Problem 1

Group Name:

UCLA VISION LAB

Group's Website:

http://www.vision.cs.ucla.edu/index.html

Group Affiliation:

UCLA VISION LAB
University of California, Los Angeles
Engineering VI Room #386
404 Westwood Plaza
Los Angeles, CA 90095
USA

Project Title:

Video Upscaling via Spatio-Temporal Self-Similarity

Project Website:

http://www.vision.cs.ucla.edu/papers/ayvaciJLCS12.pdf

People:

Professor Stefano Soatto: http://web.cs.ucla.edu/ soatto/

Project Description:

This projects proposes a new example-based video upscaling technique
that exploits self-similarity among patches of a video in both space and
time. The project proposes to encode image patches with over-complete
dictionaries constructed in a local spatio-temporal neighborhood, and es-
tablish temporal correspondence using modern optical flow techniques.
According to the people involved in this project the resulting method
performs favorably compared to the state-of-the-art in super-resolution
techniques.

Current technological developments have resulted in the rapid increase
in numbers of imaging devices as well as highresolution display units.
Nonetheless, there exists a gap between the resolution of imaging de-
vices and display capabilities. This could be discussed and addressed by
employing better imaging devices, which comes with a price. Thus, the

team behind this project is proposing a method to increase the resolution of an image sequence by taking advantages of the spatio-temporal self-similarity.

The team behind this idea proposes a method that generalizes example-based methods to video, hence taking advantages of the transitory consistency between the frames by optical flow estimation and merging the self-similar examples with nonlocal-means. The team's algorithm requires explicit motion estimation and occlusion detection unlike other video upscaling methods to achieve more precise patch correspondence, which then improves the upscaling results.

Interest Reasons:

The team behind the project is presenting a method that performs upscaling of video sequences by borrowing repeated occurrences of patches in space and time using some kind of motion estimation. This experiments show that the team's methods outperforms state-of-the-art super resolution techniques that do not exploit temporal coherence or explicit optical flow estimation. This means that in the future we can include the incorporation of external patches to complete the dictionaries when the samples from the video are insufficient. This is not just interesting as a project but this is a step toward achieving better technological advances that would help improve many experiments and projects in the field of Computer Vision.

Problem 2

Article I: Structural Constraint Data Association for Online Multi-object Tracking

Article Citation:

Yoon, J.H., Lee, CR., Yang, MH. et al. Int J Comput Vis (2019) 127: 1. https://doi.org/10.1007/s11263-018-1087-1

Main Problem:

Online two-dimensional (2D) multi-object tracking (MOT) is a challenging task when the objects of interest have similar appearances. In that case, the motion of objects is another helpful cue for tracking and discriminating multiple objects. However, when using a single moving camera for online 2D MOT, observable motion cues are contaminated by global camera movements and, thus, are not always predictable.

Overview:

To deal with unexpected camera motion, They have proposed a new data association method that effectively takes advantage of structural constraints in the presence of large camera motion. In addition, to reduce incorrect associations with failure to detection and false positive, they have developed a novel event aggregation method to integrate assignment costs computed by structural constraints. People behind this paper have also proposed to utilize structural constraints to track missing objects when they are re-detected again. By doing this, identities of the missing objects can be retained continuously. Experimental results validated the effectiveness of the proposed data association algorithm under unexpected camera motions. In addition, tracking results on a large number of benchmark datasets demonstrated that the proposed MOT algorithm performs robustly and favorably against various online methods in terms of several quantitative metrics, and that its performance is comparable to offline methods.

Strengths:

We are able to see a few new ideas being proposed in this article. These ideas are building on older methods and earlier beliefs. The point of the article is not just coming up with a solution but it is to diminish the likelihood of flawed associations with bogus results.

Weaknesses:

Although having new ideas is always a positive when it comes to performing experiments, it does look like that the team behind this paper were trying vigorously to get the detail correct. I believe by doing what was mentioned they were drifting out of the actual experiment and paying a great deal of scrutiny on minor issues.

Article II: Fast Diffeomorphic Image Registration via Fourier-Approximated Lie Algebras

Article Citation:

Zhang, M. & Fletcher, P.T. Int J Comput Vis (2019) 127: 61. https://doi.org/10.1007/s11263-018-1099-x

Main Problem:

This paper introduces Fourier-approximated Lie algebras for shooting (FLASH), a fast geodesic shooting algorithm for diffeomorphic (an isomorphism of smooth manifolds) image registration. The people behind this paper approximate the infinite-dimensional Lie algebra of smooth vector fields, i.e., the tangent space at the identity of the diffeomorphism group, with a low-dimensional, band-limited space. The main objective is for them to show that most of the computations for geodesic shooting can be carried out entirely in this low-dimensional space, which makes that the main problem as well.

Overview:

The people behind this paper are attempting to show that most of the computations for geodesic shooting can be carried out entirely in this low-dimensional space. They have come up with an algorithm and the algorithm results in dramatic savings in time and memory over traditional large-deformation diffeomorphic metric mapping algorithms, which would require dense spatial discretizations of vector fields. Now To validate the effectiveness of FLASH, they would have to run pairwise image registration on both 2D synthetic data and real 3D brain images and compare with the state-of-the-art geodesic shooting methods. Experimental results show that the algorithm dramatically reduces the computational cost and memory footprint of diffeomorphic image registration with little or no loss accuracy.

Strengths:

This article has solid mathematical evidence to back up the claims it is making. The article also has a stable tone to it as we do need witness any erroneous or illogical paths along the way. The people behind this paper really knew what they wanted to write about and they have executed that perfectly.

Weaknesses:

This article is really relying on the readers mathematical background. I do think some refresher wouldn't be a heinous plan as I found myself searching for many of the terms in this paper, compared to the first paper on this report. I expected to read more about their algorithm but we can only see some limited detail about the actual algorithm.