Computer Vision Project: Mid Evaluation Presentation

By Team 14

3D Reconstruction from Accidental Motion

Lets Break the Title

3D Reconstruction

SFM+Bundle Adjustment

Accidental Motion

- It usually takes a couple of seconds to click a picture
- During this time there is inevitable motion due to
    - Hand Shaking
    - Heart Beat

Exact Question which we Ponder upon is

If a camera were to capture a short video before and/or after the capture of a still, would it be possible to use the baseline, in other words translation, from accidental motion to reconstruct the scene

Challenges

1. First the common approach assumes good reconstruction from algebraic methods which depends on adequate baseline so algebraic methods are very unstable for accidental motion

2. The depth uncertainty is very large (Why?)

    a. Therefore previous stereo methods produce artifacts

Vague Solution

Paper suggests that we can use multiple images together to do SFM directly. Due to accidental motion, they suggest to parameterize the 3D points using inverse depth relative to a reference view. Motivation behind such parameterization is that

- It helps regularize the bundle adjustment

Moreover in the paper,

1. The authors find it experimentally good to initialize with random depth and identical camera poses
2. And Having multiple images can also help reduce uncertainty

Important Detail to Note

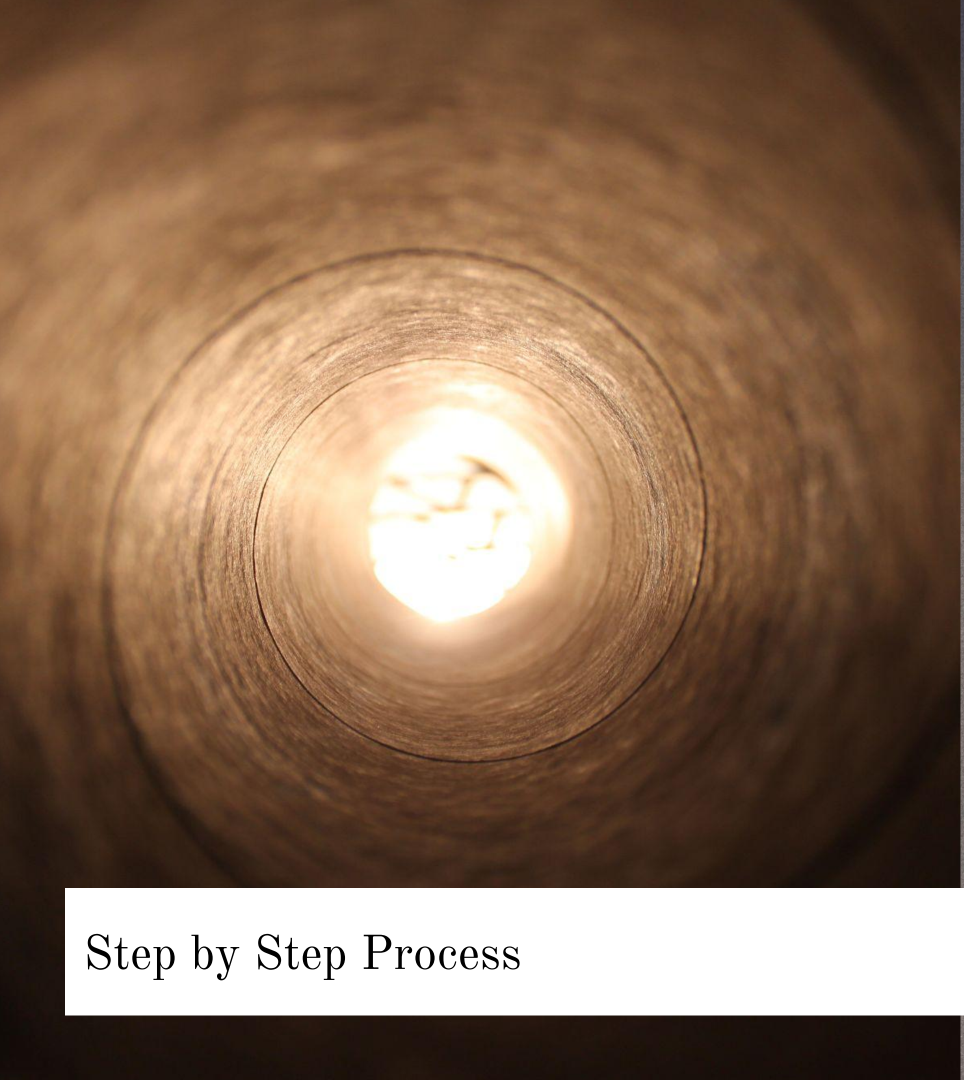
1. Since the depth map is weak and noisy, the popular first order CRF is not very effective and results in over smoothed depth maps.
   a. Low order connections cannot regularize depth effectively
2. So, the paper proposes to use long range connections and it is found that direct connections between a pixel and its bigger neighbourhood can improve the dense reconstruction in the current case
   a. It can be shown that this method can effectively regularize noisy depth maps estimated from weak data terms

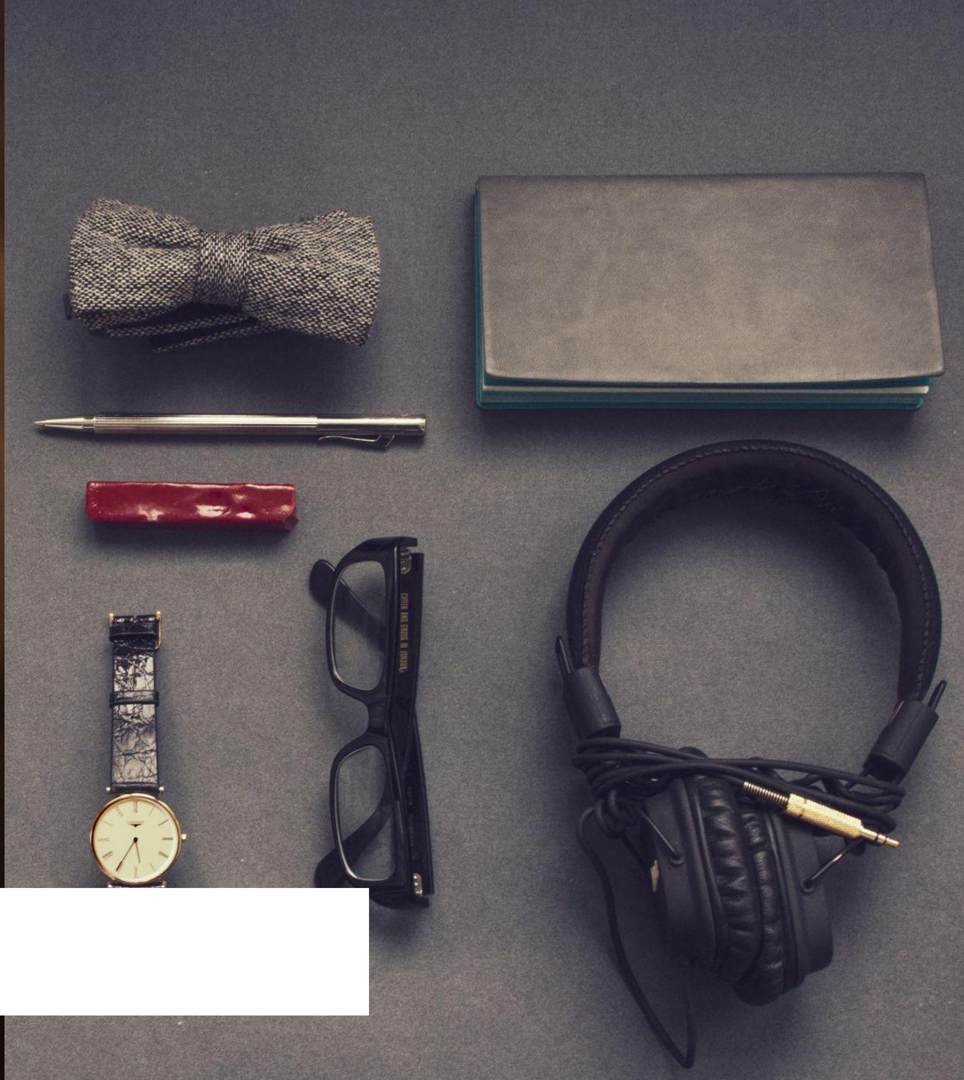
Observations Coming from Cost Function Analysis (l2)

When the camera poses are fixed, it is convex to get the depth of a feature(What Features?, Harris Ones) relative to the reference view

The cost function comes out to be of the form $((x-a)/(x-b))^2$, and basic calculus gives us the intervals of convexity.

It is convex to optimize the rotation for the points at infinity when an approximation is used.





Step by Step Process

Algorithm

1. **Initialization:**
   a. Select a reference view and initialize all camera poses with zero rotation and translation.
   b. Points are parameterized by inverse depth.
   c. Projections of 3D Points are proposed using feature tracking across images
      i. We detect corner features in the reference image.
      ii. Then we track all these corners in all the other images using KLT Feature tracker.
      iii. Maximum color gradient difference per pixel between 2 patches should be under a threshold

Algorithm (Continued)

2. **Optimization:**
    a. Optimizing Cost function of bundle adjustment with Ceres Solver
    b. Robustifiers optionally used
    c. After each optimization, we remove the points with negative depth and optimize again with remaining points

Algorithm (Continued)

3. **Dense Reconstruction**

    a. We can only get a 3D structure seen from the common viewpoint(?)
    b. Since depth signal tends to be noisy, we adopt plane sweeping together with the CRF framework to solve a smooth depth map
    c. To preserve the details while smoothing the depth map, paper proposes to use long range connection between pixels in the CRF energy function, which can pass information to a pixel effectively.eb

Scope

We aim to attempt on implementing the algorithm suggested in the paper in its entirety without leaving any detail. (Presented Until Now)
We aim to also understand the mathematics/reasons behind certain assumptions/results presented in the paper (Which we already understand to some extent)

As far as the data is concerned, there is not any particular dataset required for the task. We can record short videos as and when required

Timeline

12th March — **Project Assigned**

22nd March — **Getting Familiar**

1. Started Reading the Paper
2. Help first Meet to Discuss

26th March — **Making A Plan**

Finished the first understanding of the paper and ready for a second go to connect the dots

Decided upon what we would be doing the mid evaluation

30th March — **Mid Eval Submission**

1. Finished the presentation
2. Started writing the code, for basic input and initializations necessary

5th April — **Code+Research**

1. Progression with the code
2. Finish Understanding the references used in the paper

12th April — **Code**

1. Finish writing code
2. Debug and
3. Getting ready with first set of output

After 12th April

Finalising things, Working on Final ppt, Making last Moment additions/improvements

Thanks/Arigato!!

Questions-Queries