

Computer Vision 第三回レポート

62214327 情報工学科 4 年 中村颯

1 目的

CV2 などのライブラリを使用せずに、各線形変換 (Translation, Similarity, Affine) を自身で実装することで各線形変換の数学的構造とパラメータ推定手順を理解すること。また特徴点抽出と画像合成 (ステッチング) の流れを実装を経験することで理解すること。

2 使用画像

使用した写真は以下の 4 枚である。

それぞれ iPhone11 の外カメラを使用して撮影した。風景の撮影時にカメラの回転や拡大を使用することで、ステッチングにおける各線形変換の違いが出力から読み取れるようにした。



3 方法

Python コードで実装した．使用した特徴点抽出・マッチングライブラリは Akaze である．各線形変換はそれぞれ以下のように考えて実装した．

3.1 Translation

変換は平行移動のみとなるため，課題参考資料のヤコビアン の 2 次元変換行列を参考にしながら，以下のように行列 T を定義した．

$$T = \begin{bmatrix} 1 & 0 & tx \\ 0 & 1 & ty \\ 0 & 0 & 1 \end{bmatrix}$$

そしてこの変換を適用した．

3.2 Similarity

課題参考資料のヤコビアンの二次変換行列を参考にしながら，以下のような変換を考えた．

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} tx \\ ty \end{bmatrix}$$

そして a, b, tx, ty について最小二乗法 (Numpy.linalg.lstsq を使用) を用いて推定した．

3.3 Affine

課題参考を参考にしながら，以下のような変換式を考えた．

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 + a_{00} & a_{01} \\ a_{10} & 1 + a_{11} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} tx \\ ty \end{bmatrix}$$

そして，前項と同様に $a_{00}, a_{01}, a_{10}, a_{11}, tx, ty$ について Numpy を使用した最小二乗法で推定した．

そして，各画像の四隅を累積ホモグラフィで変換して得られる，全画像合成後の座標範囲を利用してキャンバスサイズを決定した．

最後に cv2.warpPerspective を使用して全画像を一つのキャンバスにワーピングし，各画素の RGB 値を求めた．最終的に画素の合成回数に基づいて平均を取ることで混合画像を得た．

4 結果

4.1 各画像間のマッチング結果

各画像間のマッチング結果は以下のようになり，視覚的にも特徴量抽出・マッチングに成功していることが読み取れる．しかし，特徴量が屋根や木という特定の箇所に依存している点が見受けられる．



各線形変換を利用したスティッチングの結果が以下である.

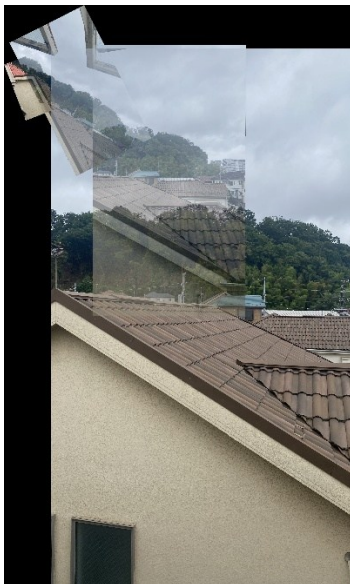
4.2 Translation

平行移動のみの推定であるため, 画像の回転やスケーリングによるずれが読み取れる.



4.3 Similarity

回転やスケーリングを考慮した線形変換のため, そのずれを改善しようとしている点は読み取れるが, やはりスティッチングに失敗していることが読み取れる.



4.4 Affine

Similarity よりもさらに精密なステッチングを行うことが出来た。これまでより最も自由度が大きいことから拡大や画像の回転にも柔軟に対応することが出来たことが出力から読み取れる。



5 議論

変換の自由度（推定する変数）が多くなるほど、柔軟に対応することができ、画像ステッチングの精度は向上するが、特徴量マッチングにおける誤差や風景の変化などの違いに対して頑健でなくなると推察される。以下で不整合を生み出す要因について挙げ、それぞれについて対策を述べる。

- 撮影位置の変化
シーン深度などに由来する視差により、歪みが生じてしまう要因となる。これの対策には、ユーザに動かないように指示することや、単一のホモグラフィでない、局所ワーピング（Zaragoza）を採用することである。
- 非平面画像
今回の変形では単一平面が仮定されているため、局所的に歪みが生じてしまう可能性がある。これも複数ホモグラフィを使用した局所ワーピングで
- レンズの歪み（特に広角レンズ）
レンズの歪みが生じる画像の歪みがステッチングに歪みを生じさせてしまう可能性がある。
複数のチェスの盤が映っている写真を使用して校正を行うことでカメラの特性を知り、調整することが出来る。
- マッチング誤差
低テクスチャ領域や照度変化により誤対応が生じ、推定行列が外れ値に引きずられる。また、特徴量が画像の特定箇所に集中してしまうと、その点を含まないス

ティッチングでは歪みが生じてしまう。これらの問題はグリッドベース特徴抽出 (Grid AKAZE) で画面全体に均質な特徴量を確保することで回避できる。

6 結論

本実験では、Translation・Similarity・Affine の各線形変換を自前で実装し、画像ステッチングに適用することで、変換の自由度が高まるほど位置合わせ精度が向上する一方、誤対応やシーン深度差への感度が増すことを確認した。具体的には、Translation では回転・スケーリングが補償できず大きなずれが残り、Similarity でこれらのある程度緩和できたものの、最終的に最も自由度の高い Affine が最良の合成結果を示した。しかし自由度の増加は外れ値に対する脆弱性を伴い、マッチング誤差や視差・レンズ歪みなどが直ちに不整合として顕在化することが議論から明らかとなった。

以上より、単一ホモグラフィによる大域的ワーピングには限界があり、今後は APAP などの局所ワーピング手法やカメラキャリブレーションを組み合わせることで、高自由度と頑健性を両立する必要がある。また、Grid-AKAZE 等による空間的に均質な特徴点抽出は、推定行列の安定化に寄与する有効な前処理と位置付けられる。これらの改善を施すことで、屋外複雑シーンに対しても高品質なパノラマ生成が可能になると結論付ける。

7 参考文献

Zaragoza, J., Chin, T.-J., Brown, M. S., & Suter, D. (2013).
As-Projective-As-Possible Image Stitching with Moving DLT.
In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2339–2346. https://www.cv-foundation.org/openaccess/content_cvpr_2013/papers/Zaragoza_As-Projective-As-Possible_Image_Stitching_2013_CVPR_paper.pdf