全天球カメラを用いた三角測量による人物ポジショニング推定

笠原　悠樹†　　横山　昌平†,††,†††

† 東京都立大学大学院 システムデザイン研究科　〒 191–0065 東京都日野市旭ヶ丘 6-6
†† 国立情報学研究所　〒 101–8430 東京都千代田区一ツ橋 2-1-2
††† 東京大学生産技術研究所　〒 153–8505 東京都目黒区駒場 4-6-1
E-mail: †kasahara-yuki1@ed.tmu.ac.jp, ††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可能である. 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. 本研究ではカメラと人の相対的な距離を取るのではなく, 絶対的な距離を取るため三角測量を用いて人の位置および向きを推定することを目的にした. 単眼カメラを用いた三角測量では 3 つの既知情報が必要であるが, 全天球カメラでの三角測量は 2 台のカメラ間の距離のみを既知情報にすることで推定することができる.

キーワード　全天球カメラ, 三角測量, 骨格検出, 位置推定, 人の向き

1　は じ め に

Google Street View[1]や AR (Argument Reality:拡張現実) や VR(Vitual 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[2]や YouTube[3]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可能である. 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な状態がある. 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, 人の向きに着目した. 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可能である. 物体の向きを検知できれば, 人流解析の精度や行動推定に応用す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.

人の向きを検出する際, 本研究では骨格検出アルゴリズムである OpenPose [8] を用いている. OpenPose で取得することができた骨格の位置情報を使うことで, 体の向きの推定を試みた. しかし, 得られた骨格の位置情報はデカルト座標系ではなく, 極座標系で表されてしまっている. これは前述した全天球カメラで撮影された動画像が球面であるものを正距円筒図法を用いて

1：https://www.google.co.jp/maps/preview
2：https://twitter.com
3：https://www.youtube.com

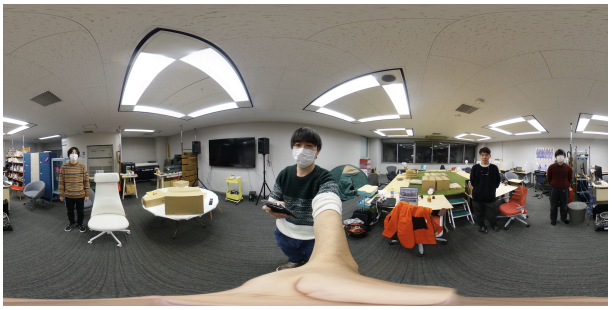

図 1　人がレンズの方を向いている全天球パノラマ画像

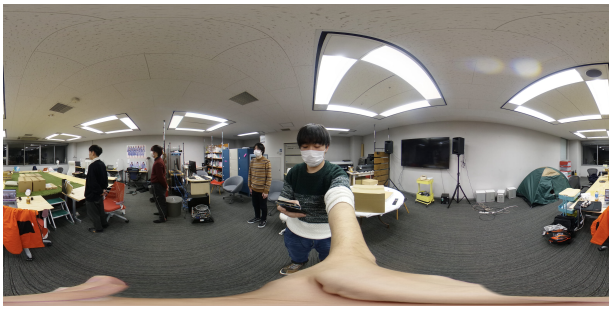

図 3　人が同じ方向を見ている全天球パノラマ画像

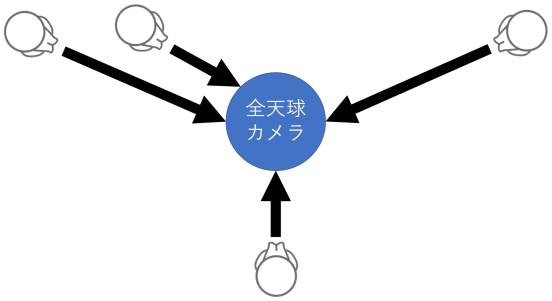

図 2　上から見た撮影風景 (人がレンズの方を向いている)

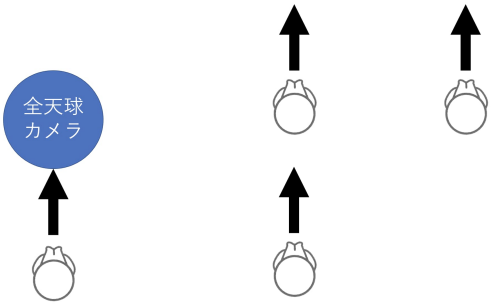

図 4　上から見た撮影風景 (人が同じ方向を見ている)

パノラマに変換しているからである．そのため, 得られる骨格の位置情報は方位角と仰角の二種類である．なので, 人とカメラの距離を推定することができれば, 極座標変換が可能となりカメラの動画像である 2 次元情報から 3 次元情報を得ることができると考えた．

カメラと人の距離を知るための情報は近年画像処理の研究が進み, 機械学習を使った深度推定などが挙げられる．しかし, これらの手法はカメラと物体の相対的な位置を知る手法であり, 絶対的な位置を知る手法ではない．なので, 複数のカメラを使うことで三角測量を行い, カメラと人との距離を推定できるのではないかと考えた．

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な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 % 以上の人数計測計測精度を実現した．

図 5　Openpose を用いて骨格検出を行った結果

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. 人のポジションを推定するには, 人の位置, 人の向きの 2 つを推定する必要があると考えた. そのため, 三角測量によって人の位置を推定する段階と得られた結果から人の向きを推定する段階の 2 つの段階に分けて今回は行う. 人の位置や向きを推定するのに本研究では OpenPose を用いた.

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 5 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

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. 既存のデータセットは単眼カメラで撮影された画像を教師データにしているため, 全天球カメラでは動画像中に歪みが発生し, 正確な検出が行われない. しかし, 歪みの発生するのは主に動画像中の高緯度領域であるため, カメラの高さを人の高さに合わせることで, 低緯度領域で検出を行うことで可能にする.

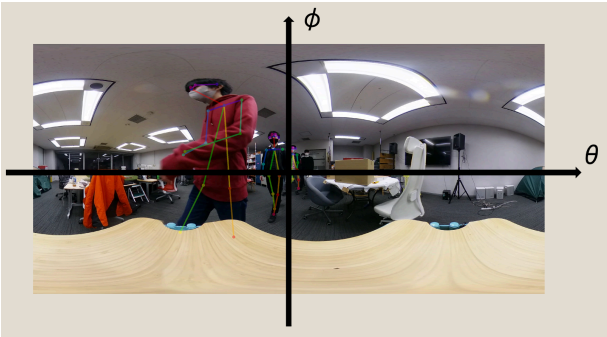

図 6　全天球カメラにおける座標軸

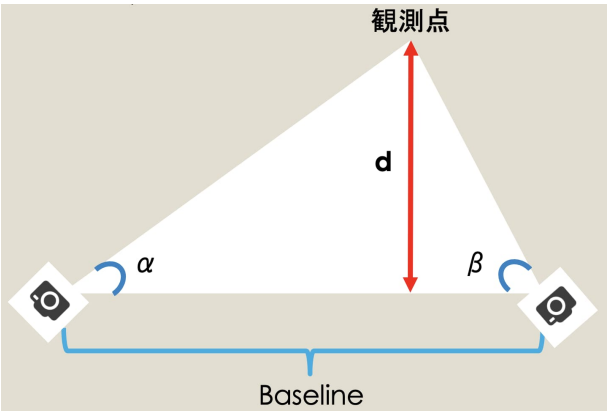

図 7　三角測量の概要

ここで, 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. 図 6 は全天球カメラにおける座標軸である.

出力された位置座標は極座標系であるから, $(\theta,\ \phi)$ である. θ は方位角, ϕ は仰角である.

3.1　Step1:三角測量を用いた人の位置推定

図 7 は, 三角測量の概要図である.

三角測量は通常単眼カメラを 2 台用いて行う手法である. 2 台のカメラ間の距離 *Baseline* とその *Baseline* とカメラと観測点を結んだ直線がなす角 α, β の 3 値必要である. これらの 3 値を使うことで, 観測点が *Baseline* に降ろす垂線の足 d を求めることができる.

$$d = \frac{Baseline}{\frac{1}{\tan\alpha} + \frac{1}{\tan\beta}} \tag{1}$$

単眼カメラでは観測点にカメラを向け角度を測る必要があるため, 観測点が静的な物でしか行えず, 動的な観測点の場合逐一カメラの向きを変えなければならなかった. しかし, 全天球カメラで行う場合, デメリットであった位置座標が極座標で表されていることがメリットとなる. 観測点が人である場合 OpenPose を用いることで, 方位角を取得することができる. それゆえ, な

す角 α, β は OpenPose を用いることで取得することができるので, $Baseline$ のみを計測することで行うことができる. また今回はカメラと観測点の距離を推定する必要があるため, その求め方は式 (2) で求めることができる.

$$distance_{camera} = \frac{Baseline \sin\beta}{\sin(\alpha+\beta)} \tag{2}$$

3.2 Step2:OpenPose を用いた人の向き推定

極座標から直交座標 (x, y, z) への変換は

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\sin\phi\cos\theta \\ r\sin\phi\sin\theta \\ r\cos\phi \end{pmatrix} \tag{3}$$

である. ここで r は極座標における中心からの距離を表すため, r は人と全天球カメラの距離である. 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. OpenPose での位置推定で観測物の骨格の位置座標をそれぞれ出力されているので, それぞれの位置座標に適応することで極座標系からデカルト座標系に変換することができる. 左肩と全天球カメラの距離を r_{left}, 右肩と全天球カメラの距離を r_{right} とすると前節の式 (2) を応用することで求めることができる. これを利用することで両肩の位置座標を極座標からデカルト座標系に直すことができる. ここで, 胴体の向きで体の向きを推定する時, 右肩の座標を A(ϕ_A, θ_A), 左肩の座標を B(ϕ_B, θ_B), とする. 体の向きベクトル $\overrightarrow{AB}$を $\boldsymbol{X}$ とすると,

$$\boldsymbol{X} = \begin{pmatrix} r_{left}\sin\phi_B\cos\theta_B - r_{right}\sin\phi_A\cos\theta_A \\ r_{left}\sin\phi_B\sin\theta_B - r_{right}\sin\phi_A\sin\theta_A \end{pmatrix} \tag{4}$$

である. ここから人の向きの角度を取得するためには, y 軸と並行の単位ベクトル $\overrightarrow{e_y}$ との内積を計算することで角度を求めることができる. 体の向きベクトル $\boldsymbol{X}$ を (X, Y) とおき, x 軸と並行の単位ベクトルを $\overrightarrow{e_x} = (1, 0)$ とすると, 求める体の向きの角度 λ は式 (5) で求められる.

$$\lambda = \arccos\frac{X}{\sqrt{X^2+Y^2}} \tag{5}$$

式 (5) の λ のとりうる範囲は, $\lambda \in [0, \pi]$ である. 体の向きは $0°$ から $360°$ であるので全ての場合を, 把握できない. しかし, $\boldsymbol{X}$ の成分の正負で判別することができる. 本来, $\frac{X}{Y} > 0$ の場合 $\lambda \in (0°, 90°)$ である. 体の向きが $180°$ を超えた場合, $\frac{X}{Y} > 0$ のとき, $\lambda \in (180°, 270°)$ である. つまり, 体の向きが $180°$ を超え, $\frac{X}{Y} > 0$ の場合式 (5) を利用すると, 出力される λ が取り得る範囲は $\lambda \in (90°, 180°)$ となる. このことを利用して, 式 (6) で体の向きを補正する必要がある.

$$\lambda' = 360° - \lambda \tag{6}$$

これらの手法を使い人の位置並びに人の向きを推定していく.

4 実験結果と考察

本研究ではリアルタイムでの実装を目的としているが, 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, 全天球パノラマ画像を入力して実験を行った. 今回の実験では $Baseline$ の距離を 4m40cm に設定した. また人の骨格が歪むのを軽減するためにカメラの高さを人の胸の位置に合わせ OpenPose の精度を下げないようにした. 人の位置を推定するために鋭角三角形, 鈍角三角形を作り推定した.

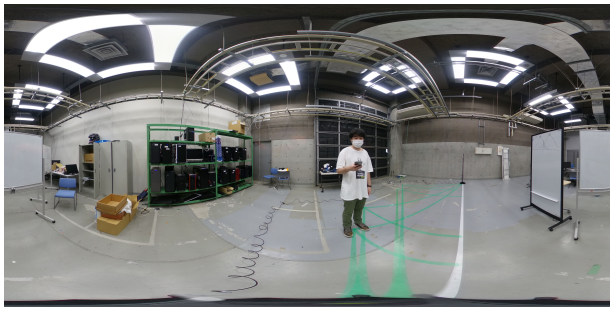

図 8 三角測量を行なった時の左側のカメラ

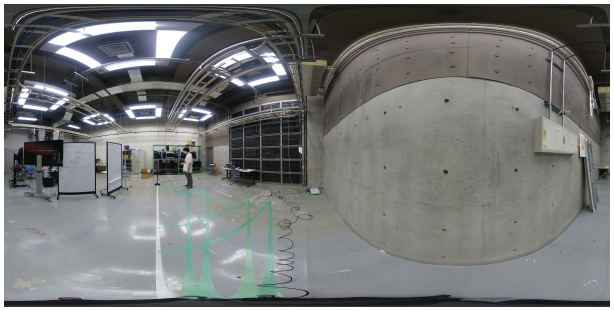

図 9 三角測量を行なった時の右側のカメラ

4.1 人の位置推定フェーズ

実際に三角測量を行ったときの画像を図 8, 9 に示す. 図 8, 9 に写っている白い線が $Baseline$ であり, 緑色の線が三角測量を行うときにあらかじめ測定した線である. 本研究はカメラと人の距離 $distance_{camera}$ のみを求めれば良いが, 三角測量を行えているか確認するために観測物が $Baseline$ に降ろす垂線の足を求めることにした. カメラと人の距離 $distance_{camera}$ は左側のカメラを基準にして求める. このような実験環境で実験を行なった.

実験の結果を表 2, 3 に示す.

表 2 垂線の足の実験結果 (位置推定)

	垂線の足	推定された垂線の足	生じた誤差
パターン 1	187 cm	190.1202 cm	3.1202 cm
パターン 2	124.5 cm	131.1538 cm	6.6538 cm
パターン 3	220 cm	218.1974 cm	-1.8026 cm
パターン 4	50 cm	64.0481 cm	14.0481 cm

パターン 1 から 3 は鋭角三角形で行い, パターン 4 は鈍角三角形で行なった. 結果として鋭角三角形の場合, 垂線の足の長さの誤差が一番大きかったのはパターン 2 の約 6.5cm となり, 全天球カメラと人の距離推定で誤差が一番大きくなったのは, パターン 3 の約 16cm となった. また, 鈍角三角形の場合は正しく推定することができず, 誤差が大きくなった. それに対し, 全

天球カメラを増やし, 鋭角三角形にした時の結果を表 4 に示す.

表 3　カメラと人の距離の実験結果 (位置推定)

	カメラとの距離	推定されたカメラとの距離	生じた誤差
パターン 1	217 cm	230.1766 cm	13.1766 cm
パターン 2	137 cm	150.3202 cm	13.3202 cm
パターン 3	350 cm	366.2878 cm	16.2878 cm
パターン 4	75 cm	160.9317 cm	85.9317 cm

表 4　鈍角三角形から鋭角三角形にしたときの実験結果 (位置推定)

カメラとの距離	推定されたカメラとの距離	垂線の足	推定された垂線の足
120 cm	127.4621 cm	86 cm	88.1273 cm

全天球カメラを新しく追加したため, *Baseline* の長さは今回 2m に変えた. *Baseline* の長さを短くしても, 誤差の大きさが変わらなかったため, 複数のカメラがある場合, 位置推定は行える.

人の位置推定について, 人と全天球カメラが作る三角形が鋭角三角形の場合は誤差が小さく推定することができたが, 鈍角三角形の場合は鋭角三角形と比べ, 垂線の足, 全天球カメラと人の距離はどちらも誤差が大きく, 推定することができなかった. これは, 三角測量が鋭角三角形の場合のみしか考えられておらず, 鈍角三角形に対応していないのが原因である. 平澤は鈍角三角形の場合も推定することができているが, センサーを用いていること, 鈍角三角形になった時に条件を変えることで推定を可能にしていた. そのため, ひとつのプログラムでは鈍角三角形を含めた三角形を測定することができない. その弱点を補うため, もう一台の全天球カメラを増やすことで解決した. 街中では複数のカメラが存在しており, カメラの台数にこだわる必要はまったくないと考えている. その結果, 表 4 にもあるように, パターン 4 に比べてカメラとの距離, 垂線の足の誤差が小さくなった.

また, 人が *Baseline* に近づくと誤差が大きくなることも確認できた. これは, 測定した角度の誤差が大きくなることが原因である. 人が *Baseline* に近づくと全天球カメラと人がなす角が小さくなる.

また, 実際のなす角が小さいと計測誤差, 推定誤差が大きく影響することが原因であると考えられる. 実験環境の大きさに限界があるため, 実際に大きさに制限がないところで実験した場合, 人が *Baseline* から遠ざかっても誤差が大きくなると予想される. これは三角測量に問題があるのではなく, OpenPose の検出限界があるからだと考えている. 本研究で使用している RICOH 社の全天球カメラは 2 つの魚眼レンズで撮影された画像を全天球カメラ側で一枚の画像を生成している. そのため, 全天球カメラで撮影された画像は通常の単眼カメラで撮影された動画像より解像度が高く, 情報量が多い. なので, 通常のカメラで撮影される場合と全天球カメラで撮影される場合だと全天球カメラで撮影される場合の方が被写体が小さくなってしまうという問題点がある. 推定したい人が小さければ小さいほど

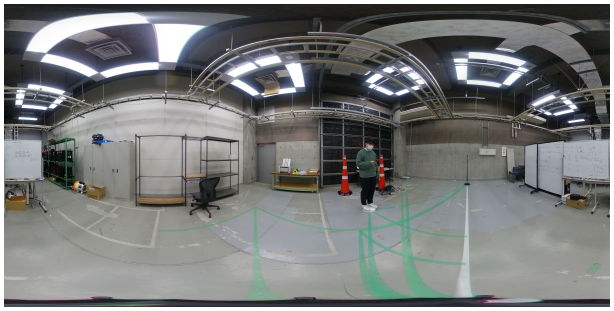

図 10　パターン 1 における 0° を向いている画像

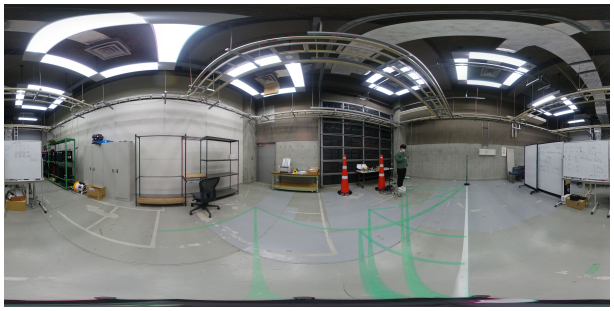

図 11　パターン 3 における 0° を向いている画像定

OpenPose の検出精度は低くなってしまう. これにより, 本研究では実験が行えなかったが, 人を *Baseline* から遠ざけてしまっても誤差は大きくなると推察することができる.

垂線の足と全天球カメラと人の距離では垂線の足の方が誤差が小さくなった. この誤差は, 測定誤差や推定された骨格の位置座標の誤差が含まれると考えられる. また, 全天球カメラと人の距離の誤差が大きくなったのは, 式 (**??**) の分母が影響していると考えられる. 式 (1) では 2 つの推定された角度を代入することのみで推定することができたが, 式 (**??**) では推定された角度を足し合わせて行なっているため, 推定された角度の誤差がより大きくなり, 全天球カメラと人の距離の誤差が大きくなったと考えられる.

4.2　人の体の向き推定

前節では胸の骨格位置座標を使うことで, 人と全天球カメラと距離を推定した. 今回は体の向きを利用するので, 右肩と左肩の骨格位置座標を利用して右肩と全天球カメラの距離推定, および左肩と全天球カメラの距離推定を行う必要がある. そのため, 位置推定を 2 回行う必要がある. 得られた位置推定を利用することで, 極座標変換を行い, 体の向きの角度を推定する.

人の位置推定が前節で観測される人と全天球カメラがなす三角形が, 鋭角三角形の場合のみ誤差が少なく推定できることを確認したため, 今回の実験では前節のパターン 1, パターン 3 で作られた三角形で人の体を向いている角度を推定することにした. 本研究では, *Baseline* と体の向きが平行かつ, 人が全天球カメラの方向を 0° とし, 人が左回転する方向を正の向きとした.

パターン 1 とパターン 3 は表 4 より, 約 3cm ほどの誤差があり, その誤差が結果にどのような影響を及ぼすのかを検証する. また, 全天球カメラと人の位置が離れている場合でも, 人の

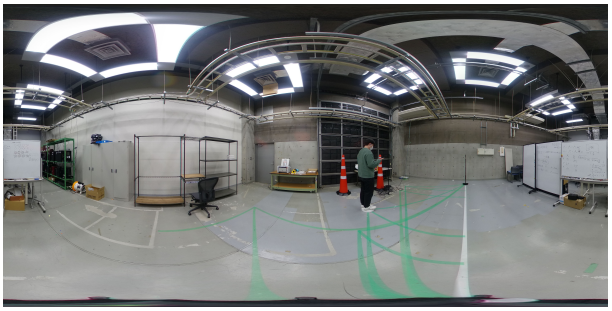

図 12　パターン 1 における 90° を向いている画像

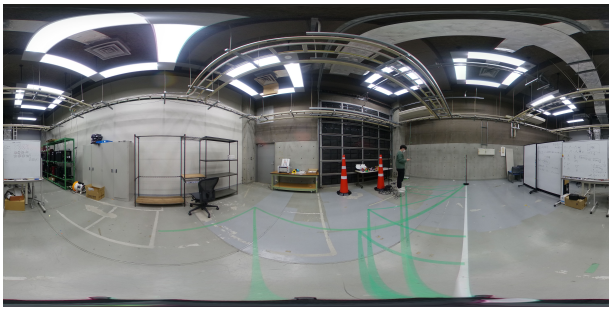

図 14　パターン 1 における 45° を向いている画像

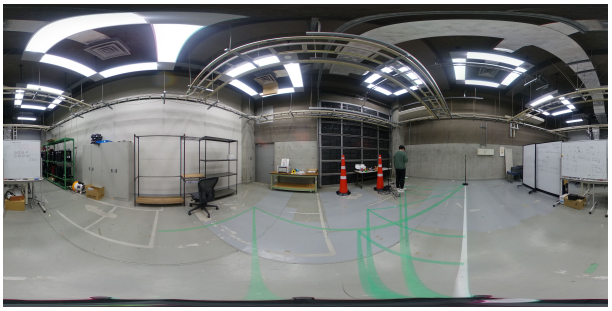

図 13　パターン 3 における 90° を向いている画像

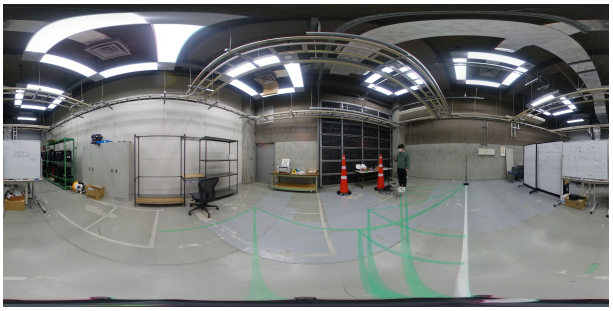

図 15　パターン 1 における 270° を向いている画像

体の向き推定に影響があるのか確認するために, パターン 1 とパターン 3 を利用した.

人の体の向き推定を行う上で, 前節のパターン 1 の三角形上で体の向きを, 0°, 45°, 90°, 135°, 180°, 225°, 270°, 315° で推定を行なった. 実験結果を表 5 に示す.

表 5　パターン 1 における角度推定の結果

	推定結果	推定結果との誤差
0°	1.3821°	1.3821°
45°	44.7295°	0.2705°
90°	91.2938°	1.2938°
135°	133.7961	1.2039°
180°	178.9315°	1.0685°
225°	224.4910°	0.5010°
270°	266.2403°	3.7567°
315°	313.7324°	1.2676°

また, 同様に前節のパターン 3 の三角形上で体の向きを, 0°, 45°, 90°, 135°, 180°, 225°, 270°, 315° で推定を行なった. 実験結果を表 6 に示す.

表 6　パターン 3 における角度推定の結果

	推定結果	推定結果との誤差
0°	358.8396°	1.1604°
45°	45.7223°	0.7223°
90°	92.1280°	2.1280°
135°	136.2401	1.2401°
180°	178.0912°	1.8888°
225°	226.8299°	1.8299°
270°	270.9283°	0.9283°
315°	314.0984°	0.9016°

パターン 1, 3 で撮影した 0°, 90° の画像を図 10, 11, 12, 13 に示す. 図 10, 11 では向いている方向が 0° だが, 画像を閲覧するだけでは 0° の方向を向いているか確認できない. これは, 正距円筒図法によって, 全天球球面画像を全天球パノラマ画像に変換しているからである. 前章の位置推定を利用し, 角度推定を行なった結果誤差が小さかったものは, パターン 1 の 45° の 44.7295° となった. 誤差が最も大きくなったのは, パターン 3 の 180° の 1.8888° となった.

位置推定が行えた人と全天球カメラが作る三角形が鋭角三角形である場合, 人の体の向きの角度の推定は距離によらず一定の誤差で推定されることになった. これは場所によらず右肩, 左肩の距離を推定することができ, 極座標変換が行えたことが確認できる. 極座標変換ができ, デカルト座標系で表された骨格の位置座標から体の向いている方向を推定することができた. 表 5, 6 より誤差が小さかったものは, パターン 1 の 45° の 44.7295° となった. 図 14 より誤差が小さくなった要因は, 画像のピクセル値を使った方位角と実際の方位角に誤差があまり生じなかったのが原因ではないかと考えている. 実際の実験結果では, 測定誤差も含み約 1.5 の誤差が生じているが, 測定誤差を含め全てが噛み合った結果誤差が小さくなったと考えている.

また, 最も大きな誤差となったのは, パターン 1 の 270° を向いているときとなる. 図 15 よりこれは推定された人の姿勢が悪いのが原因であると考えられる. 直立している条件で行なっていたが, 全天球カメラで撮影するデバイスが手元のスマートフォンであったため, 姿勢が崩れ想定していた体の向きとは違う向きを向いていると考えている.

また, 他の人物で実験を行った結果を 7, 8 に示す. 他の人物で実験を行った際, 生じた誤差はパターン 1 では平均約 2.4°, パターン 3 では平均約 2.65° となった. 先ほどの実験結果とは誤

差が大きくなってしまったが, これは最初の基準点に合わせる際に, 体の向きが正しい方向を向けていないのが原因であると考えている．そのため, 推定された値と実際の体の向きとの誤差はほとんど一定の値を出力されている．また, 他の人物にして実験を行なっても場所に関わらず, 誤差は変わらないことが確認できた.

表 7　パターン 1 における角度推定の結果 2

	推定結果	推定結果との誤差
0°	2.4100°	2.4100°
45°	48.0912°	3.0912°
90°	92.7192°	2.7192°
135°	137.3128	2.3128°
180°	177.3819°	2.6181°
225°	223.4913°	1.5087°
270°	267.9103°	2.0803°
315°	312.3130°	2.6870°

表 8　パターン 3 における角度推定の結果 2

	推定結果	推定結果との誤差
0°	357.2134°	2.7866°
45°	43.8919°	1.1081°
90°	86.9872°	3.0228°
135°	137.3128	2.3128°
180°	184.1061°	4.1061°
225°	227.3121°	2.3121°
270°	272.9281°	2.9281°
315°	317.6630°	2.6630°

表 5, 6 より結果によって誤差の大きさは異なるが, 誤差平均は約 1.5° となった．仮に 1km 歩いた場合 $\sin 1.5° \approx 0.0261769483$ なので, およそ 26m 離れることとなる．全天球カメラをセンサーとして利用する際, 1km 間隔で設置することはないと考えている．そのため, 仮に 100m 感覚に設置した場合生じる誤差は約 2.6m であり, 100m 間隔で向いている誤差を補正することによって, 人の行き先を推定する研究に応用することができるのではないかと考えている.

5　まとめと今後の課題

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．そこで全天球カメラで撮影された動画像の位置座標が極座標であることを利用して, OpenPose と三角測量を用いることで人の位置を推定することが可能となった．人の位置を相対距離でなく絶対位置で推定することができたので, 極座標変換を利用することができるようになった．人の位置推定の時に OpenPose で人の骨格を推定していたので, 人の体の向きを推定した.

その結果, 計算値はおおむね実際の距離と同じになった. その結果, 計算値は実際の距離と概ね同じであったが, パターン 1,2 ではわずかな誤差があった．しかし，全方位カメラで撮影した人物が *Baseline* に近づくと，測定した角度の誤差が大きくなり 距離の誤差も大きくなる．また, 測定角度が鈍角の場合, 距離の測定は不可能であることがわかった. これは, 三角測量では計算できないためと思われる．このような問題を解決するために, 我々は複数のカメラを使用することを検討している．実際の街には, 1 台や 2 台ではなく, たくさんのカメラがある. 複数のカメラを使用することで カメラまでの距離を正確に測定し, 3 次元情報を取得することができる．また、体の向きの誤差は人と全天球カメラがつく三角形が鋭角三角形の場合, 一定の誤差で推定することができた．誤差の平均は約 1.5° となり, 応用する研究にもよるが, 複数台の全天球カメラを使うことで体の向きを補正し, より正しい体の向きを推定できるのではないかと考えている.

今後は複数人に対応できるようにし, 街中で実際に適用できるアプリケーションの作成を目標としていきたい．本研究では一人のみで実験を行なっている．また, 実験環境も室内で行なっており, 広い場所での実験を行えていない．検出限界や推定値が大幅な誤差が出力される距離などを把握する必要があると考えている．加えて, リアルタイムでの実装も視野に入れている．これらの課題を克服することで精度の良い行動予測や人流解析ができるのではないかと考えている.

謝　　辞

本研究の一部は国立情報学研究所 CRIS 委託研究及び東京都立大学「ローカル 5G 環境を活用した最先端研究」の助成を受けたものである.

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 雅聡, 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/Jounal of the Japan Society for Precision Engineering Vol.81. No.2, 2015

[8] Cao, Zhe and Hidalgo, Gines and Simon, Tomas and Wei, Shih-En and Sheikh, Yaser, ”OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields”, IEEE Transactions on Pattern Analysis and Machine Intelligence 2021