4Q-01

Partial Convolution を用いたオプティカルフロー推定ネットワークを利用したピクセル伝播による動画補完

高橋和歩　長名優子

東京工科大学 コンピュータサイエンス学部

1 はじめに

欠損のある画像や動画をコンピュータによって補完する Inpainting という技術はこれまでに様々な手法が提案されてきている。近年では深層学習を利用して画像や動画の補完を行う研究が盛んに行われている。

画像に対する補完に関しては Partial Convolution を用いたネットワークによる画像補完 [1] が提案されている。この手法では、Partial Convolution(部分的畳み込み) を組み込んだ U-Net[2] 構造を持つネットワークを非欠損画素部から欠損画素部が推定されるように学習し、それを利用して入力された画像の欠損を補完している。入力として画素の一部が欠損した画像と欠損した画素の位置を示すバイナリマスクを入力し、出力として欠損部が補完された画像が出力されるようなものとなっている。

また、動画に対する補完に関してはオプティカルフロー推定ネットワークを利用したピクセル伝播による動画補完 [3] が提案されている。この手法はオプティカルフロー推定ネットワーク部と欠損補完部の二つで構成される。この手法ではまず、オプティカルフローマップ推定ネットワークを画素が欠損した動画からオプティカルフローマップを推定するように学習し、それを利用して妥当なオプティカルフローマップを推定している。その後、欠損補完部において推定されたオプティカルフローマップを画素を利用して伝播させることで動画の欠損部を補完している。また、ピクセル伝播によって欠損した画素すべてが補完できなかった場合には、残った欠損部を既存の画像補完技術によって補完することで動画の補完を行っている。入力として画素の一部が欠損した動画と欠損した画素の位置を示すバイナリマスクを与え、出力として欠損部が補完された動画が出力されるようなものとなっている。

本研究では、Partial Convolution を用いたオプティカルフロー推定ネットワークを利用したピクセル伝播による動画補完を提案する。提案手法ではまず、オプティカルフローの推定に Partial Convolution を用いたオプティカルフロー推定ネットワークを使用することでオプティカルフロー推定の精度を向上させる。また、オプティカルフローによるピクセル伝播で補完することができなかった画素の補完により高精度な画像補完ネットワークを使用することで精度を向上させる。

2 Partial Convolution (部分的畳み込み)

Partial Convolution は欠損画素を考慮した畳み込み処理である。入力として画素の一部が欠損したデータと欠損した画素の位置を示すバイナリマスクを与え、出力として畳み込み演算の実行結果とそれに対応するバイナリマスクが出力される。データにおける欠損部と非欠損部は欠損部を 0、非欠損部を 1 としたバイナリマスクによって表現される。畳み込み演算を実行する際にバイナリマスクを参照し、非欠損部である画素の値のみで演算を行い、出力された値を演算に用いた画素数で割ることで正規化する。また、畳み込み演算を実行する範囲の画素がすべて欠損部である場合は出力を 0 とする。加えて、バイナリマスクも出力に適するように更新を行う。畳み込み演算を実行する範囲の画素がすべて欠損部である場合のみ、その出力に対応するバイナリマスクは 0 に更新され、欠損部として扱われる。畳み込み演算を実行する範囲に一つでも非欠損部が存在している場合はその出力に対応するバイナリマスクは 1 に更新され、非欠損部として扱われる。この部分的畳み込みを重ねて実行することによって欠損部の画素は減少していくことになり最終的にはすべての画素が非欠損画素として扱われることになる。

3 Partial Convolution を用いたオプティカルフロー推定ネットワークを利用したピクセル伝播による動画補完

提案手法はオプティカルフロー推定ネットワークを利用したピクセル伝播による動画補完 [3](以下、従来手法と呼ぶ) を参考にして構成する。この手法はオプティカルフロー推定ネットワーク部と欠損補完部の二

Video Complementation by Optical Flow Estimation Network using Partial Convolution and Pixel Propageation

Kazuho Takahashi and Yuko Osana (Tokyo University of Technology, osana@stf.teu.ac.jp)

Copyright ©2022 Information Processing Society of Japan.
All Rights Reserved.

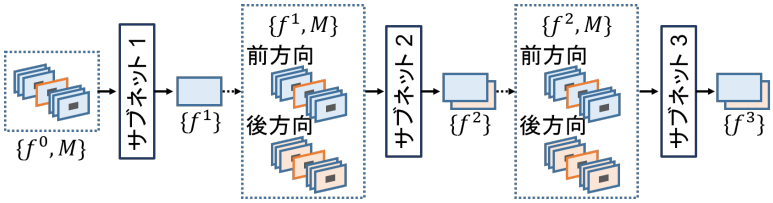

図 1: オプティカルフロー推定ネットワークの構造

つで構成される。提案手法ではまず、オプティカルフローマップ推定ネットワークを画素が欠損した動画からオプティカルフローマップを推定するように学習し、それを利用して妥当なオプティカルフローマップを推定する。提案手法では、オプティカルフロー推定ネットワークに Partial Convolution を用いるように変更する。オプティカルフローの推定を行った後、欠損補完部において推定されたオプティカルフローマップを利用して画素を伝播させることで動画欠損部を補完する。ピクセル伝播によって欠損した画素すべてが補完できなかった場合には、残った欠損部を画像補完技術によって補完する。提案手法では、この画像補完技術として Partial Convolution を用いたネットワークによる画像補完 [1] を用いるように変更する。提案手法では、従来手法 [3] と同様に、入力として画素の一部が欠損した動画と欠損した画素の位置を示すバイナリマスクを与えると、出力として欠損部が補完された動画が出力される。

3.1 オプテzィカルフロー推定ネットワーク

このネットワークでは画素の欠損した動画のオプティカルフローマップの推定を行う。入力として画素の一部が欠損した連続する画像からのオプティカルフローマップと欠損した画素の位置を示すバイナリマスクを与えられると、前方向のオプティカルフローマップと後方向のオプティカルフローマップが推定される。学習は、入力されたオプティカルフローマップとバイナリマスクから妥当なオプティカルフローマップを推定するように行う。

提案手法では、従来手法と同様に図 1 に示すようなネットワークをオプティカルフロー推定ネットワークとして用いる。図 1 における f はオプティカルフローマップを示しており、入力されたオプティカルフローマップを f^0、サブネットワーク 1 によって推定されたオプティカルフローマップを f^1 というように表している。また、M はオプティカルフローマップに対応するバイナリマスクである。このネットワークは類似する 3 つのサブネットワークから構成されている。これらのサブネットワークはそれぞれで入力されたオプティカルフローマップとバイナリマスクからオプティカルフローマップの推定を行っている。

3.2 欠損補完

欠損の補完は、オプティカルフローマップ推定ネットワークの出力を用いて行う。まず、推定されたオプティカルフローマップが妥当であるかの判断を行う。妥当であると判断されたオプティカルフローマップを用いて、すべての非欠損画素を欠損画素を埋めるように前後方向に伝播させることで補完を行う。欠損画素へ両方向の非欠損画素が伝播されている場合は、欠損画素と非欠損画素との距離に反比例する重みをつけ線形結合することによって補完する。ピクセル伝播によって補完が行われなかった画素がある場合は、Partial Convolution を用いたネットワークによる画像補完 [1] による補完を行う。補完された結果によって入力は随時更新され、以降に実行される補完において参照される。一度の伝播のみではすべての欠損画素を補完できない場合があるため、欠損部がなくなるまでピクセル伝播と画像補完が実行される。

4 計算機実験

計算機実験を行い、提案手法により動画の補完が行えることを確認した。

参考文献

[1] Guilin Liu, Fitsum A. Reda, Kevin J. Shih, Ting-Chun Wang, Andrew Tao and Bryan Catanzaro : “Image inpainting for irregular holes using partial convolutions,” https://arxiv.org/abs/1804.07723, (2022 年 1 月 5 日参照).

[2] Olaf Ronneberger, Philipp Fischer and Thomas Brox : “U-net: Convolutional networks for biomedical image segmentation,” https://arxiv.org/abs/1505.04597, (2022 年 1 月 5 日参照).

[3] Guilin Liu Fitsum A., Reda Kevin J. Shih, Ting-Chun Wang, Andrew Tao, Bryan Catanzaro, Sungho Lee, Seoung Wug Oh, DaeYeun Won, and Seon Joo Kim : “Deep flow-guided video inpainting,” https://openaccess.thecvf.com/content_CVPR_2019/papers/Xu_Deep_Flow-Guided_Video_Inpainting_CVPR_2019_paper.pdf, (2022 年 1 月 5 日参照).

Copyright ©2022 Information Processing Society of Japan.
All Rights Reserved.