

A Real-Time Road Sign Detection Using Bilateral Chinese Transform

Rachid Belaroussi and Jean-Philippe Tarel

Université Paris Est
LEPSIS, INRETS-LCPC
58, boulevard Lefebvre 75732 Paris, France
rachid.belaroussi@gmail.com
jean-philippe.tarel@lcpc.fr
www.lcpc.fr

Abstract. We present a real-time approach for circular and polygonal road signs detection¹ in still images, regardless of their pose and orientation. Object detection is done using a pairwise gradient-based symmetry transform able to detect circles and polygons indistinctly. This symmetry transform of gradient orientation, the so-called Bilateral Chinese Transform BCT, decomposes an object into a set of parallel contours with opposite gradients, and models this gradient field symmetry using an accumulation of radial symmetry evidences. On a test database of 89 images 640x480 containing 92 road signs, 79 are correctly detected (86%) with 25 false positives using the BCT approach in about 30 ms/image.

1 Introduction

Most road sign detectors use a color modelling in a connected component segmentation further validated by an appearance-based model (template matching, genetic algorithm, SVM, neural networks, classifiers): these approaches are subject to oversegmentation and missed targets. The Radial Symmetry Transform **RST** was first used for road sign detection in [1]: greyscale images are used but only 60 mph or 40 mph circular speed signs are detected. This transform suffers from several impairments: it can only detect circular shapes, and it yields to a large number of false positives. Indeed, two additional steps are required: a template matching recognition stage and a temporal filter to validate consistent candidates (which is a way to replace the color information). To cope with the lack of generality of the circular shape, Loy et al [2] developed three specific versions of the RST for octagonal, square and triangular shapes respectively. This proposed approach lacks in generality and was only tested on 15 images for each shape. More extensive results can be found in [3] where a connected component segmentation (red signs only) is further validated by a Radial Symmetry Transform **RST**: they found the RST quite sensitive to missing edge points and pre-defined object scales. It also requires a template matching validation step to remove false positives.

¹ Part of iTowns-MDCO project funded by the Agence Nationale de la Recherche.

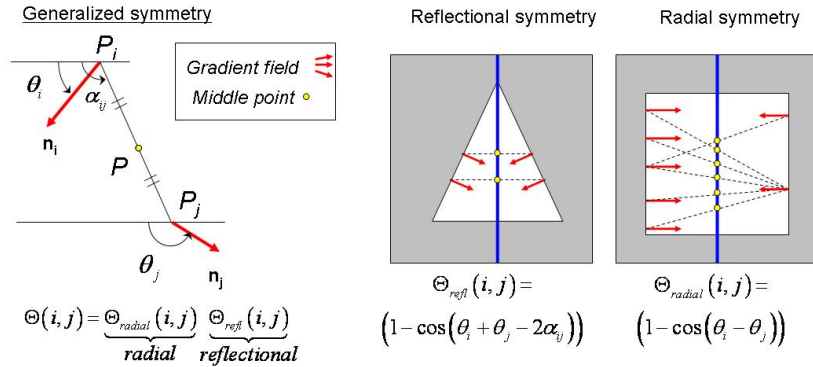


Fig. 1. Radial and reflectional symmetry between two gradient vectors \mathbf{n}_i and \mathbf{n}_j can be modelled using $(\theta_i, \theta_j, \alpha_{ij})$ parameters. In the Generalized Symmetry Transform, the symmetry magnitude is the product of two terms : Θ_{radial} and Θ_{refl} respectively for radial and reflectional symmetries.

The difficulty with RST is that it is a **mono-variate** transformation where any edge point casts its vote in several accumulators (to encompass multiple scales) independently of its neighbors. This induces a relatively high number of false positives. Introduced by Reisfeld [4], the Generalized Symmetry Transform (**GST**) uses an edge pairwise voting scheme particularly convenient for reflectional and radial symmetry detection. It is a kind of **bi-variate** Hough Transform, therefore it results in fewer false positives and it is more noise tolerant than the RST. An issue raised by the GST is that it is not specific enough as a pair of gradient vectors with no symmetry may have a significant contribution in the votes for a given point. The Chinese Transform **CT**, introduced by Milgram et al [5], avoids this impairment of the GST as demonstrated in [6]. We propose here a variant of the latter approach we named the Bilateral Chinese Transform **BCT** which focuses on pairs of gradient vectors with a strict radial symmetry. The proposed BCT is rather general and can also detect symmetry axis. In this paper, we test our BCT approach for circular and polygonal (not triangular) road sign detection in still images. Unlike the standard Chinese Transform, BCT makes no assumption on light/dark or dark/light contrast of objects to be detected. BCT works on intensity images and can detect circular, square and polygonal shapes regardless of their pose: it is an orientation free detector, i.e. robust to in-plane and small out-of-plane rotations.

We compare the performance of the standard Chinese Transform CT and the Bilateral Chinese Transform BCT in terms of road signs detection. No prior color model is used. Color information is required only to build the normalized red channel. It is then processed by a symmetry transform (CT or BCT) to detect road signs centers and their spatial extent.

We also compare BCT to a color segmentation based approach. This third algorithm is based on a red color model : regions of interest (red-like) image are processed to detect road signs centers with a standard Chinese Transform. Color

is further used to segment a road sign and to define its spatial extent using the Camshift [7] algorithm.

This paper is organized as follow: section 2 defines GST, CT and BCT operators and compare them for the task of road sign segmentation. Experimental results are discussed in section 3, and section 4 gives conclusion and perspectives of this work.

2 Radial and Reflectional Symmetry Transforms

2.1 The Original Approach of the Generalized Symmetry Transform

GST is a context-free operator, detecting symmetries if used at large scale or corners at small scale: different approaches are inspired from the GST (especially the approach of edge pairwise voting) more specific to a class of symmetry or an object shape.

We are interested in the symmetry magnitude defined by the GST: it is an accumulator of contribution of all pairs of edge points. For each point P of the image, a set of voters is defined:

$$\Gamma(P) = \left\{ (i, j) \mid \frac{P_i + P_j}{2} = P \right\} \quad (1)$$

The accumulator of the GST defining the symmetry magnitude at points P is:

$$Accu(P) = \sum_{(i,j) \in \Gamma(P)} C(i, j) \quad (2)$$

where $C(i, j)$ is the contribution of a pair of points (P_i, P_j) to the votes of its middle point P . $C(i, j)$ is a product of: a distance weighting function $D(i, j)$, a phase weighting function $\Theta(i, j)$ and a function r of the gradient magnitudes:

$$C(i, j) = D(i, j) \Theta(i, j) r_i r_j \quad (3)$$

r_i can be a logarithmic function of the gradient magnitude: $r_i = \log(\|\mathbf{n}_i\|)$. $D(i, j)$ is a distance weighting function decreasing with $\|P_i P_j\|$, typically:

$$D(i, j) = \frac{1}{\sqrt{2\pi\sigma}} \exp^{-\frac{\|P_i - P_j\|}{2\sigma}} \quad (4)$$

where σ depends on the target spatial extent. $\Theta(i, j)$ formula is given in fig. 1.

Several versions of the GST has been proposed with different distance or phase weighting functions, or a different feature space Γ depending on the application. GST has been applied to face and facial features detection [8], gait recognition or car license plate detection [9], and a noise tolerant version using negative votes [10] has recently been proposed. In the field of reflectional symmetry, [11] proposed to cast the vote of a pair of edge points (P_1, P_2) in the (r, θ) space (parameters of the line passing through the midpoint of $[P_1 P_2]$), using a Standard Hough Transform weighted by a function of the angle between edges orientation. Their approach is used to model the fingers and grasping gestures.

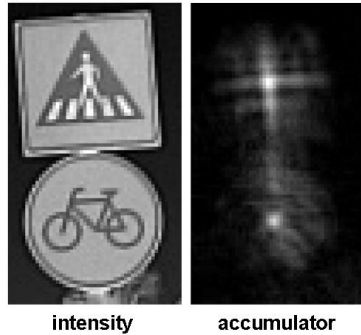


Fig. 2. Standard CT accumulators for a rectangle and a circular shape: both symmetry axis (horizontal and vertical) are detected, as well as the center of the circle

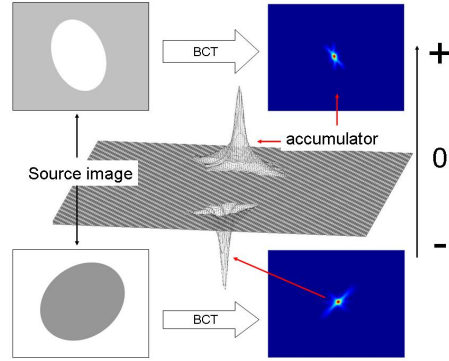


Fig. 3. Bilateral CT accumulator in ellipses cases: BCT is in-plane rotation free

2.2 The Chinese Transform CT

For the case of radial symmetry, another approach has been recently proposed, the so-called Chinese Transform (CT) [5], and successfully applied to facial features or fingers detection [6]. The Chinese Transform CT is related to the GST, yet it is quite different as it does not make use of an explicit reflectional symmetry formulation. The main distinction between these two transforms is the phase weighting function $\Theta(i, j)$. Like the GST, the feature space of the CT is the image, and a pair of edge pixels votes for their middle. The difference lies in the conditions required to form a pair. A special case of the CT is the Fast Circle Detection FCD [12] for iris detection. The CT, like the FCD, uses the assumption that the object to detect is bright on a dark background, so that the pairs of edge points to consider are reduced to couples (P_1, P_2) with gradient pointing at each other. However, this assumption of convergent gradients is verified in the case of iris or eyes (generally), but not in the case of road sign detection.

The Chinese Transform only makes use of a radial symmetry term, as the FCD, but with a condition on the alignment of gradient vectors taking into account possible reflectional symmetry. As for the GST, a scale parameter defines the area of influence of each pixel.

$$\Theta(i, j) = \underbrace{W_\beta(\theta_i - \alpha_{ij})}_{\text{Restricted alignment}} \underbrace{W_\delta(|\theta_i - \theta_j| - \pi)}_{\text{Radial symmetry}} \tag{5}$$

where W_R is a top-hat function with a length of R :

$$W_R(x) = 1 \quad \text{if} \quad |x| < R, \quad 0 \quad \text{elsewhere} \tag{6}$$

δ is a small quantity ensuring that numerically $|\theta_i - \theta_j| = \pi$, so that only edge points with opposite direction (radial symmetry term) cast their vote to their

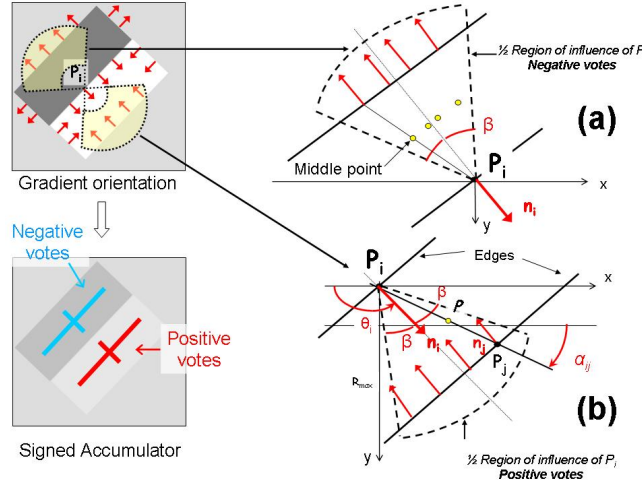


Fig. 4. Bilateral Chinese Transform. (a) In the case of a dark object on a clearer background, a pair of edge points with opposite (and diverging from each other) gradient orientation ($\mathbf{n}_i, \mathbf{n}_j$) results in a negative contribution in the accumulator. (b) A pair of edge points with a radially symmetric gradient orientation ($\mathbf{n}_i, \mathbf{n}_j$) pointing towards each others results in a positive vote for their middle point (bright/dark rectangle).

middle point. The first term of the phase weight is actually a condition on the alignment of $\mathbf{P}_i\mathbf{P}_j$ and \mathbf{n}_i . For $\beta = 0$ the CT is equivalent to the FCD algorithm. As shown in fig. 2, the CT can model both circular and polygonal shapes. However, it supposes that the object to detect has light/dark contrast.

An overview of the image processing used to feed the CT symmetry transform is given in fig. 5. The spatial extent of objects with high symmetry magnitude makes use of a second accumulator as explained in the next section.

2.3 Bilateral Chinese Transform BCT

To take into account the case of object with divergent edge gradients, see fig. 3 and fig. 4, a modification of the phase weighting function of the CT is proposed:

$$\Theta(i, j) = \left(\underbrace{W_\beta(\theta_i - \alpha_{ij})}_{\text{light/dark}} - \underbrace{W_\beta(\theta_i - \alpha_{ij} - \pi)}_{\text{dark/light}} \right) \times \underbrace{W_\delta(|\theta_i - \theta_j| - \pi)}_{\text{Radial symmetry}} \quad (7)$$

where δ is the precision with which the assertion $|\theta_i - \theta_j| = \pi$ is numerically verified: $\delta = \frac{2\pi}{N}$ if the gradient orientation is quantified on N values. α_{ij} is the orientation of vector $\mathbf{P}_i\mathbf{P}_j$ with respect to the horizontal axis (see convention on figure 1). β is a limit angle between \mathbf{n}_i and $\mathbf{P}_i\mathbf{P}_j$. It is the tolerance on the alignment of gradient vectors \mathbf{n}_i and \mathbf{n}_j and it delineates the region of influence of P_i point as shown in figure 4. A small value of β would fit to strictly circular

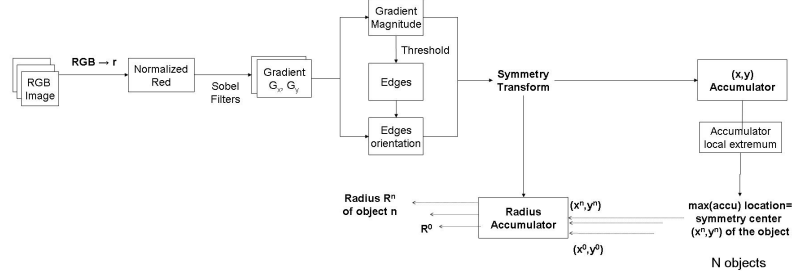


Fig. 5. Overview diagram (CT or BCT). The color image is converted in the red normalized channel which is further processed by the symmetry transform (BCT or CT) to detect symmetrical objects (both center and spatial extent).

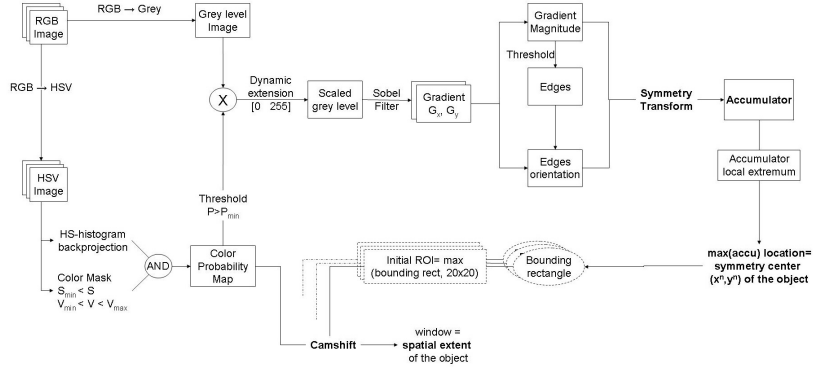


Fig. 6. CT+Camshift processing. A red color model is built in the Hue-Saturation plane of HSV. It is backprojected onto the color image: grey level of red-like area processed by CT for object centers detection and a Camshift segmentation is performed.

shape whereas $\beta = \pi/2$ can be used for more general case yet increasing the number of voters and so the processing load of the algorithm and the number of false positives. Regarding the distance weighting function, we find it more convenient and faster to use thresholds on distance between points:

$$D(i, j) = W_{R_{max}}(\|P_i - P_j\|) - W_{R_{min}}(\|P_i - P_j\|) \quad (8)$$

where R_{min} and R_{max} are respectively the minimum and maximum width of a road sign. That is to say, we try to detect objects within the range of size $[R_{min}, R_{max}]$. The accumulator is incremented as a percentage of the product of $\Theta(i, j)$ and of the gradient magnitude functions $r_i = \log(1 + \|\mathbf{n}_i\|)$:

$$Accu(P) = \sum_{(i,j) \in \Gamma(P)} D(i, j) \Theta(i, j) r_i r_j \quad (9)$$

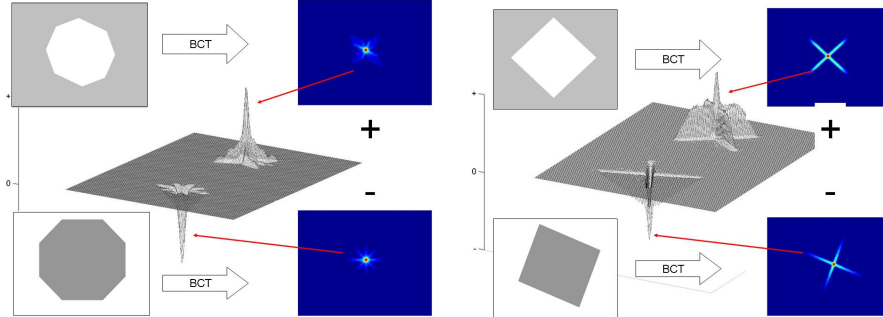


Fig. 7. Bilateral Chinese Transform in cases of polygonal shapes: source image and accumulator. Light/dark object edges (top) results in a positive contribution to the accumulator; the dark/light polygon at bottom has negative votes in the accumulator. The case of out-of-plane rotation is illustrated on the right with a diamond (top) and a rectangle (bottom).

For the case of polygonal shape with four or more sides, we use $\beta = \frac{\pi}{8}$ in our experiments, gradient orientation being quantized on 8 direction (Freeman code).

As shown in figure 5, the BCT transform has two output accumulators, one for the center of objects and the other one recording their radius. The center of a target can be extracted from the accumulator 9. To define the corresponding spatial extent, an another accumulator registering the distance between voters $\|P_i - P_j\|$ is built during the voting process:

$$Radius(P) = \frac{1}{N} \sum_{(i,j) \in \Gamma(P)} \|P_i - P_j\| / 2 \quad (10)$$

with $N = Card(\Gamma(P))$. The spatial extent is then an arithmetic average of the set of voters, and therefore the scale is not discretized. That is an advantage over methods like [1] using a multi-scale approach (one accumulator per scale).

3 Experimental Results

3.1 A Pose Free Detector for Circle and Polygons

The approach of the BCT dispenses with the bright/dark contrast assumption made in the CT by enabling the object to detect to have dark/bright or bright/dark contrast. Furthermore, as mentioned in [10], the use of a signed accumulator strengthens the BCT by making it noise tolerant. Figure 3 illustrates the signed accumulators obtained with the BCT over two elliptical shapes: it is worth noticing that the BCT is also orientation free. This can especially be seen in figure 7 illustrating the case of polygons.

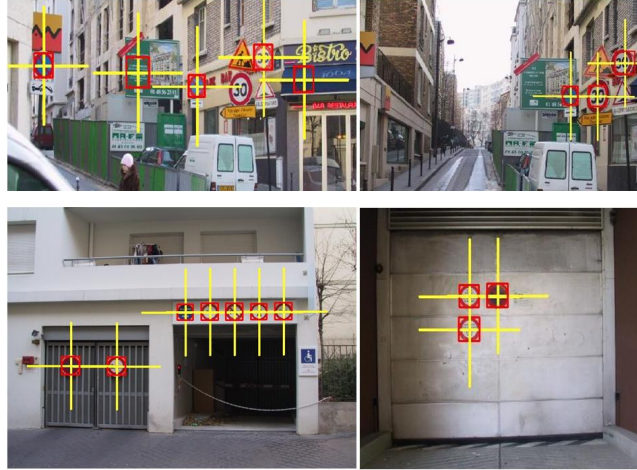


Fig. 8. Road sign segmentation examples using the Bilateral Chinese Transform

3.2 Performances

We used a database of 89 images of streets of Paris with cluttered background (see fig. 8), containing 81 circular red signs and a total of 92 road signs (11 blue or yellow signs, rectangular or circular). Both the CT and the BCT algorithms are tested in the same condition using the same image processing steps as illustrated in fig. 5. Compared to the BCT algorithm, the CT algorithm is less efficient, as can be seen in curves of fig. 9. Figure (a) plots the correct detection rate CDR versus the false detection rate FDR, and figure (b) plots the dice coefficients :

$$\text{CDR} = \frac{\text{TP}}{\text{P}} \quad \text{FDR} = \frac{\text{FP}}{\text{Number of images}} \quad \text{Dice coeff.} = \frac{2\text{TP}}{\text{TP} + \text{FP} + \text{P}}$$

where TP = Number of True Positive, FP = Number of False Positives and P= Total Number of signs. Table 1 indicates the performances of the algorithms for a given point of the ROC curves: the BCT is able to detect 79 signs out of 92 with 25 false positives whereas the standard CT could only detect 69 road signs with 24 false positives. As the normalized red channel is processed, red signs are likely to have a light/dark contrast and can be detected by a standard CT. Road signs with another color, especially blue signs, are likely to have a dark/light contrast: most of them are detected by the BCT unlike to a standard CT. The dice curve achieves a maximum value of 82% for the BCT compared to a maximum value of 77% for the CT : with a difference of 5% in the dice coefficient, the BCT significantly improves the CT algorithm.

We also compared the BCT with a color model based approach illustrated in fig. 6: a red 2D-histogram in the Hue-Saturation plane is built from an external database. The red color model is backprojected to find regions of interest (ROI). Gradient field of these ROI is computed and feeds the input of a standard Chinese

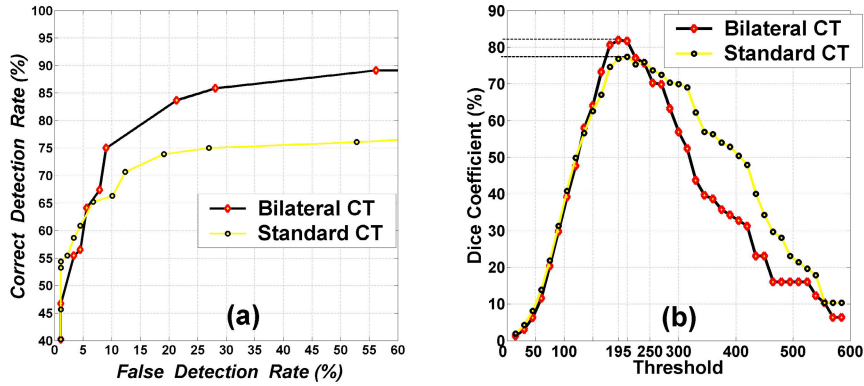


Fig. 9. Comparison of the performances of Bilateral CT (black) and Standard CT (yellow), over a test set of 89 images with 92 signs. (a) ROC curves plotting the Correct detection rate (%) versus the False positive rate. (b) Dice coefficient (%) versus accumulators threshold.

Transform. The spatial extent of objects with high symmetry magnitude are further segmented using the Camshift [7]. The CT+Camshift algorithm is faster but misses more road signs than the BCT as reported by table 1. Amongst the 81 red signs, 75% are detected with 22 false positives, in about 10ms/frame on a standard PIV@1,2GHz laptop. This illustrates the weakness of the color selection since the red color model is built with a different camera than the one used for the test bench. A lot of red signs do not have the colors of the H-S histogram, and cannot be detected.

The BCT on the other hand avoids the issue of color constancy by using the normalized red channel: $r = \frac{R}{R+G+B}$. This channel is less dependant on illumination conditions. Moreover, it enables us to detect signs with any color as a blue is likely to appear darker than its background, and a red is usually brighter. Also yellow and green signs would generally be contrasted enough to be detected in this channel.

Table 1. Road sign detection performances on a 89 urban images database

Approach	Correctly detected	Number of false positives	Number of targets	Processing Time (avg)
Bilateral Chinese Transform	79 (86%)	25	92	30 ms/frame
Standard Chinese Transform	69 (75%)	24	92	30 ms/frame
CT with Camshift on red signs	61 (75%)	22	81	10 ms/frame

4 Conclusion and Perspectives

We presented a road sign segmentation system using an efficient symmetry detector, the Bilateral Chinese Transform. It can process a 640x480 image in about 30 ms, with a high detection rate: 86% with 25 false positives over a set of 89 images containing 92 road signs with different colors and shapes (except triangular signs). It performs better than the standard CT which detects only 75% of the signs with 24 false positives. We compared it to an approach using a color segmentation step: the BCT is a bit slower, but it is more efficient and it is not limited to a particular color. The BCT algorithm is more general than the RST as it can detect either circle, square, diamond and polygons. It is also more precise than the GST because it focuses exclusively on pairs of points with radial symmetry. The next step in our approach is to add a recognition stage in order to remove the remaining false positives that usually occur on symmetrical object like cars, windows, logos, signboard, or pedestrians. The BCT is less adequate in the case of triangular signs. Indeed, the triangle center location is quite fuzzy using this approach. To cope with this issue, we developed a specific geometrical transformation for this case [13].

References

1. Barnes, N., Zelinsky, A., Fletcher, L.: Real-time speed sign detection using the radial symmetry detector. *Transactions on Intelligent Transportation Systems* (2008)
2. Loy, G., Barnes, N.: Fast shape-based road sign detection for a driver assistance system. In: *Intelligent Robots and Systems, IROS* (2004)
3. Foucher, P., Charbonnier, P., Kebbous, H.: Evaluation of a Road Sign Pre-detection System by Image Analysis. In: *VISAPP* (2), pp. 362–367 (2009)
4. Reissfeld, D., Wolfson, H., Yeshurun, Y.: Context Free Attentional Operators: the Generalized Symmetry Transform. *Int. J. of Computer Vision* (1995)
5. Milgram, M., Belaroussi, R., Prevost, L.: Multi-stage Combination of Geometric and Colorimetric Detectors for Eyes Localization. In: Roli, F., Vitulano, S. (eds.) *ICIAP 2005*. LNCS, vol. 3617, pp. 1010–1017. Springer, Heidelberg (2005)
6. Belaroussi, R., Milgram, M.: A Real Time Fingers Detection by Symmetry Transform Using a Two Cameras System. In: Bebis, G., Boyle, R., Parvin, B., Koracin, D., Remagnino, P., Porikli, F., Peters, J., Klosowski, J., Arns, L., Chun, Y.K., Rhyne, T.-M., Monroe, L. (eds.) *ISVC 2008, Part II*. LNCS, vol. 5359, pp. 703–712. Springer, Heidelberg (2008)
7. Bradsky, G.: Computer Vision Face Tracking For Use in a Perceptual User Interface. *Intel Technology Journal* (1998)
8. Hayfron-Acquah, J.B., Nixon, M.S., Carter, J.N.: Automatic gait recognition by symmetry analysis. *Pattern Recognition Letters* (2003)
9. Kim, D.S., Chien, S.I.: Automatic car license plate extraction using modified generalized symmetry transform and image warping. In: *ISIE* (2001)
10. Park, C.-J., Seob, K.-S., Choib, H.-M.: Symmetric polarity in generalized symmetry transformation. *Pattern Recognition Letters* (2007)
11. Li, W.H., Zhang, A.M., Kleeman, L.: Fast Global Reflectional Symmetry Detection for Robotic Grasping and Visual Tracking. In: *ACRA* (2005)
12. Rad, A.A., Faez, K., Qaragozlou, N.: Fast Circle Detection Using Gradient Pair Vectors. In: *Digital Image Computing: Techniques and Applications* (2003)
13. Belaroussi, R., Tarel, J.-P.: Angle Vertex and Bisector Geometric Model for Triangular Road Sign Detection. In: *Workshop on Applications of Computer Vision* (2009)