Tracking People in Video Sequences by Clustering Feature Motion Paths

Adam Gudyś[1,2], Jakub Rosner[1,2], Kamil Wereszczyński[1], Jakub Segen[1], Marek Kulbacki[1*]

1. Polish-Japanese Institute of Information Technology, Warsaw, Poland
2. Silesian University of Technology, Gliwice, Poland
**mk@pjwstk.edu.pl*

INTRODUCTION

Tracking people's motion is an important issue in intelligent video analysis. It is applied e.g., for:

- counting people in a monitored area,
- determining time spent in the area by individual persons,
- identifying patterns in a human traffic.

It can be also used as a preliminary step for more detailed analyses performed by video surveillance systems, such as learning and recognition of human actions. The presented algorithm is based on feature paths, i.e., sequences of matching feature points across consecutive video frames, and can be considered as an extension of the method introduced previously (Segen and Pingali, 1996). It employs the fact that paths corresponding to a motion of a single object are close to each other spatially and overlap in time. Therefore, the paths are clustered in a way that each cluster gathers elements corresponding to a single person.

METHODS

The core of the presented tracking system are feature motion paths. Feature points are calculated with a use of SURF algorithm (Bay et al., 2006) and are more robust than contour corners used in the previous method (Segen and Pingali, 1996). Points extraction is done in the motion areas obtained as a result of foreground-background separation.

As the system is intended to operate in real-time, the clustering is updated incrementally in every frame as feature paths extend. The assignment of path P to cluster Z can be done when distance $D(P,Z)$ between them meets some conditions. The distance is a combination of two components:

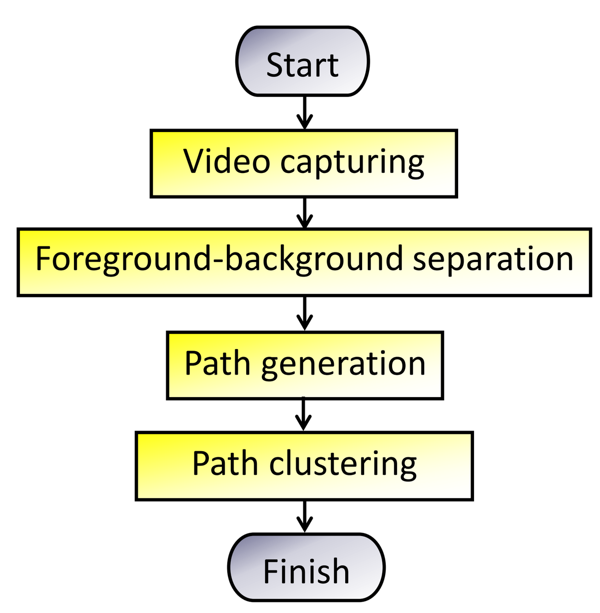

Fig 1. Data processing pipeline.

a) spatial: $$d(P,Z) = \frac{1}{L}\left(\sum_{i=1}^{L}\left(\hat{P}_x^i - \hat{Z}_x^i\right)^2 + \sum_{i=1}^{L}\left(\hat{P}_y^i - \hat{Z}_y^i\right)^2\right),$$

b) tangential: $$d\tau(P,Z) = \frac{\sum_{i=1}^{L}\left(\hat{P}_u^i - \hat{Z}_u^i\right)^2 + \sum_{i=1}^{L}\left(\hat{P}_v^i - \hat{Z}_v^i\right)^2}{\min\left(\sum_{i=1}^{L}\left({\hat{P}_u^i}^2 + {\hat{P}_v^i}^2\right), \sum_{i=1}^{L}\left({\hat{Z}_u^i}^2 + {\hat{Z}_v^i}^2\right)\right) + 2L},$$

with $\hat{Z}$ and $\hat{P}$ being sequences of Z and P points overlapping in time (L is the overlap length). The components (x,y) and (u,v) are, respectively, spatial and tangential coordinates of consecutive path/cluster points. The algorithm can be adapted to particular conditions by adjusting thresholds on $d(P,Z)$ and $d\tau(P,Z)$ distances.

Every N frames clusters are scanned for possible merges which is especially beneficial at the beginning of a motion when there are few short paths and there is a significant probability of several paths from the same object to be assigned to different clusters.

Input:
$\mathbb{P}$ – set of existing paths, $\mathbb{Z}$ – set of existing clusters, n – previous frame number
$d_{\text{assign}}, d\tau_{\text{assign}}$ - spatial and tangential thresholds in path assignment

for each $P = (P^1, P^2, \ldots, P^n) \in \mathbb{P}$:
  extend P by point P^{n+1}
  if P is assigned to some $Z \in \mathbb{Z}$:
    update Z
  else
    for each $Z \in \mathbb{Z}$:
      calculate distance $D(P,Z)$ - the combination of $d(P,Z)$ and $d\tau(P,Z)$
    end for
    select Z_{min} as the cluster minimising $D(P,Z)$
    if $d(P,Z_{\text{min}}) < d_{\text{assign}}$ and $d\tau(P,Z_{\text{min}}) < d\tau_{\text{assign}}$:
      assign P to Z_{min}
    else
      create new cluster Z_{new} containing path P, insert Z_{new} into $\mathbb{Z}$
    end if
  end if
end for

Table 1. Pseudocode of the path assignment – the main part of the clustering algorithm. The cluster merging is done analogously, with a use of greedy heuristics. For each cluster from $\mathbb{Z}$ it finds the closest cluster from following ones, and joins those two clusters if spatial and tangential distances meet assumed conditions.

EXPERIMENTS

The experiments were performed on simulated as well as real-life data. Synthetic video sequence consisted of several circles moving on a two dimensional plane, each circle having a single path attached to its centre. The aim was to check how altering algorithm parameters (spatial and tangential distance thresholds) affect clustering results. The correctness of the algorithm was verified on a short video sequence captured on Bytom Market Square.

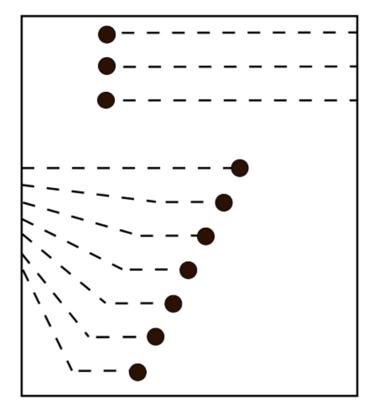

Fig 2a. Synthetic dataset consisting of moving circles, each with a path attached to its centre (a dashed line).

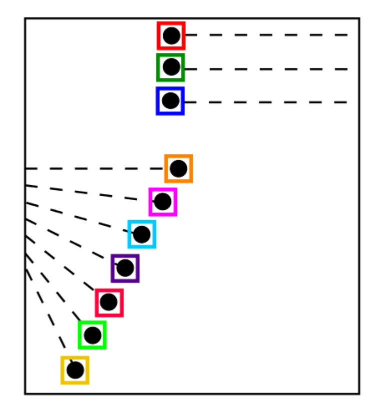

Fig 2b. Setting thresholds to 0 results in each path falling into a separate cluster (a rectangle).

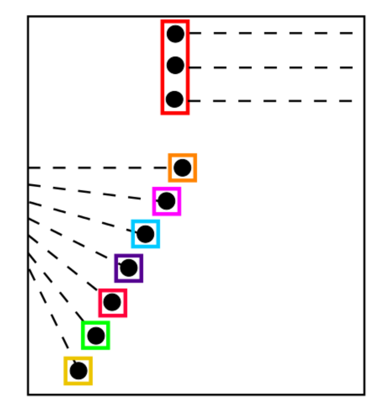

Fig. 2c Increasing spatial distance threshold in path assignment allows parallel paths to be clustered together.

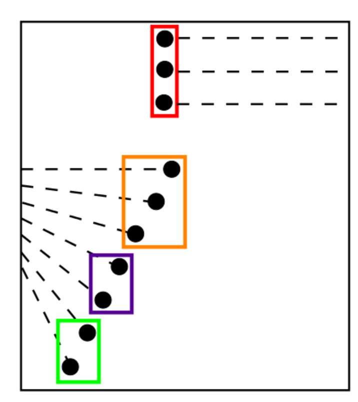

Fig 2d. Non-zero value of tangential distance threshold in path assignment allows non-parallel paths to be clustered.

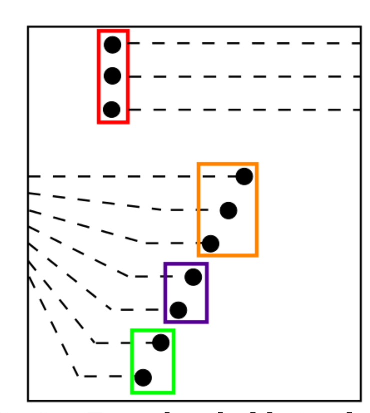

Fig 2e. Zero thresholds in cluster merging procedure prevent clusters from being joined.

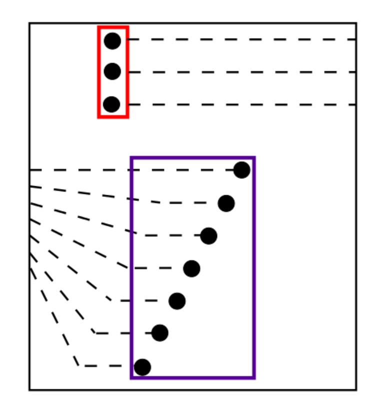

Fig 2f. Non-zero thresholds in the merging procedure allow existing clusters to be joined in further video frames.

Fig 3. Path clustering procedure applied on a video sequence from Bytom Market Square. Red lines indicate paths, colour bold lines join cluster centres over consecutive frames, rectangles are cluster contours in the current frame. Too low distances between silhouettes prevents them from being separated by the clustering algorithm (right-bottom figure).

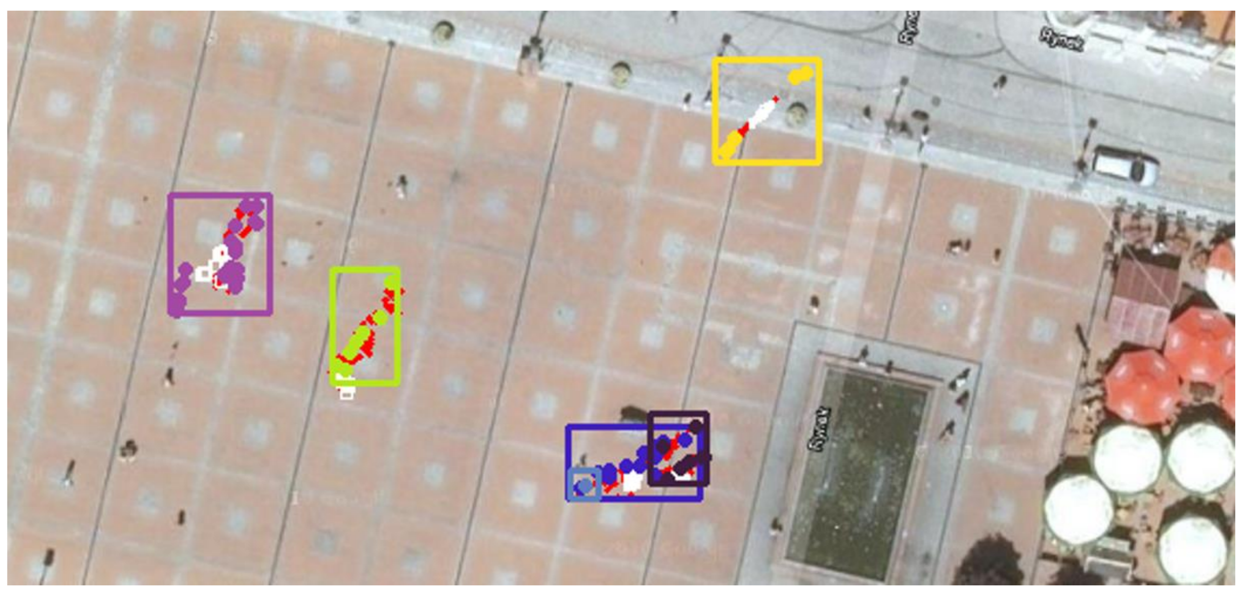

Fig 4. If camera is calibrated, clusters can be projected to the world reference system allowing precise tracking of people's movement.

SELECTED REFERENCES

1. Andriyenko, A., Schindler, K., Roth, S.: Discrete-Continuous Optimization for Multi-Target Tracking. Proc of CVPR 2012, 1926-1933 (2012)
2. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: Speeded up robust features. Proc. Of ECCV 2006, 404–417 (2006)
3. Benfold, B., Reid, I.: Stable Multi-Target Tracking in Real-Time Surveillance Video. Proc. of CVPR 2011, 3457-3464 (2011)
4. Milan, A., Schindler, K., Roth, S.: Detection- and Trajectory-Level Exclusion in Multiple Object Tracking. Proc of CVPR 2013, 3682-3689 (2013)
5. Moeslund, T.B., Hilton, A., Kr¨uger V.: A survey of advances in vision-based human motion capture and analysis. Comput. Vis. Image Und. 104(2) 90–126 (2006)
6. Rabaud, V., Belongie, S.: Counting crowded moving objects. Proc. of CVPR 2006, 705–711 (2006)
7. Segen, J., Pingali, S.G.: A camera-based system for tracking people in real time. Proc. of ICPR 1996, 63–67 (1996)

CONCLUSIONS

A novel method is presented for tracking people's motion in a video sequence, based on clustering feature paths. After separating a non-stationary foreground from a stationary background, motion paths are constructed from feature points extracted using the SURF detector. This approach is superior to the earlier clustering method based on characteristic points of bounding contours, when the SURF-based features are more stable or numerous than the characteristic points of contours, which is often the case. The effectiveness of the algorithm has been tested on synthetic and real video sequences. The processing and memory requirements were carefully balanced, which enabled processing of a 1080p video sequence in real time, on a dual-processor workstation.

This project has been supported by Applied Research Programme of the National Centre for Research and Development (project ID 178438 path A).