

顔検出と人領域の主成分分析を用いた視覚に基づく小型無人航空機の人追従手法の提案
日大生産工(院) ○野村 拓矢 日大生産工 岡 哲資

1 背景

近年, 小型無人航空機(以降 UAV)が測量, インフラ点検, 空撮, 農薬散布, 宅配, 無人監視, 人の撮影などの分野で活用されるようになってきた. それに伴い, UAV に関する研究も活発に行われている. UAV の活用分野の中でも, 無人監視や人の撮影の分野では, UAV に人を追従させる必要がある.

UAV に人を追従させる手法として, 人の持つ携帯電話やタブレット端末の GPS と, UAV 自身に搭載された GPS を利用する手法がある. この手法では, 人が GPS を所持していなければならないため, 不特定多数の人に使用することはできない. また, 人の向きは分からないため, 人の正面や背面・側面に移動しながら追従することはできない. さらに, 屋内で使用することはできない.

GPS を用いずに人を追従する Higuchi らの手法がある 1). Higuchi らは特定の色の服を人に着用させ, 色抽出と Particle Filter を用いることによって人を追従した. また, 人と UAV に電子コンパスを装着することにより, UAV に対する人の向きを求め, 人の周囲の任意の方向への移動を実現した. しかしこの手法では人に特定の色の服を着用させ, 電子コンパスを持たせる必要があるため, 不特定多数の人に使用することはできない.

Higuchi らは, 人に電子コンパスを装着せず, 特定の色の服を着用させて色抽出と Particle Filter によって人を追従し, 人の周囲を旋回することを試みた 2). この手法では, 人の胴体の地面からの高度の仮定値と, UAV の高度, UAV のピッチ角を用いて人と UAV の間の距離を推定している. しかし, これらの値にはノイズが含まれており, 距離の正確な推定も人の周囲の旋回もできない. また, 特定の色の服を着用させる必要があるため, 不特定多数の人に使用することはできない.

人に特定のデバイスを持たせずに画像のみを用いて人を追跡する手法がある 3). Lekkala らによるこの手法では, Tracking-Learning-Detection アルゴリズム 4)を用いて人矩形のトラッキングを行い, 人矩形面積と人矩形の座標を特徴として抽出している. 抽出した特徴を PD 制御し, 追跡を行っている. しかし, この手法では人領域を人が事前に指定する必要があるため, 不特定多数の人に用いることはできない. また, 人矩形の面積は人の正面と側面では異なり, 姿勢の変化にも影響を受けやすいため, 一定の距離を保ちながら人の周囲を旋回することは難しいと考えられる.

Imamura らは, 機械学習により人を含む矩形領域を検出し, その座標値と幅, カメラの画角を用いて三角測量法で人のカメラ座標内三次元座標を推定した 5). しかし, 矩形の座標値と幅は人の姿勢に影響を受けやすい. また, 人の正面と側面でそれらの値が変わるため, 一定の距離を保ちながら人の周囲を旋回することは難しいと考えられる.

以上を踏まえ, 本稿では新しい人追従手法を提案する. 提案手法では, 人領域点群の重心, UAV と人の間の距離に比例する人の写り方や姿勢の変化にロバストな特徴, 人の正面で最大値を取り側面で最小値を取る特徴, さらに顔の有無を画像から抽出し, 特徴を PD 制御と場合分けによる手法によって制御する. これにより, 常に人と一定の距離を保ちながら人を画像の中心にとらえるだけでなく, 人の正面, 背面, 側面への移動を試みる.

正面・背面・側面に移動しながら人を追従することには様々なメリットがある. 無人監視では, 怪しい人を追跡する場合, その人の顔をより撮影しやすくなる. 犯人の特定や, 犯罪の抑止の観点から, 怪しい人の顔を撮影することは重要である. UAV のジェスチャ操作や, ロボットペットとしての UAV など, ヒューマンロ

Vision-based human tracking method for an unmanned aerial vehicle using face detection and principal component analysis of human region

Takuya NOMURA and Tetsushi OKA

ボットインタラクションの場面では、UAV が自律的に人の正面に移動することで、よりスムーズで自然なインタラクションを可能とする。

2 人追従手法

提案人追従手法は特徴抽出処理と特徴制御処理から構成されている。図1に、人追従手法の処理とデータの流れを示す。

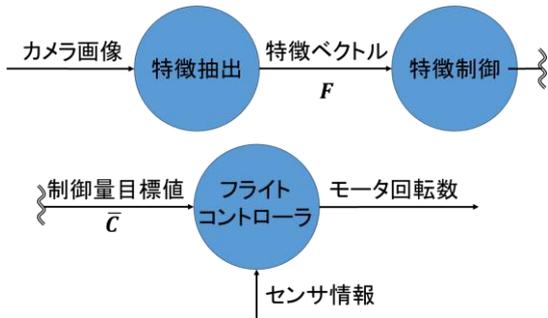


図1 人追従手法の処理とデータの流れ

特徴抽出処理では、特徴制御で用いる特徴ベクトル F を画像から抽出する。

特徴制御処理では、特徴ベクトル F を制御し、フライトコントローラに渡すための制御量の目標値ベクトル \bar{c} を計算する。フライトコントローラは、各種センサからの情報と \bar{c} を統合し、UAV の姿勢を安定させながら \bar{c} を満たすようなモーターの回転数を決定する。なお、コントローラは提案手法外であるため、本論文での説明は割愛する。

3 特徴抽出処理

特徴抽出処理では、画像から人領域点群を抽出し、画像と人領域点群を用いて特徴ベクトル F を計算する。図2に、人領域点群抽出の様子を示す。

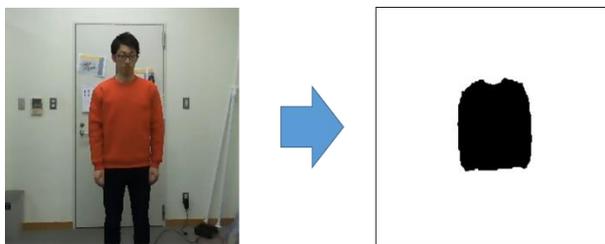


図2 人領域点群抽出

なお、画像から人領域点群を抽出する手法については提案手法外であるため、本論文での説明は割愛するが、レーザーレンジファインダーや深度カメラ、サーモグラフィーを用いる手法

や、人に特定の色の服を着用させ、色抽出を施すことにより抽出する手法が考えられる。

特徴ベクトル F は、式(1)に示すように5つの特徴量から構成されている。

$$F = (p_x, p_y, d, r, f) \quad (1)$$

3. 1 から 3. 3 では、それぞれの特徴量について説明する。

3. 1 $p_x \cdot p_y$

p_x, p_y は人領域点群の重心座標である。常に人を画像の中央に捉えるために p_x, p_y を用いる。式(2)に示す人領域点群の二次元座標から構成される行列 A の平均ベクトルの x 成分を p_x 、 y 成分を p_y とする。

$$A = \begin{pmatrix} x_1 & \cdots & x_n \\ y_1 & \cdots & y_n \end{pmatrix} \quad (2)$$

3. 2 $d \cdot r$

d は A の分散共分散行列の第1固有値である。また、 r は第1固有値と第2固有値の比率である。 d は、UAV が人に近づき、画像内の人領域が大きくなるほど値が大きくなる。そのため、人と UAV の間の距離を一定に保つために d を用いる。 r は、人を正面もしくは背面から見たときに最大値を取り、側面から見たときに最小値を取る。そのため、UAV を人の正面や背面、側面などに移動させるために r を用いる。

$d \cdot r$ の求め方を説明する。初めに、式(3)に示すように、 A の分散共分散行列 V を求める。

$$V = \begin{pmatrix} \frac{\sum_{i=1}^n (x_i - \mu_1)^2}{n-1} & \frac{\sum_{i=1}^n (x_i - \mu_1)(y_i - \mu_2)}{n-1} \\ \frac{\sum_{i=1}^n (y_i - \mu_2)(x_i - \mu_1)}{n-1} & \frac{\sum_{i=1}^n (y_i - \mu_2)^2}{n-1} \end{pmatrix} = \begin{pmatrix} S_{xx} & S_{xy} \\ S_{yx} & S_{yy} \end{pmatrix} \quad (3)$$

ここで、 $\mu_1 \cdot \mu_2$ はそれぞれ A の平均ベクトルの x 成分と y 成分である。次に、式(4)に示す V の固有方程式を解いて固有値を求め、 d とする。

$$\det \begin{pmatrix} S_{xx} - \lambda & S_{xy} \\ S_{yx} & S_{yy} - \lambda \end{pmatrix} = 0 \quad (4)$$

最後に、第1固有値と第2固有値の比率を計算し、 r とする。

3.3 f

f は顔の有無を示す二値の変数である。 f は、人の顔が画像に映っているとき1、そうでないとき0を取る。UAVが人の正面と背面どちらを飛行しているのかを推定するために f を用いる。人の顔が画像に映っているかどうかは、画像に対して顔検出を施すことにより推定する。

4 特徴制御処理

特徴制御処理では、 \mathbf{F} を特徴量の目標値 $\bar{\mathbf{F}}$ に近づけるための制御量の目標値ベクトル $\bar{\mathbf{C}}$ を求める。制御量ベクトル \mathbf{C} はUAVのヨー角速度 $\dot{\psi}$ 、上昇速度 \dot{z} 、ピッチ角 θ 、ロール角 ϕ から構成される。

4.1 $d \cdot p_x \cdot p_y$ 制御

UAVと人の間の距離を一定に保つため、 \bar{d} に近づけるように d を制御する。ここで、 \bar{d} は保ちたい距離によって任意の値に設定する。また、人を画面の中央に捉えるため、 $\bar{p}_x \cdot \bar{p}_y$ に近づけるように $p_x \cdot p_y$ を制御する。ここで、 $\bar{p}_x \cdot \bar{p}_y$ はそれぞれ画像の中央 x 座標と y 座標である。

$d \cdot p_x \cdot p_y$ を $\bar{d} \cdot \bar{p}_x \cdot \bar{p}_y$ に近づけるために、本研究ではPD制御を用いる。式(5)、(6)、(7)に、 $d \cdot p_x \cdot p_y$ を制御するためのPD制御式を示す。

$$\theta = K_d^P e_d + K_d^D \dot{e}_d \quad (5)$$

$$\dot{\psi} = K_{p_x}^P e_{p_x} + K_{p_x}^D \dot{e}_{p_x} \quad (6)$$

$$\dot{z} = K_{p_y}^P e_{p_y} + K_{p_y}^D \dot{e}_{p_y} \quad (7)$$

ここで、 $e_d \cdot e_{p_x} \cdot e_{p_y}$ は $d \cdot p_x \cdot p_y$ と $\bar{d} \cdot \bar{p}_x \cdot \bar{p}_y$ の偏差である。 $\dot{e}_d \cdot \dot{e}_{p_x} \cdot \dot{e}_{p_y}$ は $e_d \cdot e_{p_x} \cdot e_{p_y}$ の1階時間微分であり、 $K_d^P \cdot K_{p_x}^P \cdot K_{p_y}^P$ と $K_d^D \cdot K_{p_x}^D \cdot K_{p_y}^D$ はそれぞれの特徴量の比例ゲインと微分ゲインである。

4.2 $r \cdot f$ 制御

UAVを人の正面・背面・側面に移動させるため、 $\bar{r} \cdot \bar{f}$ に近づけるように $r \cdot f$ を制御する。ここで、 $\bar{r} \cdot \bar{f}$ はUAVを移動させる位置によって変化する。図3に、UAVの位置と $r \cdot f$ の値の関係を示す。

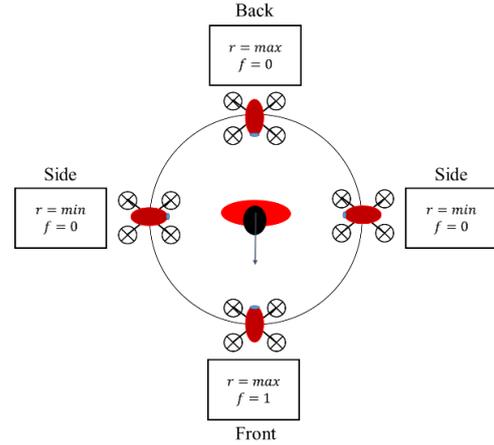


図3 UAVの各位置における $r \cdot f$ の値

図3に示すように、UAVが人の正面を飛行しているとき、 r は最大値を取り、 f は1になる。背面を飛行しているときは、 r が最大値を取り、 f は0になる。側面を飛行しているときは r が最小値を取り、 f は0になる。 $\bar{r} \cdot \bar{f}$ は、図3に示した各位置における $r \cdot f$ に設定する。ここで、 r の最大値、最小値は事前に測定する。

$r \cdot f$ の制御は、 r を \bar{r} に近づける r 制御と、正面・背面を判定する処理、人の反対側に回り込む処理を組み合わせで行う。また、UAVを人の正面・背面・側面に移動させるためには、常に人と一定の距離を保ち、人を画像の中央に捉え続けなければならない。そのため、 $d \cdot p_x \cdot p_y$ の制御を並行して行う。

4.3 r 制御

r 制御では、 r の増減を確認しながらロール角を左右に傾けることで、 r を最大化もしくは最小化する。人の正面・背面に移動する場合は r を最大化し、人の側面に移動する場合は r を最小化する。図4に、人の正面・背面に移動する際の r 制御の疑似コードを示す。

```

 $\phi_c = e_r \cdot K_r$ 
While
  If  $\dot{r} < 0$ 
    Negative++
  Else
    Negative--
  End if
  If Negative > 4
    Sign = -1
    Negative = 0
  Else
    Sign = 1
  End if
   $\phi_c = \text{sign} \times \text{sgn}((\phi_p)) \times e_r \times K_r$ 
   $\phi_p = \phi_c$ 
End while

```

図4 r 制御疑似コード

ここで、 ϕ_c , ϕ_p は現在のロール角と1フレーム前のロール角である。 ϕ_c が正の値を取った時 UAVは右に傾き、負の値を取った時左に傾く。 e_r は \bar{r} と r の偏差である。また、 K_r は値を調整するためのゲインである。 \dot{r} は r の1階時間微分であり、傾きを示す。Negativeは \dot{r} が負になった回数をカウントするための変数であり、Signはロール角の符号を格納するための変数である。Sgn関数は引数の符号が正の時1を返し、負の時-1を返す関数である。

初めに、ロール角を右に傾ける。次に、 r が連続で減少した時、ロール角の符号を反転させる。これを繰り返すことで、 r を増加させ、 \bar{r} に近づける。

人の側面に移動させる場合は、正面・背面に移動させる時とは逆に r が連続で増加した時にロール角の符号を反転させる。

4. 4 位置推定

$d \cdot p_x \cdot p_y \cdot r$ が $d \cdot \bar{p}_x \cdot \bar{p}_y \cdot \bar{r}$ に十分近づいたとき、UAVが目標の位置まで移動したとみなす。しかし、正面・背面は $d \cdot p_x \cdot p_y \cdot r$ だけでは区別できないため、 f を用いてUAVの位置を推定する。

R が十分大きく、 $f = 0$ の状態が一定時間続いた場合背面を飛行していると推定する。逆に、 $f = 1$ の状態が一定時間続いた場合は正面を飛行していると推定する。

4. 5 回り込み

正面・背面間の回り込みと側面間の回り込みは、 r が一定の値を満たすまでロール角を左に傾け続けることで行う。図5に、回り込みと r の関係を示す。

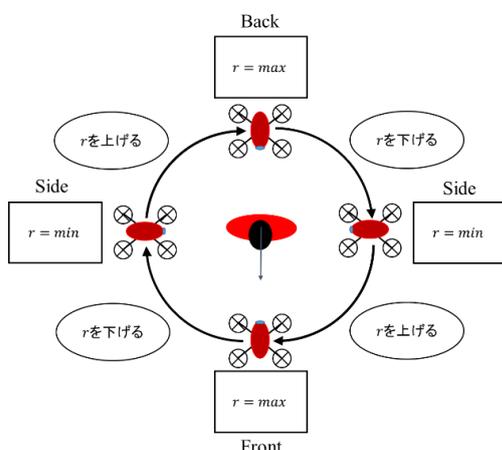


図5 回り込みと r の関係

図5に示すように、正面・背面間の回り込みは、 r が一度最小値まで下がってから最大値まで上がるのを確認するまでロール角を左に傾け続けることで行う。側面間の回り込みは、 r が一度最大値まで上がってから最小値まで下がるのを確認するまでロール角を左に傾け続けることで行う。

5 おわりに

本稿で提案した追跡手法では、画像から5つの特徴を抽出し、抽出した特徴をPD制御と場合分けによって制御することで人とUAVの距離を一定に保ち、人を画像の中央に捉えながら人の正面・背面・側面に移動する。今後は、提案手法を実装し、実機を用いて評価実験を行うことにより、提案手法の有効性、特性について考察し、既存手法との比較を行う。

「参考文献」

- 1) Keita Higuchi, Tetsuro Shimada, and Jun Rekimoto. "Flying sports assistant: external visual imagery representation for sports training." *Proceedings of the 2nd Augmented Human International Conference*. ACM, (2011).
- 2) Keita Higuchi, Yoshio Ishiguro, and Jun Rekimoto. "Flying eyes: free-space content creation using autonomous aerial vehicles." *CHI'11 Extended Abstracts on Human Factors in Computing Systems*. ACM, (2011) p. 561-570.
- 3) Kiran Kumar Lekkala, and Vinay Kumar Mittal. "Simultaneous aerial vehicle localization and Human tracking." *Region 10 Conference (TENCON), 2016 IEEE*. IEEE (2016) p. 379-383.
- 4) Zdenek Kalal, Krystian Mikolajczyk, and Jiri Matas. "Tracking-learning-detection." *IEEE transactions on pattern analysis and machine intelligence* vol. 34, no. 7 (2012) p1409-1422.
- 5) Yusuke Imamura, Shingo Okamoto, and Jae Hoon Lee. "Human Tracking by a Multi-rotor Drone Using HOG Features and Linear SVM on Images Captured by a Monocular Camera." *Proceedings of the International MultiConference of Engineers and Computer Scientists*. Vol. 1 (2016).