Global Motion Estimation

An indirect, multiscale and robust approach

Samuele Conti   Filippo Daniotti

Department of Information Engineering
and Computer Science

April 20, 2022

Table of Contents

Introduction

In this work we are going to present:

- a broad introduction to the problem of global motion estimation
- an overview of the theoretical foundations of our solution
- the framework we developed from the aforementioed literature
- an analysis of the performances of our solution
- a brief discussion of the limitations and possible improvements

Table of Contents

Global motion estimation

The motion in videos can be analized as the displacement of pixel per pair of frames

Is usually the combination of two different motions:

- the actual motion of objects in the scene
- the egomotion of the camera

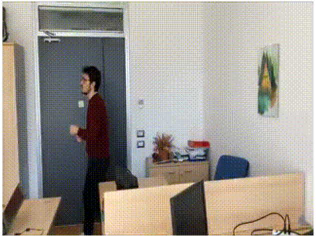


Figure: Local motion vs. egomotion

The problem

Global motion estimation aims to extract the camera global motion patterns in a video

Motion models

Motion models describe camera motion patterns through equations

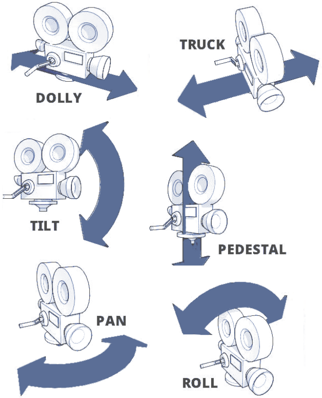


Figure: Different models describe different types of camera motion

The displacement of a pixel across frames as a function of some parameters

$$p' = \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix} = Rp + T \quad (1)$$

E.g. for the **affine model** (our choice) parameter vecotr is $a = [a_0, a_1, a_2, b_0, b_1, b_2]$

Estimating parameters

There are two main approaches, both of which aim to minimize a prediction error

- **direct methods** work in the pixel domain
- **indirect methods** work with motion fields

For our solution we opted for **indirect methods**

Indirect paramter update

The parameter vector a is updated by choosing the parameters that minimize the dissimilarity of the $gt(x)$ and the prediction $MM(x, a)$ for a point x

$$a = \arg\min_{a} \sum_{p} E(gt(p) - MM(p, a)) \tag{2}$$

Table of Contents

Overview

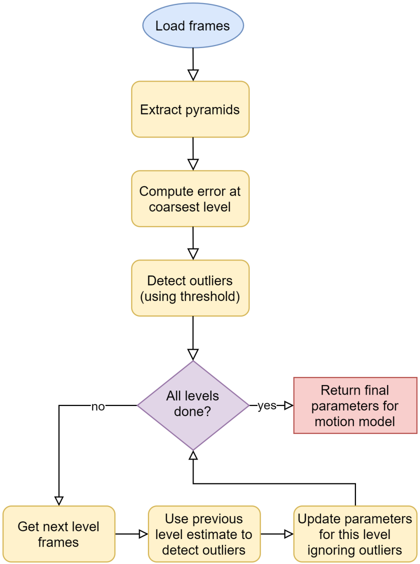


Figure: Flowchart of our framework

Implementation choices:

- Affine model
- Indirect method
- Block-based motion estimation
- Multiscale
- Robust

Block-based motion estimation

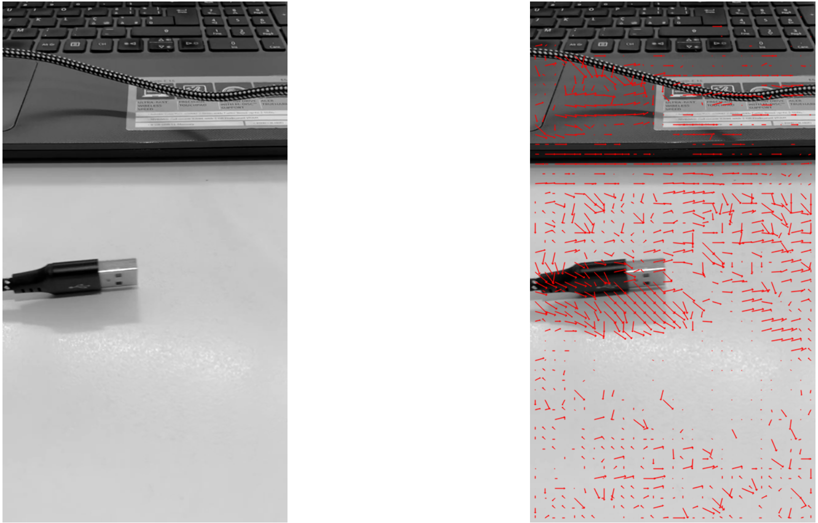
BBME algorithms perform motion estimation and return the motion field to use as ground truth for indirect parameter estimation

Chosen algorithms

We implemented 4 well-known BBME algorithms:

- exhaustive search
- three-step search
- 2D log search
- diamond search

Exhaustive search



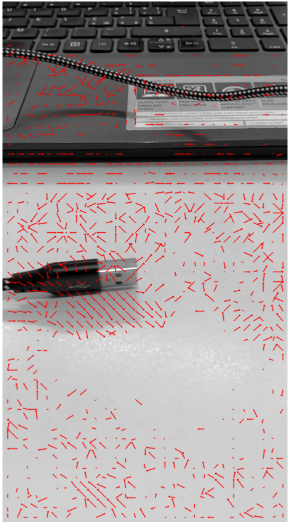
(a) Target frame

(b) Needle diagram

Figure: Motion field result of our EBBME implementation

Three-step search

(a) Target frame

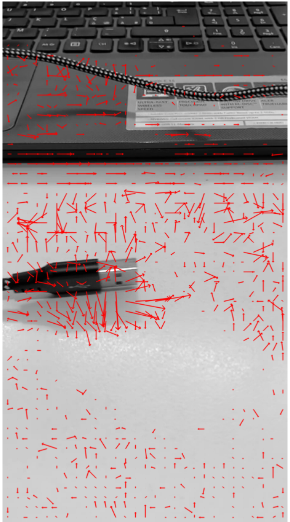


(b) Needle diagram

Figure: Motion field result of our TSS implementation

2D Log search

(a) Target frame



(b) Needle diagram

Figure: Motion field result of our TDLS implementation

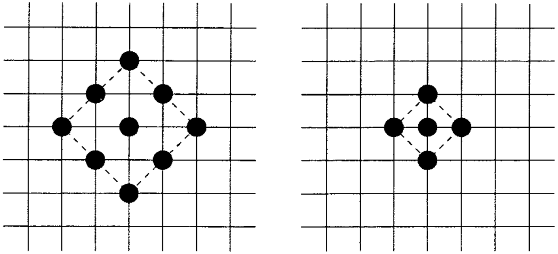
Diamond search

BBME algorithm presented in Zhu and Ma, 2000.

Better performances thant classic algorithms (TSS) and comparable to latest ones (NTSS), yet more efficient in computation.

It uses two different diamond-shaped search patterns:

- large diamond search pattern (LDSP)
- small diamond search pattern (SDSP)



(a) Large diamond search pattern (LDSP) (b) Small diamond search pattern (SDSP)

Figure: The shape of LDSP and SDSP

DS algorithm

Diamond search procedure

For each block:

1. place LDSP in the center of the block and compute DFD for all of the 9 displacements;
    - if the minimum DFD is in the center position, then go to 3
    - else, go to 2
2. reposition LDSP in the minimum DFD of the previous operation and recompute DFD;
    - if the minimum DFD is in the center position, then go to 3
    - else, repeat
3. place SDSP, compute DFD and return coordinates of the minum DFD block as motion vector

DS algorithm visualized

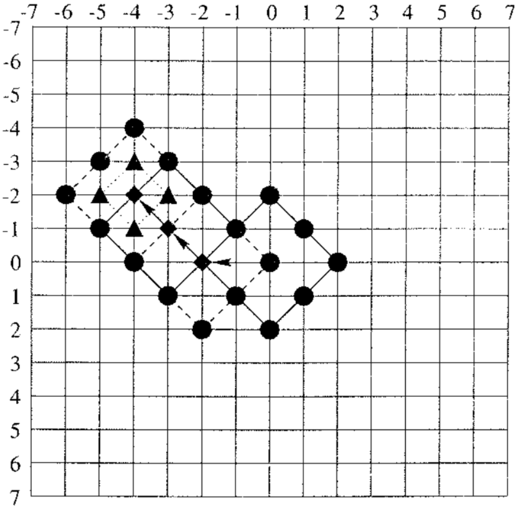
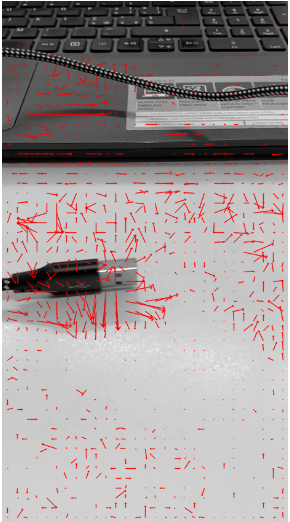


Figure: Example of a full execution cycle of diamond search

DS results

(a) Target frame



(b) Needle diagram

Figure: Motion field result of our DS implementation

Affine model prediction error

The affine model computes the displacement of a given pixel wrt parameters a as

$$d(p, a) = A[p]\ a \tag{3}$$

where $A[p]$ is an intermediate matrix

$$A[p] = \begin{bmatrix} 1 & x & y & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x & y \end{bmatrix}$$

Fitting error

The fitting error computation gets rewritten as

$$E = \sum_{p} |d(p, a) - d(p)|^{P} \tag{4}$$

Affine model parameter estimation

We are interested in the parameters that minimize the fitting error $\rightarrow \nabla_a E = 0$

Parameter update

Assuming $p = 2$, we get

$$a = \left(\sum_p A[p]^T A[p] \right)^{-1} \left(\sum_p A[p]^T d(x) \right) \tag{5}$$

Refinement

- each point p can be weighted
- parameter vector a can be split: $a_x = [a_1, a_2, a_3]$ and $a_y = b_1, b_2, b_3$
  - reduce the overall complexity

$$a_x = \left(\sum_p w(p) A_x[p]^T A_x[p] \right)^{-1} \left(\sum_p w(p) A_x[p]^T d_x(x) \right) \tag{6}$$

Pseudocode

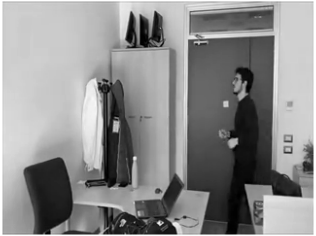
Algorithm 1: High-level pseudocode of our solution.

```
Input: prev_frame, cur_frame
Output: a    // parameter vector

prev_pyrs = get_pyr(previous_frame)
cur_pyrs = get_pyr(current_frame)
a = first_estimation(prev_pyrs.next(), cur_pyrs.next())
foreach l in levels do
    prev_frame = prev_pyrs.next()
    cur_frame = cur_pyrs.next()
    ground_truth_mfield = BBME(prev_frame, cur_frame)
    estimated_mfield = affine(a)
    outliers = detect_outliers(ground_truth_mfield, estimated_mfield)
    a = minimize_error(prev_frame, cur_frame, outliers)
end
return a
```

Table of Contents

Compensation

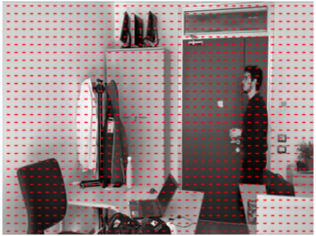


(a) Previous frame

(b) Previous frame

(c) Compensated frame

(d) Absolute difference between current and previous frame



(e) Global motion field estimated by our procedure

(f) Absolute difference between current and compensated frame

Frame Title

Video	Properties	Average	Variance	Maximum	Minimum
`pan240.mp4`	Fast motion	22.724	5.125	27.802	17.981
`coastguard_qcif.mp4`	Two objects moving, background moving	22.733	2.194	26.875	15.158
`foreman.mp4`	Big object moving, still background	19.677	18.443	30.436	11.746
`numeri_del_piero.mp4`	Medium object moving, moving background	19.072	13.642	47.722	16.323

Table: PSNR result on a vunch of samples

Table of Contents

Summary

In this work, we have presented:

- a broad introduction to the problem of global motion estimation
- an in-depth explanation of the theoretical foundations of our solution
- the framework we developed from the aforementioed literature
    - the BBME algorithms chosen
    - the computation of the parameter for the affine model
    - the removal of outliers
    - the implementation of the multiscale approach
- an analysis of the results of our solution
    - qualitative results, using motion compensation
    - quantitative results, comparing PSNR values