

1. はじめに

特定物体認識やパノラマ画像生成などで用いられるキーポイントマッチングは、見えの変化に対する頑健性が求められる。キーポイントマッチングの代表的な手法であるSIFTは、回転やスケール変化に対して頑健なマッチングが可能である。視点変化に対する頑健性が高いAffine-SIFT (ASIFT) [1] は、入力画像に対して複数回のアフィン変換を行い、視点合成画像を生成してSIFTにより特徴量記述を行う。しかし、アフィン変換は計算コストが高く、それをオンラインで考える視点変化分を行うため、処理に時間を要するという問題がある。この問題を解決するために、本研究では特徴記述子をフィルタで表現し、畳み込み処理で特徴量記述を行う。また、フィルタの視点合成を事前に行うことで、オンライン時のアフィン変換を行わずに、視点合成に基づいた特徴量記述を行う。しかし、視点合成により生成された全てのフィルタを畳み込み処理することは非効率である。そこで、本研究ではテンソル分解を用いて視点合成フィルタ群をコンパクト化する手法を提案する。

2. 従来手法

従来手法としてASIFTと特異値分解によるフィルタ群のコンパクト化について述べる。ASIFTは視点変化が起きた画像間のマッチングを行うために、入力画像に対してアフィン変換により複数の視点合成画像を生成し、生成した全ての画像に対してSIFT特徴量を記述する。これにより、視点合成に基づいた特徴量記述をする。しかし、実行時に計算コストの高いアフィン変換を行うため処理に時間がかかるという問題がある。

視点合成により生成したフィルタ群をコンパクトに表現する方法として、特異値分解(SVD)を用いる手法が提案されている[2]。図1に示すように、全てのフィルタをベクトル化して行をフィルタサイズ、列をフィルタの総数とする行列Lを生成してSVDによる次元削減を行う。しかし、SVDが十分に計算コストを削減をするには、1つの軸に対して大幅な特徴の削減を行う必要がある。そのため、精度が大幅に低下する場合がある。

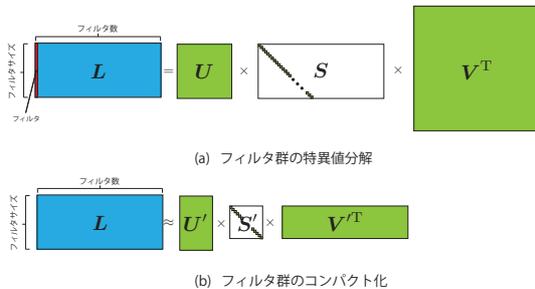


図1：特異値分解によるコンパクト化

3. 提案手法

提案手法は、特徴記述子を線形フィルタとして用いる。このフィルタをオフライン処理でアフィン変換することで、視点合成に基づいた特徴量記述を行う。また、視点合成フィルタ群をテンソル分解によりコンパクト化することで、様々な成分に対して次元削減を行うことが可能であり、精度を維持しながら大幅な計算コストの削減が期待できる。

3.1 特徴記述子の線形フィルタ表現

提案手法は、ORB記述子とGLOH記述子を線形フィルタの畳み込みで表現する。ORBは2点間の輝度差を特徴量とする。輝度差の計算は2点の片方に“+1”、もう一方に“-1”、それ以外に“0”を代入した線形フィルタの畳み込みで代用する。また、輝度差を比較する2点の中心に $\sigma = 1.2$ のガウシアンフィルタを乗算する。生成したフィルタ $\omega_{orb}^{d=1,2,\dots,256}$ の例を図2(a)に示す。GLOHはパッチ画像内を対数極座標に基づいて領域分割を行い、領域ごと

に16方向の勾配強度を抽出し特徴量とする。線形フィルタは各領域に“+1”、それ以外に“0”を代入したものとす。生成したフィルタ $\omega_{gloh}^{d=1,2,\dots,17}$ の例を図2(b)に示す。

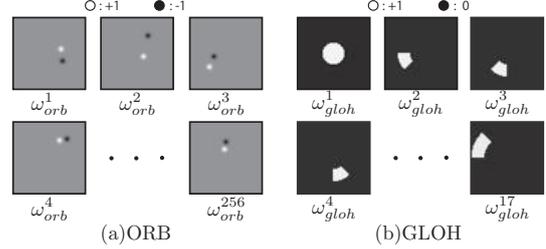


図2：線形フィルタ

3.2 視点変化に対応したキーポイントマッチング

提案手法は視点変化に対応した特徴量記述を行うために、フィルタに対してアフィン変換を行う。このとき、スケールパラメータとカメラ軸の回転パラメータはSIFTのスケールとオリエンテーションで代用するため、アフィン変換を行うホモグラフィ行列Hは式(1)で近似可能である。

$$H = \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \quad (1)$$

$t > 1$ は緯度に相当する傾きパラメータであり、 $\phi \in [0, \pi)$ は経度の相当する回転パラメータである。アフィン変換で生成された45視点分のフィルタを視点合成フィルタと呼び、ORBフィルタと視点合成フィルタを合わせて視点合成フィルタ群と呼ぶ。提案手法におけるキーポイントマッチングを図3に示す。パッチ画像AはORBフィルタを用いて特徴量記述を行い、パッチ画像Bは視点合成記述フィルタ群を用いて特徴量記述を行う。これにより、視点変化に対応した特徴量記述を行う。

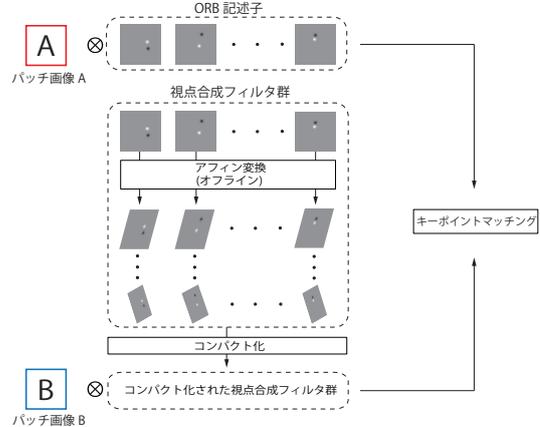


図3：提案手法におけるキーポイントマッチング

3.3 視点合成フィルタ群のテンソル分解によるコンパクト化

視点合成記述フィルタ群は256次元 \times 45視点=19,200枚となる。これを全て畳み込み処理することは非効率であるため、テンソル分解でコンパクト化を行う。テンソル $W \in \mathbb{R}^{M_1 \times M_2 \times M_3 \times M_4}$ を図4のように構築し、Tuckerモデルに分解する。分解手法にはthe higher-order singular value decomposition(HOSVD)を用いる。HOSVDは次元軸毎にテンソルを行列化(モード)してSVDによる次元削減を行う。

$$W \approx G \times_1 U_1 \times_2 U_2 \times_3 U_3 \times_4 U_4 \quad (2)$$

次に、行列 $U_1 \in \mathbb{R}^{M_1 \times m_1}$ とパッチ画像のベクトル $I \in \mathbb{R}^{1 \times M_1}$ は予め計算を行い、ベクトル $U_1 I$ を生成する。特徴量記述は式(3)で行う(図5)。

$$F = G \times_1 U_1 I \times_2 U_2 \times_3 U_3 \times_4 U_4 \quad (3)$$

このテンソル $F \in \mathbb{R}^{1 \times i_2 \times i_3 \times i_4}$ の各要素が特徴量となる。通常のSVDによる次元削減では、十分な計算コストを実現するために1つの成分に対して大幅な次元削減を行う必

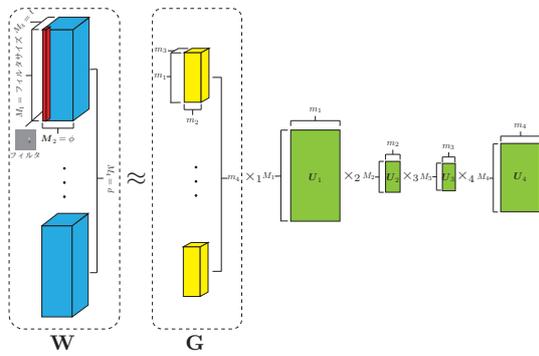


図 4: テンソルの構築と分解

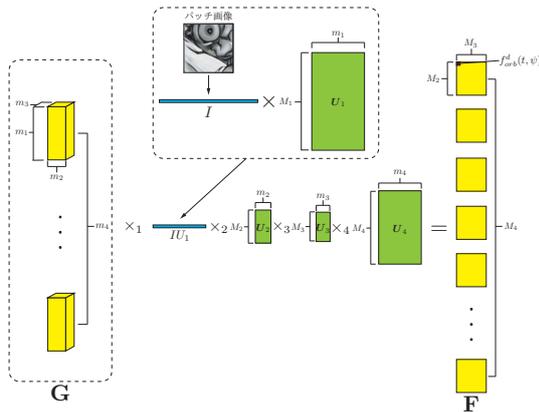


図 5: 提案手法による特徴記述

要がある。しかし、テンソル分解では4つのモードにSVDを行うことで、4つ成分それぞれで精度が低下しないように次元削減を行う。そのため、精度低下を抑えつつ計算コストの削減が実現できる。

4. 評価実験

提案手法の有効性を示すために以下の評価実験を行う。

1. モード毎に次元削減した精度の評価実験
2. 同等な圧縮率での精度比較
3. 大規模パッチ画像データセットでの評価実験

4.1 実験概要

評価実験1, 2にはRegion Detector Evaluation DatasetsのUnderground(射影変化), Grace(射影変化)を用いる。評価にはキーポイントマッチングのマッチング精度(正解点/対応点)と計算コストを用いる。計算コストは特徴量記述の処理における繰り返し処理回数である。

評価実験3にはHpatchesデータセットを用いる。Hpatchesはパッチ画像ペアのpositiveとnegativeの分類タスク, パッチ画像の対応付けタスク, 大規模パッチ画像群からクエリパッチ画像を検索するタスクの3つのタスクを評価できるデータセットである。

4.2 モード毎に次元削減した精度の評価実験

モード毎に次元削減の特異値の寄与率を変化させて、精度への影響を調査する。変化させるモード以外の寄与率は90%とする。モード毎に次元削減した際のマッチング精度を図6に示す。寄与率40%のときにパラメータtは精度が維持されているのに対して、サンプリングペアは精度が10%以上低下している。以上より、提案手法は寄与率をパラメータ毎に設定することで精度に影響が小さい次元を多く削減することができる。

4.3 同等な圧縮率での精度比較

特徴記述子ごとに同等な圧縮率(次元削減後の計算コスト/次元削減前の計算コスト)まで次元削減を行い精度を比較する。特徴記述子はORBとGLOHを用いる。表1にマッチング率の比較を示す。提案手法は評価実験1より各モードの精度が維持される寄与率を特徴記述子ごとに調査し用いる。SVDは提案手法と同等な圧縮率となる基底数を特徴記述子ごとに決定して用いる。また、アフィン部分空

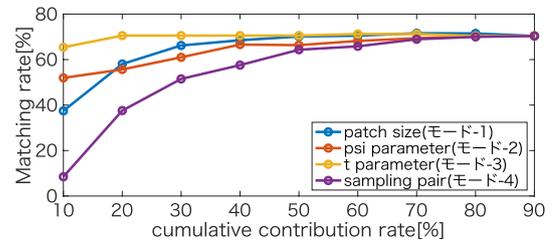


図 6: モード毎のマッチング精度 (ORB)

間特徴量であるASR[3]を用いた精度比較も行う。表1の結果より、同等の圧縮率の場合にマッチング率はORBが約27%精度向上した。また、GLOHはORBよりも高精度であるが、圧縮率は低いことがわかる。また、アフィン部分空間による高精度化も確認でき、提案手法のGLOH(ASR)はSVDと比較して、同等の圧縮率で約9%精度向上した。

表 1: 同等の圧縮率での比較

提案手法	指標	ORB		GLOH	
		-	ASR	-	ASR
提案手法	マッチング率 [%]	64.07	93.58	77.83	93.30
	圧縮率 [%]	17.9	17.9	20.8	20.8
SVD	マッチング率 [%]	47.65	72.27	77.85	84.51
	圧縮率 [%]	18.0	18.0	20.8	20.8

4.4 大規模パッチ画像データセットでの評価実験

Hpatchesデータセットを用いてパッチ画像の対応付けタスクでの比較実験を行う。また、Hpatchesデータセットではパッチ画像のオリエンテーションに $[-30^\circ, 30^\circ]$ の範囲で誤差が設定されているため、提案手法にカメラ軸回転パラメータを追加し、評価を行う。図7は対応付けタスクでの精度である。大きな視点変化が起きたHard, Toughの画像において最も高精度な対応付けが可能であることを確認できた。

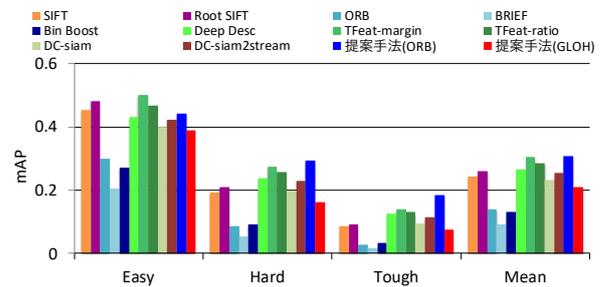


図 7: Hpatches による精度評価 (対応付けタスク)

5. おわりに

本研究では視点合成フィルタを用いたマッチングとテンソル分解によるコンパクト化を提案した。テンソル分解で適切な次元削減を行うことで、同等の圧縮率の時にSVDよりも約27%精度向上することが確認できた。今後はテンソル分解に適したフィルタの調査を行う。

参考文献

- [1] M. Jean-Michel, *et al.*, "ASIFT: A new framework for fully affine invariant image comparison", SIIMS, Vol. 2, No. 2, pp. 438-469, 2009.
- [2] T. Hasegawa, *et al.*, "Multiple-Hypothesis Affine Region Estimation With Anisotropic LoG Filters", ICCV, 2015.
- [3] Z. Wang, *et al.*, "Affine subspace representation for feature description", ECCV, 2014.

研究業績

- [1] 河合康平等, "テンソル分解を用いた視点合成フィルタ群のコンパクト化", 画像センシングシンポジウム, 2018.
- [2] K. Kawai, *et al.*, "Compactification of affine transformation filter using tensor decomposition", IEEE International Conference on Image Processing, 2018.