



不可視の色振動を用いた AR マーカによる カメラ位置推定の基礎検討

松本晟¹⁾, 阿部知史¹⁾, 平木剛史¹⁾, 福嶋政期²⁾, 苗村健^{1), 2)}

1) 東京大学 大学院情報理工学系研究科 (〒 113-8656 東京都文京区本郷 7-3-1, {akiram, abe, hiraki}@nae-lab.org)

2) 東京大学 大学院情報学環 (〒 113-8656 東京都文京区本郷 7-3-1, {shogo, naemura}@nae-lab.org)

概要: スマートフォンなどに搭載されるカメラで AR マーカを認識することで、カメラの位置を推定する技術が普及している。しかし、認識を容易にするためマーカ提示面自体の表現が制限されるという問題がある。そこで、ディスプレイの各画素の色度を高速で振動させて重畳した不可視の AR マーカからスマートフォンの位置を推定する手法を提案する。本稿では、位置が 1 mm 程度の誤差で推定可能であることを確認した。

キーワード: 位置推定, ディスプレイ, スマートフォン, 色振動, ユビキタスコンピューティング

1. はじめに

ディスプレイ-スマートフォン間の連携の幅を広げるために、ディスプレイに対するスマートフォンの位置を手軽に取得できることが望ましい。これに関して、ディスプレイに表示したマーカをスマートフォンに搭載されるカメラで撮影する手法が用いられてきた [1]。しかし、ディスプレイに可視のマーカを表示すると人間にとって意味のない領域が映像上に現れることになり、コンテンツの見た目を損なうという問題を抱える。そこで、画像の局所特徴量を用いる手法 [2] や、特殊な装置を使って不可視の AR マーカを埋め込む手法 [3] などが提案されている。これらは、映像の見た目を損なわない一方、前者は常に特徴点が豊富な映像をディスプレイに提示する必要があり、後者は赤外線プロジェクタといった特殊な装置を必要とする。

筆者らは、輝度は一定のまま色度を振動させる色振動によりリフレッシュレート 60 Hz の一般的なディスプレイでも不可視情報の伝送を可能にする研究を行ってきた [4]。そこで、本研究では特殊な装置を用いずに、色振動を用いて不可視の AR マーカをディスプレイ上に並べて表示し、それらをスマートフォンのカメラで読み取ることでスマートフォンの位置を推定する手法を提案する。本稿では、提案するシステムのプロトタイプを実装し、スマートフォンがディスプレイに対して水平な条件下での位置推定精度を評価した。

2. 関連研究

2.1 ディスプレイ-スマートフォン間の位置推定

ディスプレイとスマートフォンの位置関係を推定する手法として AR マーカのような可視パターンを用いる手法がある [1]。これは高精度で位置を推定できるが、人間にとって全く意味のないコード領域が場所をとることになる。そのため、ディスプレイの映像の見た目を損なってしまう。

この問題を解決するため、様々な手法が提案されている。Chan らは、可視光プロジェクタで映像を投影し、赤外線プロジェクタで不可視の AR マーカを投影する手法を提案している [3]。これにより、映像を損なわずにマーカを埋め込むことができるが、赤外線プロジェクタのような特殊な装置を用いる必要がある。一方、THAW は色の勾配を用いたマーカをスマートフォンがある位置に局所的に表示することで位置情報を伝送している [5]。このシステムは、一般に普及している装置のみで実現できる。ディスプレイとスマートフォンを接触させた状態においてはスマートフォンがマーカを隠す。ゆえに、人間にはマーカは見えない。しかし、スマートフォンがディスプレイから離れた状態ではマーカが見えてしまうため映像の見た目を損ねてしまうという問題がある。Virtual Projection はディスプレイの映像の特徴量を用いてスマートフォンの位置を推定する [2]。これは、位置推定のためだけの情報を提示する必要はないが、同じものが連続しているような映像や単色の映像などを用いることは困難である。

2.2 不可視の色振動

輝度を固定して色度を変化させた場合、色度変化の時間周波数が 25 Hz を超えたあたりから人間はその中間色のみを知覚する [6]。本稿では、25 Hz を超えた色度変化を「不可視の色振動」と定義する。筆者らはリフレッシュレート 60 Hz の一般的なディスプレイの映像に不可視の色振動を用

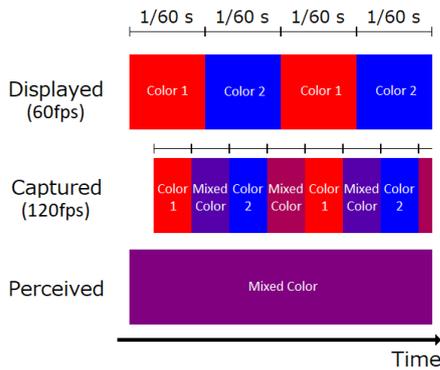


図 1: 上段: ディスプレイ, 中段: スマートフォン, 下段: 人間から見た色の時間系列 (文献 [4] より)

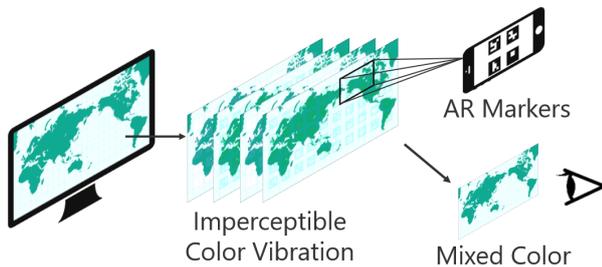


図 2: 提案システム

いて不可視情報を付加した。これにより、映像コンテンツを損なわずに情報を伝送可能なシステムを実現した [4]。

ディスプレイ上の映像の各画素に関して、元の色から、同一の輝度を持つ 2 色を生成し、それらを交互に表示することで不可視の色振動を提示する。色振動のあり/なしを画素ごとに定め、QR コードの黒/白に対応させることでディスプレイの映像に QR コードを埋め込むことができる。受信に関しては、スマートフォンで 120 fps の動画を撮影すると、図 1 のようにディスプレイの 1 フレームのみを撮影したフレームが 1 フレームおきに現れる。撮影した各フレームに対して 1 フレームおきに差分をとり画像処理を行うことで QR コードの読み取りを行う。本稿では、この手法を応用し、ディスプレイに対するスマートフォンの位置の取得を行った。Yamamoto ら [7] が同様の研究を行っているが、これはカメラがディスプレイのほぼ全域を写している必要がある。本手法はカメラがディスプレイの一部のみを写している場合でも適用可能である。

3. 提案手法

図 2 に提案するシステムの概要を示す。本システムは、リフレッシュレート 60 Hz の一般的なディスプレイとスマートフォンのみで構成される。まず、不可視の色振動により多数の AR マーカをディスプレイの映像に埋め込む。この不可視の AR マーカをスマートフォンで画像処理により認識し、ディスプレイに対する位置を推定する。ただし、本稿ではスマートフォンとディスプレイの位置関係が水平であることを前提とする。

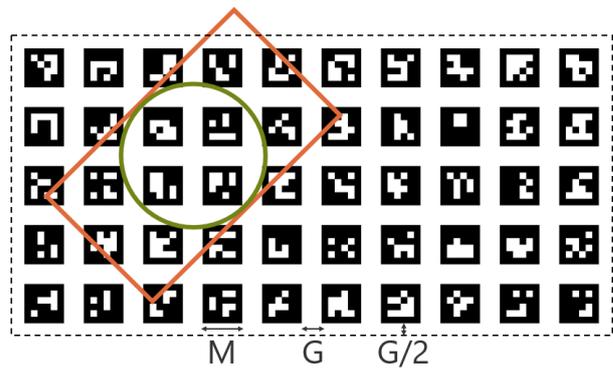


図 3: マーカの配置, スマートフォンが写す領域 (橙), カメラ中心を中心に回転しても常に写る領域 (緑), スマートフォンのカメラ中心が動く範囲 (破線内部)

スマートフォンがディスプレイのどこを撮影していても AR マーカが写るようにするため、映像に図 3 のように並べた AR マーカを埋め込む。このとき、AR マーカの大きさと並べ方によりディスプレイとスマートフォンの距離関係が決定する。AR マーカの 1 辺の長さを M , AR マーカの間隔を G とする。スマートフォンのカメラ中心は図 3 の破線で示された領域内を動くものとする。破線で示された領域とは各辺から一番近いマーカの縁までの距離が $G/2$ となるような長方形およびその内部である。スマートフォンのカメラの対角画角を θ , 動画のアスペクト比を $m:n$ ($m \geq n$), ディスプレイとスマートフォンの距離を L とする。このときスマートフォンが写す範囲は、ディスプレイ上で対角線の長さが $2L \tan \frac{\theta}{2}$ の長方形 (図 3 の橙) となる。辺の長さは $m \frac{2L \tan \frac{\theta}{2}}{\sqrt{m^2+n^2}}$, $n \frac{2L \tan \frac{\theta}{2}}{\sqrt{m^2+n^2}}$ となる。ここで、スマートフォンをどう平行移動, 回転させても少なくとも 1 つはマーカを写すために必要なディスプレイとスマートフォンの距離を考える。図 3 の緑の円のよう、長方形の中心を中心とし長辺に接する円の中に常にマーカがあればよい。これに必要な円の直径の最小値は 2×2 のマーカを囲う正方形の対角線の長さである。よって、スマートフォンが写す範囲内に AR マーカが必ず入るための条件は

$$n \frac{2L \tan \frac{\theta}{2}}{\sqrt{m^2+n^2}} \geq \sqrt{2}(2M+G) \quad (1)$$

であり、よって

$$L \geq \frac{\sqrt{2}(2M+G)\sqrt{m^2+n^2}}{2n \tan \frac{\theta}{2}} \quad (2)$$

であれば必ずスマートフォンにマーカが写ることになる。所望の距離関係に応じて M と G を設定すればよい。ディスプレイ上のマーカの位置をあらかじめスマートフォンが保持することで、検出された AR マーカから位置推定を行うことができる。

4. 実験

提案システムによってスマートフォンの位置が推定可能であることを確認するための実験を行った。真値と定めた位置と実際にスマートフォンがある位置のずれを考えない

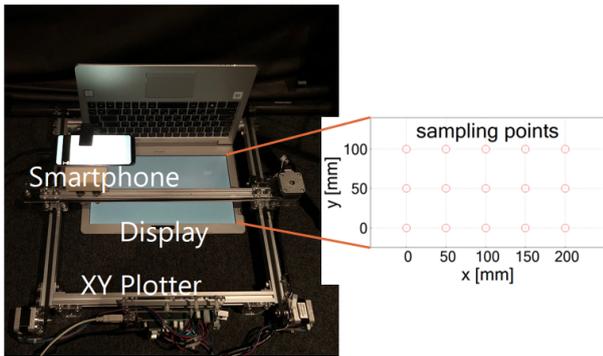


図 4: 実験の様子と測定位置。スマートフォンとノートパソコンのディスプレイの距離は 140 mm.

ようにするため、ある位置からのスマートフォンの相対位置を推定した。

4.1 条件

図 4 に実験環境を示す。映像を表示するデバイスとして 165 dpi の 13.3 インチノートパソコンのディスプレイ (MB-S250XN1-EX3, マウスコンピューター) を用い、撮影するデバイスとしてスマートフォン (Galaxy S8, Samsung) を用いた ($\theta \approx 70^\circ$, $m = 16$, $n = 9$)。このスマートフォンで 120 fps 動画を撮影する際にピントが合う最小距離が 100 mm 程度であったため、 $L \geq 100[\text{mm}]$ であれば必ずマークが写るようにマークの配置を行った。映像として、 1920×1080 で、グレー ($(R, G, B) = (128, 128, 128)$) 単色の画像に対して $M = 120[\text{px}]$, $G = 60[\text{px}]$ で 10×5 個の AR マークを埋め込んだ映像を表示した。この際、AR マークの作成には、ArUco [8, 9] を利用した。また、不可視の色振動に用いた色は $(R, G, B) = (158, 117, 131)$ と $(R, G, B) = (85, 138, 125)$ であった。実際に用いた 2 枚の画像を図 5 に示す。組立式暗室 (ADR-F2, アズワン) 内で、ディスプレイから垂直 140 mm の位置で XY プロッタを用いてディスプレイの左上部のある 1 点を始点にスマートフォンを 50 mm 刻みでディスプレイに対して水平な平面上をディスプレイの辺と平行に移動させ、 5×3 点で測定した。各位置で 10 回測定し、始点での測定値を引くことで、平面内の始点に対するスマートフォンの相対位置を測定した。また、AR マークによる位置推定そのものの誤差の確認のため、一般的な可視のマークを表示して同様の実験を行った。なお、ここでの可視のマークとは、白色背景 ($(R, G, B) = (255, 255, 255)$) に黒の単色 ($(R, G, B) = (0, 0, 0)$) の AR マークを図 5 と同様に並べた映像である。

4.2 結果

本稿では、平面内での真値と測定値の間の距離を誤差と定義する。図 6 に可視のマークを表示した場合 (振動なし) と不可視の色振動を用いた場合 (振動あり) それぞれでの誤差の箱ひげ図を示す。不可視の色振動を用いた場合の各位置での誤差の中央値は 0.90 mm であった。また、誤差の最大値は 3.53 mm であった。可視のマークを表示した際の誤差の中央値は 0.39 mm で、誤差の最大値は 0.72 mm であっ

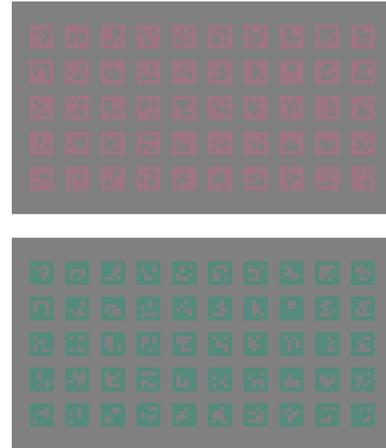


図 5: 実験に用いた画像。この 2 枚を 60 fps で切り替えて表示した。

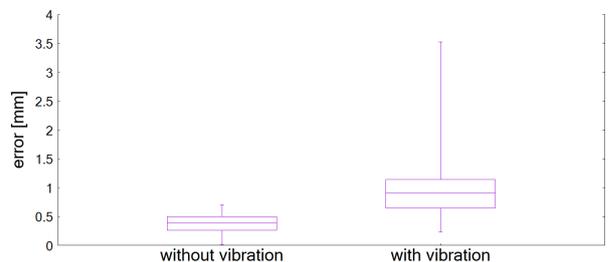


図 6: 結果のグラフ。誤差は真値と測定値の間の距離。可視のマークの条件を「振動なし」、不可視のマークの条件を「振動あり」と定義した。

たため、色振動のためにマークの輪郭が歪んで検出されていると考えられる。しかし、これはアプリケーションの実装には十分な精度であると言える。

5. 結論

本稿では、ディスプレイに対するスマートフォンの位置を推定するため、不可視の色振動を用いて AR マークを映像全体に並べる手法を提案し、スマートフォンとディスプレイの位置関係が水平のもと、実際に位置を推定可能であることを実験により確認した。

今後の展望として、スマートフォンの傾きに対応することや、スマートフォンとディスプレイの距離に応じて動的にマークの大きさを変更すること、提案システムを用いたアプリケーションの制作などが考えられる。

謝辞 本研究は、JSPS 科研費 JP16H01739, JSPS 特別研究員奨励費 JP17J04216 の助成を受けたものです。

参考文献

- [1] N. Pears, D. G. Jackson and P. Olivier: "Smart Phone Interaction with Registered Displays", IEEE Pervasive Computing, 8, 2, pp. 14–21 (2009).
- [2] D. Baur, S. Boring and S. Feiner: "Virtual Projection: Exploring Optical Projection As a Metaphor for Multi-device Interaction", Proceedings of the

- SIGCHI Conference on Human Factors in Computing Systems, CHI '12, pp. 1693–1702 (2012).
- [3] L.-W. Chan, H.-T. Wu, H.-S. Kao, J.-C. Ko, H.-R. Lin, M. Y. Chen, J. Hsu and Y.-P. Hung: “Enabling Beyond-surface Interactions for Interactive Surface with an Invisible Projection”, Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology, UIST '10, pp. 263–272 (2010).
- [4] 阿部 知史, 荒見 篤郎, 平木 剛史, 福嶋 政期, 苗村 健: “不可視の色振動を用いた2次元パターンによるディスプレイ-カメラ間通信の基礎検討”, 日本バーチャルリアリティ学会第22回大会, 2B2-02 (2017).
- [5] S. Leigh, P. Schoessler, F. Heibeck, P. Maes and H. Ishii: “THAW: Tangible Interaction with See-Through Augmentation for Smartphones on Computer Screens”, Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction - TEI '15, pp. 89–96 (2015).
- [6] Y. Jiang, K. Zhou and S. He: “Human visual cortex responds to invisible chromatic flicker”, *Nature Neuroscience*, **10**, pp. 657–662 (2007).
- [7] G. Yamamoto, L. Sampaio, T. Taketomi, C. Sandor, H. Kato and T. Kuroda: “Imperceptible On-Screen Markers for Mobile Interaction on Public Large Displays”, *IEICE Transactions on Information and Systems*, **E100.D**, 9, pp. 2027–2036 (2017).
- [8] S. Garrido-Jurado, R. M. noz Salinas, F. Madrid-Cuevas and M. Marín-Jiménez: “Automatic generation and detection of highly reliable fiducial markers under occlusion”, *Pattern Recognition*, **47**, 6, pp. 2280–2292 (2014).
- [9] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas and R. Medina-Carnicer: “Generation of fiducial marker dictionaries using Mixed Integer Linear Programming”, *Pattern Recognition*, **51**, pp. 481–491 (2016).