

1.はじめに

Deep Convolutional Neural Network(DCNN)をベースとした2-Channel Network[1]は、2つのパッチ画像間の類似度を出力するネットワークであり、輝度値等の変化に対して頑健である。しかし、テンプレートマッチングを利用する際、2-Channel Networkは畳み込み処理に時間を要するため、疎密探索を導入して計算コストを削減する必要がある。本研究では、2-Channel Networkによるテンプレートマッチングにおける探索の効率化として、疎密探索を用いない手法を提案する。

2.2-Channel Network

2-Channel Networkは、入力したペア画像の類似度を出力するネットワークである。ペア画像を1枚の2チャンネル画像として畳み込み層に入力することで、画像間の幾何変化に頑健な類似度計算が可能である。2-Channel Networkの学習に用いるポジティブペアは、テンプレートと対応した位置のパッチ画像の組とし、ネガティブペアはテンプレートと対応していない位置のパッチ画像の組である。式(1)のペア画像 $\{i|1, 2, \dots, N\}$ を入力した際の学習誤差 E_i には、ヒンジ誤差を用いる。

$$E_i = \max(0, 1 - O_i T_i) \quad (1)$$

ここで、 N はミニバッチのサイズ、 O_i は2-Channel Networkが output した類似度、 T_i は教師ラベルである。2-Channel Networkはヒンジ誤差を用いており、曖昧な類似度を出力したサンプルの誤差で重みを更新する。そのため、類似度が突発的に高くなるため、疎な探索においては、誤対応が発生するという問題がある。

3.提案手法

提案手法は2-Channel Networkを用いて、疎な探索後にパラボラフィッティングにより、対応点を算出する。従来の、ヒンジ誤差を用いた2-Channel Networkでは、疎な探索の時に突発的な出力となるため、パラボラフィッティングに向いていない。そこで、本研究では、パラボラフィッティングに適した類似度を学習できる誤差関数を設計する。

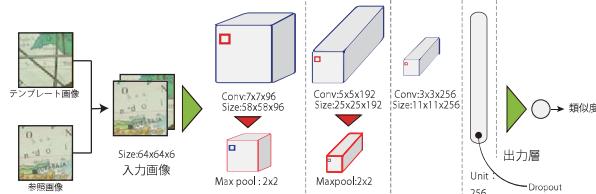


図1：提案手法の構造

3.1.2-Channel Networkの学習方法

2-Channel Networkによる類似度の教師ラベル T_{sim} は、ポジティブペアに1、ネガティブペアに-1を与える。提案手法では、ポジティブとネガティブの判定には、ヒンジ誤差を使用し、ポジティブペアの類似度の変動を表現するために、重なり率 T_{over} を教師信号として、L2ノルム2乗誤差により誤差を算出する。これらの処理から、式(2)のように誤差関数を設計する。

$$E_i = \begin{cases} (O_i - T_{over})^2 & (O_i > 0.5) \text{ and } (T_{sim} = 1) \\ \max(0, 1 - O_i T_{sim}) & \text{otherwise} \end{cases}$$

ここで、重なり率 T_{over} は0から1の範囲で値をとっているため、出力される類似度 O_i をシグモイド関数に通す。曖昧な出力をするペア画像とネガティブペアを学習する際は、従来のヒンジ誤差と同様の処理を行う。

3.2.疎な探索とパラボラフィッティング

対応点の探索時に、探索の移動量を大きくすると、正確な対応点の位置を推定するのが困難になる。そのため、パラボラフィッティングにより、ピーク周辺の類似度から対応点を推定することで、正確な対応点の位置を推定できる。

4.評価実験

提案手法の有効性を調査するために、NCCと従来の2-Channel Network、提案手法に対して、疎密探索と疎な探索をしたときの処理時間と対応点の誤差を評価する。学習と評価にはMiddlebury Stereo Dataset[2]を使用し、学習に19セット、評価には別13セット使用する。ここで、左右のカメラから得られる2つのシーン画像を1セットとする。使用的な画像サイズは 64×64 のカラー画像である。学習係数は1、学習の更新回数は50万回である。

4.1.実験結果

図2に、対応点との推定誤差と各手法の処理時間を示す。提案手法は、従来の2-Channel Networkより対応点の推定精度が優れており、NCCと同等の精度である。そして、NCCより処理速度は高速であり、従来の2-Channel Networkと同等の速度である。以上のことから提案手法は、効率的かつ高精度にテンプレートマッチングが可能であることがわかる。

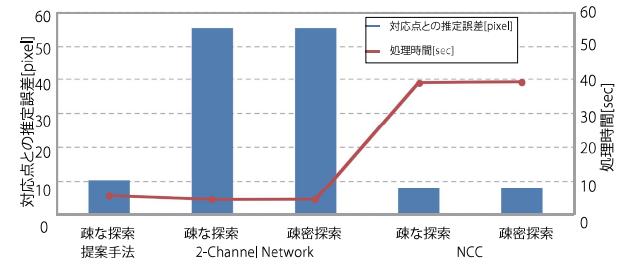


図2：各手法の比較

図3に、提案手法と従来の2-Channel Networkのテンプレートマッチング結果と、算出した対応点におけるパラボラフィッティングの誤差を示す。図3の赤枠と赤線が推定された対応位置、緑枠と緑線が真値の対応位置である。従来の2-Channel Networkでは、誤対応しているのに対して、提案手法は真値の対応位置と同じ位置が推定できている。これらの結果から、従来の2-Channel Network 提案手法は、類似度の変動が緩やかになり、パラボラフィッティングに適した類似度がoutputできていることが確認できる。

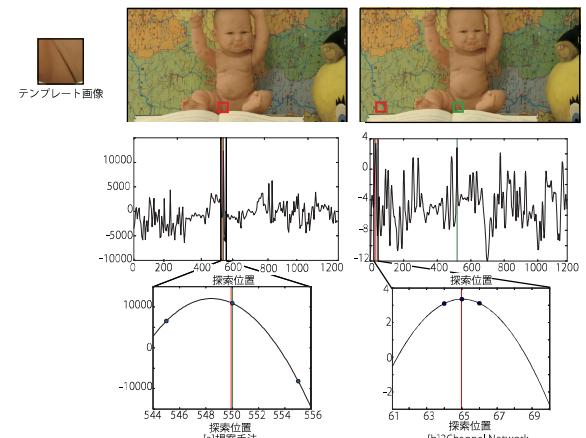


図3：各手法のペアマッチング例

(2)おわりに

本研究では、2-Channel Networkによるテンプレートマッチングで疎な探索を導入することにより、効率的な探索を実現した。今後は、ステレオマッチングにおけるサブピクセル推定への応用を検討する。

参考文献

- [1] S.Zagoruyko,*et al.* “Learning to Compare Image Patches via Convolutional Neural Networks”, CVPR, 2015.
- [2] D. Scharstein,*et al.* “Middlebury Stereo Datasets” <http://vision.middlebury.edu/stereo/data/>