3D CV: Research Proposal

Apoorv Agnihotri, 16110020, apoorv.agnihotri@iitgn.ac.in

November 17, 2018

Introduction

This proposal is basically an implementation of the research paper 'Temporal Epipolar Regions' [1] and further highlighting the importance of the said paper. The paper in view basically looks at unconstrained set of images of a perticular scene, i.e. the images which are not captured with a uniform time interval between them as well as with different viewpoints. What we want to do is to find the correspondence of the moving features in such a set of images.

This paper looks at a novel method of using epipolar lines to predict approximately linear moving features.

In this paper we try to use epipolar geometry to divide the image into regions. These regions are termed Temporal Epipolar Regions (TERs). These TERs can be *valid* or *invalid*, which is another way of saying that a region contains the trajectory of the moving feature or not. The basic idea is to reduce the search space of the feature matching algorithms to increase the efficiency and the correctness of matched features. This paper helps us achieve this using TERs.

The method described in this paper is different from what is usually found in video domain. This is because of the fact that we have gradual changes in viewpoint as well as uniformly divided frames (with respect to time) in the case of videos.

Significance

- The significance of this proposal is introduction to a all new problem of locating moving points in a *CrowdCam* settings.
- This praper introduces the technique of using epipolar lines to find out *Temporal Epipolar Regions (TERs)* that are *valid*. This helps in reducing the computations required by a whole class of algorithms by only working on the *valid* TERs.
- This method can be used in conjunction with other algorithms for finding correspondence between still images [2, 3] as TERs are used to reduce the search space for feature matching.
- Up-til now we had algorithms that needed at-least 5 correspondence to be known, including all the calibration matrices of every camera to reconstruct the 3D trajectory of the motion of interest. This is also known as trajectory reconstruction. [4, 5].
- This paper uses epipolar lines specially by locating the moving features in an image instead of the usual use of epipolar lines to find the location of the static features.

Objectives

- Get a working implementation of the above algorithm and open-source it.
- Try to find some correspondences that this paper can have in relation with self driving cars.
- Additional extraction of information available in *CrowdCam* images. There are studies done already that focus on ordering these *CrowdCam* images in time[6, 7] and space[8].

Methodology

- First we need to find the temporal order (order of the *CrowdCam* images w.r.t. time), if not available.
- Calculating the fundamental matrices for each of the image in which we want to search for correspondence w.r.t. the images in which we have been given correspondences.
- Calculation of valid TERs on each of the image in which we have not been given the correspondences (by computing epipolar lines and using a lookup table).
- Next the paper now focuses on improving the feature matching algorithms by using these regions.
- We can run the feature matching algorithms in the valid TERs for using the reduced search space to our benefit. We can also provide something called as *forgiveness parameter*, for the cases when we have non linear motions. What this parameter does is increases the size of TERs by some pixels provided as the paramter.

Assumptions

- Temporal order σ of the image has been already been calculated using [7] or has been provided already.
- We have been given the correspondences between the point of interest in at least 3 of the images, I_a, I_b, I_c.
- We are required to find the valid region in the image I_d, where this correspondence for the point of interest can lie.
- We are provided the fundamental matrices between the required set of images.

Steps

- Calculate the features in all the 4 Images (using SIFT[9]).
- Find the Fundamental Matrices of I_d w.r.t. I_a, I_b, I_c, if not already provided.
- Find the epipolar lines in I_d w.r.t. the given point of interest in I_a, I_b, I_c.
- Use the look-up table and find the TERs that are valid given the temporal order σ.
- Black out the regions that are marked invalid to visualize the valid regions.

Expected Outcomes

- Implementation of a part of this paper, up-til finding the valid TERs, given 4 images, 3 with correspondences and 4 image where we need to mark the valid temporal epipolar regions.
- Deeper learning of camera properties, multiple viewpoints and the use cases of epipolar lines in practice.

Results

- Code assumes that d^{th} image is the image whose valid TERs we are interested in.
- To support any arbitrary image to be the target image, one would need to change the code to calculate the corresponding fundamental matrices. Other than that, valid regions would still give valid regions, with change in argument *temporal_order* in the function *get_valid_regions*.
- See the README for image results.

Future Work

- Forgiveness metric to make the valid regions more robust to a nonlinear motion of the point of interest.
- Forgiveness metric to make the valid regions more wide in the case of narrow TERs.
- Convert the code more modular to accept any fundamental matrices that user might want to give.

References

[1] M. Dar and Y. Moses. Temporal epipolar regions. pages 1220–1228, June 2016.

[2] Michael Calonder, Vincent Lepetit, Christoph Strecha, and Pascal Fua. Brief: Binary robust independent elementary features. pages 778–792, 2010.

[3] Marius Muja and David G. Lowe. Fast approximate nearest neighbors with automatic algorithm configuration. In *In VISAPP International Conference on Computer Vision Theory and Applications*, pages 331–340, 2009.

[4] S. Avidan and A. Shashua. Trajectory triangulation: 3d reconstruction of moving points from a monocular image sequence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(4):348–357, April 2000.

[5] Jeremy Yirmeyahu Kaminski and Mina Teicher. A general framework for trajectory triangulation. *Journal of Mathematical Imaging and Vision*, 21(1):27–41, Jul 2004.

[6] Tali Basha, Yael Moses, and Shai Avidan. Photo sequencing. In Andrew Fitzgibbon, Svetlana Lazebnik, Pietro Perona, Yoichi Sato, and Cordelia Schmid, editors, *Computer Vision – ECCV 2012*, pages 654–667, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[7] Gagan Kanojia, Sri Raghu Malireddi, Sai Chowdary Gullapally, and Shanmuganathan Raman. Who shot the picture and when? In George Bebis, Richard Boyle, Bahram Parvin, Darko Koracin, Ryan McMahan, Jason Jerald, Hui Zhang, Steven M. Drucker, Chandra Kambhamettu, Maha El Choubassi, Zhigang Deng, and Mark Carlson, editors, *Advances in Visual Computing*, pages 438–447, Cham, 2014. Springer International Publishing.

[8] Hadar Averbuch-Elor and Daniel Cohen-Or. Ringit: Ring-ordering casual photos of a temporal event. 34:1–11, 05 2015.

[9] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, Nov 2004.