。これは、投影したキーボードを人差し指でタッチし、握るという一連の動きをスマートフォンの背面カメラで撮影し、入力とする。他にも、サムズアップした手をカメラで撮影し、トラッキングすることで、手をポインティングデバイスとして使う Ishii らの手法 [3] などがある。これらではカメラと撮影対象との距離を大きく取ることが通常である。

一方で、近接した対象物を撮影し、入力として用いる手法も存在する。例えば、スマートフォン HMD において、フロントカメラで目の動きをトラッキングし、ポインティ

¹ 芝浦工業大学

² NTT ドコモ

^{a)} ma20062@shibaura-it.ac.jp

^{b)} wataruyamada@acm.org

^{c)} manabehiroyuki@acm.org

ングを行う Hakoda らの手法 [4] や、スマートフォンの背面に設置したグリッドを指でタッチし、その指の動きを鏡越しに認識する Wong らの Back-Mirror[5]、カメラで手のジェスチャを撮影し、タッチパネル入力の補助を行う Jie らの手法 [6] などである。

これらはカメラと被写体までの距離が近いことによる接近型の入力と言え、スマートフォンの近傍に新たな入力機能を追加している。一般に、カメラと被写体間の距離が短くなると、得られる画像がボケてしまい、高精度な特徴点抽出は難しくなる。そのため、前述の接近型の入力であっても、ある程度の距離を必要としていた。この距離を更に詰め、超接近型の入力を実現できれば、省スペースでの実装が可能となり、より多くの新しいインタラクションや応用が考えられる。

超接近型の入力を実現した例として、カメラを用いてスマートフォンにトラックポイントを追加する Yamada らの CamTrackPoint[7] がある。これは指をカメラに触れさせて撮影すると、指を通過した赤い光による輪郭が取得できるため、それをトラッキングすることで超接近型の入力を可能としている手法である。しかしこの手法は、ボケた画像をそのまま入力に用いるため、2次元の入力しか行えなかった。また、複数のカラーマーカを組み込んだ透明弾性体をカメラの上に配置し、弾性体を握る、弾くといった3次元的な力を検出する渡辺らの手法 [8] がある。この手法は複数のマーカを同時に認識し、それを基に力を検出する。これは立体的にマーカを配置する必要があるため形状が限られるため、適用範囲がせまく、汎用的な入力とは言えなかった。他にも Watanabe らは、3D プリンタで印刷したオブジェクトをスマートフォンケースに固定し、オブジェクト内部で反射させた光を機械学習することで、オブジェクトの変形や把持を入力として用いる手法を提案している [9]。この手法は、オブジェクトとケースを完全に一体化させるため、カメラで撮影したいときにケースごと外さなければならないという欠点があった。

このように、現在実現されている超接近型の入力手法は新しい入力機能を追加しているが欠点も大きく、まだ十分に検討されたとは言えない。

3. 提案手法

ピンホールカメラは近傍から遠方までどこにでもピントが合う特性を持っている。スマートフォン内蔵カメラの表面にピンホールを貼り付けることで、その特性を利用し、被写界深度を高めることができる。被写界深度が高ければ近接した対象であっても高精細に撮影できるため、高い精度でのトラッキングが可能である。我々の提案手法では、カメラ近傍にビジュアルマーカを配置し、姿勢推定を行う。そうすることで、x, y, z と、roll, pitch, yaw の6次元の情報が獲得でき、ビジュアルマーカを貼り付けたコン

ローラを用いた多次元の入力が可能になる。

3.1 予備調査

ここで、予備調査として、フラッシュライトの有無と被写体までの距離でどの程度精細さや明るさが変化するかを調査した。図 1, 図 2 に結果を示す。カメラからの距離は近いものから 3mm, 5mm, 10mm, 15mm, 20mm, 25mm, 30mm, 45mm であり、被写体には日本銀行券の 1000 円札を利用した。明るさの調節のため、ISO とシャッタースピードは手動で細かく変えている。フラッシュライトはそのままで指向性が強すぎるため、紙を用いて拡散させている。

図 1 では ISO を 6400 に固定し、シャッタースピードは $\frac{1}{30}$ から $\frac{1}{5}$ の幅で細かく調整している。ISO をかなり高めに設定しているにも関わらず、全体的に暗い。(e) や (f) などを見ると、明るさこそ足りないものの、10mm, 15mm など接近した状態でも精細であり、特徴点検出は十分可能であることが分かる。また、(g) 以降では暗くなってしまい、ほとんど特徴点が取得できないことが分かる。一方でフラッシュライトを点灯した図 2 は全体的に明るく高精細な画像が取得できている。特に (f) は ISO800, シャッタースピード $\frac{1}{15}$ とかなり余裕のあるパラメータで撮影出来ている。そこから更に (g), (h) と近づくと光量が不足するため、ISO を大幅に引き上げ、シャッタースピードを遅くする必要があった。

これらの画像はフラッシュライトの恩恵を十全に受けられる環境で撮影しており、実際はここより数段階明るさが低下する。例えば、後述するアナログパッドでは、アナログパッド筐体が光を遮蔽してしまい、マーカに照射する光量が低下する。また、シャッタースピードを遅くすると、フレームレートにも影響が出る。入力手法であるため、30fps 前後を目指せる程度のシャッタースピードである必要がある。多少暗くなっても影響が少なく、かつ 30fps を目指せる環境として、フラッシュライトを点灯した上で、マーカとカメラ間の距離は 10mm 前後とした。図 2f はシャッタースピード $\frac{1}{15}$ 秒で撮影しているため一見 30fps を出せないように見えるが、ISO が 800 とかなり余裕があるため、これを引き上げることで調整が可能である。

3.2 キャリブレーション

本手法はマーカをウィジェットに貼付することを想定している。しかし、マーカをウィジェットの回転中心に全く誤差なく貼り付けることは不可能であるため、ウィジェットを回転させたときにマーカの座標にズレが生じる。マーカを回転させた場合、マーカ位置はある点を中心とした円周上に位置することになる。そのズレを無視できる程度に小さくするためには、キャリブレーションが必要である。

ウィジェット中心とマーカ中心が一致していない状態で



図 1 ピンホールを貼付した状態で千円札を撮影した。フラッシュライトはつけていない。ISO 感度はかなり上げているが、全体的に暗い。ピンホールをつけることで明度が落ちることが分かる。



図 2 フラッシュライトを点灯したうえで撮影した。暗くはなるが、ISO とシャッタースピードの調整で修正できる範囲に収まった。かなり近い距離でも精細に見えていることが分かる。

マークを回転させた時に、マーク中心およびマーク回転角がどのようになるかを数式で表現し、その式で用いられているパラメータを観測値から推定する。キャリブレーションを適用することで、少なくとも無視できる程度までズレが小さくなるのが期待される。よって今回の場合は、まず yaw 回転のみを行った際の xy と yaw を取得し、誤差が最小になるようなパラメータを求める。ウィジェット位置は、ここで得られたパラメータから推定する。

[1] では、この手法に基づいて実装した装置をシステム上で確認することで、想定された入力が可能であることを示したが、実際の稼働量とシステムで推定した稼働量がどの程度一致しているのか等の精度評価は行っていなかった。今回は実際に 3 次元の入力が可能な装置を実装し、それぞれの軸に対する精度評価を行った。

4. 実装

提案手法を実現するために、水平面上での 2 次元の稼働および yaw 回転が可能なアナログパッドと、一般的なスマートフォン (Google Pixel 3) を用いた実装を行った。

本手法には、マークを認識し、入力へと変換するソフトウェアと、カメラから確認できる位置にマークが貼り付けられる程度の大きさの可動部を持つウィジェットの両方が必要となる。

まず事前準備として、スマートフォンの背面カメラに、薄いゴム製の遮光シートで作成したピンホールを貼付した。ピンホールを介するため、カメラに入射する光量が不足する。明るさを補うためフラッシュライトの点灯が必要になるが、そのままだと指向性が強すぎ、カメラに近い部分で白飛びが発生し、ライトから遠いエリアのコントラスト向上に寄与しなくなってしまう。そのため紙を用いて拡散させている。

アナログパッドは背面が観察できるように改造し、1 辺が 2mm の AR マークを貼付した。スマートフォンの背面に厚さ 10mm のアクリル板を固定し、その上に作成したアナログパッドを配置した。

ソフトウェア構成は、AndroidStudio と OpenCV で開発した、カメラを起動し、ISO やシャッタースピードなどの設定や取得画像の解析が可能なシステムを、Unity で作成したテスト・デモ環境から呼び出す形になっている。カメラパラメータは、予備調査を基に fps を落とさずはっきりと見える用に調整し、ISO6400、シャッタースピード $\frac{1}{33}$ とした。また、画像処理に時間がかからないよう画像解像度を幅 640 高さ 480 まで低下させている。

また、実装した入力が有効なアプリケーションとして、操作中も画面全体を見たいが、これまでの入力ではオクルージョンや煩雑な動作の繰り返しが発生してしまうものが挙げられる。例えば、絵画鑑賞アプリケーションや、地図アプリケーションである。どちらも細かい部分を拡大して見る、遠くから俯瞰するといった行為が繰り返しよく発生するアプリケーションである。タッチパネルを用いた入力では、拡大縮小にピンチ操作が割り当てられていることが多く、その場合動作中、画面の大部分が手によるオクルージョンが発生する。今回の実装ならば、オクルージョンを発生させることなく、拡大・縮小・移動を実行できるため有効であると言える。今回は挙げた例のうち、絵画鑑賞アプリケーションを作成した。アナログパッドを上下左右に動かすことで対象の絵を動かすことができ、yaw を左回転させると拡大、右回転させると縮小する。

4.1 キャリブレーション

図 3 に、アナログパッドの xy を変えず、yaw 回転だけした際の xy 座標を示す。黒点がキャリブレーションを行っ

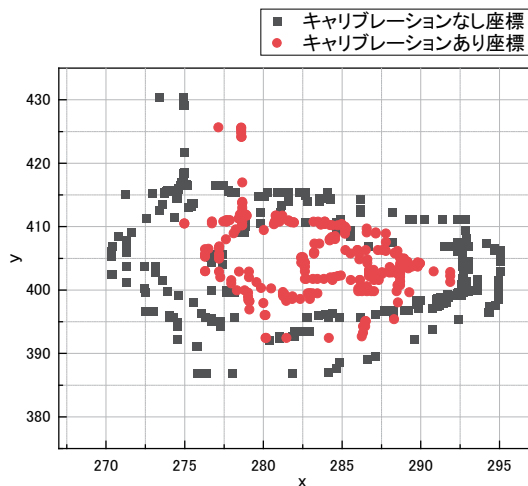


図 3 キャリブレーション有無それぞれの xy 座標

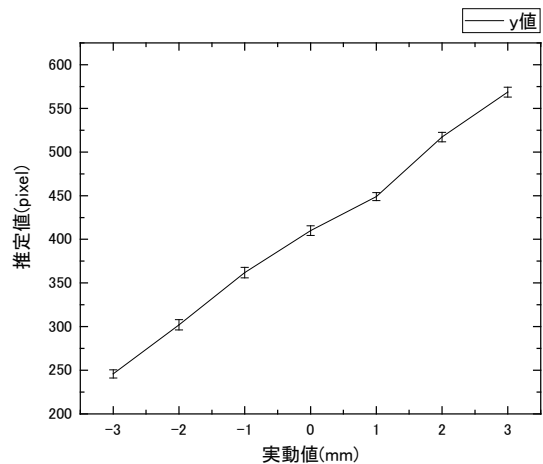


図 5 y を実際に動かした値とシステム上の値を比較した。エラーバーは標準偏差である。

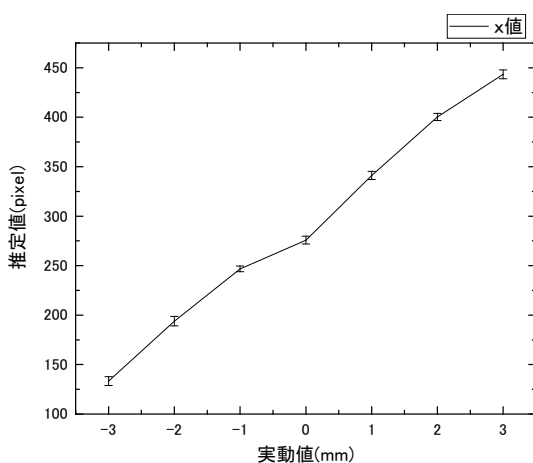


図 4 x を実際に動かした値とシステム上の値を比較した。エラーバーは標準偏差である。

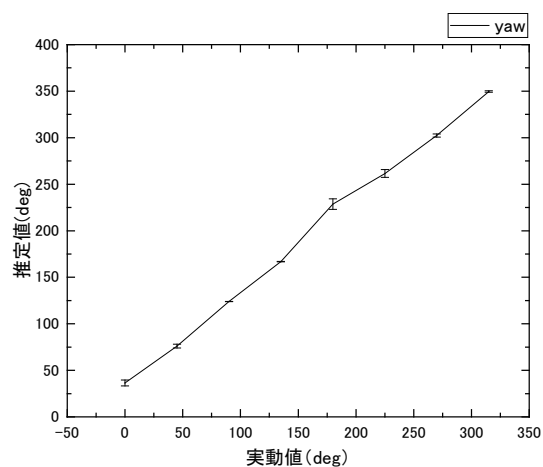


図 6 yaw を実際に動かした値とシステム上の値を比較した。エラーバーは標準偏差である。

ていない状態で計測を行ったものであり、赤点はキャリブレーション後の結果である。

キャリブレーションが完全にうまく機能しているとは言えないが、未キャリブレーションの黒に比べ、キャリブレーション後は全体的に中心に寄っており、誤差が小さくなるような補正がかかっていることがわかる。

5. 評価

どの程度の精度で入力が可能なかを示すため、アナログパッドを 1mm ずつずらしながら計測し、x, y の平均と標準偏差を求める実験及び、水平には動かさずに 45° ずつ回転させ yaw 平均と標準偏差求める実験を行った。アナログパッドの可動域は無入力の標準位置から ± 3mm であったため、x 方向と y 方向それぞれに 6 つずつ計測ポイントがある。

図 4 に x 方向にずらした際のグラフを示す。横軸が実際の移動値であり、縦軸がシステム上の数値である。グラフの点は平均値であり、エラーバーが標準偏差を表している。図 5 は y 方向の、図 6 は yaw の結果である。このグ

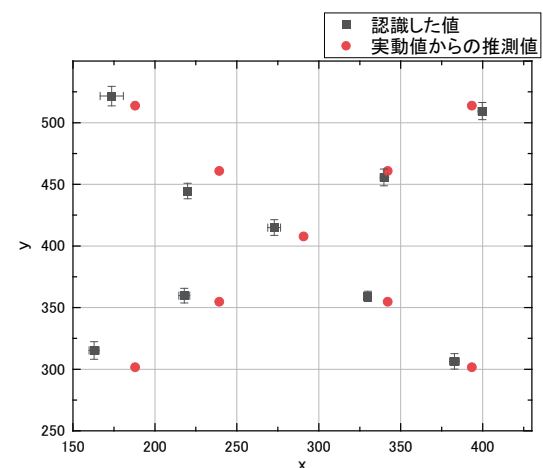


図 7 システムが認識した斜め方向の座標と水平方向の近似式を基に算出した座標。エラーバーは標準偏差である。

ラフから、少なくとも xy に関しては 6 段階の入力が可能であることが、yaw に関しては 8 段階以上の入力が可能であることがわかる。

また、このデータから、実際の移動量 x,y を基にシステ

ム上の移動量 $estx, esty$ を求めるための近似式

$$estx = 51.33x + 290.64$$

$$esty = 53.06y + 407.79$$

を得た。得られた水平方向の近似式から、斜め方向の座標推定を行い、実際に動かした座標との比較を図 7 に示す。アナログパッドの形状による制限から計測は±2mmで行っている。xの増加方向の先にフラッシュライトがあり、図の右から左に向かうにつれて暗くなっていく。この図から、フラッシュライトから遠ざかるほど数値が不安定になることが分かる。また、原点を含め全体的にxがマイナス方向にズレていることに関して、改善が必要であることがわかる。

5.1 作成したアプリケーション

作成した絵画鑑賞アプリケーションにて、実際に入力として活用できることを確認した。起動した直後はブレることもあるが、その後はyaw回転を行っても水平移動することもなく安定して拡大・縮小操作が行えた。

6. 結論

本稿では、ピンホールを貼付したスマートフォンカメラの近傍に、マーカを設置し読み取ることで、リアルタイムな多次元入力を可能とする手法を提案した。また、アナログパッドを用いて、その手法に基づいたx,y,yawの3次元入力が可能なアプリケーションを実装し、精度の評価を行った。評価の結果、x,yはそれぞれ少なくとも6段階の入力が可能であること、yawは少なくとも8段階の入力が可能であることが分かった。

一方で、マーカの検出精度に関しては調査しきれていない。体感では特定の位置、角度で検出精度が落ちる組み合わせがある。具体的にどの組み合わせで認識落ちが発生し、通常時と比べてどの程度悪いのかを調査する必要がある。また、提案手法は6次元の入力であるとしているが、z, roll, pitchの実装は行っていないため、それらを使った実装も行い、精度評価を行う必要がある。

参考文献

- [1] 田中智泰, 山田 渉, 真鍋宏幸: スマートフォンカメラとピンホールを用いた多次元入力手法, *インタラクション2020 論文集*, pp. 912–913 (2020).
- [2] Dotenco, S., Götzelmann, T. and Gallwitz, F.: Smartphone input using its integrated projector and built-in camera, *International Conference on Human-Computer Interaction*, Springer, Cham, pp. 124–133 (2014).
- [3] Ishii, A., Adachi, T., Shima, K., Nakamae, S., Shizuki, B. and Takahashi, S.: FistPointer: Target Selection Technique Using Mid-Air Interaction for Mobile VR Environment, *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, New York, NY, USA, Association for Computing Machinery, p. 474 (online), DOI: 10.1145/3027063.3049795

- (2017).
- [4] Hakoda, H., Yamada, W. and Manabe, H.: Eye Tracking Using Built-in Camera for Smartphone-Based HMD, *Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology*, New York, NY, USA, pp. 15–16 (online), DOI: 10.1145/3131785.3131809 (2017).
- [5] Wong, P. C., Fu, H. and Zhu, K.: Back-Mirror: Back-of-Device One-Handed Interaction on Smartphones, *SIGGRAPH ASIA 2016 Mobile Graphics and Interactive Applications*, New York, NY, USA, Association for Computing Machinery, (online), DOI: 10.1145/2999508.2999512 (2016).
- [6] Song, J., Sörös, G., Pece, F., Fanello, S. R., Izadi, S., Keskin, C. and Hilliges, O.: In-Air Gestures around Unmodified Mobile Devices, *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, New York, NY, USA, Association for Computing Machinery, p. 319–329 (online), DOI: 10.1145/2642918.2647373 (2014).
- [7] Yamada, W., Manabe, H. and Ikeda, D.: CamTrackPoint: Camera-Based Pointing Stick Using Transmitted Light through Finger, *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, New York, NY, USA, Association for Computing Machinery, p. 313–320 (online), DOI: 10.1145/3242587.3242641 (2018).
- [8] 渡部陽一, 佐藤克成, 牧野前野 隆司: 光学式力測定手法を用いた携帯型タッチパネル端末用入力デバイスの提案, *インタラクション'12 予稿集*, pp. 521–526 (2012).
- [9] Watanabe, C., Cassinelli, A., Watanabe, Y. and Ishikawa, M.: Generic Method for Crafting Deformable Interfaces to Physically Augment Smartphones, *CHI '14 Extended Abstracts on Human Factors in Computing Systems*, New York, NY, USA, Association for Computing Machinery, p. 1309–1314 (online), DOI: 10.1145/2559206.2581307 (2014).