Noname manuscript No.
(will be inserted by the editor)

A Direct Approach for Object Detection with Catadioptric Omnidirectional Cameras

Ibrahim Cinaroglu · Yalin Bastanlar

Received: date / Accepted: date

Abstract In this paper, we present an omnidirectional vision based method for object detection. We first adopt the conventional camera approach that uses sliding windows and Histogram of Gradients (HOG) features. Then, we describe how the feature extraction step of the conventional approach should be modified for a theoretically correct and effective use in omnidirectional cameras. Main steps are modification of gradient magnitudes using Riemannian metric and conversion of gradient orientations to form an omnidirectional sliding window. In this way, we perform object detection directly on the omnidirectional images without converting them to panoramic or perspective images. Our experiments, with synthetic and real images, compare the proposed approach with regular (unmodified) HOG computation on both omnidirectional and panoramic images. Results show that the proposed approach should be preferred.

Keywords Catadioptric omnidirectional cameras · object detection · human detection · car detection · vehicle detection

1 Introduction

Detecting certain objects with cameras is an important task for many research and application areas such as visual surveillance, ambient intelligence and traffic analysis. Last decade has witnessed significant advances in object detection both in terms of effectiveness and processing time. Quite a variety of approaches have been proposed for object detection. A major group in these studies uses the sliding window approach in which the detection task is performed via a moving and gradually growing search window. A significant performance improvement was obtained with this approach by employing HOG (Histogram of Oriented Gradients) features. Inspired by SIFT (Scale Invariant Feature Transform) [17], Dalal and Triggs [7] proposed to use HOG for the feature extraction step and they used SVM (Support Vector Machines) for the classification step. Later on, this technique was enhanced with part based models [10] and with pyramid HOG features and Intersection Kernel SVM [18]. More recently, it was shown that using combinations of features outperforms the approaches that use a single type of feature [24]. For a detailed summary and comparison of methods, specific to pedestrian detection, we refer readers to [9].

Omnidirectional cameras provide 360 degree horizontal field of view in a single image (vertical field of view varies). If a convex mirror is placed in front of a conventional camera for this purpose, then the imaging system is called a catadioptric omnidirectional camera. An example image can be seen in Fig. 3. Despite its enlarged view advantage, so far omnidirectional cameras have not been widely used for object detection. This is partly due to the resolution disadvantage. However recent omnidirectional cameras have adequate resolution to detect objects that cover a small part of the image. Another reason is that the conventional camera methods should be mathematically modified to be used with omnidirectional cameras. As described in Section 2, previous studies in this direction were focused on SIFT.

In a study on object recognition with omnidirectional cameras [25], a mobile robot is given the images of several objects in the environment and it is asked to recognize these objects. Actually, the omnidirectional

This work was supported by the TUBITAK project 113E107.

Ibrahim Cinaroglu · Yalin Bastanlar
Computer Engineering Department, Izmir Institute of Technology, Izmir, Turkey
E-mail: {ibrahimcinaroglu,yalinbastanlar}@iyte.edu.tr

image is warped into a cylindrical panoramic image before matching with the images of the objects using SIFT. In [2], objects in an indoor office environment are classified with a generative model, where the system is first trained with annotated images from the same environment. In [13], authors use Haar features to perform face detection with catadioptric omnidirectional cameras. Instead of modifying the feature extraction step, they convert the omnidirectional images into panoramic images and directly use the conventional (perspective) camera technique. In a similar manner, panoramic images are used in [14] for human detection.

A human tracking method for omnidirectional cameras is proposed in [23]. As a part of the proposed algorithm, HOG features are computed. However, a rectangular rotating and sliding window is used with no mathematical modification for the omnidirectional camera.

In this paper, we propose a modification for the conventional approach to tackle object detection directly on catadioptric omnidirectional images. That is, our method does not require the conversion of the omnidirectional images to panoramic or perspective images. Apart from the advantage of eliminating the image conversion step, the detection performance of the proposed method is superior as given in experiments section.

To our knowledge, the proposed method is the first that mathematically modifies an object detection approach to be effectively used for omnidirectional cameras. A second contribution is that we construct an omnidirectional image dataset with annotated humans, cars and vans and it can be downloaded from our website [1]. We believe this dataset will be useful to the community for omnidirectional vision based object detection research.

The organization of the paper is as follows. In Section 2, we explain why our approach is theoretically correct. We adopt HOG+SVM [7] approach for object detection and as explained in Section 3, we modify the HOG feature extraction step for catadioptric omnidirectional cameras. Our experiments, given in Section 4 were held for human, car and van detection. Their results indicate that the adaptation of HOG features improves the performance when compared to the unmodified HOG computation, i.e. rotating rectangular windows. We also compare our method with object detection on panoramic images converted from omnidirectional ones and conclude that the proposed method is superior especially for objects with a width/height ratio <2.5.

This paper is an extended version of our previous work [6], which included experiments with a limited image dataset and considered only human detection.

[1] http://cvrg.iyte.edu.tr/

2 Processing of omnidirectional images

Due to their non-linear imaging geometry, working with omnidirectional cameras requires geometric transformations. At first sight, converting an omnidirectional image to a panoramic or several perspective images may seem to be a practical solution. However, it has two major drawbacks: The conversion, which is a non-linear warping, can be computationally expensive for large video frames especially when an omnidirectional image is converted to numerous perspective images to properly fit sliding windows. More importantly, the interpolation required by the image warping introduces artifacts that affect the detection performance.

Among a small number of omnidirectional object detection studies (cf. Section 1), none of them developed a method peculiar to omnidirectional cameras. On the other hand, last decade witnessed some effort on computing SIFT features in omnidirectional images. Starting from [8], researchers tried to avoid warping omnidirectional images and instead they assumed a unitary sphere S^2 as the underlying domain of the image function. When these studies (which consider the convolution step of SIFT) are examined, several approaches can be observed. Below, we describe these approaches briefly.

1. The simplest approach would be backprojecting the image onto a sphere surface S^2 and convolving it with a spherical Gaussian function G_S[5]. Since this approach requires resampling of the whole image, authors in [8] project the kernel G_S into image plane instead of backprojecting the image onto S^2, and the convolution is carried directly on the image plane. This avoids image resampling but since the mapped Gaussian kernel changes at every image location it leads to an adaptive filtering. This computational complexity makes the solution unsuitable.
2. Another approach processes omnidirectional images on the sphere after an inverse stereographic projection [12]. Scale space is computed with Gaussian kernels on the sphere, while, the convolution is performed using the spherical Fourier transform. It was stated in [3] and [16] that this operation leads to aliasing issues due to bandwidth limitations.
3. The processing on the sphere is achieved through a suitable differential operator that adapts to the non-uniform resolution, while using the original image pixel values. In [4], scale space representation is computed using the heat diffusion equation and differential operators (Laplace-Beltrami operators) on the non-Euclidean (Riemannian) manifolds. Moreover, authors in [3] tested this approach by evaluating the matching performance of SIFT. Lastly, au-

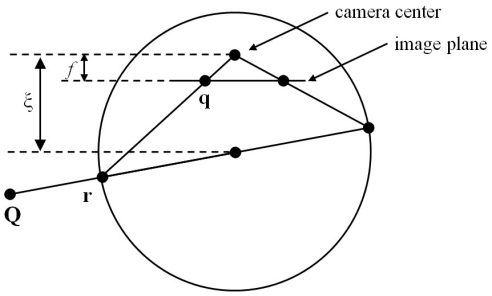


Fig. 1 Projection of a 3D point onto the image plane in the sphere camera model.

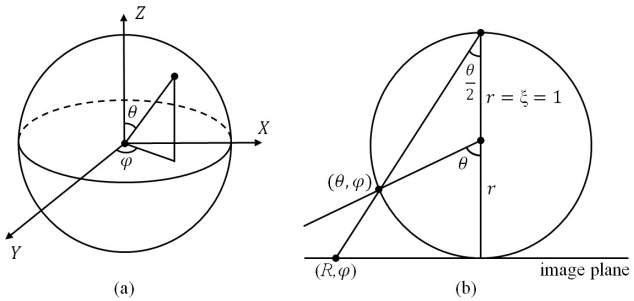


Fig. 2 (a) A 3D point on the sphere is represented by two angles (θ, φ). (b) Consider the unitary sphere ($r = 1$). Image plane is placed at the south pole ($f = 2$). A 3D point is first projected onto the sphere surface and then projected onto the image plane, where in this case $\xi = 1$.

thors in [20] compared the original SIFT with the version modified by Laplace-Beltrami operators on the Riemannian manifolds and observed that the modified version has a better performance. Later, this approach was extended to radially distorted images as well [16] and also generalized to any camera to produce camera invariant features [22].

Exploiting the experience gained by the summarized work, we compute the gradients on Riemannian manifolds (as in [3] and [4]) and adapt the gradient magnitude computation step (Section 3.1) of our algorithm accordingly. Since our study aims object detection, we also modify the gradient orientations to form an omnidirectional sliding window (Section 3.2).

3 The proposed HOG computation

In the sliding window based object detection approach, a window is moved horizontally and vertically on different scales of an image. No rotation is applied since there is an assumed orientation of the object, for instance pedestrians should be upright. In a similar manner, to detect objects in omnidirectional images, we rotate the sliding window around the image center. In addition, to achieve a mathematically correct detection method, we modify the image gradients. The operations that we perform can be divided into two steps:

1. Modification of gradient magnitudes using Riemannian metric.
2. Conversion of gradient orientations to form an omnidirectional (non-rectangular) sliding window.

3.1 Modification of gradient magnitudes using the Riemannian metric

3.1.1 Sphere camera model

We use the sphere camera model [11] which was introduced to model central catadioptric cameras. The model comprises a unit sphere and a perspective camera. The projection of 3D points can be performed in two steps (Fig. 1). The first one is the projection of point $\mathbf{Q}$ in 3D space onto a unitary sphere, resulting in point $\mathbf{r}$, and the second one is a perspective projection from the sphere surface to the image plane, resulting in point $\mathbf{q}$. This model covers all central catadioptric cameras with varying ξ.

A point on the sphere $\mathbf{r} = (X, Y, Z)$ can also be represented by two angles (θ, φ), the former is the vertical angle and the latter is the azimuth (Fig. 2a). In para-catadioptric camera (the ones using a paraboloidal mirror) $\xi = 1$. If we place the image plane at the south pole (which only differs the scale), $f = 2r = 2$ and the perspective projection within the sphere model corresponds to the stereographic projection (Fig. 2b).

There are several methods to perform sphere camera model calibration [21,19]. We used [19] since a MATLAB toolbox is provided with it. In our experiments we used a para-catadioptric camera (ξ=1). Focal length f is the distance to the image plane. For a para-catadioptric camera this is also equal to the distance between image center and any point that is at the same horizontal level with the camera center. Along with ξ and f, image center coordinates (c_x, c_y) are used to modify the gradient magnitudes as explained in Section 3.1.2.

3.1.2 Differential operators on Riemannian manifolds

Let us briefly describe how the differential operators on the Riemannian manifolds are defined. Suppose M denotes a parametric surface on $\Re^3$ and g_{ij} denotes the Riemannian metric that encodes the geometrical properties of the manifold. In a local system of coordinates x^i on M, the components of the gradient are given by

$$\nabla^i = g^{ij} \frac{\partial}{\partial x^j} \tag{1}$$

where g^{ij} is the inverse of g_{ij}.

A similar reasoning is used in [3] and [20] to obtain the Laplace-Beltrami operator, which is the second order differential operator defined on and used for scale space representation for SIFT. In this paper, we are working on the first derivatives. Let us briefly go over the para-catadioptric case and derive the metric that allows us to compute the derivatives on the sphere directly using the image coordinates.

Consider the unitary sphere S^2 with radius=1 (Fig. 2a). A point on S^2 is represented in Cartesian and polar coordinates as

$$(X, Y, Z) = (\sin\theta\sin\varphi, \sin\theta\cos\varphi, \cos\theta) \tag{2}$$

The Euclidean line element in Cartesian coordinates, dl, can be expressed in polar coordinates as

$$dl^2 = dX^2 + dY^2 + dZ^2 = d\theta^2 + \sin^2\theta d\varphi^2 \tag{3}$$

The stereographic projection of the sphere model sends a point on the sphere (θ, φ) to a point in polar coordinates (R, φ) in the image plane (plane $\Re^2$), for which φ remains the same and $\theta = 2\tan^{-1}(R/2)$ in a para-catadioptric system (Fig. 2b).

Using the identities, $R = \sqrt{x^2+y^2}$, $\varphi = \tan^{-1}(y/x)$ the line element reads

$$dl^2 = \frac{16}{(4+x^2+y^2)^2}(dx^2 + dy^2) \tag{4}$$

giving the Riemannian inverse metric

$$g^{ij} = \frac{(4+x^2+y^2)^2}{16} \tag{5}$$

With this metric, we can compute the differential operators on the sphere using the pixels in the omnidirectional images. In particular, norm of the gradient reads

$$|\nabla_{S^2} I|^2 = \frac{(4+x^2+y^2)^2}{16}|\nabla_{\Re^2} I|^2 \tag{6}$$

We see that the para-catadioptric gradients are just the scaled versions of the gradients in Euclidean domain. Therefore, we multiply our gradients with metric g^{ij}.

At the center of the omnidirectional image, $(x, y) = (0, 0)$, Riemannian and Euclidean gradients are the same. At an image location where $\sqrt{x^2+y^2} = 2$, which corresponds to a 3D point at the same horizontal level with the sphere center (mirror focal point), the Riemannian metric is equal to 4. Therefore, the gradients are magnified as we move from the center to the periphery of the omnidirectional image.

The Riemannian metric for other catadioptric systems (with varying ξ) are derived in [20].

3.2 Conversion of gradient orientations for omnidirectional sliding window

After the image gradients are obtained with Riemannian metric, we convert the gradient orientations to form an omnidirectional (non-rectangular) sliding window. The shape of the omnidirectional sliding window varies according to the size and location of the object in the omnidirectional image. As depicted in Fig.3, a car close to the camera is severely bent. However, a window covering the car at a distance is close to a rectangle. The difference can not be represented with a scale ratio, therefore we are not able to train one object model for detection in omnidirectional images. Since it did not seem plausible to train many omnidirectional HOG models, we chose to train our object models with perspective image datasets. Gradients in the sliding window should be computed as if a perspective camera is looking from the same viewpoint.

Fig. 3 Two cars in the omnidirectional image are indicated with black frames. The one close to the camera covers a larger area and it should be searched with a more bent sliding window, the other one is far away and it should be search with a more straight sliding window

Fig. 4a shows a half of a synthetic para-catadioptric omnidirectional image (400x400 pixels) where the walls of a room are covered with rectangular black and white tiles. Conventional HOG result of the marked region (128x196 pixels) in this image is given in Fig. 4b where the gradient orientations are in accordance with the image. However, since these are vertical and horizontal edges in real world, we need to obtain vertical and horizontal gradients. Fig. 4d shows converted gradients for the region marked in Fig. 4c, which is an example of the proposed HOG computation.

To obtain the gradients in Fig. 4d from the image in Fig. 4c, we performed a transformation from polar to Cartesian coordinates without using any camera calibration information. Both gradient orientations and gradient magnitudes in the proposed HOG window are computed from the omnidirectional image using bilinear interpolation with backward mapping. While transforming coordinates, the height and width of rectangular area in Fig. 4d are kept equal to the thickness and center arc length of the doughnut slice marked in Fig. 4c respectively.

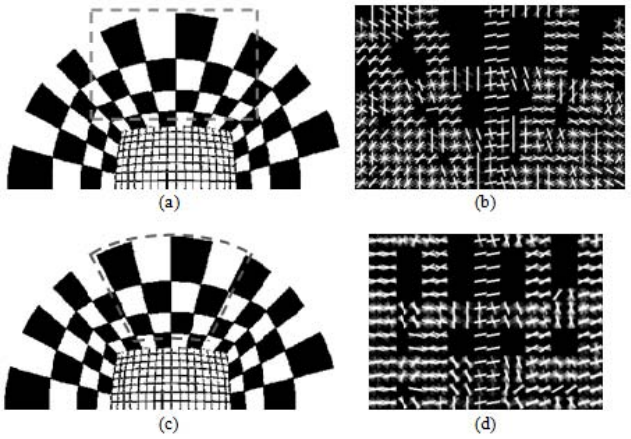


Fig. 4 Description of how the gradients are modified for an omnidirectional sliding window. Result in (b) is the regular HOG computed for the region marked with dashed lines in (a). Modified HOG computation gives the result in (d) for the region marked in (c). Vertical and horizontal edges in real world produce vertical and horizontal gradients in the modified version.

4 Experiments

Our experiments consider the detection of standing humans, cars and vans. For human detection, we trained a 128x64 model using INRIA person dataset as described in [7]. For car detection, we trained a 40x100 model using UIUC [1] and Darmstadt [15] sets together totalling 602 car side views. The model trained for van detection is 40x100 as well. For this object type, we constructed a database of 107 images containing vans viewed from either side. While training all object models, the number of the negative samples in the dataset were increased by collecting so-called 'hard-negatives'. These are the false-positive detections of the initial model that was trained with the original positive and negative samples.

4.1 Evaluation of the proposed HOG computation using synthetic omnidirectional images

Let us first compare the results of the proposed and the regular (unmodified) HOG computation. Since the computed HOG features are given to an SVM trained with an image dataset of corresponding object type, we aim to obtain higher SVM scores with the proposed omnidirectional HOG computation.

We artificially created 210 omnidirectional images containing humans, following an approach similar to [12]. Images in INRIA person dataset are projected to omnidirectional images using certain projection angle and distance parameters. Fig. 5 shows an example omnidirectional image, where the regular HOG window (rectangular, 128x64 pixels) and the proposed omnidirectional HOG window (non-rectangular) are shown. The HOG features computed with the two window types are compared with their resultant SVM scores. Since the locations of projections in these images are known, no search is needed for this experiment. However, vertical position of the window affects the result. For both approaches, we chose the position that gives the highest mean SVM score. Table 1 summarizes the result of the comparison, where we see that the mean score (also minimum, maximum and quartiles) for the proposed approach is higher than that of regular HOG window.

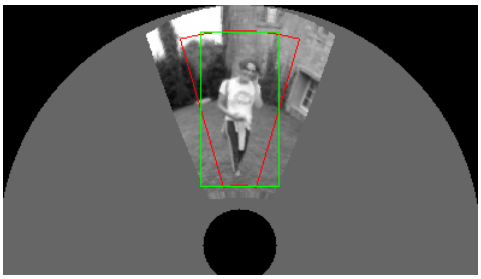


Fig. 5 Depiction of the regular HOG window (green rectangle) and the proposed window (red doughnut slice) on an omnidirectional image artificially created by projecting a perspective image from INRIA person dataset.

Table 1 Comparison of the regular and proposed HOG window by their SVM scores for human detection

	Min. score	Lower quart.	Mean score	Upper quart.	Max. score
Regular HOG	-1.01	1.16	1.69	2.20	3.21
Proposed HOG	-0.42	1.51	1.93	2.45	3.64

Table 2 Comparison of the regular and proposed HOG window by their SVM scores for car detection

	Min. score	Lower quart.	Mean score	Upper quart.	Max. score
Regular HOG	-1.81	-0.38	-0.09	0.24	1.17
Proposed HOG	-1.55	-0.17	0.19	0.55	1.79

For synthetic car images, 602 perspective car images from UIUC [1] and Darmstadt [15] datasets are projected to omnidirectional images. 40x100 pixel regular HOG computation and the proposed non-rectangular HOG window are compared in Table 2. The result is in accordance with the human detection experiment: mean SVM score, together with minimum, maximum and lower/upper quartiles, for the proposed approach is higher than the regular method.

4.2 Experiments of human detection in real images

In this subsection, we present the results for a set of images taken with our catadioptric omnidirectional camera. We compared the proposed HOG computation not

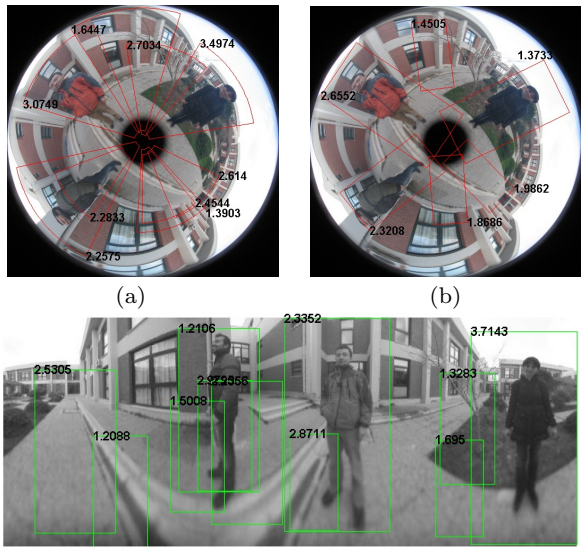


Fig. 6 Human detection results on an omnidirectional image with SVM scores (at upper left corners) greater than 1. (a) Proposed sliding windows. (b) Regular sliding and rotating windows. (c) Regular sliding windows on panoramic image.

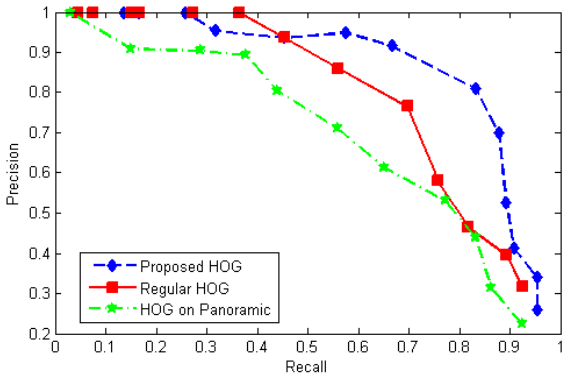


Fig. 7 Precision-Recall curves to compare the proposed HOG computation, the regular HOG and HOG after panoramic conversion approaches for human detection. The data points in the curve correspond to the varying threshold values for the SVM score, which change from 0 to 5. As the threshold increases, all approaches reach Precision=1.

only with the regular HOG window, but also with the approach that first converts the omnidirectional image to a panoramic image and then performs regular HOG computation. Although it was explained in Section 2 that working on panoramic images is not a theoretically correct approach, if the performance of detection on panoramic image is high it can still be considered as an option for practical applications.

Fig. 6 shows the results for one of the images in the dataset. Positive detections, after non-maximum suppression, are superimposed on the images with the proposed HOG window, the regular HOG window on omnidirectional image and HOG after panoramic conversion. The corresponding SVM score of each window is given at the upper left corner. Since a fixed size object (128x64) is searched in gradually resized versions of the original image, different sizes of detection windows seen in the figure correspond to detected objects in different scales. Since the feet of the body is very close to the blind spot of the camera and 128x64 human object model has a 16-pixel margin around the body, the best scoring windows usually exceed to the blind spot. The motion of the omnidirectional sliding window is based on polar coordinates. Each time, it turns by a fixed angle around the center and when the circle is completed, radius is changed. For the proposed HOG window, 64 is the length of the center arc and 128 is the thickness of the doughnut slice. For a fair comparison, the number of windows checked is equalized for all three approaches.

For the humans in Fig. 6, the average SVM scores for the proposed HOG, the regular HOG and HOG on panoramic image approaches are 2.94, 2.11 and 2.41 respectively. To evaluate the overall performance of these three approaches, we plot precision-recall curves for the whole dataset which consists of 30 real omnidirectional images taken in different scenes including indoor and outdoor environments (Fig. 7). A total of 66 humans were annotated in these images. The larger the area under the curve, the better the performance of the algorithm. One can observe that the performance of the proposed HOG computation is better than the others. Only for a limited range regular HOG performs better. When recall >0.5, the proposed approach is distinctively superior.

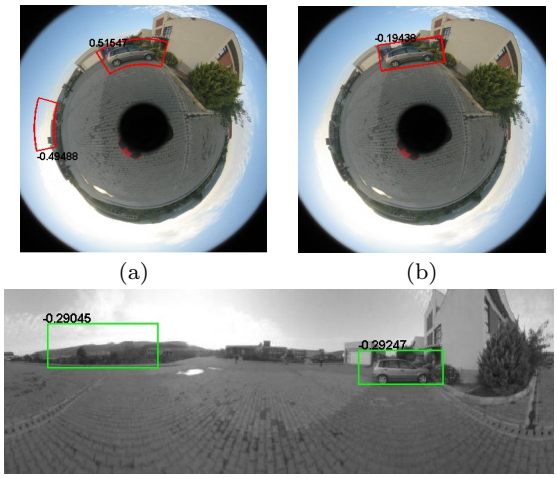
A detection window is considered to be a True-positive if it overlaps an annotation by 50% (following the advice in [9]), where the overlap is computed as

$$O = \frac{area(detectionwindow \cap annotation)}{area(detectionwindow \cup annotation)} \quad (7)$$

For a fair comparison, the annotations are separately prepared for the mentioned three methods. Annotations of the proposed HOG approach are doughnut slices (e.g. Fig.6a), annotations of the regular HOG approach are rectangles rotating around the omnidirectional image center, and annotations of HOG on panoramic image approach are upright rectangles. While annotating, a margin is left around the object to be in accordance with the training set images.

4.3 Experiments of car detection in real images

We repeated the comparisons between the evaluated methods for car detection. Fig. 8 shows the results for a single image as an example. For the overall performance



(a) (b) (c)

Fig. 8 Results of car detection on an omnidirectional image with SVM scores (at upper left corners) greater than -0.5. (a) Proposed sliding windows. (b) Regular sliding/rotating windows. (c) Regular sliding windows on panoramic image.

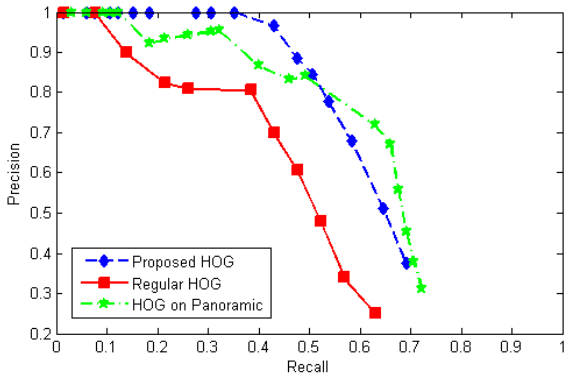
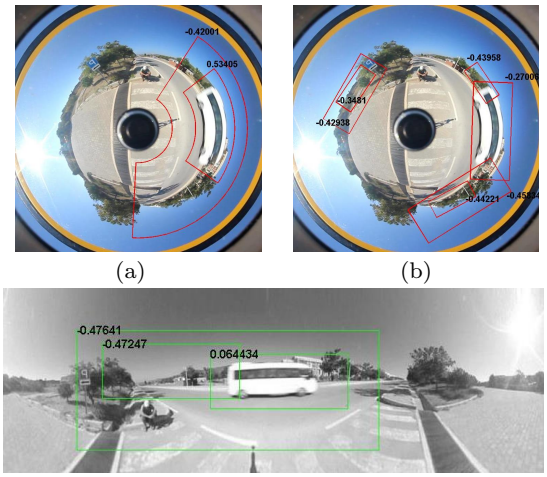


Fig. 9 Precision-Recall curves to compare the proposed HOG computation, the regular HOG and HOG after panoramic conversion approaches for car detection. The data points in the curve correspond to the varying threshold values for the SVM score, which change from -1.0 to 1.5.

comparison of the proposed HOG computation, the regular HOG computation and HOG after panoramic conversion approaches, we plot precision-recall curves (Fig. 9) for our dataset that includes 50 real images containing a total of 65 annotated cars.

When we compare the results in Fig. 9 with the ones in Fig. 7, one observation would be that now the proposed method is better than the regular HOG everywhere. This is due to the fact that car is a wider object than human. The regular HOG computation is affected more as the width/height ratio of the object model increases because it tries to fit a rectangle to the object in the omnidirectional image, which is bent more.

A second observation would be the increased performance of detection on panoramic image. It is now comparable to the proposed method. This can also be explained by the fact that car has a 'wide' model with



(a) (b) (c)

Fig. 10 Results of van detection on an omnidirectional image with SVM scores (at upper left corners) greater than -0.5. (a) Proposed sliding windows. (b) Regular sliding/rotating windows. (c) Regular sliding windows on panoramic image.

a width/height ratio of 2.5. It is harder for taller object models, like standing humans, to maintain the original width/height ratio in panoramic images. Since the panoramic image is created on a cylindrical surface rotating around the viewpoint, as we move down on the surface, same amount of viewing angle starts to cover a larger height in the image. This can be observed in the lower parts of Fig. 6c.

4.4 Experiments of van detection in real images

As a third object type, we performed experiments on van detection. Fig. 10 shows the results for a single image. For this image we observe that all three methods has a true-positive detection, however score obtained with the proposed method (Fig. 10a) is higher than the score obtained on panoramic image (Fig. 10c) which is relatively higher than the score with regular HOG on omnidirectional image (Fig. 10b). Precision-Recall curves in Fig. 11 show overall performance comparison for our dataset that includes 50 real images containing a van each. We used 57 other van images as a positive training set.

This time, the proposed approach is consistently better than HOG on panoramic approach. Regular HOG approach again has the worst performance since the vans we work on are wide objects similar to cars. One can also observe that Recall=1 can be reached for low thresholds for all three approaches. This is explained by the fact that test and training images are chosen from the same dataset that we built. However for car detection experiment, the training images were from a publicly available dataset.

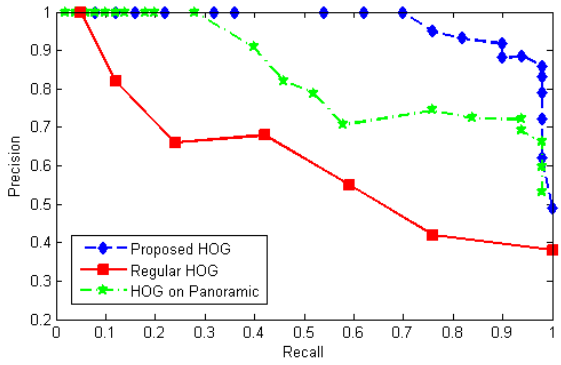


Fig. 11 Precision-Recall curves to compare the proposed HOG computation, the regular HOG and HOG after panoramic conversion approaches for van detection.

.

5 Conclusions

We aimed to perform object detection directly on the omnidirectional images. As a base, we took the HOG+ SVM approach which is one of the popular object detection methods. After describing how the feature extraction step of the conventional method should be modified, we performed experiments to compare the proposed method with the regular HOG computation on omnidirectional and panoramic images. Results of the experiments indicate that the performance of the proposed approach is superior to the regular approach. The performance of regular HOG on panoramic image is partially comparable to the proposed approach for objects that have high width/height ratio (such as cars). Having a high width/height ratio is an advantage for detection on panoramic image but a disadvantage for applying regular HOG on omnidirectional images. One should also note that the detection on panoramic images has the disadvantage of requiring image conversion beforehand.

In this work, we concentrated on HOG features for object detection. However, other features, especially the ones based on image derivatives can be modified in a similar fashion for a theoretically correct and effective use in omnidirectional cameras.

References

1. Agarwal, S., Roth, D., Learning a sparse representation for object detection, European Conference on Computer Vision (ECCV), (2002).
2. Amaral, F. H., Costa, A. H. R., Object Class Detection in Omnidirectional Images, Workshop de Visao Computacional (WVC), (2009).
3. Arican, Z., Frossard, P., OMNISIFT: Scale Invariant Features in Omnidirectional Images, IEEE 17th International Conference on Image Processing (ICIP), (2010).
4. Bogdanova, I., Bresson, X., Thiran, J.P., Vandergheynst, P., Scale Space Analysis and Active Contours for Omnidirectional Images. IEEE Transactions on Image Processing, 16(7), 1888-1901, (2007).
5. Bulow, T., Spherical diffusion for 3D surface smoothing, IEEE Transactions on PAMI, 25, 1650-1654, (2004).
6. Cinaroglu, I., Bastanlar, Y., A Direct Approach for Human Detection with Catadioptric Omnidirectional Cameras, IEEE Conference on Signal Processing and Communications Applications (2014).
7. Dalal, N., Triggs, B., Histograms of Oriented Gradients for Human Detection, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2005).
8. Daniilidis, K., Makadia, A., Bulow, T., Image Processing in Catadioptric Planes: Spatiotemporal Derivatives and Optical Flow Computation, International Workshop on Omnidirectional Vision (OmniVis), (2002).
9. Dollar, P., Wojek, C., Schiele, B., Perona, P., Pedestrian Detection: An Evaluation of the State of the Art, IEEE Transactions on PAMI, 34(4), 743-761, (2012).
10. Felzenszwalb, P., McAllester, D., Ramanan, D., A Discriminatively Trained, Multiscale, Deformable Part Model, Computer Vision and Pattern Recognition (CVPR), (2008).
11. Geyer, C., Daniilidis, K., A unifying theory for central panoramic systems and practical applications, European Conference on Computer Vision (ECCV), (2000).
12. Hansen, P. Corke, P., Boles, W., Daniilidis, K., Scale Invariant Features on the Sphere. IEEE International Conference on Computer Vision (ICCV), (2007).
13. Iraqui, A., Dupuis, Y., Boutteau, R., Ertaud, J., Savatier, X., Fusion of omnidirectional and PTZ cameras for face detection and tracking, International Conference on Emerging Security Technologies, (2010).
14. Kang, S., Roh, A., Nam, B., Hong, H., People detection method using GPUs for a mobile robot with an omnidirectional camera, Optical Engineering 50(12), 127204, (2011).
15. Leibe, B., Leonardis, A., Schiele, B., Combined object categorization and segmentation with an implicit shape model, Proc. of the Workshop on Statistical Learning in Computer Vision, (2004).
16. Lourenco, M., Barreto, J.P., Vasconcelos, F., sRD-SIFT: Keypoint Detection and Matching in Images with Radial Distortion, IEEE Trans. on Robotics, 28(3), (2012).
17. Lowe, D., Distinctive image features from scale invariant keypoints, International Journal of Computer Vision (IJCV), 60, 91-110, (2004).
18. Maji, S., Berg, A.C., Malik, J., Classification using Intersection Kernel Support Vector Machines is Efficient, Computer Vision and Pattern Recognition (CVPR), (2008).
19. Mei, C., Rives, P., Single Viewpoint Omnidirectional Camera Calibration from Planar Grids, in: Proc. of International Conference on Pattern Recognition (ICPR), (2007).
20. Puig, L., Guerrero, J. J., Scale Space for Central Catadioptric Systems: Towards a Generic Camera Feature Extractor, Int. Conf. on Computer Vision (ICCV), (2011).
21. Puig, L., Bastanlar, Y., Sturm, P., Guerrero, J., Barreto, J., Calibration of central catadioptric cameras using a DLT-like approach, International Journal of Computer Vision (IJCV), 93(1), (2011).
22. Puig, L., Guerrero, J. J., Daniilidis, K., Scale Space for Camera Invariant Features, IEEE Trans. PAMI, (2014).
23. Tang, Y., Li, Y., Bai, T., Zhou, X., Human Tracking in Thermal Catadioptric Omnidirectional Vision, Int. Conf. on Information and Automation (ICIA), (2011).
24. Walk, S., Majer, N., Schindler, K., Schiele, B., New Features and Insights for Pedestrian Detection, IEEE Conf. Computer Vision and Pattern Recognition, (2010).
25. Wang, M.L., Lin, H.Y., Object Recognition from Omnidirectional Visual Sensing for Mobile Robot Applications, IEEE Int. Conf. on Systems, Man and Cybernetics, (2009).