Tracking Based Solely on Location of Interest Points

Gabriel Dauphin

L2TI, University of Paris 13, France

Abstract

Two algorithms are proposed for tracking a region of interest. Given a fairly small number of spatial points at each frame and a high level of outliers, a Bounding Box (BB) is drawn around the region of interest. The first algorithm, called DMX, searches the BB having the greatest density. The second algorithm, called BHM, searches the BB whose theoretical 2D random distribution matches best with the 2D empirical random distribution. An estimate of the error is also available for BHM.

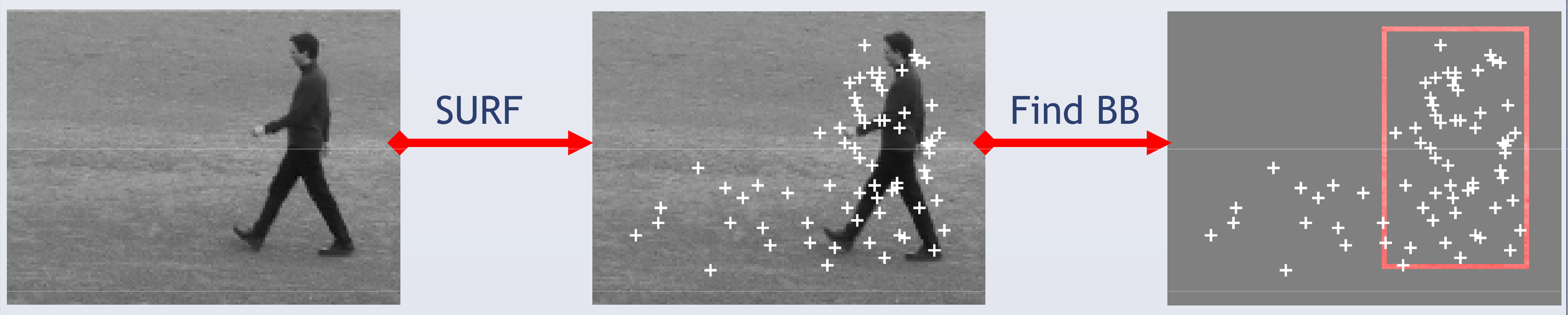
These two algorithms are tested on synthetic data and real data: BHM has better performance on synthetic data, whereas DMX has better performance on real data and could be used in real-time applications.

As information related to a feature can be represented as a set of points, these algorithms may prove useful to computer vision designers in feature selection or in data fusion (fusion of different tracks derived from different features).

1. Introduction

The main proposition is to achieve tracking in two steps: first find interest points, and then derive the Bounding Box (BB) from these interest points. Interest points are here extracted with the SURF algorithm, but there is a wide way of collecting information and of choosing the features in order to extract relevant spatial points. This unified framework should also enable designers achieving information fusion and feature selection.

Motivation:
- Low complexity
- Sufficient information
- Non stationary



Challenge on Synthetic Data:

* Points are drawn from a mixture of two uniform distributions.

$$M_n = (1-\chi_n)U_n + \chi_n W_n$$

$$\text{where } \begin{cases} U_n \to U([0,1]\times[0,1]) \\ W_n \to U([x_m, x_M]\times[y_m, y_M]) \\ \chi_n \to B(1,p) \qquad P(\chi_n = 1) = p \end{cases}$$

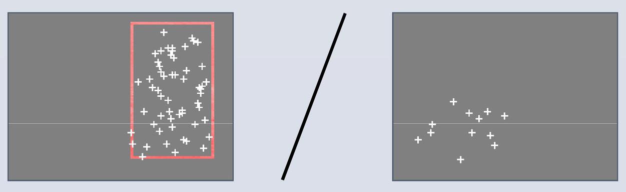
=>2.,3.,4.

Challenge on Real Data:

* Tracking on the KTH database to find the BB of the walking person.

=>5.,6.

2. First Algorithm: DMX[1] (Density Maximisation)



Objective Function:

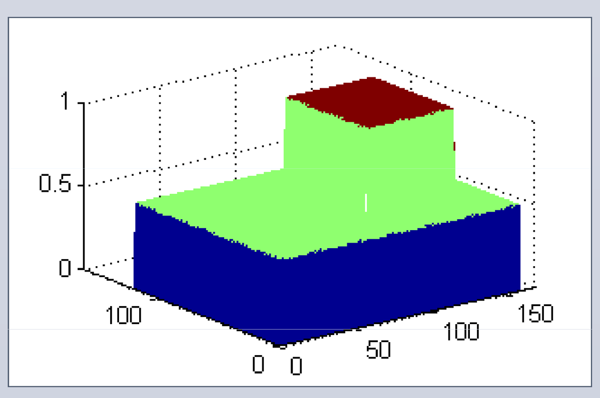
* Maximisation of the ratio of the density of points inside the BB to the density of points outside the BB.

$$[x_m, y_m, x_M, y_M] = \arg\max_I \left[\frac{\#\{n | (x_n, y_n) \in I\}}{\#\{n | (x_n, y_n) \notin I\}} \frac{S - S_I}{S_I} \right]$$

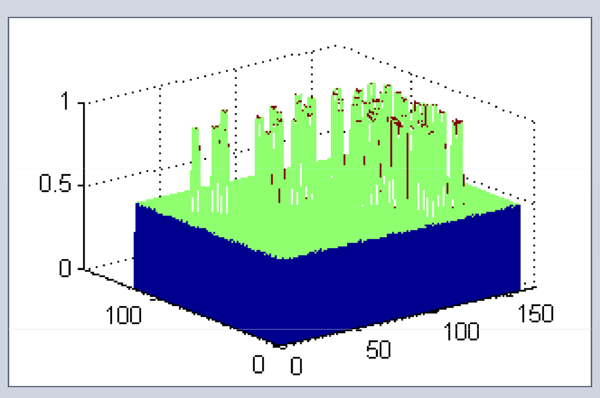
Method of Optimization: Tree Traversal Algorithm.

	The DMX Algorithm
1:	Select the MAX BB.
2:	Find the best BB by moving inwards each side of the BB.
3:	Select the best BB.
4:	Repeat 2 and 3 as long as the BB is not empty.
5:	Select the best of all BB (w.r.t. the obective function).

3. Second Algorithm: BHM[1] (Bivariate Histogram Matching)



Theoretical 2D Distribution



Empirical 2D Distribution

Objective Function:

* Best matching between the theoretical 2D cumulative distribution and the empirical 2D cumulative distribution.

$$J = \frac{1}{4N}\sum_n \left(HT^{++}(x_n, y_n) - HN^{++}(n)\right)^2 + \ldots$$

Similar formulas along the three other causality definitions

$$\text{where } \begin{cases} HT^{++}(x,y) = (1-p)xy + pF_{xm,xM}(x)F_{ym,yM}(y) \\ F_{sm,sM}(s) = \frac{1_{[sm,sM]}(s)}{s_M - s_m}(s - s_m) + 1_{[sM,1]}(s) \\ HN^{++}(n) = \#\{k | (x_k \le x_n) \text{ and } (y_k \le y_n)\} / N \end{cases}$$

An estimate of the error is available.

Method of Optimization: Levenberg Marquardt algorithm.

4. Results

Processing time and error are shown for DMX and BHM. The error is the average distance in pixels between the extreme points of the two BB. Better precision is achieved with higher processing time.

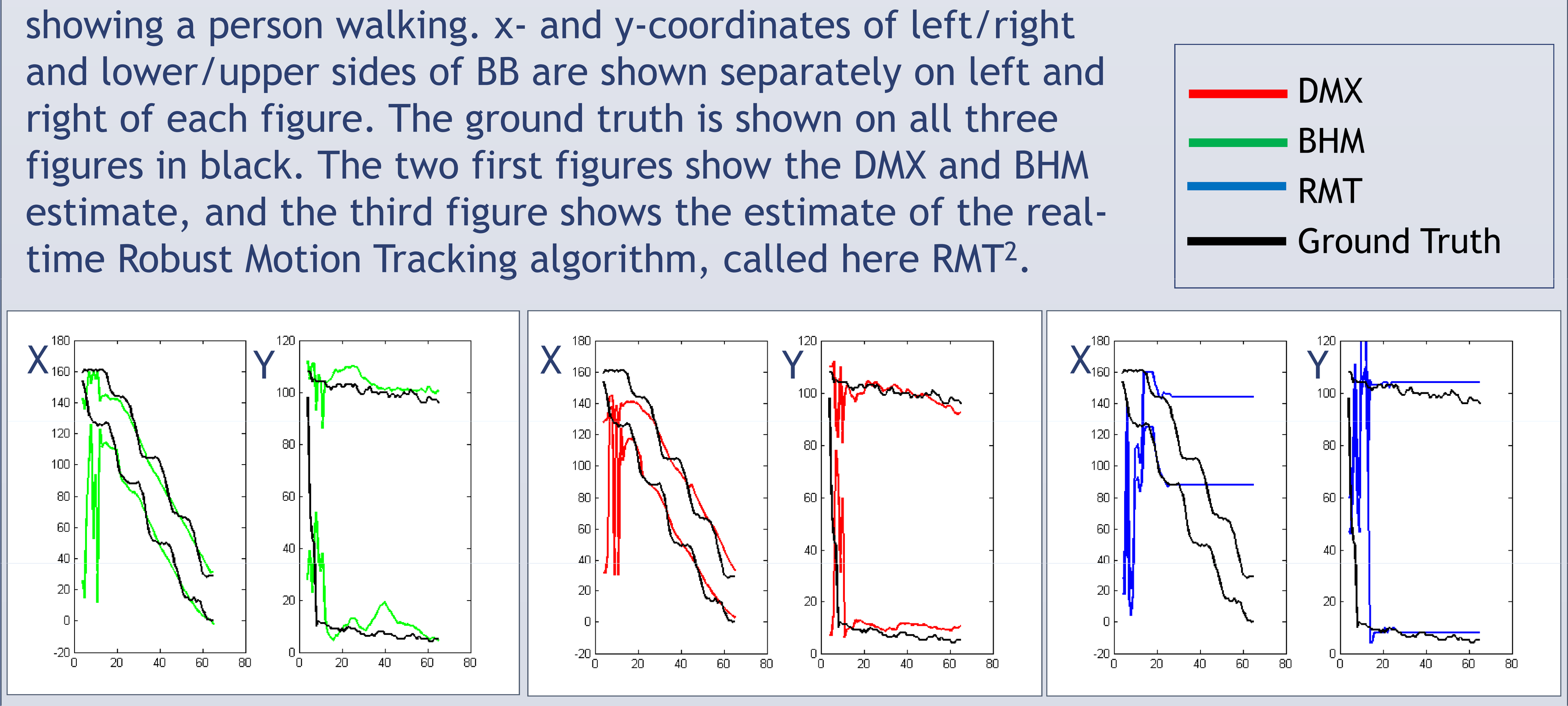
Algorithm	Time Per Frame	Error
DMX	1.7×10^{-2}	24.0
BHM	2.8×10^{-1}	15.4

5. Application on the KTH Database

		The Tracking Algorithm
Extraction of Interest Points	1:	Computation of interest points at each frame with the SURF algorithm.
	2:	Use the RANSAC algorithm to find the dominant motion.
	3:	Subtraction of each frame with the preceding shifted frame (w.r.t. dominant motion).
Temporal Segmenta-tion	4:	Computation of the new set of interest points, named Mn.
	5:	Computation of the standard deviation of x- and y-coordinates of Mn.
	6:	HMM processing .
	7:	Optimization of parameters (to have a small number of shots).
	8:	Temporal morphological filtering of the temporal segmentation.
Estimating the BB	9:	Application of the DMX and the BHM to the Mn-points in the first shot to find the BB.
	10:	Linear correction of BB with a true BB from a given frame.
	11:	Kalman smoothing.

6. Results

The following figures show the estimate of the BB as a function of time on one shot of a video extracted from the KTH database showing a person walking. x- and y-coordinates of left/right and lower/upper sides of BB are shown separately on left and right of each figure. The ground truth is shown on all three figures in black. The two first figures show the DMX and BHM estimate, and the third figure shows the estimate of the real-time Robust Motion Tracking algorithm, called here RMT[2].



x- and y-coordinates of left/right and lower/upper sides of the BB as a function of time.

7. Conclusion

This work attempts to illustrate how spatial points can be used as a unifying framework to represent data extracted from any feature. Such a framework would help doing information fusion and feature selection. Two algorithms are proposed, they transform a set of points into an estimate of the BB. DMX finds the BB for which the density of points inside the BB is highest while the density of points outside is the smallest. BHM finds the BB whose theoretical 2D cumulative distribution matches best the empirical 2D cumulative distribution. BHM has better results on synthetic data, however DMX has better results on real data.

This discrepancy between the performance on synthetic data and on real data raises important questions. Should the distribution inside the BB and the neighbourhoods be less deterministic? Should there be more constraints on how spatial points are extracted from features? Should there be a greater number of points, less outliers; should these points follow more precisely a given random distribution?

8. References

1 : http://www-l2ti.univ-paris13.fr/~dauphin/

2 : C. Stauffer and W.E.L. Grimson, “Learning patterns of activity using real-time tracking.” PAMI, 22(8), pp.747-757, 2000. [Online]http://www.cs.berkeley.edu/flw/tracker/