

Video-Based Bicycle Detection in Underground Scenarios

Vítězslav Beran, Adam Herout, Ivo Řezníček
Brno University of Technology
Faculty of Information Technology
Department of Computer Graphics and Multimedia
Božetěchova 2, 612 66 Brno, CZ
beranv, herout, ireznice@fit.vutbr.cz

ABSTRACT

Automatic surveillance systems are an important emerging application of object detection algorithms in video. The nature of such systems implies several requirements on the used algorithms. Also, searching for less usual objects (in contrast to frontal human faces, car masks, etc.) is required, such as detection of bicycles. It appears that detection of such objects cannot be solved by just applying a standard statistical or other general detector, but by constructing a specialized detector composed of several standard image processing and object-detection techniques combined together ad hoc. A detector of bicycles in video data from standard low-resolution CCTV surveillance system is presented in this contribution.

Bicycle detection approach covered by this paper aims to cope with highly-noisy low-resolution data, to use simple image-processing methods and to work in real time. Although the method itself does not constitute a generally usable object detector, it covers several interesting aspects which can be re-used in tasks similar to the given one. Low-level features extracted from the video used for wheel-candidate classification are described in detail. The system is applied and evaluated on real data and the results are discussed.

Keywords

Video surveillance, object detection, bicycle detection

1. INTRODUCTION

Contemporary surveillance systems monitoring public places provide massive amounts of data. There exist several common scenarios such as cross-roads, highways, parks and public transport places. Each of the scenarios presents a different type of issues in video-processing point of view.

The presented work aims to cope with the data from underground scenarios, especially from the subway stations, and to detect bicycles.

The task of searching for bicycles in general positions does not seem to be solved in the past. Bicycle detection is used in traffic lights control, but the means of detection are typically induction loops placed in the road surface. Detection based on image processing has also been scarcely explored by some researches (e.g. [Zhi03]), but the algorithms rather

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

detect any object in the road and do not attempt to distinguish the bicycles of a special category.

General object detection algorithms can also deal with bicycle class detection and classification [PASCAL05].

Another research of searching bicycles in video sequences from colour CCTV cameras is bicycle theft detection. We can imagine a bicycle parking place and a CCTV camera which captures this place. This approach is based on identifying the bicycle in the video sequence and detecting the person who dropped the bicycle somewhere in parking place. System stores features of person who dropped the bike and associates it with the current bike. If someone tries to pick up the same bike later, system compares him with a person who dropped this bike and is able to generate alarm. This approach is more thoroughly described in [Dan07].

Besides the theft detection, the bicycle detection improves the control of underground safety. Most of the underground stations forbid to enter with bicycle and automatic bicycle detection decrease the work load of security guards.

The following section shows the structure of the system and introduces the main logic blocks. The

features used to classify bicycle candidates are described in section 3 in detail. The data used for method evaluation are more specified in section 4 together with experiment results. The last section concludes the achievements and proposes possibility for future work.

2. SYSTEM STRUCTURE

The designed system processes the video stream frame-by-frame and provides the position information and probability of possible bicycles.

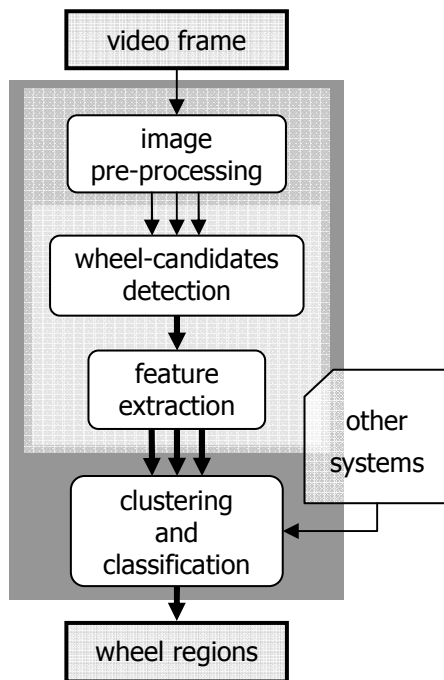


Figure 1. Schema of bicycle's wheel detection procedure.

Because the bicycle detection system works in parallel to other systems processing the same video stream, the design also incorporates some information from other systems. Although it is not crucial for the method itself, it is depicted in the system structure also.

The method pre-processes each frame several times with different parameters. That gives a set of slightly different images. Those images are then one-by-one searched for wheel candidates for which the features are extracted. This repetitive process returns a number of candidates that are finally clustered and classified. The final classification is improved by information from the other systems. In the case of bicycle detection in subway scenario, the information about active regions of interest is used to validate the classification.

Image Pre-Processing

Our method for bicycle detection basically looks for wheel-like shapes. Analysis of the actual data from the scenario (subway access corridors, platforms) showed that it would be impossible to look for the whole shape of a bicycle, because the bicycle can often be partially covered by a human figure. On the other hand, the shape of a wheel of reasonable dimensions does not appear in the videos just except for the case of bicycles. The analysis of the data also allows us to use a prior knowledge about the supposed colour of the wheels. Most of bicycles have black wheels, so the pre-processing step is designed to emphasize the dark grey patches of a particular tone.



Figure 2. Data examples (video frames and zoomed positive and negative samples).

Firstly, the input video frame is morphologically eroded and smoothed. The smoothing step helps to reduce the influence of harsh compression noise. The algorithm then generates images for detection of wheel candidates by computing the distance from dark-grey-colour model with different grey intensities. The dark-grey-colour model is implemented using a thresholding step, morphology closing and finally computing the distance from given grey value. This process is summarized in the following steps:

1. preprocessing: erosion, smoothing
2. colour model image created
3. several reference intensity images with different thresholds

Several image layers are generated by thresholding the color model image (see Figure 3). The wheel candidates are then detected on each layer.



Figure 3. Candidates from different layers.

The information about the candidate layer is used in final clustering. The more candidates are detected on the same location and different consecutive layers the higher candidate probability is.

Detection of Wheel-Candidates

The image with emphasized and thresholded dark patches is used to detect wheel candidates based on contours of clustered areas of constant colour [Son99] which are suspected to be insides of the bicycle wheels and are further examined to either confirm or reject this hypothesis. A huge amount of initially detected contours is reduced using filtering by allowed size range (dimensions in image axes and contour area) or width/height ratio. The next step is to find out, whether the contour has the shape of a wheel. For that task the ellipse-fitting algorithm [Dan07] is used. Having the contour described by an affine transform matrix, further filtering can take place. The error, the difference between the original contour and the fitted ellipse, from the ellipse-fitting algorithm serves as one of the candidate features (see section 3 below).

Clustering and Classification

The method takes into account several measurements. Besides the features described in the next section, also the candidate stability is used to increase its robust classification.

The stability feature is evaluated from the several sets of wheel-candidates obtained from appropriate image. The candidates are clustered using equivalency computation [Cor04] based on the Euclidean distance between samples' positions and sizes.

The final candidates are then represented by the clusters where the feature values, poses and sizes are computed as the mean of all cluster samples. The final candidates are classified by summing their weighted feature values and by comparing the final feature value to the main threshold. The main threshold is one of the tuning parameters of the method. The feature weights were set experimentally.

3. WHEEL FEATURES

Designing features for presented method, we experimented with standard ones, e.g. template matching and also with some we made up for the task purpose itself, e.g. ellipse profiling. Besides those, which are presented in more details in following subsections, we computed several others, such as the contour curvatures, length, contour length-area ratio, etc. Due to the low resolution of the data and the

wheel sizes, those features have no use for classification.

Multi-Scale Template Matching

When detecting wheels using template matching, there arises a problem of different wheel widths. We solve the problem by utilizing several wheel templates with different circle widths (see example templates on Figure 4).



Figure 4. Set of templates covered different wheel width.

The first step before the matching takes place is to normalize the candidate patch. We use the ellipse parameters to compute affine transformation matrix. The normalized patch is then resampled to several different scales. The template matching is then executed for all templates and all scaled patches. Because the templates and also the scaled patches are normalized, we then get the maximum of all matches as the wheel feature.

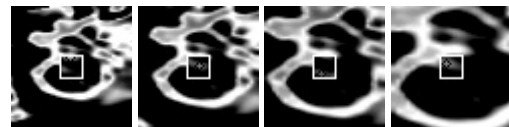


Figure 5. Rescaled sample with the results of matching showed in the rectangular area.

The detector robustness is improved by looking for the best match not only over the scale-space, but also in spatial domain. The white rectangles in the middle of subfigures on Figure 5 show the position change range. The pixel intensities in the rectangles correspond to matching results for particular template position and size.

Ellipse Profiling

The ellipse-profiling is the principal method we designed and develop. The method gives the most robust and distinctive features out of all others used by our system.

The main idea of the ellipse profiling method is to analyze the image intensity profiles defined by ellipses with different parameters. The image intensity values represent the probability given by dark-grey-colour model computed in pre-processing stage. For each wheel candidate, the algorithm generates several ellipses with increasing size. Then the image intensity values corresponding to ellipse profiles are summed up.

Each candidate region is described by affine transformation matrix. The transformation is used to

compute the elliptic profile and read image pixel values below the profile. The equation of the elliptic profile is:

$$ep(\mathbf{A}, r) = \int_0^{2\pi} I \left(\mathbf{A} \begin{bmatrix} r \cos t \\ r \sin t \end{bmatrix} \right) dt$$

where \mathbf{A} is a affine transformation matrix describing the candidate region (position, rotation and scale), r is ellipse profile size and I is the color probability image.

We then search for maximal difference between consecutive summed values. The maximal difference represents the biggest change of ellipse profiles. Such change can occur, when one ellipse minimally overlaps the analyzed candidate image patch and the consecutive ellipse overlaps it maximally.

$$r_{\max} = \max_r ep(\mathbf{A}, r)'$$

The maximal difference found is the next candidate feature:

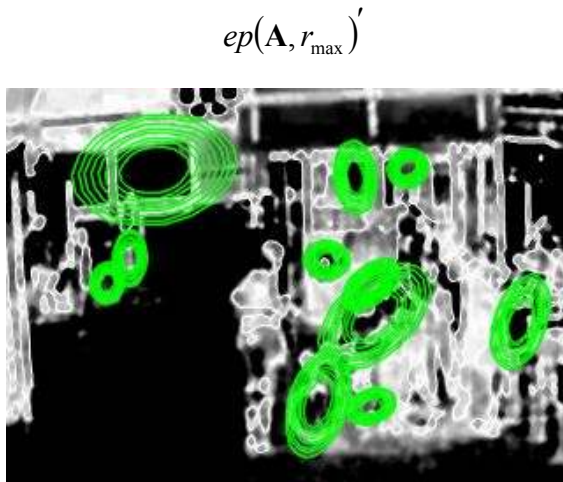


Figure 6. Elliptic profiles on several wheel candidates.

We improve the robustness of the method by evaluating more features by changing the ellipse positions within some small range. The candidate feature is then the best one.

4. EVALUATION

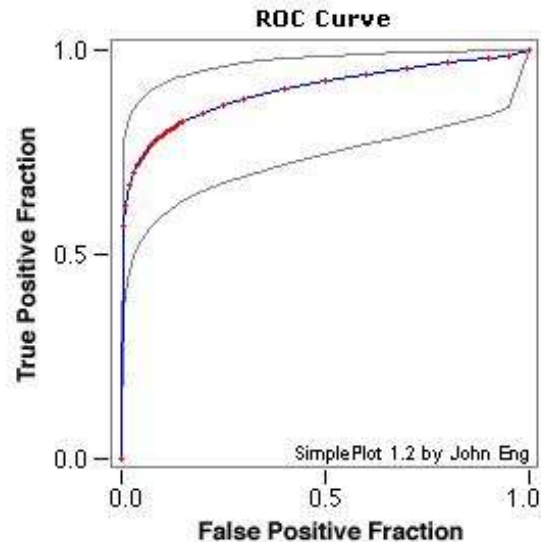
The presented method is designed for video data from underground scenarios such as subway stations. Such data are noisy due to the high compression and the resolution is low. On the other hand, the lighting conditions are absolutely stable and the setting of the scene (where people can appear, stable objects in the scene etc) is reliably predefined. The precision of the method is highly sensitive to its parameters setting, which, in the constant conditions, is not a serious handicap. Also the method itself merely detects

bicycle wheels. This limiting factor does not appear to be too restrictive: every appearance of anything similar to a wheel in the sample data (dozens of hours) really was a part of a bicycle.

The detector performance was evaluated on images from video of Roma undergrounds. The image resolution 720x576 using format 24bits/pixel, RGB.

Since the system is based on the wheel detector, the detector's ability and properties are crucial for the overall performance. For the wheel detector evaluation, about 60 still images containing bicycles were taken and manually annotated that gives about 120 wheels – positive samples.

The ROC curve below was generated by processing the wheel detector on each image and the wheel candidates were compared to the ground truth data. The ROC curve describes the trade-off between true positive and false positive detection rates. The gray curves are 95% confidence interval of the fitted ROC curve. The ROC curve was generated by [ROC].



7. ROC curve of the wheel detector.

In real applications, other supporting information is taken into account. Also, when the detector is used in the system for bicycle detection in the subway scenario, the moving object information, active regions of interest knowledge and temporal information from several consecutive frames improve the overall ability of the system.

Our system is designed to work on CCTV data, so instead of classifying each frame whether there is a bicycle or not, the whole video shot is classified. The Figure 8. shows the application screenshot in real environment.

The bicycle detector is sensitive to wheel-like objects that do not necessary need to be a bike. Such false positives are baby carriages, logos on clothes or shiny

bald spots. Such candidates are not properly distinguished by the system.

The presented system was integrated and evaluated in the surveillance system called CareTaker. The system is fully functioning and running in Roma and Torino underground.

5. CONCLUSION

We have presented a method for bicycle detection for underground scenarios. This method is optimized and targeted to surveillance systems with CCTV cameras, i.e. on low resolution low quality systems with requirement of real-time or close-to-real-time operation, but accepting compromises in reliability of results. Besides use of the common approaches, we designed a new method for similarity measure between cycle in raster image and ellipse. We designed a system to run experiments and evaluate our approach and performed preliminary evaluation of the algorithm.



8. Screenshot of bicycle detection application running in real environment.

Current work is on designing a generalized algorithm to cope with any shape that could be described by parametric curve and evaluating its robustness and stability.

6. ACKNOWLEDGEMENTS

This work has been supported by the CareTaker, IST EU project number 027231 and the “Centre of Computer Graphics” (CPG-LC06008), Czech Ministry of Education, Youth, and Sports grant.

7. REFERENCES

[Dan07] Damen, D., Hogg, D. Associating People Dropping off and Picking up Objects. In British Machine Vision Conference, 2007.

[Son99] Šonka, M., Hlaváč, V. and Boyle, R. Image Processing, Analysis and Machine Vision. PWS, ISBN 0-534-95393-X, 1999.

[Fit99] Fitzgibbon, A. W., Pilu, M. and Fisher, R. B. Direct least-squares fitting of ellipses. In IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(5), 476—480., 1999.

[Cor04] Cormen, H. T. 2001. Introduction to Algorithm. MIT Press, ISBN 0262032937.

[Zhi03] Zhijun, Q, et al. 2003, The study of the detection of pedestrian and bicycle using image processing. In Intelligent Transportation Systems, 2003. Proceedings. 2003 IEEE Volume 1, Issue , 12-15 Oct. 2003 Page(s): 340 - 345 vol.1

[PASCAL05] M. Everingham, A. Zisserman, C. Williams, L. Van Gool, M. Allan, C. Bishop, O. Chapelle, N. Dalal, T. Deselaers, G. Dorko, S. Duffner, J. Eichhorn, J. Farquhar, M. Fritz, C. Garcia, T. Griffiths, F. Jurie, D. Keysers, M. Koskela, J. Laaksonen, D. Larlus, B. Leibe, H. Meng, H. Ney, B. Schiele, C. Schmid, E. Seemann, J. Shawe-Taylor, A. Storkey, S. Szedmak, B. Triggs, I. Ulusoy, V. Viitaniemi, and J. Zhang. In *Selected Proceedings of the First PASCAL Challenges Workshop*, LNAI, Springer-Verlag, 2006

[ROC] Eng John. ROC-Analysis, Web-based Calculator for ROC Curves, Johns Hopkins University, Baltimore, Maryland, USA. <http://www.rad.jhmi.edu/jeng/javarad/roc/JROCFITi.html>

