

PSO とガウス過程を用いた屋内ドローンの最適視点選択 Optimal viewpoint selection for indoor drones using PSO and Gaussian Processes

横松大政 関山浩介*

名城大学理工学部メカトロニクス工学科

Department of Mechatronics Engineering, Science and Technology, Meijo University

Abstract: 本研究では、屋内ドローンを用いて構図バランスの整った写真撮影における三次元空間での最適視点選択手法を構築した。屋内ドローンの三次元空間探索手法には、ガウス過程と PSO 探索を組み合わせることで、短い時間の中で効率的にヒューリスティックな最適視点の予測を可能にした。また、写真の構図バランスの評価は、被写体の配置と理想配置をそれぞれ二次元ガウス分布で表し、その類似度を KL 情報量から定量化する構図評価手法を用いた。

1 はじめに

近年、ドローンの小型化、低価格化によるホビー分野で撮影用ドローンが注目されている。これまで、空撮映像はコストをかけて、ヘリコプターを手配して撮影されてきたが、ドローンの登場によって撮影業界では変革が起こり、高品質な空撮映像を安価に撮影することが可能になった [1]。しかし、美的感性に基づいた視点選択や撮影のタイミングは、熟練した操縦者によって決める必要がある。

そこで本研究では、屋内ドローンを用いて構図バランスの整った写真撮影における三次元空間での最適視点選択手法を構築した。屋内ドローンの三次元空間探索手法には、ガウス過程と PSO 探索を組み合わせることで、短い時間の中で効率的にヒューリスティックな最適視点の予測を可能にした。撮影写真の構図評価は被写体の画面内での配置状況を二次元ガウス分布で表現し、あらかじめ設定した被写体の理想配置を表す二次元ガウス分布との類似度を Kullback-Leibler 情報量 (KL 情報量) から定量化する手法を提案した [2]。

2 最適視点探索の手法

本システムの構成を図 1 に示す。本システムは、「自己位置推定 (Position Estimation) モジュール」、「構図評価 (Composition Evaluation) モジュール」、「視点選択 (Viewpoint Selection) モジュール」、「飛行制御 (Flight Control) モジュール」、「画像処理 (Image Processing) モジュール」の 5 つのモジュールから構成されている。

本システムの流れを大まかに説明する。画像処理モジュールではドローンのカメラから映像を取得する。被

写体の検出には YOLO (You Only Look Once) を用いる。画像処理モジュールで取得したカメラ情報をもとに自己位置推定、構図評価を行う。自己位置推定は Visual SLAM を用いる。構図評価は、YOLO から検出した被写体のカメラ画面内での座標と大きさを取得し、その情報から構図評価を行う。視点選択モジュールでは、構図評価モジュールで得られた構図評価値と自己位置推定モジュールで作られた SLAM マップによるガウス過程を行い、構図マップを作成する。この時に、PSO をガウス過程と組み合わせることで探索の多様性を維持している。作成した構図マップに基づいて経路計画をし、その情報をもとに飛行制御を行い、最適視点で撮影を行う。被写体検出後は、対象となる被写体をドローンカメラ画面からフレームアウトしないために視覚フィードバックを行なっている。

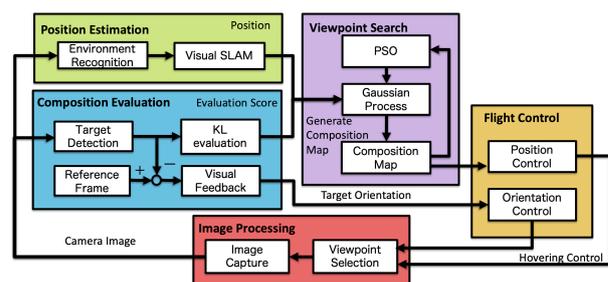


図 1: 自律撮影のシステムアーキテクチャ

3 ORB-SLAM を用いた屋内ドローンの自己位置推定

3.1 SLAM の選定

屋内環境では GPS による自己位置推定が不安定なため、「非 GPS 環境下での自己位置推定」が必要であ

*連絡先: 名城大学理工学部メカトロニクス工学科
〒468-8502 名古屋市中天白区塩釜口 1 丁目 501
E-mail: sekiyama@meijo-u.ac.jp

る [3]. そこで, SLAM(Simultaneous Localization and Map-*ping*) を用いる. SLAM とは自己位置推定と環境の地図生成を同時に行うアルゴリズムのことである. 本研究では, ドローンに内蔵されているカメラで自己位置推定が可能な Visual SLAM を用いた. また, Visual SLAM の中でも ORB SLAM[4] を用いる.

3.2 ORB-SLAM の座標変換

しかし, ここで問題点として, ORB SLAM で推定される位置・姿勢がローカル座標系であるため床からの高さが分からない. この問題に対し, 本研究ではカメラ視点を原点としたローカル座標系から床面を原点としたグローバル座標系に変換するために, 補助的に AR マーカを用いる.

図 2 に実験時の AR マーカの位置・座標系を示す.

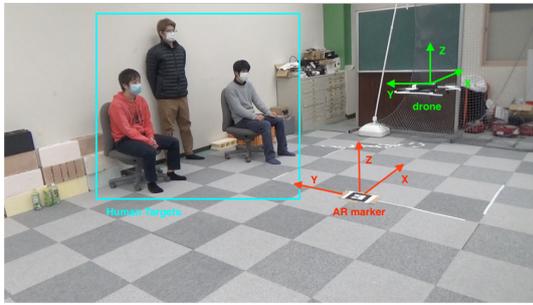


図 2: AR マーカの位置・座標系

カメラ視点を原点としたローカル座標系 ${}^{local}\mathbf{p} = [X_L, Y_L, Z_L]^T$ から, グローバル座標系 ${}^{global}\mathbf{p} = [X_W, Y_W, Z_W]^T$ への変換を式 (1) に示す.

$$\begin{bmatrix} X_W \\ Y_W \\ Z_W \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X_L \\ Y_L \\ Z_L \\ 1 \end{bmatrix} \quad (1)$$

これを簡略化すると,

$${}^{global}\mathbf{p} = {}^{global}_{local}\mathbf{T} {}^{local}\mathbf{p} \quad (2)$$

と表せる. ${}^{global}_{local}\mathbf{T}$ は回転行列 \mathbf{r} と並進ベクトル \mathbf{t} によるローカル座標系からグローバル座標系への同次変換行列を表している. 以下のような座標変換を逐次更新することでグローバル座標系での正しい位置・姿勢を用いてドローン制御を行なっている.

4 KL 情報量を用いた構図評価

ドローンカメラ内の被写体を画像のピクセル座標における 2次元ガウス分布として表現する. そして, いくつかの被写体の配置は, 2次元混合ガウス分布として表

現することができる. この 2次元混合ガウス分布と, あらかじめ設定した被写体の理想配置の 2次元混合ガウス分布との類似度を, Kullback-Leibler divergence(KL 情報量) から定量化し評価する. 本研究では, この構図評価手法に基づいて自律撮影を行う.

被写体の中心座標 $\boldsymbol{\mu} = (x, y)^T$ は x 軸と y 軸の平均であり, 幅と高さは分散共分散行列 $\boldsymbol{\Sigma}$ によって表現する. 画面内の任意位置 $\mathbf{a} = (x, y)^T$ での被写体 n の存在確率密度は次式である.

$$N(\mathbf{a}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{2\pi|\boldsymbol{\Sigma}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{a}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{a}-\boldsymbol{\mu})} \quad (3)$$

被写体が複数存在する場合, 被写体の配置状況は混合ガウス分布で表現する. 被写体 k の重みは g_k とし, 確率密度関数が $\sum_{k=1}^K g_k N(\mathbf{a}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ というような式で表せる. ここで, $\sum_{k=1}^K g_k = 1, 0 < g_k < 1, \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_k$ である.

配置評価は画面内の配置分布 $P(\mathbf{a}; \mathbf{X}, \boldsymbol{\Sigma})$ と参照構図分布 $Q(\mathbf{a}; \mathbf{X}', \boldsymbol{\Sigma}')$ の KL 情報量から 2つの確率分布の類似度を定量化し評価する.

$$D^c = \iint_H Q(\mathbf{a}; \mathbf{X}', \boldsymbol{\Sigma}') \log \frac{P(\mathbf{a}; \mathbf{X}, \boldsymbol{\Sigma})}{Q(\mathbf{a}; \mathbf{X}', \boldsymbol{\Sigma}')} dH \quad (4)$$

ここで H は写真のピクセル画面の範囲である. D^c は分離度を表すため, 小さければ評価値が高い.

5 最適視点探索

5.1 ガウス過程による評価値予測

PSO は基本分散探索であるが, PSO のメリットである広域的な探索が失われてしまう. そこで, ガウス過程と組み合わせることでこのメリットを補う. ガウス過程とは回帰問題の一種であり, 予測のあいまいさを表現できることが特徴として挙げられる. 本システムの自律撮影においては, PSO によって決められた計測点での位置と評価値からガウス過程を行い, 構図マップを作成している. ガウス過程の入力はドローンの位置 $I = {}^{global}\mathbf{p} = [X_W, Y_W, Z_W]^T$ とその時の構図評価値 $J = D^c$ である. 出力は評価値の期待値と分散とその時の位置である. 本システムではカーネルは RBF カーネル (動径基底関数) 式 (5) を用いた.

$$k(I, J) = \exp(-\beta \|I - J\|^2) \quad (5)$$

評価値の期待値 m と分散 σ は以下のように求めることができる.

$$m(x_{N+1}) = \mathbf{k}^T C_N^{-1} \mathbf{t} \quad (6)$$

$$\sigma^2(x_{N+1}) = \mathbf{k}^T C_N^{-1} \mathbf{k} \quad (7)$$

ここで, C_N は共分散行列である.

5.2 PSO によるランダム探索

実際の写真撮影の際には、撮影プロセスに制約があり、探索空間には局所解が存在するため、短い時間で探索空間内を広く探索し大域的な視点選択手法の提案が必要である。探索領域内を効率的かつヒューリスティックな予測をすることで、大域内で収束時間の早い視点選択を実現する。そのために PSO (粒子群最適化法) を用いる。探索の対象となる目的関数が与えられた時、複数の粒子が互いに情報を共有しながら最適解を求めて探索空間内を動き回る。複数の粒子の更新式としては、以下のようなものが基本である [5]。

$$v_j^i(n+1) = wv_j^i(n) + \rho_1(p_j^i - x_j^i(n)) + \rho_2(l_j^i - x_j^i(n)) \quad (8)$$

$$x_j^i(n+1) = x_j^i(n) + v_j^i(n) \quad (9)$$

ただし、 $i = 1 \sim N, j = 1 \sim M, N$ は粒子の個数、 M は次元である。 w, ρ_1, ρ_2 はパラメータであり、適当な乱数によって与えられる場合が多い。 $\mathbf{p}^i \equiv (p_1^i, \dots, p_N^i)$ をパーソナルベストと言い、 i 番目の粒子の現在までの最良値をあたる粒子位置を指す。また、 $\mathbf{l}^i \equiv (l_1^i, \dots, l_N^i)$ をローカルベストと言い、 i 番目の粒子の近傍のパーソナルベストの最良値を指す。

5.3 構図マップによる PSO の更新

PSO の次の候補地点を決めるために、ガウス過程による KL 情報量の予測分布の期待値 m 、標準偏差 σ を用いて、構図評価関数 M を作成する。 $0 < M < 1$ で正規化するためにシグモイド関数を用いる。構図評価関数 M は、

$$E_{KL} = \frac{m}{\sigma} \quad (10)$$

$$M = \frac{1}{1 + \exp\{\alpha \times (E_{KL} - \beta)\}} \quad (11)$$

で表す。ここで、 α, β はシグモイド関数の係数であり、本研究では、 $\alpha = 0.1, \beta = 10$ とする。評価関数 M は、大きいほど KL 情報量の評価値の予測が小さく、分散が小さいことを表す。そのため、良い構図が得られることが期待できる点を表すことができる。実際に実験した際の環境とガウス過程により作成された 3 次元構図マップを図 3, 4 に示す。

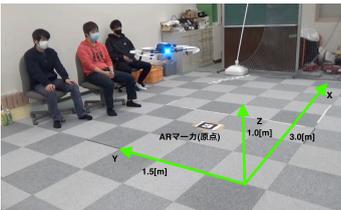


図 3: 実験環境

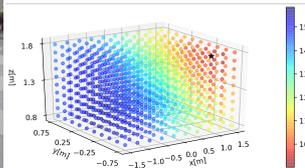


図 4: 3 次元構図マップ

ドローンの探索範囲 ($x: -1.5 \sim 1.5$ [m], $y: -0.75 \sim 0.75$ [m], $z: 0.8 \sim 1.8$ [m]) の空間を $50 \times 50 \times 50$ 個の点 (図 4 は $10 \times 10 \times 10$ で簡略化) に分割し、その位置の評価値をそれぞれ色で表現する。色は青～赤に変化していき、赤色が強いほど評価値が高い地点を表現する。3 次元構図マップだと、探索空間の構図評価全体が見にくいので、これを構図評価関数 M を用いて XY 平面に射影した 2 次元構図マップを作成する。2 次元構図マップを図 5 に示す。

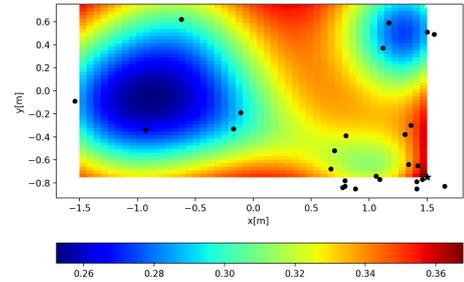


図 5: 2 次元構図マップ

式 (11) の評価関数 M によって作成された構図マップの最大値 M_{max} 付近に、次の位置の更新をする。期待値の大きさに応じてランダム探索する範囲を調整する。ランダム探索する範囲の大きさ B は次の式となる。

$$B = \frac{1}{1 + \exp\{\alpha \times (M_{max} - \beta)\}} \quad (12)$$

ばらつき B は $0 \sim 1$ の値をとる。評価マップの最大値をとる位置を中心として半径 B の円を考え、その中からランダムに次の候補地点を選択する。評価値が最大となる地点ではなくその周辺を探索することで、局所解に陥らない PSO 探索が可能になる。 M_{max} が大きい、つまり良い評価値が得られる可能性が高くなるほど、ばらつき B が小さくなるように設計した。このようにして、PSO とガウス過程を同時に行い、お互いに強化していくことで、より短時間で効率的に最適視点を探索する。

6 最適視点選択手法の実機実験

6.1 実験設定

提案したガウス過程と PSO による 3 次元視点選択手法が有効であることを実機実験を通じて検証する。実験用のドローンは parrot 社の BEBOP2 を用いる。最適視点とする参照構図は、画面を縦に 4 等分、横に半分にした時の交点に被写体の中心が配置する構図 (図 6) とする。実験環境、SLAM の座標系と原点を図 3 に示す。実験の被写体配置は、3 人が静的に座っている配置とする。本実験では、探索範囲を SLAM 座標系の原点から ($x: -1.5 \sim 1.5$ m, $y: -0.75 \sim 0.75$ m, $z: 0.8 \sim 1.8$ m)

と設定し、撮影までのプロセスの時間制約を 100[s] とする。

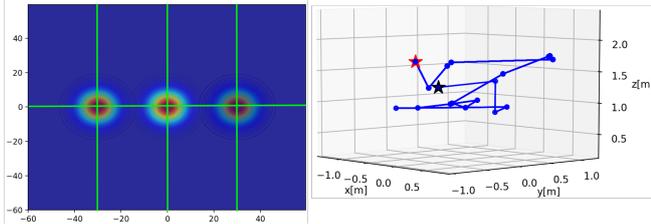
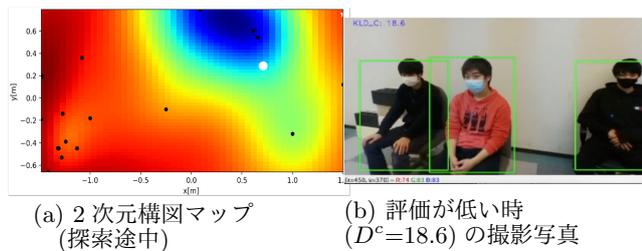


図 6: 参照構図

図 7: ドローン軌跡

6.2 実験結果・考察

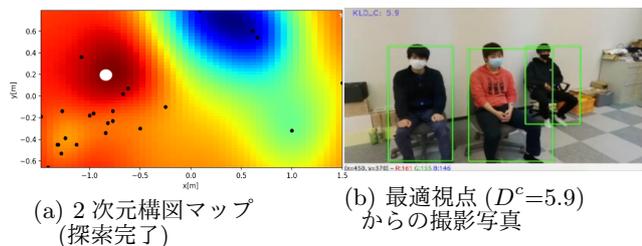
図 7 はドローンの撮影中の探索点の軌跡である。3次元空間を局所解に陥らずランダムに探索していることがわかる。図 8(a) は三次元の構図評価を xy 平面に落とし込んだ構図マップ (探索途中) である。赤色が濃いほど構図評価が高く、青色が濃いほど構図評価が低いことを表す。また、白い丸は図 8(b), 9(b) の写真が撮影された時のドローンの位置を示す。図 8(b) は探索途中の評価値が低い時の写真を示す。図 9(a) は探索完了時の構図マップを示す。図 9(b) は最適視点からの撮影写真を示す。表 1 に撮影時のドローンの位置と KL 値をまとめたものを示す。図 8(b) より図 9(b) の方が参照構図に近い写真となっているので、KL 値が良い結果となった。以上の実験結果からドローンが三次元視点探索をし、作成した構図マップの最適位置に移動し撮影できたことがわかる。



(a) 2次元構図マップ (探索途中)

(b) 評価が低い時 ($D^c=18.6$) の撮影写真

図 8: 探索開始 40[s] の実験写真



(a) 2次元構図マップ (探索完了)

(b) 最適視点 ($D^c=5.9$) からの撮影写真

図 9: 探索完了 100[s] の実験写真

表 1: 撮影写真時のドローンの位置と KL 値

	時間 [s]	ドローンの位置 [m]	評価値
探索途中	40[s]	$(x,y,z) = (0.76,0.38,0.94)$	18.6
最適視点	100[s] 以降	$(x,y,z) = (-0.81,0.21,1.14)$	5.9

7 まとめと今後の展望

本実験では、屋内ドローンを用いて構図バランスの整った写真撮影における三次元探索空間での最適視点選択手法を構築した。ガウス過程と PSO を組み合わせることで、三次元空間という広い探索範囲を 1 分間という制約の中で視点選択を行うことを可能にした。今後の展望として、ドローンが最適視点を探索中に被写体の人数が増えたり配置が変化した場合 (動的被写体) に対応できる視点選択手法を構築する。また、SLAM マップをドローン自らが探索して作成したり、被写体やドローンの自己位置がロストした場合などエラーへの対策を行なっていく。

謝辞

本研究の一部は JSPS 科研費 16H02880 の助成を受けたものである。

参考文献

- [1] Kai Lan, Kosuke Sekiyama, Autonomous robot photographer with KL divergence optimization of image composition and human facial direction. Robotics and Autonomous Systems, 2019, Vol.111, pp. 132-144.
- [2] 野波健蔵, 「ドローン技術の現状と課題およびビジネス最前線」. 情報管理, 2017, vol.59, no.11, pp.755-763.
- [3] R. Murartal and P. Yoonseok et al. ORB-SLAM: a Versatile and Accurate Monocular SLAM System. IEEE Transactions on Robotics, 2015, Vol.31, No.5, pp.1147-1163.
- [4] 斎藤 利通, 「粒子群最適化と非線形システム」 IEICE Fundamentals Review, 2011, Vol.5, No.2, pp.155-161.