

# 全天球カメラを用いたリアルタイム行列検知

笠原 悠樹<sup>†</sup> 横山 昌平<sup>††</sup>

<sup>†</sup> 東京都立大学 システムデザイン学部 〒191-0065 東京都日野市旭ヶ丘6丁目6番

<sup>††</sup> 東京都立大学 システムデザイン研究科 〒191-0065 東京都日野市旭ヶ丘6丁目6番

E-mail: <sup>†</sup>kasahara-yuki1@ed.tmu.ac.jp, <sup>††</sup>shohei@tmu.ac.jp

**あらまし** 近年、全方位が撮影できる全方位カメラと呼ばれるカメラが普及している。この全天球カメラで撮影された全天球パノラマ動画を用いて、実空間を解析した。全天球パノラマ動画は、通常のカメラで撮影された動画と同じ形式で視聴することができる。しかし、球面で撮影された動画を平面に変換しているため歪みが生じてしまう。全天球カメラを用いた研究は行われており、物体検出や物体の位置推定が挙げられる。そこで、我々は次のステップとして物体の状態を検知することが必要だと考えた。本研究では人の向きに着目して行列を検知した。その手法として物体検出アルゴリズム Yolov4 を用い、人群を検出する。その人群中にいる人の向きを検出し、直線判定を行う。人群が直線であればあるほど行列の評価値を上げるという手法である。

**キーワード** 全天球カメラ、行列検知、物体検出、骨格検出

## 1 はじめに

近年、Google Street View<sup>1</sup>や AR (Argument Reality:拡張現実) や VR(Virtual Reality:仮想現実) を用いた視聴を体験するための撮影デバイスとして、全天球カメラと呼ばれている全方位を記録できるカメラが用いられている。この全天球カメラは RICOH 社の THETA シリーズをはじめ、様々な種類の全天球カメラが販売されている。その中でも安価な全天球カメラは、一般ユーザーの手に届きやすく、Twitter<sup>2</sup>や YouTube<sup>3</sup>に全天球カメラで撮影された動画像が共有されている。

全天球カメラで撮影された動画像は、全天球球面動画像と全天球パノラマ動画像の 2 種類がある。今回はその中でも全天球パノラマ動画を使って実空間を解析した。全天球パノラマ動画の特徴として、全方位の風景を通常のカメラで撮影された動画と同じ形式で視聴することができる。しかし、一般的に全天球カメラで撮影された動画像は、全天球球面動画像を正距円筒図法という投影法によって全天球パノラマ動画像に変換されるため、全天球パノラマ動画像は歪んでしまう。

この問題に対して、林田ら [1] は物体検出アルゴリズム You Only Look Once version3 (YOLOv3) を用いて、全天球カメラで歪んでしまった全天球パノラマ動画像でも物体検出を行えることを示した。また、増田ら [2] は、全天球カメラで撮影された全天球パノラマ動画に対して、物体検出アルゴリズム YOLOv3 を用いて、歪んだ場所でも物体の位置推定を行った。このことから、歪んでいる全天球パノラマ動画像は既存の物体検出アルゴリズムを用いて、物体検出、位置推定を行えることが言える。

これらの研究から、次の課題として、実空間における物体の状態を検知することであると考えた。物体の状態にはさまざま

状態がある。動いている物体の状態、止まっている状態などがあげられる。動いている物体の状態を研究するとなると人流解析、止まっている物体の状態を研究するとなると混雑度推定などが挙げられる。

本研究では、物体の向きに着目した。物体の向きは、動いている物体、止まっている物体どちらにもアプローチすることができる。物体の向きを検知できれば、人流解析の精度や人の行動推定に応用することができる。本研究は行列中の人の向きは同一方向を向いているという仮説を立て、物体の向きと結びつきが強いと考えられる行列検出を行った。行列の先頭は、イベントや人気のお店、商品が存在している可能性が高い。この行列を検出することでカメラからイベントの盛り上がり度、お店、商品の人気度がリアルタイムで得られることが可能である。

しかし、全天球カメラで撮影された全天球パノラマ動画像で得られた人物の向いている方向は、画像から得られる方向と実空間における向いてる方向とは違う。例として、人がレンズの方を向いている全天球パノラマ画像図 1 と人が同じ方向を向いている全天球パノラマ画像図 1 に示す。また、それぞれ上から俯瞰した簡易図を図 2, 4 に示す

図 1 は、図 2 の方法で撮影された全天球パノラマ画像である。図 1 は画像上では同じ向きを向いているが、実空間の人の向きは図 2 の黒矢印であり、違う方向を向いている。これは通常のカメラでも生じている問題だが、全天球カメラでは全方位を撮影しているため、実空間における物体の向きと動画像の向きの差異が大きく出てしまう。

図 3 は、図 4 の方法で撮影された全天球パノラマ画像である。図 3 は画像上では人が異なる方向を向いているが、実空間の人の向きは図 4 の黒矢印であり、同じ方向を向いている。図 1, 3 から、全天球カメラで撮影された全天球パノラマ動画像における人の向きの検出は、一つの課題であると言える。

図 5 のように行列に並んでいる人たちは、同じ方向を向いて並んでいることが多い。人の向きを推定することが可能であれば、

1 : <https://www.google.co.jp/maps/preview>

2 : <https://twitter.com>

3 : <https://www.youtube.com>



図 1 人がレンズの方を向いている全天球パノラマ画像



図 3 人が同じ方向を見ている全天球パノラマ画像

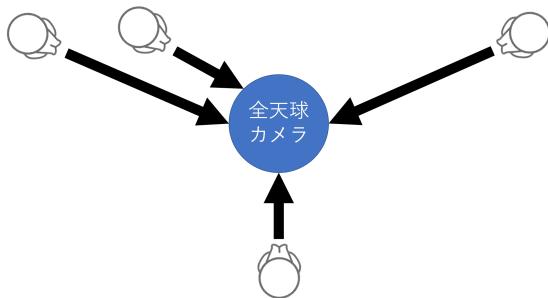


図 2 上から見た撮影風景 (人がレンズの方を向いている)

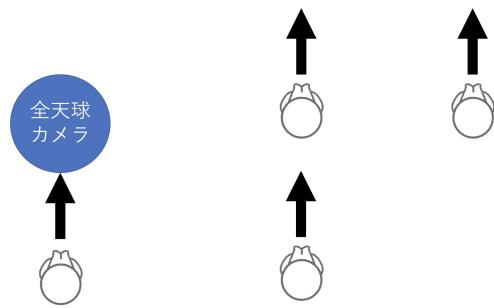


図 4 上から見た撮影風景 (人が同じ方向を見ている)

行列を検出可能であると考えた。しかし、前述したように実空間では同じ方向を向いていても、全天球パノラマ動画像では異なる方向を向いてしまう。さらに、実空間では人が列になっており、物体検出アルゴリズムのみでの行列の検出は精度が低い。

全天球パノラマ動画像の位置座標は、極座標で表され、実空間の直交座標とは異なる。そのため、人が全天球パノラマ動画像を見た時に、錯覚を起こしてしまう。だが、その位置座標は極座標から直交座標に変換することで、実空間における人の向き推定を行うことが可能である。

以上のことから、我々は行列を検知するシステムを作成した。行列を検知する方法として You Only Look Once version4 (以下 YOLOv4 と略す) を用いて、行列を検出する。しかし物体アルゴリズムだけでは、正確に行列と人混みを認識することができない。そこで骨格検出アルゴリズム Openpose を用いて、人骨格を検出することで、人の向いている方向を検出することが可能である。人の向いている方向をベクトル化し、内積を計算することで、他人の向いている方向と角度の差が算出される。この差が小さければ小さいほど、行列である評価値をあげるというシステムである。この手法を用いることで、行列判定を行えると考えた。

本論文の構成は以下の通りである。2章では、関連研究について述べる。3章では、提案手法を述べる。4章では、実験結果と考察について述べる。5章では、まとめと今後の課題について述べる。

## 2 関連研究

本章では関連研究について述べる。本研究と同様に全天球カメラを用いた研究は様々な分野で行われている。その中でも全

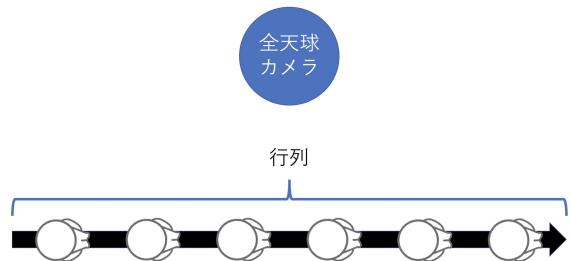


図 5 全天球カメラで撮影される行列

天球カメラを用いた物体検出の研究を紹介する。

林田ら [1] は、全天球カメラで配信された動画を用いて、物体検出を行った。既存の物体検出アルゴリズム YOLOv3 の教師データに歪んだ画像を入れることで、精度を保ちつつリアルタイムでの物体検出を可能にした。

物体検出の応用として、増田ら [2] は、全天球カメラを用いて、駒の位置推定を行った。物体検出アルゴリズム YOLOv3 で駒の位置座標を検出し、その駒の推定される実空間での位置座標に変換させることに成功した。

浅沼ら [3] は、Deep Convolutional Neural Network を用いて人検出を行った。誤検出率  $10^{-3}$  における未検出率が最良の 28.2% を達成している。しかし、頭部画像のみを学習させたため、全身画像を含め学習させた検出精度の検証も必要であると述べている。

井上ら [4] は、全天球カメラを用いて物体を検出し、リアルタイムで視覚障害者にその情報を伝達するシステムを開発した。全天球カメラのレンズを中心に 8 面に分けるキューブマップを

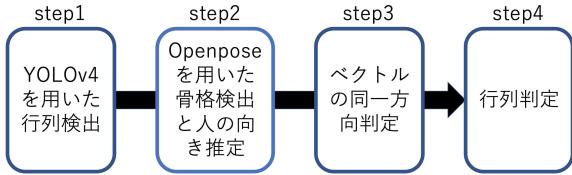


図 6 提案手法の流れ

用いて、通常のカメラで撮影されたような歪みの少ない動画を生成して物体検出を行った。しかし、物体検出に 8 面キューブマップを用いているため、処理時間が遅くなってしまった。

また、本研究では全天球カメラを用いて行列を検出する。しかし、全天球カメラだけでなく、通常のカメラでも行列を検出する研究が調べたところ存在しなかった。ここでは、混雑度推定、人流解析を行っている研究を紹介する。

庭川ら [5] は、監視カメラの映像を用いて、撮影場所における群衆の時間的かつ空間的な混雑度を画像解析により計算する手法を提案した。空間的な混雑度の分布とその時間的な変動を機械的に可視化することを可能にしたが、動き検出結果や集計結果と実際の人数との対応や、人数が一定数を超えると動き領域面積と人数との相関関係が飽和状態になることが予想される点について、課題が残ってしまった。

馬場ら [6] は、監視カメラの位置にステレオカメラを設置し、画像処理を行うことで人流解析を行った。監視カメラには人と人が重なって人が見える隠蔽問題が発生してしまう。しかし、開発した人流計測アルゴリズムでは、見えなくなってしまった人を過去と現在のフレームで相関度の高い組み合わせを探索することで、98 % 以上の人数計測精度を実現した。

川下ら [7] は、設置場所を限定せず、簡易に用いることが可能であるインスタント人流センサを提案した。ステレオカメラから得られる画像と視差画像を用いて、カメラの外部パラメータと背景差分の閾値を自動で推定し、混雑度合いに応じて、パーティクルフィルタを用いた手法、KLT (Kaneda-Lucas-Tomasi feature tracker)、ボロノイ分割を用いた手法のそれぞれを使い分けて人流を測定した。

### 3 提案手法

この章では本研究の提案手法について述べる。図 6 に提案手法の流れを示す。行列を検知するにあたって、まず YOLOv4 で物体検出をする。次に Openpose で人の骨格を検出して、検出された骨格の座標から人の向きを推定する。その後、行列中の人の向きが同じであればあるほど、評価値を高くするといったものである。

#### 3.1 Step1: YOLOv4

YOLOv4 [8] は、深層学習アルゴリズムを用いた物体検出手

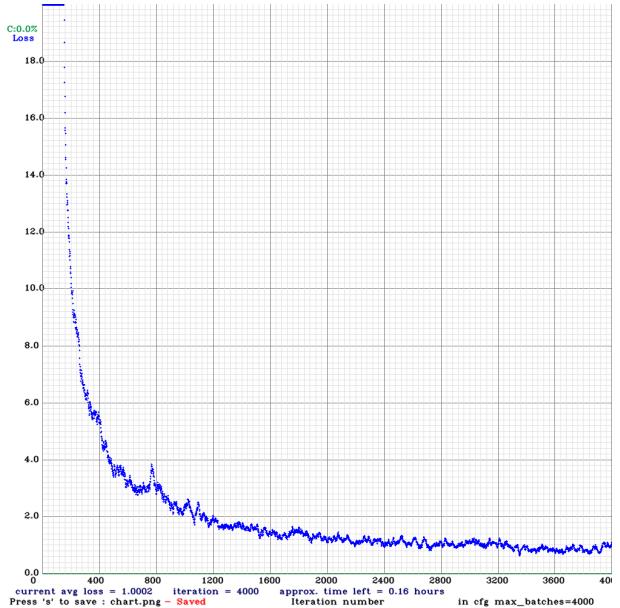


図 7 学習時に生じた損失

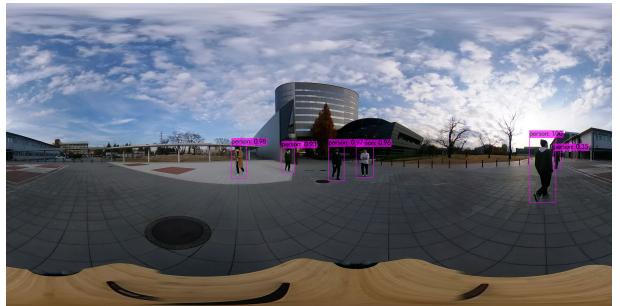


図 8 学習回数 4000 回の学習データを用いた人検出

法であり、処理速度が早く、認識が高い手法である。深層学習アルゴリズムである YOLOv4 は、画像全体から特徴量を自動で抽出し学習、検出を行うもので、物体検出手法の中でも画像認識を回帰問題として扱っているものである。ラベル付けした教師データを使い、Bouding Box の位置を学習し、信頼度とその物体が存在する条件付き確率の積から物体の存在領域を確率で抽出するアルゴリズムである。

YOLOv4 は用意された学習済みの教師データを用いることが可能であるが、本研究では全天球パノラマ画像を 500 枚用意し、それらを教師データにした。この step では人の位置座標を知ることを目的としているので、検出する対象は人のみにした。学習させる際に学習させた回数と損失を示すグラフが示される。このグラフの損失が収束していれば十分に学習されている状態である。しかし、学習をやり過ぎると過学習をしてしまい損失が増えてしまう。このことから学習回数を単に増やせば良いものでない。今回は、最大学習回数を 4000 回にし、損失のグラフを見て学習データを選ぶことにした。学習させた時に出力された損失を表すグラフを図 7 に示す。

図 7 を見ると 2800 回で損失が収束している。そのため 4000 回でも過学習を生じていないことから学習回数 4000 回の学習データを用いて物体検出を行い、人の位置座標を取得する。こ



図 9 Openpose を用いて骨格検出を行った結果

の学習回数 4000 回の学習データを用いて人を検出した図を図 8 に示す。

### 3.2 Step2:Openpose を用いた人の向き推定

Openpose は、深層学習アルゴリズムを用いた骨格検出手法であり、リアルタイムで処理することが可能である。入力画像を使って、部位の位置をエンコードする Part Confidence Maps と、部位間の関連度をエンコードする Part Affinity Fields を作る。そして Part Affinity Fields をもとに Bipartite Matching し、結果を出力する。

Openpose は人物の関節を検出することが可能である。図 9 は web カメラで骨格検出を行った。検出した関節は 18 種類あり、json ファイルに検出された 18 種類の関節の位置座標が配列として格納される。検出された関節は表 1 に示す。

表 1 配列に格納された関節の種類

0	鼻	9	右膝
1	胸	10	右足首
2	右肩	11	左腰
3	右肘	12	左膝
4	右手首	13	左足首
5	左肩	14	右目
6	左肘	15	左目
7	左手首	16	右耳
8	右腰	17	左耳

YOLOv4 と同様に Openpose は教師データを自分で用意することが可能であるが、今回は既存のデータセット用いて進めていった。Openpose で検出された関節はその動画像における位置座標を出力することが可能である。出力された位置座標は極座標系であるから、 $(\phi, \theta)$  である。 $\phi$  は方位角、 $\theta$  は仰角である。極座標から直交座標  $(x, y, z)$  への変換は

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \sin \theta \cos \phi \\ r \sin \theta \sin \phi \\ r \cos \theta \end{pmatrix} \quad (1)$$

である。 $r$  はカメラから物体までの距離だが、カメラから物体までの距離を動画像から取得するのは不可能であるため、式 2 を用いて変換する。

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \tan \phi \\ \frac{\tan \theta}{\cos \phi} \\ 1 \end{pmatrix} \quad (2)$$

本研究では、Z 軸の向きは考慮しないので、 $(x, y)$  成分を使う。顔の向きで行列を検出する時、右目の座標を  $A(\phi_A, \theta_A)$ 、左目の座標を  $B(\phi_B, \theta_B)$  とする。胴体の向きで行列を検出する時、顔の向きと同様に、右肩の座標を  $A(\phi_A, \theta_A)$ 、左肩の座標を  $B(\phi_B, \theta_B)$  とする。 $\vec{AB}$  を  $A$  とすると、

$$A = \begin{pmatrix} \tan \phi_B - \tan \phi_A \\ \frac{\tan \theta_B}{\cos \phi_B} - \frac{\tan \theta_A}{\cos \phi_A} \end{pmatrix} \quad (3)$$

である。ここから人の向きのベクトルを取得するためには、 $A$  の法線ベクトル  $N$  を求めれば良い。法線ベクトル  $N$  は

$$N = \begin{pmatrix} -\frac{\tan \theta_B}{\cos \phi_B} + \frac{\tan \theta_A}{\cos \phi_A} \\ \tan \phi_B - \tan \phi_A \end{pmatrix} \quad (4)$$

または

$$N = \begin{pmatrix} \frac{\tan \theta_B}{\cos \phi_B} - \frac{\tan \theta_A}{\cos \phi_A} \\ -\tan \phi_B + \tan \phi_A \end{pmatrix} \quad (5)$$

で求められる。法線ベクトルが二つ出てきてしまうのは、人の向いている方向と反対の方向のベクトルも法線ベクトルであるからである。どちらを使っても問題ないが、(4) の法線ベクトルを用いて進める。

### 3.3 Step3: 同一方向判定

前節で述べた法線ベクトルを用いて、内積を計算することでベクトルのズレを計算することが可能である。法線ベクトル  $N$  を  $(a, b)$  とおき、他の人の向いているベクトルを  $B = (x, y)$  とすると、

$$\cos \lambda = \frac{ax + by}{\sqrt{a^2 + b^2} \sqrt{x^2 + y^2}} \quad (6)$$

$\lambda$  の範囲は  $0^\circ \leq \lambda \leq 180^\circ$  であるので、 $\cos \lambda$  の取る値の範囲は  $-1 \leq \cos \lambda \leq 1$  である。 $\cos \lambda$  の値が大きければ大きいほど、ベクトルのなす角度が小さい。つまり、人が直線上にいることが導き出される。

### 3.4 Step4: 行列判定

step3 で算出された  $\cos \lambda$  の平均値を用いて、行列判別を行っていく。行列中に  $n$  人いると、算出される  $\cos \lambda$  の総数は  $nC_2$  個である。これより、平均評価値  $M$  は (7) で求められる。

$$M = \left( \frac{1}{nC_2} \sum_{k=1}^{nC_2} \cos \lambda_k + 1 \right) \div 2 \quad (7)$$

$\frac{1}{nC_2} \sum_{k=1}^{nC_2} \cos \lambda_k$  は、 $[-1, 1]$  であるので、 $M$  の取り得る範囲は、 $[0, 1]$  である。 $M$  は人の向きのずれを示す  $\cos \lambda$  の平均値に 1 を足して、その後 2 で割るというものである。

以上の Step を踏まえて、行列を検知していく。



図 10 カメラのレンズに向かって人が並んでいる画像



図 11 カメラのレンズに向かって並んでいる人の骨格検出

#### 4 実験結果・考察

本研究ではリアルタイムでの実装を目的としていたが、実験ではライブストリーミング動画を入力するのではなく、全天球パノラマ動画を入力して評価を行った。今回の実験では3本の入力動画を用いた。

**入力動画 1:** カメラのレンズに向かって人が並んでいる動画。  
**入力動画 2:** カメラのレンズに対して人が横方向に並んでいる動画。

**入力動画 3:** 人が円形になって談笑している動画

である。また、撮影は5人で行った。カメラのレンズに向かって人が並んでいる画像を図10、カメラのレンズに対して人が横方向に並んでいる画像を図13、人が円形になって談笑している画像を図15に示す。

また、入力された動画に骨格検出を行った結果を図11、14、16に示す。それぞれの動画に対して結果と考察を行っていく。

##### 4.1 カメラのレンズに対して人が縦方向に並んでいる動画の行列判定

カメラのレンズに対して人が縦に並ぶように動画を撮影し、その動画を入力した。図11は入力動画に対して評価指数が最大となった骨格検出の結果である。5人の撮影で行ったが、3人の骨格しか検出することができなかった。これは、人が重なってしまったこと、骨格検出アルゴリズムの精度が原因であると考えられる。図11で検出された骨格を用いて算出された人の向きのベクトルと評価指数を表2に示す。

評価指数は0.99を超えており、人の向きのズレの平均はおよそ $10^\circ$ になっている。 $xy$ 座標で計算していることや、パノラマ座標から直交座標に変換するときにプログラムの数値計算で誤

表 2 算出された人の向きベクトルと評価指

1人目の向きベクトル ( $x$ )	$\begin{pmatrix} -0.10845550286497828 \\ -0.009253843777562165 \end{pmatrix}$
2人目の向きベクトル ( $y$ )	$\begin{pmatrix} -0.20331613369061274 \\ 0.0020761955729072523 \end{pmatrix}$
3人目の向きベクトル ( $y$ )	$\begin{pmatrix} -0.8711693236974073 \\ -0.21967540985374967 \end{pmatrix}$
評価指	0.991580570106837



図 12 方位角が $-\pi$ と $\pi$ を示す黒線

差が生じることを考えると満足のいく結果になった。

しかし、評価指数が低くなってしまう事例も見られた。図12に示す黒線の付近に骨格が存在すると、骨格の位置座標が $10^{16}$ 以上の値を返してしまう。この黒線は全天球カメラの側面である。これは、カメラの側面の方位角は $\pi$ もしくは $-\pi$ なので変換公式を用いると骨格の位置座標がとても大きくなってしまうのが原因であると考えられる。そのため、向きベクトルに大きな数値がある場合、式(6)の分母が大きくなってしまい、 $\cos\lambda$ の値が0に近い値になってしまい、評価指数が0.50前後なってしまう。しかし、本研究の目的は直線状の行列を検出することなので、全天球カメラを移動させることでこの問題は解決することが可能である。

##### 4.2 カメラのレンズに対して人が横方向に並んでいる動画の行列判定

カメラのレンズに対して人が縦に並ぶように動画を撮影し、その動画を入力した。図14は入力動画に対して評価指数が最大となった骨格検出の結果である。

左二人の骨格が一人の骨格と認識されている。これは左から一番目の人の左肩が左から二番目の人と重なってしまい骨格が正確に検出されなかったのではないかと考えられる。図14で検出された骨格を用いて算出された人の向きのベクトルと評価指

指数を表3に示す。

人が横方向に並んでいる状態でも評価指

数が0.99を超えることを確認した。

また、縦方向に並んでいる状態と同様に評価指

数が下がってしまう事例が起こった。原因も全く同じで人の向きベクトルが大きいと評価指

数が0.50前後になってしまった。

2つの実験から骨格検出の精度を向上させるために全天球画像を用いた学習を行うこと、全天球カメラの横側に存在している人の骨格の位置座標を検討することが課題であることが分かった。



図 13 カメラのレンズに対して人が横方向に並んでいる画像



図 14 カメラのレンズに対して横方向に並んでいる人の骨格検出

#### 4.3 人が円形に談笑している場合

人が円形に談笑している場合にも提案手法を適用してみた。今研究は直線の行列検知なので評価指標が低ければ直線検知は成功したと言える。円形に並んでいる状態で骨格検出を行ったものを図 16 に示す。この時の骨格の位置座標と評価指標を表 4 に示す。

このパノラマ動画像での解析した結果、評価指標が約 0.52～0.65 を出力した。これは、人の向きのズレが約 74° から 88° であることを示す。4.1, 4.2 節の評価指標より低い数値となった。このことから人が直線で並んでいる状態に実装したアルゴリズムを適用すると評価指標が高くなり、それ以外の人群に適用すると例外を除き、評価指標は低くなることが分かった。

### 5 まとめ・今後の展望

本研究では、先行研究から実空間における物体の状態を検知することが、人の向きに着目した。全天球カメラで撮影された全天球パノラマ動画像は、実空間で向いている方向とは違う方向に表示してしまう。これは全天球パノラマ動画像が、全天球



図 15 人が円形になって談笑している画像



図 16 円形で談笑している人の骨格検出

表 4 算出された人の向きベクトルと評価指標

1 人目の向きベクトル	$\begin{pmatrix} x \\ y \end{pmatrix}$	$\begin{pmatrix} -0.07395756886369838 \\ -0.0045363790051060615 \end{pmatrix}$
2 人目の向きベクトル	$\begin{pmatrix} x \\ y \end{pmatrix}$	$\begin{pmatrix} -0.6926570284433753 \\ -0.8976266517923013 \end{pmatrix}$
3 人目の向きベクトル	$\begin{pmatrix} x \\ y \end{pmatrix}$	$\begin{pmatrix} -0.0710638049602938 \\ -0.015552343666047408 \end{pmatrix}$
4 人目の向きベクトル	$\begin{pmatrix} x \\ y \end{pmatrix}$	$\begin{pmatrix} -0.09736236954734634 \\ 0.12011290922584039 \end{pmatrix}$
5 人目の向きベクトル	$\begin{pmatrix} x \\ y \end{pmatrix}$	$\begin{pmatrix} 0.023139504655534604 \\ 0.01651630954902572 \end{pmatrix}$
評価指標		0.559055588515173

球面動画像を正距円筒図法で変換されているからである。そこで、全天球パノラマ動画像の位置座標を、極座標から直交座標に変換し、人の向きを推定した。人の向きを推定することができたので、実空間の行列を検知するシステムを作成した。

そのシステムの流れは、既存の物体検出アルゴリズム YOLOv4 を用いて行列を検出し、行列中の人の骨格を Openpose を用いて検出する。検出された骨格から人の向いている方向を推定し、その向いている方向が直線であればあるほど、検出した行列の評価指標を上げるというものである。

実験では人が縦方向に並んでいる行列と横方向に並んでいる行列を解析した。結果はどちらの状態でも評価指標が 0.99 を超えた。しかし、課題として全天球カメラの側面に存在している人の骨格の位置座標の値が大きくなってしまい、評価指標が 0.50 前後の値しか出力しない事例が見られた。また、直線状ではない人の集合体に対して解析した結果、評価指標が約 0.52～0.65 となり直線に並んだ人群と差が生じた。このことから本研究で実装したアルゴリズムは、直線に並んだ人群中に適用すると評価指標が高く、それ以外の人群中では評価指標が低い結果になるこ

表 3 算出された人の向きベクトルと評価指標

1 人目の向きベクトル	$\begin{pmatrix} x \\ y \end{pmatrix}$	$\begin{pmatrix} -0.08157645740342542 \\ 0.0036562413720656883 \end{pmatrix}$
2 人目の向きベクトル	$\begin{pmatrix} x \\ y \end{pmatrix}$	$\begin{pmatrix} -0.010879302716581274 \\ -0.0008726744710094159 \end{pmatrix}$
3 人目の向きベクトル	$\begin{pmatrix} x \\ y \end{pmatrix}$	$\begin{pmatrix} -1.2464142291716604 \\ -0.19300661854550794 \end{pmatrix}$
4 人目の向きベクトル	$\begin{pmatrix} x \\ y \end{pmatrix}$	$\begin{pmatrix} -0.3347630439039755 \\ -0.08131559051909604 \end{pmatrix}$
評価指標		0.992834577769814

とが分かった。

今後の展望として骨格検出アルゴリズムの精度を向上させるために、用意されていたデータセットを用いるのではなく、全天球カメラで撮影された画像をデータセットに入れて学習させることを行っていきたい。人が重なって骨格検出されなかった事例が見られた。また、全天球カメラの側面にいる人の骨格の位置座標の見直しをしていくことも課題である。全天球カメラの側面にいる人がいる場合、全天球カメラの位置を変えることで検出することが可能である。しかし、全天球カメラが固定されている場合や今後直線状ではない行列を検出する時に回避する術がない可能性がある。なので、二つの課題点を改善し、行列判定の精度を向上させたい。

また、本研究では直線の行列に着目して検知を行った。しかし、大人数で行列を作るとすると蛇状に並んでいる行列やくの字で並んでいる行列など直線に並んでいる行列のみではない。それらの行列に対して今の提案手法では行列かどうか判断できない。全ての行列が検知可能であるような手法を考えていくのと同時に、既存のアプリケーションに対して、人の向きを検出をして行列を検出していきたい。行列を検出することで、行列とイベントなどの相関が得られる可能性が考えられる。相関が得られるならば、そのような研究も今後行っていきたい。

## 文 献

- [1] 林田 和磨, 横山 昌平 ”全天球カメラにより配信される正距円筒図法動画からのリアルタイム人物検出”, DEIM2020 C6-2(day2 p19)
- [2] 増田 純也, 横山 昌平 ”全天球カメラを用いた騎乗物体に対する位置推定”, DEIM2020 C6-3(day2 p20)
- [3] 浅沼 仁, 岡本 一志, 川本 一彦 ”特徴学習による全方位画像からの人検出”, 知能と情報 (日本知能情報ファジィ学会誌) Vol.27, No.5, pp.813-825 (2015)
- [4] 井上 慶彦, 岩村 雅一, 黄瀬 浩一 ”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—”, Technical Report 20, 大阪府立大学, may 2018
- [5] 庭川 誠, 恩田 寿和 ”監視カメラ映像中の局所的な動き検出とイベント累積による時間的かつ空間的な混雑度調査”, IEEJ Trans. IA, Vol.124, No.10, 2004
- [6] 馬場 賢二, 榎原 孝明, 湯淺 裕一郎, ”画像処理による人流計測システム”, 東芝レビュー Vol.61 No.12 (2006)
- [7] 川下 雄大, 柴田 雅聰, 増山 岳人, 梅田 和昇, ”ステレオカメラを用いた簡易な人流計測システムの構築”, 精密工学会誌/Journal of the Japan Society for Precision Engineering Vol.81. No.2, 2015
- [8] Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao, ”YOLOv4: Optimal Speed and Accuracy of Object Detection”, arxiv:2004.10934vl [cs.CV] 23 Apr 2020