

全天球カメラを用いた単一カメラシェアリングシステムの開発

八木 優成[†] 横山 昌平^{††,†††}

[†] 東京都立大学システムデザイン学部 〒191-0065 東京都日野市旭が丘 6-6

^{††} 東京都立大学システムデザイン研究科 〒191-0065 東京都日野市旭が丘 6-6

^{†††} 東京大学生産技術研究 〒153-8505 東京都目黒区駒場 4-6-1

E-mail: [†]yagi-yusei@ed.tmu.ac.jp, ^{††}shohei@tmu.ac.jp

あらまし 本研究では離れた地点からの撮影を手軽にすることを目的に、環境に設置された1台のカメラを複数人でシェアして同時に撮影を行うシステムを開発した。近年スマートフォンの普及とともに手軽に撮影が行える「自撮り」の文化が発達してきた。しかし離れた距離からの撮影は機材や撮影者が必要なため、自撮りのように手軽に行うことができない。そこで我々は頻繁に撮影が行われる場所にあらかじめカメラを設置し、利用者が自由に使用できるシステムを開発した。カメラで撮影した画像から各ユーザを中心とした画像を切り出して利用するため、複数ユーザで同時使用することができる。またカメラには全天球カメラを使用し、より広範囲の撮影を可能とした。さらに本システムの実現のため、全天球パノラマ画像における Multi Object Tracking アルゴリズムを提案する。

キーワード 全天球カメラ, 物体追跡, 物体検出

1 はじめに

近年スマートフォンの普及とともに、「自撮り」の文化が発達してきた。自撮りとは撮影者が手に持ったカメラで自らを撮影する行為である。自撮りは特別な撮影器具が不要であり、スマートフォンのみで手軽に撮影ができる。観光地での記念撮影や友人、家族との集合写真などで幅広く行われており、撮った写真を SNS に投稿する者も多い。また海外でも「selfie」として日本国内同様に親しまれている。しかし自撮りには撮影方法の特性からカメラの画角や被写体との距離に限界がある。一方で離れた地点から撮影するには、三脚などの撮影機材または撮影者が必要であり、自撮りのように手軽に撮影することができない。そこで本研究では離れた地点からの撮影を手軽にすることを目的に、環境に設置された一つのカメラを複数人でシェアして撮影を行うシステムを開発した。このシステムによりユーザはカメラを持参することなく、スマートフォンを用いて認証するだけで、あらかじめ環境に設置されたカメラを使用して撮影を行うことができる。またカメラで撮影した画像から各ユーザを中心とした画像を切り出して送信することで、複数のユーザで同時撮影を行うことができ、1台のカメラを疑似的に複数台のカメラとして使うことができる。さらに本システムではカメラ部分に全天球カメラを用いることで、より広範囲の撮影、多人数での同時利用を可能とした。

一般的に全天球カメラで撮影した画像は正距等円図法で全天球パノラマ画像に変換され、保存、配信される。これは通常の画像の保存、送信方式を使うことができるためである。全天球パノラマ画像は全景が一度に確認できる一方で、図1に示すように高緯度領域に幾何学的な歪みが発生する。そのため高緯度領域に映る物体は通常のカメラで撮影した場合と外観が大きく異なる。そこで本研究ではカメラが被撮影者と同程度の高さに

設置され、全天球パノラマ画像において歪みの少ない低緯度領域に被撮影者が映り込むことを前提とした。ただし被撮影者を中心とした全天球パノラマ画像に変換することで、歪みを軽減することは可能である。

また本システムではユーザが自分の姿を確認しながら撮影できるようにする必要がある。この機能の実現には撮影した画像内に映った人物を検出し、フレーム間で同一の人物を判定する Multi Object Tracking アルゴリズム (MOT アルゴリズム) を用いる。しかし全天球パノラマ画像における MOT アルゴリズムの研究はあまり行われていない。そこで我々は通常の動画における MOT アルゴリズムである DeepSORT [1] の応用を考えた。DeepSORT は物体検出ベースの MOT アルゴリズムであり、検出した物体の位置情報と外観特徴量を用いてトラッキングを行う。構造はシンプルでありながら精度が高く、処理が高速であり、オンライン処理が可能なアルゴリズムである。しかし DeepSORT は通常の平面動画を対象としており、全天球パノラマ画像ではそのまま使用することはできない。これはパノラマ画像上の距離が、実空間上の距離と異なるためである。図2に示すように、カメラの方位角0度を跨いで二人の人物が並んでいた場合、全天球パノラマ画像上では図3のように離れた位置に存在するように映る。また物体が方位角0度を跨いで移動した場合、別の物体としてトラッキングされてしまう。そこで本研究では DeepSORT をベースとした全天球パノラマ画像における MOT アルゴリズムを開発した。このアルゴリズムでは方位角方向に拡張した全天球パノラマ画像で検出を行い、IoU により同一物体の重複検出を削除する。さらに全天球パノラマ画像における検出物体の位置座標を、カメラを中心とした世界座標系に変換してトラッキングに用いることにより、方位角0度を跨いだトラッキングを可能とした。

また本システムでは、撮影した画像上で検出したトラック部分を切りだし、対応するユーザへ送信するため、トラックと



図 1 高緯度領域で人が歪んだ全天球パノラマ画像

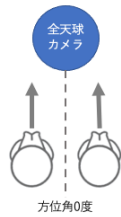


図 2 上から見た撮影風景 (2 名が方位角 0 度を跨いで並んでいる)



図 3 2 名が方位角 0 度を跨いで並んでいる全天球パノラマ画像

ユーザが所持しているスマートフォンを結びつける必要がある。我々はこの機能に QR コードによる認証を用いた。しかし全天球パノラマ画像の高緯度領域における歪みにより、QR コードは正常に検出されない可能性がある。そこで我々は 1) 光の反射や撮影時のブレなどが存在しない仮想的作成した全天球パノラマ画像。2) 実際に全天球カメラを用いて撮影した全天球パノラマ画像において、QR コードの認識精度を検証した。この結果、全天球パノラマ画像において低緯度領域であれば QR コードは正常に認識できることがわかった。

本論文の構成は以下のとおりである。2 章では本研究で利用した技術及びその活用に関する研究について述べる。3 章では開発したシステムについて説明し、4 章では実験結果と考察について述べる。まとめと今後の課題を 6 章にまとめる。

2 関連研究

全天球カメラを用いた物体検出は多数の手法が提案されてい

る。林田ら [2] は既存の物体検出アルゴリズムである You Only Look Once version 4 (YOLOv4) [3] を用い、全天球パノラマ画像における物体検出を提案した。高緯度領域で歪んだ物体を学習データに含め、低緯度領域とは別の物体として学習させることで、従来の物体検出手法と比べ、処理速度は変わらずにより広い範囲での物体検出を可能とした。井上ら [4] は全天球カメラで撮影した画像をカメラを中心とした 8 面キューブマップに投影し、物体検出を行う手法を提案した。この手法では画像の全景が 8 つの画像に分割して投影されるため、各画像の歪みは小さくなる。しかし 8 つの画像それぞれに対して物体検出を行うため、処理時間が大きくなってしまった。本研究ではリアルタイム性が重視されるため、全天球パノラマ画像に対して直接物体検出を行うこととした。

また全天球カメラを用いて実空間上の物体の状態を予測する技術も研究されている。増田ら [5] は全天球カメラを用いた物体の位置推定を行なった。全天球パノラマ画像から物体を検出し、その物体の実空間上での位置座標を推定した。さらに誤差補正を行うことにより、補正前と比較して精度を約 2.5 倍向上させた。笠原ら [6] は全天球パノラマ画像上での行列検知を行なった。物体検出アルゴリズムである YOLOv4 に加え、骨格検出アルゴリズムである Openpose を用いることにより、人の向いている方向を推定した。また複数人の向きを推定し、角度を比較することで、行列の判定を行なった。

通常の動画におけるトラッキング手法は多数提案されており、Bewley ら [7] は物体検出ベースの MOT アルゴリズムである SORT を提案した。検出された物体と既にトラッキングしている物体をマッチングするために Intersection over Union (IoU) を用いた。IoU とは複数の領域がどの程度重なっているかを示す値である。処理が軽量なため、非常に高速な処理が可能であるが、物体の位置情報のみを使うため、追跡物体の ID が入れ替わる ID スイッチが起こりやすい。Wojke ら [1] は物体の位置情報及び外観特徴量を用いてトラッキングを行う DeepSORT を提案した。この手法ではまず位置情報と外観特徴量からなるベクトルをマハラノビス距離によりマッチングする。さらにマッチングしなかった物体に対して SORT によるマッチングを行う。外観特徴量を用いたことにより従来手法と比べ ID スイッチが小さくなった。また比較的単純なアルゴリズムであるため、応用しやすい。本研究では全天球カメラへの応用を考え、実装の単純さ、精度、速度のバランスの取れた DeepSORT を使用した。

また魚眼カメラや全天球カメラを用いた物体追跡についても研究が行われている。Nishimura ら [8] は天井に設置した魚眼カメラによる人物追跡を行なった。魚眼カメラで撮影した画像では歪みが発生し、検出したバウンディングボックスに背景などの検出対象以外の部分が多く含まれる。そこで検出したバウンディングボックスを回転させ、可能な限り背景が含まれない画像を生成した。また画像から物体の位置情報を世界座標系に変換することで、歪みの影響を受けにくいトラッキング手法を提案した。Premachandra ら [9] は全天球カメラを用いた交差点における移動物体の検出及び追跡を行なった。処理は全天球



図 4 方位角 0 度を跨いで人が存在する場合の検出結果



図 5 方位角方向に 1.5 倍に拡張した画像での検出結果

パノラマ画像上で行い、撮影中の画像と事前に用意した背景画像との差分を用いた物体検出手法を提案した。またカルマンフィルタを用いて予測した物体の位置と新たに検出した物体の位置を比較し、マッチングさせることでトラッキングを行った。この手法は処理が高速なものの、背景画像を事前に用意する必要がある。またトラッキングは全天球パノラマ画像の方位角 0 度を越えて移動する物体は考慮されていない。

3 提案手法

本研究ではまず全天球パノラマ画像におけるトラッキング手法として、通常の動画における MOT アルゴリズムである DeepSORT をベースとしたトラッキング手法を提案する。続いて提案したトラッキング手法を用い、単一の全天球カメラを複数のユーザで同時に使用し、撮影を行うシステムを開発する。

3.1 全天球パノラマ画像におけるマルチオブジェクトトラッキング

本手法では通常の画像における MOT アルゴリズムである DeepSORT を全天球パノラマ画像で使えるように応用した。まず各フレームにおいてトラッキング対象となる物体を検出する。本研究では物体検出アルゴリズムである YOLOv4 を検出に用いた。YOLOv4 は深層学習を用いた物体検出手法であり、処理速度が速く、認識精度が高い。検出を行う際、通常の画像とは異なり、全天球パノラマ画像では方位角 0 度を跨いで物体が存在する場合、図 4 のように 1 つの物体が 2 つの異なる物体として認識される可能性がある。そこで我々は図 7 に示すように、方位角方向に 1.5 倍に拡張した画像で物体認識を行った。これにより全天球パノラマ画像の方位角 0 度付近の物体においても全景を検出することができる。また、方位角 0 度付近に存在する検出物体に対し、各検出物体との IoU を計算することで、同一物体の検出の重複を削除した。



図 6 IoU により重複物体を削除した検出結果

続いてトラッキングにおいて、DeepSORT では画像上の検出物体の位置、アスペクト比、高さ、及び検出物体の外観特徴量を用いる。しかし全天球パノラマ画像では物体の位置座標 (x, y) が実世界上の物体の位置と対応していない。このため方位角 0 度を越えてユーザが移動した場合、別の人物と認識され、トラックが割り当てられてしまう。そこで我々は式 1 により全天球パノラマ画像上の位置座標 (x, y) を全天球カメラを中心とした球面座標系 (r, θ, ϕ) に変換し、そこから式 2 により世界座標系 (x', y', z') に変換した。

$$\begin{pmatrix} r \\ \theta \\ \phi \end{pmatrix} = \begin{pmatrix} \frac{W}{2\pi} \\ \frac{2\pi x}{W} \\ \frac{\pi}{2} - \frac{\pi y}{H} \end{pmatrix} \quad (1)$$

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} r \cos(\phi) \cos(\theta) \\ r \cos(\phi) \sin(\theta) \\ r \sin(\phi) \end{pmatrix} \quad (2)$$

W, H はそれぞれ全天球パノラマ画像の幅、高さである。位置情報以外のパラメータは DeepSORT と同様のものを使用した。この変換により、物体の位置情報は全天球パノラマ画像上の位置座標ではなく、実空間に即した位置座標になるため、方位角 0 度を跨いだトラッキングが可能となる。

3.2 単一カメラシェアリングシステムの開発

我々は提案した MOT アルゴリズムを用いて複数ユーザで単一の全天球カメラを共有して同時に撮影を行うシステムを開発した。このシステムでは全天球パノラマ画像に映る人物を切り出し、対応するユーザのスマートフォンに送信することで複数ユーザでの同時使用を可能とする。

システムの概要を図 7 に示す。まずカメラで撮影している画像をもとに画像処理サーバが人物を検出し、トラッキングを行う。システムを利用するユーザはスマートフォンを用いて web サーバにアクセスし、一意のユーザ ID を示す QR コードを受け取る。ユーザは受け取った QR コードをカメラに読み込ませる。画像処理サーバは QR コードを検出し、QR コードに最も近い未認証のトラックと QR コードに示されたユーザ ID を結びつける。これにより web サーバにアクセスしているユーザとトラックの認証を行う。画像処理サーバは撮影したフレームから認証された各トラックの画像を切り出し、web サーバを通じてユーザに配信する。

3.2.1 QR コードとトラックの認証

本システムではカメラで撮影しているトラックをどのユーザ

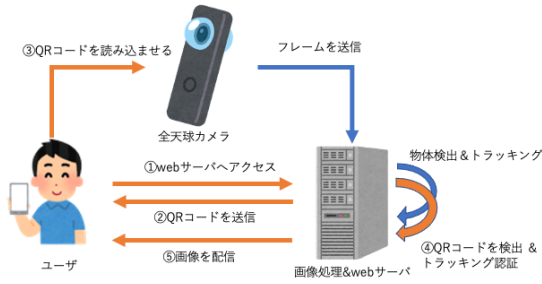


図7 開発したシステムの概要

に配信するかを判別するため、webサーバにアクセスしているユーザとトラックを紐付ける必要がある。そこで我々はQRコードによる認証を考えた。スマートフォンに一意のユーザIDを示すQRコードを表示し、カメラで読み込んだ際に最も近い位置のトラックに紐づける。式1、式2により変換したQRコードとトラックの世界座標をそれぞれ p_{qr} 、 p_t とする。式3によりQRコードとトラック間のコサイン類似度を求め、比較した。

$$\cos(p_{qr}, p_t) = \frac{p_{qr} \cdot p_t}{\|p_{qr}\| \|p_t\|} \quad (3)$$

3.2.2 スマートフォンへの配信

サーバからスマートフォンへの配信はWebRTCを用いた。WebRTCはコンピュータ間で音声、動画、データの共有を可能にするオープンソースフレームワークである。通信はP2Pで行われ、非常に低遅延である。また双方向通信が可能なため、オンライン会議システムなどに利用されている。本システムではサーバからスマートフォンへの一方方向の配信であるが、構成がシンプルであることと低遅延であることからWebRTCを採用した。

4 実験結果

開発したシステムではQRコードを全天球カメラに読み込ませることで認証を行うが、1章で説明した通り、全天球パノラマ画像は高緯度領域に幾何学的な歪みが発生するためQRコードを読み取れない可能性がある。そこで本研究では予備実験として全天球カメラによるQRコードの読み取り精度を検証した。続いて開発したシステムを実際に使用した際の結果について述べる。本実験では全天球カメラにRICOH社のTHETA Z1を使用した。

4.1 全天球パノラマ画像におけるQRコードの認識精度評価

QRコードはその特性上、全天球パノラマ画像で生じる歪みに対して弱いと考えられる。そこでまず全天球パノラマ画像の高緯度領域における歪みのQRコードに対する影響を検証した。また実空間では光の反射や撮影時のブレなどが影響すると考えられる。そのため我々は1) 仮想的に作成した全天球パノラマ画像、2) 実空間上で撮影した全天球パノラマ画像の2つの条件においてQRコードの認識実験を行った。

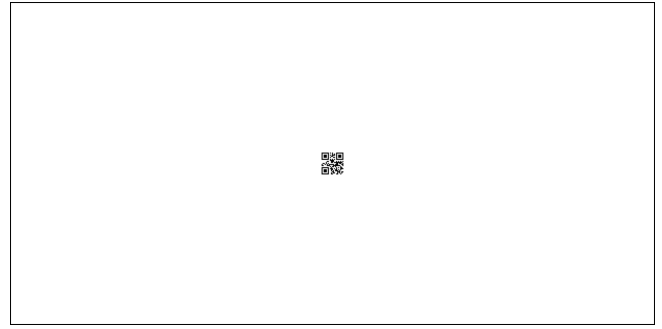


図8 背景画像にQRコードを合成した画像

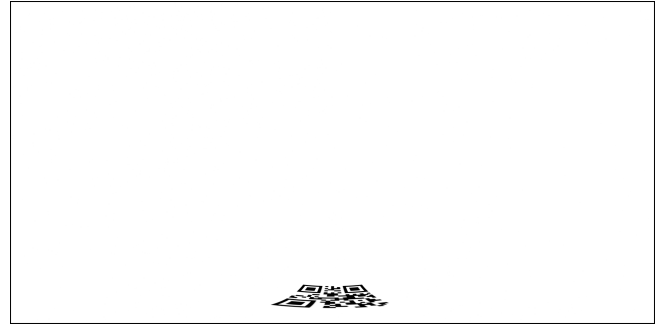


図9 変換された全天球パノラマ画像

本実験で用いたQRコードはversion1、誤り訂正レベルHとした。またQRコードの検出にはPythonのライブラリであるpyzbarを用いた。

4.1.1 仮想的に作成した全天球パノラマ画像

まず全天球パノラマ画像の各領域にQRコードが存在する画像のデータセットを作成した。このデータセットの作成は以下の方法で行なった。

(1) 単色背景画像の中央にQRコード画像を合成する(図8)

(2) 生成された画像を球面にマッピングし、仰角、方位角を調整する

(3) 球面マッピングした画像から全天球パノラマ画像を作成する(図9)

図10は作成したデータセットにおけるQRコード検出結果をQRコードの位置に関連付けて記録したものである。右側のカラーバーはQRコードの検出精度を示している。この結果から分かるように、QRコードは緯度方向±20度を超えると歪みの影響を受け、正常に検出できないことがわかった。一方低緯度領域ではほぼ確実に認識できている。方位角0度付近の検出精度が低くなっているのは、QRコードが途切れてしまったためだと考えられる。

4.1.2 実空間上で撮影した全天球パノラマ画像

仮想的に作成した全天球パノラマ画像の場合と同様に、全天球パノラマ画像の各領域にQRコードが存在するようなデータセットを作成した。この実験ではQRコードをスマートフォンに表示し、THETA Z1で撮影を行った。またカメラとスマートフォンの距離は20~30cmとした。

図11は実空間上で撮影したデータセットにおけるQRコー

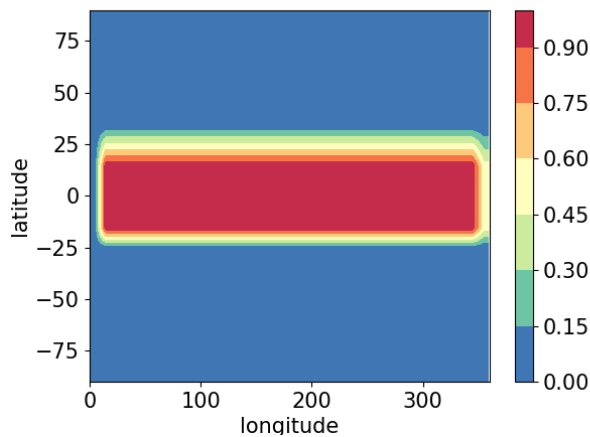


図 10 緯度・経度に対する仮想環境上の QR コードの検出精度

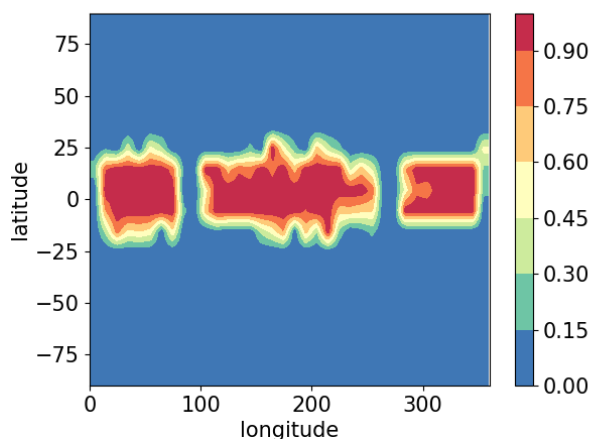


図 11 緯度・経度に対する実空間上の QR コードの検出精度

ドの検出結果である。仮想的に作成した全天球パノラマ画像同様、緯度方向は± 20 度程度までの範囲で検出が可能であった。また方位角方向 90, 270 度付近で検出率が低くなっているのは全天球カメラの死角が含まれるためであると考えられる。今回実験で用いた RICOH 社の THETA Z1 は 2 つの魚眼レンズを合わせた作りになっており、2 つのレンズの境目には死角が存在する。このため図 12 に示すように、QR コードが一部途切れてしまったことが原因であると考えられる。また低緯度領域でも検出精度にムラがあるが、これは仮想空間とは違い、実空間では光の反射や撮影時のブレが存在するためだと考えられる。検出前に前処理を行うなどして、改善の余地がある。

4.2 カメラシェアリングシステムの評価

我々は開発したシステムを実際に使用し、評価を行った。使用したケースは 1) 3 人が屋内空間に存在し、全員がシステムを使用する。2) 4 人が屋内空間に存在し、2 人がシステムを利用する。の 2 パターンである。

4.2.1 3 人が屋内空間に存在し、全員がシステムを使用する場合

このパターンでは 3 人のユーザが順番に認証を行い、システ



図 12 THETA の死角によって途切れた QR コード

ムの利用を行なった。実験の間、ユーザは 9 回交差したが、ID スイッチは一度も発生しなかった。また 2 人のユーザは認証から実験終了まで正常にトラッキングが行われていたが、1 人のユーザは途中でトラッキングが終わってしまった。これはユーザが座った状態から立ち上がり、急に移動を始めた状態であった。外観特徴量と位置座標の急激な変化が発生したことから、MOT アルゴリズムの閾値を超え、移動前の物体と移動後の物体が別のトラッキングとして認識されてしまったものと考えられる。トラッキングが失敗すると、スマートフォンへの画像送信が止まってしまうが、ユーザは再度 QR コード認証を行うことで、撮影を再開することができた。またユーザはカメラから最大で約 10m 離れたが、正常にトラッキングを行うことができた。また実験中の処理速度は 4.23FPS であった。

4.2.2 4 人が屋内空間に存在し、2 人がシステムを使用する場合

実際のシステム使用を想定した場合、カメラの撮影範囲内にはシステムを使用しないユーザも存在すると考えられる。そこで 4 人がカメラの撮影範囲に存在し、そのうち 2 人がシステムを利用する状態で実験を行った。実験の間、ユーザは 10 回交差したが、ID スイッチは一度も発生しなかった。また 1 人目のユーザは正常に認証を行うことができたが、2 人目のユーザは QR コードが別のユーザとマッチングしてしまった。これは QR コードとユーザのマッチングを最短距離で行っていたためだと考えられる。QR コードの認証の際にはカメラに近づく必要があるため、ユーザの検出面積による閾値を設けることや、骨格検出を用いてスマートフォンを持っているユーザを判定するなど、QR コードとトラッキングのマッチングに工夫が必要であることがわかった。また実験中の処理速度は 4.22FPS であった。

5 まとめと今後の課題

本研究では全天球パノラマ画像上における MOT アルゴリズムを提案した。物体検出の際に拡張した全天球パノラマ画像を用い、さらに IoU により重複した検出を削除した。また DeepSORT をもとに、位置情報を世界座標系に変換することで、全天球パノラマ画像上の方位角 0 度を跨いだトラッキングを可能とした。さらに提案した MOT アルゴリズムを用いて単一全天球カメラを複数人でシェアするシステムを開発した。

実験の結果から、開発したシステムはビデオ撮影は難しいも

の、写真撮影には十分な処理速度があることがわかった。また数名での使用はほぼ問題なく行えるが、多数の人物が存在する空間ではユーザ認証に改良が必要であることがわかった。

今後の課題としては、まず QR コード以外のユーザ認証方法の提案が挙げられる。今回の用いた QR コード認証は全天球パノラマ画像の歪みの影響を受けやすいため、認証範囲が限られる。またカメラから数十 cm の距離まで近づけることを前提としている。代わりの手法として、スマートフォンを振るなどのユーザの動作を用いて認証を行う手法を考えている。またユーザのトラッキング精度を上げるために、DeepSORT 以外の MOT アルゴリズムの全天球パノラマ画像への応用を考えている。

謝 辞

本研究の一部は JSPS 科研費 19K11982 の助成を受けたものです。

文 献

- [1] Nicolai Wojke, Alex Bewley, and Dietrich Paulus. Simple online and realtime tracking with a deep association metric. In *2017 IEEE international conference on image processing (ICIP)*, pages 3645–3649. IEEE, 2017.
- [2] Kazuma Hayashida, Junya Masuda, and Shohei Yokoyama. Real-time human detection using spherical camera for web browser-based telecommunications. In *Proceedings of the 13th International Conference on Management of Digital EcoSystems*, pages 1–7, 2021.
- [3] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*, 2020.
- [4] 井上慶彦, 岩村雅一, 黄瀬浩一, et al. 全方位カメラを用いた物体検出とトラッキング—視覚障害者支援システムの実現に向けて—. *研究報告コンピュータビジョンとイメージメディア (CVIM)*, 2018(20):1–6, 2018.
- [5] Junya Masuda, Kazuma Hayashida, and Shohei Yokoyama. Position estimation for objects in the high latitude region of 360° panoramic images. In *Proceedings of the 13th International Conference on Management of Digital EcoSystems*, pages 146–152, 2021.
- [6] 笠原悠樹 and 横山昌平. 全天球カメラを用いたリアルタイム行列検知. In *DEIM Forum*, 2021.
- [7] Alex Bewley, Zongyuan Ge, Lionel Ott, Fabio Ramos, and Ben Upcroft. Simple online and realtime tracking. In *2016 IEEE international conference on image processing (ICIP)*, pages 3464–3468. IEEE, 2016.
- [8] Hitoshi Nishimura, Naoya Makibuchi, Kazuyuki Tasaka, Yasutomo Kawanishi, and Hiroshi Murase. Multiple human tracking using an omnidirectional camera with local rectification and world coordinates representation. *IEICE Transactions on Information and Systems*, 103(6):1265–1275, 2020.
- [9] Chinthaka Premachandra, Shohei Ueda, and Yuya Suzuki. Detection and tracking of moving objects at road intersections using a 360-degree camera for driver assistance and automated driving. *IEEE Access*, 8:135652–135660, 2020.