ROBUST MOTION SEGMENTATION FROM PAIRWISE MATCHES

Federica Arrigoni and Tomas Pajdla

Czech Institute of Informatics, Robotics and Cybernetics – Czech Technical University in Prague

This MATLAB directory contains the data and code developed in the following papers:

[1] *Robut motion segmentation from pairwise matches*. Federica Arrigoni and Tomas Pajdla. International Conference on Computer Vision (ICCV), 2019.

[2] *Motion segmentation via synchronization*. Federica Arrigoni and Tomas Pajdla. Workshop on Autonomous Navigation in Unconstrained Environments (AUTONUE), International Conference on Computer Vision Workshops (ICCVW), 2019.

In order to run the code, the following packages are needed:

- RPA http://www.diegm.uniud.it/fusiello/demo/rpa/

- VLFeat http://www.vlfeat.org/

- GPCA http://vision.jhu.edu/code/

Contact: Federica Arrigoni, Czech Technical University in Prague, Federica.Arrigoni@cvut.cz

USAGE

The "Mode" method developed in [1] is implemented in the function *segment_mode.m*. The script *test_ICCV19.m* permits to run it on the data from [1].

The "Synch" method developed in [2] is implemented in the function *segment_synch.m*. The script *test_AUTONUE19.m* permits to run it on the data from [2].

The script *Demo.m* permits to run the "Mode" method on raw images. The "Mode" method usually performs better than the "Synch" method, therefore it is the recommended choice. If you want to run it on different data, please check the quality of SIFT matches before performing motion segmentation, and use different descriptors if needed. The quality of two-frame segmentations produced by RPA should also be checked before running *segment_mode.m*.

DATA

This directory contains the following image sequences: Penguin, Flowers, Pencils, Bag, Bears, Pen, Pouch, Needlecraft, Biscuits, Cups.

Two additional sequences (Food, Tea) are available at:

https://drive.google.com/open?id=1a-8YdwuiGt1SRjrYETY_0Txu5GwdOkpX

https://drive.google.com/open?id=1QdZvI6bZSq-T7tGGKnGMkyDv2qXI-nTA

For each sequence, the file *data.mat* contains the following variables:

- *d* is the number of motions;

- *ncams* is the number of cameras;

- *dim* is a *ncams* x 1 vector such that *dim(i)* contains the number of SIFT points in image *i*;

- *m* is the total number of image points: *m=sum(dim)*;

- *cumDim* is the cumulative sum of elements in *dim*;

- *imnames* is a *ncams* x 1 array such that *imnames(i).name* contains the name of image *i*;

- *format_img* is the format of the images in the sequence;

- *SIFT* is a 1 x *ncams* array such that *SIFT{i}.desc*, *SIFT{i}.locs* and *SIFT{i}.scale* contain the descriptors, locations and scales of SIFT keypoints in image *i*;

- *scale* is a quantity by which the images are resized in order to speed up SIFT computation (*scale=1* means no scaling, *scale=0.5* means that images are resized by 50%);

- *min_match* is an integer such that all the points having less than *min_match* matches are removed (*min_match=1* means that isolated points are removed; *min_match=2* means that points that are matched in one image only are removed...);

- *pairwiseEst* is a *ncams x ncams* array such that *pairwiseEst{i,j}.ind1* and *pairwiseEst{i,j}.ind2* contain the indices of corresponding points in images *i* and *j*; note that only *pairwiseEst{i,j}* with *j>i* is computed;

- *Z_pairwise* is a representation of pairwise matches via permutation matrices; in particular, *Z_pairwise(1+cumDim(i):cumDim(i)+dim(i),1+cumDim(j):cumDim(j)+dim(j))* is a partial permutation matrix that encodes the matches between images *i* and *j*;

- *labels_pairwise* is a *ncams x ncams* array such that *labels_pairwise{i,j}* contains the labels (computed with RPA) of corresponding points in images *i* and *j*; the zero label corresponds to outliers/mismatches; note that only *labels_pairwise{i,j}* with *j>i* is computed;

- *Z* is a representation of two-frame segmentations (computed with RPA) via binary matrices; in particular, *Z(1+cumDim(i):cumDim(i)+dim(i),1+cumDim(j):cumDim(j)+dim(j))* is a binary matrix that encodes the labels of corresponding points in images *i* and *j*.

For each sequence, the file *gt_labels.mat* contains the following variable:

- *labels_gt* is a *m* x 1 vector containing (manually computed) ground-truth labels of points in all the images.

CITATION

If you use the "Mode" method and/or the Penguin/Flowers/Pencils/Bag/Bears sequences, please cite [1].

If you use the "Synch" method and/or the Pen/Pouch/Needlecraft/Biscuits/Cups/Tea/Food sequences, please cite [2].

ACKNOWLEDGEMENTS

This work was supported by the European Regional Development Fund under the project IMPACT (reg. no CZ.02.1.01/0.0/0.0/15_003/0000468).