

広域移動する人物の映像解析のための自動追従ドローン

山下 浩豊^{1,a)} 満上 育久^{1,b)}

概要: 本研究では、広範囲移動する人物の映像解析を想定した、人物を一定距離・一定方向から撮影し続ける自動追従ドローンを提案する。既存の人物追従ドローンでは、画像中での人物全身の矩形検出情報から機体を制御するため、追従対象人物の姿勢状態（向き）までは認識されず、人物の向きを考慮した追従は行われていない。提案システムでは、ドローンの単眼カメラから撮影される映像に対して OpenPose による人物の 2 次元姿勢推定を適用し、さらにそこから人物の向きを得るために 3 次元姿勢推定を行い、それらに基づいてドローン制御を行う。実機を用いた実験により姿勢推定精度および制御性能を評価し、提案システムの有用性を確認する。

1. はじめに

画像・映像内の人物の行動や振る舞いを理解する人物映像解析は、コンピュータビジョン分野における重要な研究対象の 1 つである。人物検出・追跡、顔認識、姿勢推定等は、犯罪捜査、購買行動分析、スポーツ解析、セキュリティ認証など、我々のほぼ全ての生活空間内でこの人物解析技術が応用されつつある。

この実社会で応用される人物映像解析では、防犯カメラに代表されるような固定カメラの映像が主な対象となる。このような固定カメラ映像において、人物を遠くから撮影する場合、その人物の観測可能範囲は大きくなるが、動作や表情の詳細度は相対的に少なくなる。一方、人物を近くから撮影した場合、詳細な情報は得られるが、観測範囲は狭い範囲に限定される。このようにカメラが固定設置されている限り、撮影される人物の詳細度と観測可能範囲の広さの間にはトレードオフの関係が存在する。したがって、従来の固定カメラを用いた人物映像解析では、移動人物を継続して高解像度に撮影することができず、解析の適用範囲もその固定カメラの撮影エリアの狭い範囲に限定されていた。このような単眼固定カメラの問題に対して、複数固定カメラから撮影される映像を活用することで、人物映像処理の性能の向上を図る試みも行われているが、このような複数カメラ映像においても観測範囲は依然として有限である。

近年、身近なデバイスになりつつあるドローンは、この

問題を解決できる可能性を秘めている。実環境上を自由に移動できるドローンを使用して、特定の人物の側から撮影し続けることが可能となれば、人物の移動範囲の広さに関わらず高解像度に撮影できるため、映像解析の適用範囲や処理精度の向上が期待でき、マラソン選手の運動解析や犯罪者・不審者追跡による犯罪捜査など様々な分野への応用が考えられる。

このような映像解析を実現するためには、ドローンは解析に適する人物映像を取得する必要がある。例えば、顔検出・認証処理では、人に対する撮影向きが正面から背後に遷移するにつれて精度が低下するため、人物を前方から撮影し続ける必要がある。また、姿勢推定や歩容認証の研究分野では、それらの処理精度が人物の観測方向によって異なるため [1]、最適な撮影方向を調べる研究 [2], [3] なども行われている。したがって、人物の観測方向は映像解析において重要な調整要素である。

そこで本研究では、人物に対して一定距離を保ちながら、かつ、システム自身が人物の向きを把握し、一定方向から観測し続けるドローンを開発する。提案システムでは、ドローンの単眼カメラから撮影される映像に対して人物の 2 次元姿勢推定を適用し、さらにそこから人物の向きを得るために 3 次元姿勢推定を行う。この 2 次元姿勢・3 次元姿勢の関節点に基づいて、人物の歩行方向を推定し、ドローンと人物の相対的な距離・方向を保つ制御を行う。実機を用いた実験において、姿勢推定の精度および制御性能を評価し、提案システムの有用性を示す。

2. 関連研究

人物追従ドローンはすでに実用化が進み、DJI 社の

¹ 広島市立大学
Hiroshima City University

a) yamashita@sys.info.hiroshima-cu.ac.jp

b) mitsugami@hiroshima-cu.ac.jp

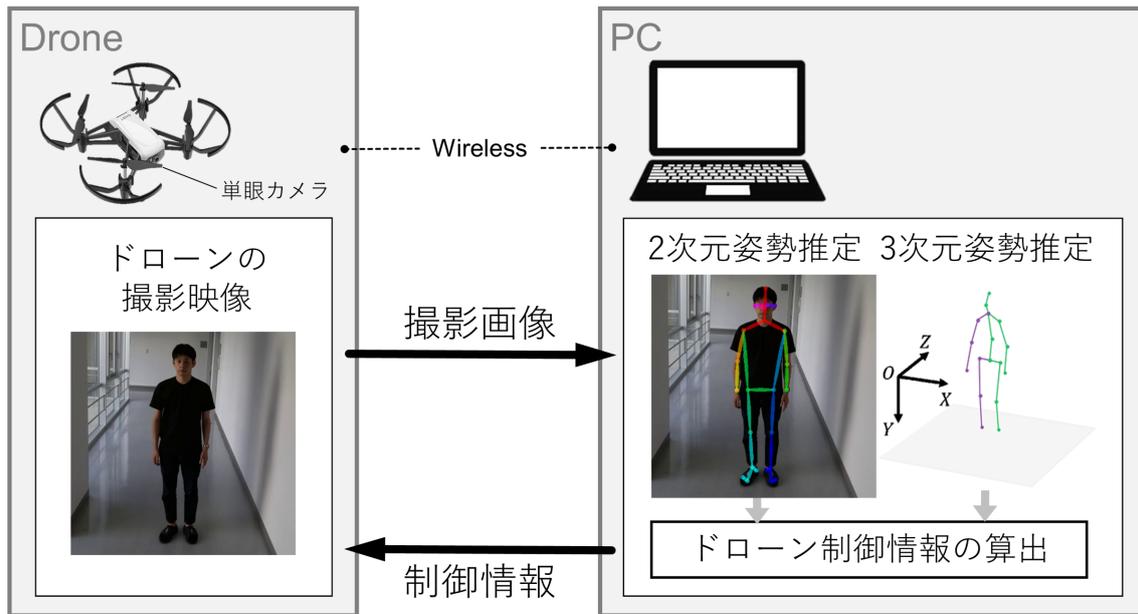


図 1 人物追従ドローンのシステム構成. ドローンの撮影画像に対して, PC 側で人物検出処理を行い制御情報を算出してドローンに送信する. このサイクルを繰り返すことで人物追従を行う.

MAVIC 2 PRO[4] や Skydio 社の Skydio 2[5] などの自動追尾機能を搭載したドローンは, 観光先やスポーツの空撮などに利用されている. また, ドローンで人物追従を行う試み [6], [7] や犯人や不審者を自動で検出・追跡するドローン [8], スポーツ解析用のドローン [9] などが存在する. しかし, これらの人物追従ドローンに共通する点は, ドローンが撮影対象人物に対して比較的遠い地点から観測している. したがって, これにより撮影される映像は俯瞰映像に近いものとなり, 人物の詳細情報を多く含むような映像の取得には不向きである.

また, Pestana らは, GPS 非搭載ドローンを使用して, 撮影画像内の動的物体 (人物を含む) を OpenTLD[10] で検出し, それに対して一定距離を保ちながら追従を行う手法 [11] を提案している. さらに, Mel らは, 追従時の静的障害物の衝突回避を想定して, ドローンが人物の移動軌跡を追従するシステム [12] を提案している. この手法においても, 人物歩行経路を予測するために人物を常に観測できるよう, その人物に対して一定範囲内の距離を保ちながら追従している. しかし, これらの人物追従ドローンは, 追跡人物の検出処理において, その人物の全身が画像中のどの位置に存在するのかを矩形検出しているが, 人物の姿勢状態までは認識されない. したがって, ドローン自身は人物に対する観測方向を把握できない.

これに対して, Nageli らは複数ドローンから得られる映像から人物の 3 次元姿勢を推定する手法 [13] を提案している. この手法では複数ドローンから得られる映像から, その撮影人物の 3 次元姿勢を推定しつつ, 同時に推定関節点に基づいて機体の位置姿勢の制御情報を算出し, 制御を

行っている. しかし, この手法では, 追跡人物に装着されるマーカーに基づいた 3 次元姿勢から追従を行ってが, このマーカーによる人物検出処理において, 例えば, 犯人不審者追跡のような場合に追従対象者に装着できない. したがって, 追跡対象人物へのマーカーの装着を要求することは, 実利用を考えた場合に好ましくない.

我々の提案手法では, ドローンから撮影される単眼画像から 3 次元姿勢を推定することによって, 人物に対して比較的近い一定距離を保ちながら, かつ, システム自身が人物をどの方向から観測しているのかを把握し, 一定の向き (人物の前方や背後など) から観測し続ける自動追従ドローンシステムを提案する.

3. 人物追従ドローンシステム

3.1 システム構成

本システムでは, Ryze Technology 社の Tello [14] を使用する. この屋内用小型ドローンは Python プログラミングでの機体制御や画像処理を行うことができ, Python 用の様々なライブラリと併用することで高度な画像処理を可能とする. また, 機体前方に内蔵された単眼カメラ (以下, ドローンカメラ) によって撮影可能で, その画像解像度は 960×720 (pixel) である.

人物追従ドローンのシステム構成を図 1 に示す. ドローンは無線 LAN のアクセスポイントとなっており, そのネットワークに PC が接続されている. ドローンカメラから撮影された画像は随時 PC に伝送され, PC 側で取得画像に対して人物姿勢推定が行われる. 推定された人物姿勢から算出されたドローンの制御入力値は, ネットワークを通じ

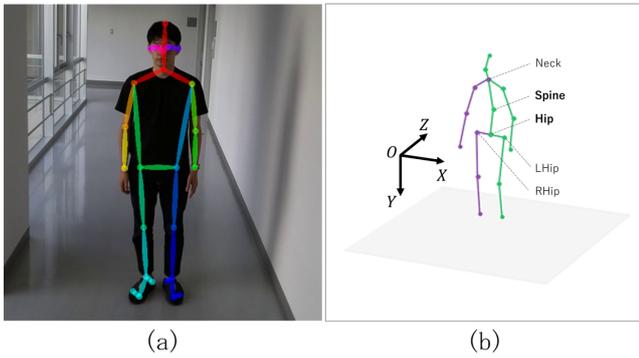


図 2 (a)OpenPose の 2 次元姿勢推定, (b)Martinez らの手法による 3 次元姿勢推定

て随時ドローンに伝送される。この処理により、ドローンが人物をリアルタイムに追従する。

3.2 人物姿勢推定

本システムで使用する 2 次元・3 次元姿勢推定手法の適用例を図 2 に示す。本研究では、ドローンカメラの画像中から人物の位置およびその姿勢を獲得するために、CNN ベースの人物姿勢推定器である OpenPose [15] を用いる。OpenPose には異なるデータセットから構築された推定学習モデルが複数存在し、各モデルによって推定関節数、推定精度、推定速度がそれぞれ異なるが、本システムでは、BODY_25B を使用する。これにより、ドローンカメラの RGB 画像内に写る人物の合計 25 個の関節点の 2 次元座標 $B = \{b_{(JointID)}\}$ が獲得できる。

さらに、人物の向きを得るため、Martinez らが提案した 3 次元姿勢推定手法 [16] を適用する。これは、2 次元関節点群 B を入力として、その 3 次元座標群 $P = \{p_{rm(JointID)}\}$ を推定する手法であるが、ここで想定される骨格モデルは関節点が 17 個であり、上述の BODY_25B とは一部異なる。本研究では、両モデルに含まれる 15 個の関節については $b_{(JointID)}$ をそのまま用い、不足している 2 つの関節点 (Hip, Spine) については、他の関節点から以下のとおり算出して用いる。

$$b_{Hip} = \frac{b_{LHip} + b_{RHip}}{2} \quad (1)$$

$$b_{Spine} = \frac{b_{Hip} + b_{Neck}}{2} \quad (2)$$

なお、上式から得られる b_{Hip} は、3 次元空間において関節点 LHip と RHip の中点である Hip を、2 次元画像中の中点として算出しており射影幾何学的に不正確である。また、関節点 Spine は Hip と Neck の中点として算出しているため、射影幾何学的な不正確性に加えて、そもそも胴体の屈曲を捉えていない。しかし、本研究においてドローン制御のために必要となるのは後述の両肩の関節点 $p_{LShoulder}, p_{RShoulder}$ のみである。Martinez らの 3 次元姿勢推定手法 [16] では、2 次元関節点が 1 つでも欠落しているとすべての関節点の 3 次元位置が得られないため、ここ

では便宜上このように近似的に Hip と Spine を算出している。

3.3 ドローン制御

本システムでは、前節の手續きで推定された 2 次元・3 次元関節点に基づいてドローンを PID 制御する。以下の 3.3.1, 3.3.2 節でその詳細を述べる。

3.3.1 人物を一定距離から観測する制御

OpenPose により推定される 2 次元関節点に基づいて、ドローンが人物との相対的距離を維持する制御を行う。ただし、この推定姿勢情報から人物のメトリックな距離情報を獲得するのは本質的に不可能である。もし追従対象者の身長情報が得られれば距離情報が推定可能だが、人物ごとに身長の入力が求められるのは、実利用を考えた場合に面倒である。そこで提案システムでは、メトリックな距離情報を推定するのではなく、追従対象人物の画像内での大きさを一定にする制御を行うことで相対的距離を維持させる方法を採用する。

この制御において、画像内の人物の大きさ情報としてどの関節点を用いるかが制御の安定性に大きく影響する。例えば、手や足の関節点をもとに、その間の長さを一定にするよう制御した場合、人物が一定速度で直線歩行しても、これらの関節位置は常に上下に移動するため、画像内の変動が大きくなり、制御が不安定になる。また、例えば鼻・手・足の関節点を使用する場合、人物をある一定の向きから観測したときに、自己遮蔽により検出不可能となり、制御が行えなくなる。したがって、ドローン制御に用いる関節点は、人の歩行動作中の位置の変動が小さく、かつ、任意方向から撮影した場合でも安定して取得されることが望ましい。以上の考察のもと、提案手法では首 (Neck) と腰 (Hip) を採用する。

ドローンカメラの画像 (解像度 960×720 (pixel)) 座標上に Neck および Hip の目標点をそれぞれ (480, 210), (480, 330) と定義しておき、人物追従時に撮影画像から推定されるこれら 2 つの関節点位置をこの目標点に近づけるように PID 制御を行う。制御入力値は、前後移動速度 (pitch), 上下移動速度 (throttle), 回転角速度 (yaw) の 3 つである。pitch 操作はドローンが人物との相対的距離を維持し、throttle・yaw 操作は人物を常に画角中央に収めることに寄与する。

3.3.2 人物を一定方向から観測する制御

前節の制御だけでは、追従対象者から一定距離を保つようなドローン制御はできるものの、その人物を一定の方向から観測することはできない。この方向に関する制御を実現するために、前節の制御に加えて、3 次元姿勢情報に基づく制御を行う。

提案システムでは、人物の左肩および右肩の 3 次元位置 $p_{LShoulder}, p_{RShoulder}$ から人物の向き情報を獲得する。具

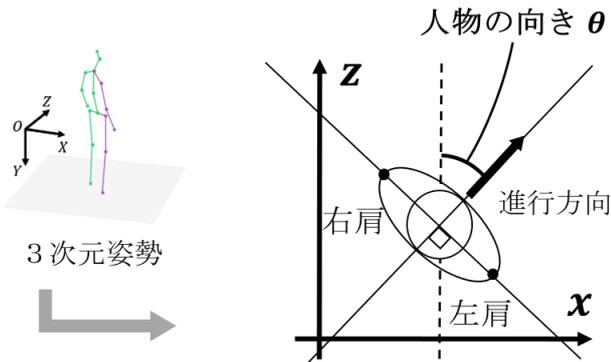


図 3 人物の向き の推定方法

体的には、図 3 に示すように、XZ 平面上で両肩の関節点を通る直線の法線方向 θ を人物の向きとする。なお、厳密には人物の向きは実世界の水平面上で定義されるべきで、そのためには 3 次元姿勢の胴体の方向などから水平面の方角を求めるべきである。しかし、2 次元画像からの 3 次元姿勢推定は本質的に不良設定問題であり、実際に Martinez らの手法によって得られる 3 次元の関節点においても特にカメラの光軸方向の位置精度には不安定性が認められる。そのため、そのような関節点から得られる実世界の水平方向も不安定になりやすい。上述の計算方法の場合、カメラ座標系の XZ 平面で水平面を近似していることに相当するが、ドローンは移動中には機体が傾くため多少誤差が大きくなるものの、基本的に水平を維持しながら飛行するため、XZ 平面を用いる方が平面方向の安定性は高いと考えられる。

制御においては、前節の制御則に加えて、上述の θ を目標値に近づけるようにドローンの左右移動速度 (Roll) を制御する。例えば、人物を背後を観測し続けたい場合は、 θ の目標値を 0 とする。

4. システム性能評価

開発した人物追従ドローンの性能評価を行う。本システムでは、姿勢情報の関節点に基づいて制御を行っている。そこで、その計測値の精度に関する実験から、推定姿勢情報の計測値の安定度を評価する。さらに、その計測値の推定精度のもとで、実際の環境での人物追従実験を行い、制御性能の有用性を評価する。

4.1 計測値の精度評価

4.1.1 距離に関する評価

提案システムでは、2 次元姿勢の首・腰の関節点間距離に基づいて人物との相対的距離を維持する制御を行っている。そこで、人物追従におけるこれらの関節点の安定度を調査する。被験者がトレッドミル上を歩行し、ドローンを固定した状態からその人物を撮影する。観測方向は人物に対して真後ろと真横とする。このとき、歩行動作における

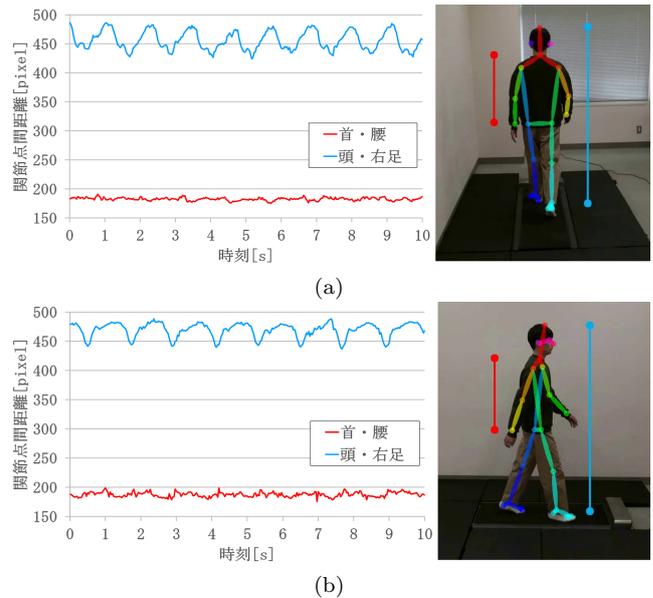


図 4 歩行動作中の関節点間距離の変化 (a) 真後ろ (b) 真横. 赤線は提案手法で使用される 2 次元姿勢の首・腰の関節点間距離、青線は頭・右足の関節点間距離を示す。

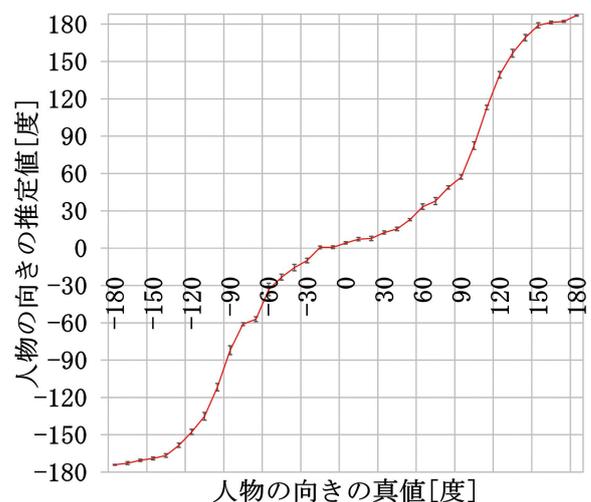


図 5 人物の推定向き。人物に対して背後から観測した場合を 0 度、左側の真横からの観測した場合を 90 度、真正面から観測した場合を 180(-180) 度とする。

首・腰間距離の変動を調査する。

各観測方向における歩行動作中の関節点間距離の変化を図 4 に示す。同図には、頭と右足の 2 点間距離の変動も示している。歩行動作における首と腰の長さは、歩行動作や姿勢推定誤差による小さな変化はあるが、位置の変動が比較的小さい。一方で、頭・右足間距離は、特に足元は画像中で大きく上下するため、変動が比較的大きい。したがって、提案システムで参照している首・腰の関節点は追従時の指標として適していることが確認できる。

4.1.2 方向に関する評価

本システムにおけるドローンに対する人物の向き の推定精度を評価する。評価実験では、人物の向きを -180 度～180 度の範囲を 10 度刻みに撮影した人物画像に対して向

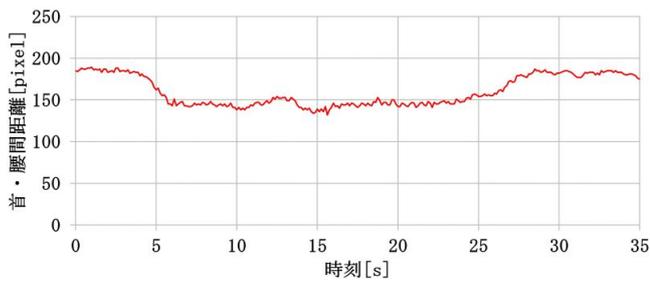


図 6 人の直進歩行動作における 2 次元姿勢の首・腰間距離の変化

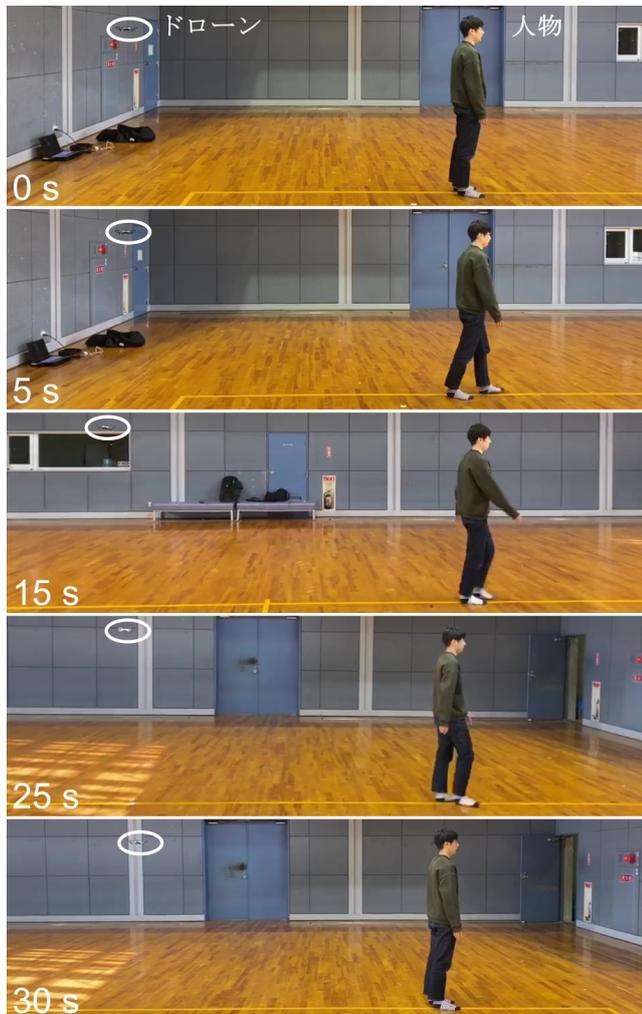


図 7 直進歩行時の人物とドローンの位置。人の停止状態と歩行のいずれの動作に対しても相対的な距離を維持しながら追従する。

きを推定する。なお、評価用画像は各方向 10 枚である。

各観測方向における人物の推定向きを図 5 に示す。0 度、 ± 90 度、 ± 180 度付近（すなわち、人物の真正面、真後ろ、真横方向）においては推定される向きはほぼ正確である一方、その他の向きでは一貫したずれ方をすることが確認できる。特に、0 度および ± 180 度付近ではグラフの傾きが水平に近くなる傾向が認められる。これは、これらの角度付近では多少人物向きが変わっても推定値が変化しない、すなわち向きの変化に鈍感であることを示している。

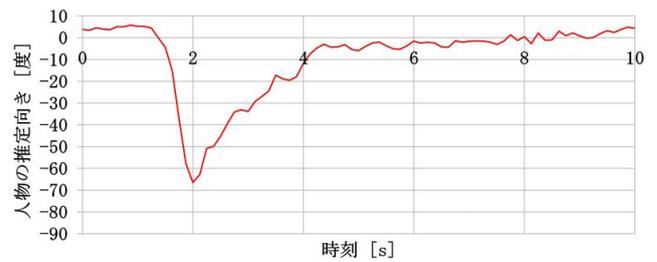


図 8 人の回転動作における人物の推定向きの変化



図 9 人の回転動作における人物の推定向きの変化（ドローン視点映像）

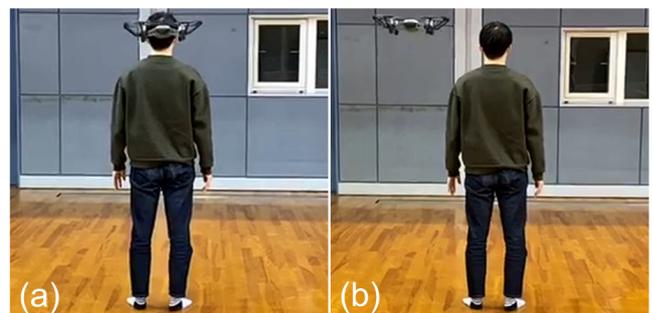


図 10 (a) 制御性能実験の成功例。人物の背後から正確に観測できている。(b) 制御性能実験の失敗例。人物の背後から少しずれた位置から観測する。

4.2 制御性能評価

4.2.1 距離に関する評価

人物の歩行動作に対して、ドローンがどの程度相対的な距離を保ちながら追従するのかを評価する。はじめに停止状態の人物が直線歩行を開始して再度停止する動作に対して、ドローンの制御性能を調査した。

歩行動作中における 2 次元姿勢の首・腰関節間距離の変化を図 6、追従対象人物とドローンの位置の変化を図 7 に示す。追従対象者が歩行動作を開始すると、ドローンはその運動変化に対して多少遅延するが、安定した追従を示した。人物とドローンの距離は歩行動作中 (5 秒～25 秒) にやや大きくなっているものの、人物の停止状態と歩行動作

のいずれの動作においても常に一定範囲内の観測距離を維持しながら追従することが確認できた。

4.2.2 方向に関する評価

人物の回転動作に対するドローンの制御性能を測るため、停止状態の人物がその場で90度回転した際のドローンの挙動を調査した。

人物の回転動作における推定向きの変化を図8、ドローンから撮影される人物の向きの変化を図9に示す。追従対象者が回転動作を開始すると、ドローンはその運動変化に対して多少の遅延はあるが、人物の位置を中心に半円を描きながら再びその人物の背後を観測する位置に適切に移動する様子が確認された(図10(a))。なお、ドローンが人物の回転動作に対して背後とは少しずれた位置で停止状態になり、その場で観測を継続する場合が確認された(図10(b))。これは、4.1.2で述べたように3次元姿勢推定が0度付近では向きの変化に鈍感であることに起因するものであると考えられる。

5. おわりに

本研究では、広域移動人物の映像解析を想定した、人物を一定距離・方向から撮影し続ける自動追従ドローンの開発を行った。本システムでは、ドローンカメラの撮影映像に対して人物の2次元・3次元姿勢を推定し、距離に関する制御では、2次元姿勢のうち、首と腰の関節点をドローン画像内の目標点に一致するよう制御し、また、方向に関する制御では、推定3次元姿勢のうち、両肩の関節点から人物の向きを算出し、その値を目標値に近づけるような制御する。実際に直進歩行や向き変化を行う人物をドローンに自動追従させる実験を行い、その有用性を確認した。

今後の展望として、まずはシステムの制御性能の向上が挙げられる。現在のシステムでは、人物の向きを3次元姿勢情報のみから推定しているが、人物の真正面や真後ろから捉える場合に向き推定精度が低下することがあったことから、今後は人物の移動軌跡情報も活用し、より高精度な向き制御を目指す。また、最適な制御パラメータの決定も実用に向けた重要な課題の一つである。また本稿では、ドローンが人物の背後から観測するように実装したが、犯人・不審者追跡を目的とした場合、ドローンがその人物の顔を捉えるために、正面から撮影し続ける制御が必要となる。これには、Murakamiらが提案する頭部運動から注視方向を推定する手法[17],[18]等を適用することで、人物の進行方向の予測による将来の移動軌跡を考慮した制御を検討する。

謝辞 本研究はJSPS科研費JP18H03312の助成を受けたものです。

参考文献

- [1] S. Yu, D. Tan, T. Tan, "Modelling the Effect of View Angle Variation on Appearance-Based Gait Recognition," Proc. 7th Asian Conference on Computer Vision, Vol.1, pp.807-816, 2006.
- [2] S. Kiciroglu, H. Rhodin, S. Sinha, M. Salzmann, P. Fua, "ActiveMoCap: Optimized Drone Flight for Active Human Motion Capture," Computer Vision and Pattern Recognition, 2019.
- [3] Y. Shigeki, F. Okura, I. Mitsugami, K. Hayashi, Y. Yagi, "Directional Characteristics Evaluation of Appearance-Based Gait Recognition," IPSJ Transactions on Computer Vision and Applications, Vol.10, No.10, 2018.
- [4] MAVIC 2 PRO, <https://www.dji.com/jp/mavic-2/>
- [5] Skydio 2, <https://www.skydio.com/>
- [6] C. Bian, Z. Yang, T. Zhang, H. Xiong, "Pedestrian Tracking from an Unmanned Aerial Vehicle," IEEE International Conference on Signal Processing, 2016.
- [7] B. Risse, M. Mangan, L. D. Pero, B. Webb, "Visual Tracking of Small Animals in Cluttered Natural Environments Using a Freely Moving Camera," IEEE International Conference on Computer Vision, pp.2840-2849, 2017.
- [8] 高橋亮輔, 藤ノ木俊宏, 稲数幸祐, 小布施大志, 太田豊, 「世界初! 民間防犯用ドローン」, 電気学会誌, Vol.136, No.9, pp.623-626, 2016年.
- [9] 渡邊紀文, 木浦豊治, 有村勇紀, 糸田孝太, 大森隆司, 「ドローンを利用したサッカー選手間の意図共有過程の分析」, 第34回ファジィシステムシンポジウム, 2018年.
- [10] Z. Kalal, K. Mikolajczyk, J. Matas, "Tracking-learning-detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.34, No.7, pp.1409-1422, 2012.
- [11] J. Pestana, J. L. Sanchez-Lopez, S. Saripalli, P. Campoy, "Computer Vision Based General Object Following for GPS-denied Multirotor Unmanned Vehicles," American Control Conference, 2014.
- [12] D. D. Mel, K. A. Stol, J. A. D. Mills, B. R. Eastwood, "Vision-Based Object Path Following on a Quadcopter for GPS-Denied Environments," International Conference on Unmanned Aircraft Systems, 2017.
- [13] T. Nægeli, S. Oberholzer, S. Pliiss, J. Alonso-Mora, O. Hilliges, "Flycon: Real-time Environment-independent Multi-view Human Pose Estimation with Aerial Vehicles," ACM Transactions on Graphics, Vol. 37, No. 6, 2018.
- [14] Tello, <https://www.ryzerobotics.com/jp/tello>
- [15] Z. Cao, T. Simon, S.-E. Wei, Y. Sheikh, "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [16] J. Martinez, R. Hossain, J. Romero, and J. J. Little, "A simple yet effective baseline for 3d human pose estimation," IEEE Conference on Computer Vision, pp.2659-2668, 2017.
- [17] J. Murakami, I. Mitsugami, "Gaze from Head: Gaze Estimation without Observing Eye," The 5th Asian Conference on Pattern Recognition, 2019.
- [18] J. Murakami, I. Mitsugami, "VR-based Eye and Head Motion Collection for Modeling Their Coordination," IEEE 8th Global Conference on Consumer Electronics, 2019.