
Détection des panneaux de signalisation routière par accumulation bivariée

Rachid Belaroussi, Jean-Philippe Tarel

Université Paris Est – LEPSIS, INRETS-LCPC
85, boulevard Lefebvre, 75732 Paris cedex 15, France
rachid.belaroussi@gmail.com, jean-philippe.tarel@lcpc.fr

RÉSUMÉ. Nous présentons une méthode géométrique utilisant l'orientation du gradient pour la détection de la signalisation verticale dans des images fixes, indépendamment de leur position et de leur orientation. La détection est réalisée par une transformation de type accumulateur de Hough bivariée, fondée sur l'utilisation de paires de points avec des contraintes sur leurs gradients. Les panneaux circulaires et polygonaux (non triangulaires) sont détectés par la transformation chinoise bilatérale TCB. Cette transformation est rapide et ne fait pas de distinction entre les cercles et les polygones 4 côtés ou plus. Le cas des panneaux triangulaires est traité par la transformation en sommet et bissectrice TSB, capable de détecter précisément les bissectrices et les sommets d'un triangle. Les performances de la TCB et de la TSB sont estimées sur plusieurs bases d'images de scènes urbaines.

ABSTRACT. We present a real-time approach for circular and polygonal road signs detection in still images, regardless of their pose and orientation. Object detection is performed using a Hough like transform based on accumulation on pairs of points with their gradients. Circular and polygonal (non triangular) traffic signs are detected by the so-called Bilateral Chinese Transform BCT. Triangular traffic signs are processed by the Vertex Bisector Transform VBT. These two approaches are tested on several databases of urban scene.

MOTS-CLÉS : transformation en symétrie généralisée, transformation en symétrie pour les coins, transformation en sommet et bissectrice, transformation chinoise, contours orientés, panneaux routiers, temps réel, transformée de Hough.

KEYWORDS: generalized symmetry transform, corner symmetry, vertex and bisector transform, bilateral chinese transform, oriented edges, traffic sign, real-time, Hough transform.

DOI:10.3166/TS.27.265-296 © 2010 Lavoisier, Paris

Extended abstract

Road sign identification in images is an important issue, especially for vehicle safety and road management applications. It is usually tackled in three stages: detection, recognition and tracking, and evaluated as a whole. To progress towards better algorithms, we focus in this article on the first stage of the process, namely road sign detection. We focus our work on a feature-based approach to build geometrical models of various kind of shapes: triangle on the one hand, square, diamond, octagon and circle on the other hand.

We introduce two cumulative bivariate transformations:

- the Bilateral Chinese Transform BCT, tailored for the detection of circular and polygonal road signs with four or more sides.
- The Vertex and Bisector Transformation VBT, dedicated to the detection of triangular traffic signs.

In the BCT approach (Belaroussi *et al.*, 2009b), object detection is done using a pairwise gradient-based symmetry transform able to detect circles and polygons indistinctly. The BCT seeks the symmetry of shape in the gradient orientation: it 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. Figure 7 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 8 illustrating the case of polygons.

To illustrate the prevalence of a geometrical model over color based methods, we first perform a test on circular and rectangular road signs. On a test database of 89 images 640×480 containing 92 road signs (81 are red, the remaining 11 are blue or yellow), 79 are correctly detected (86%) with 25 false positives using the BCT approach in about 30 ms/frame. On the same database, 61 road signs (75% out of the 81 red signs) are correctly detected by the color-based detector with 22 false positives in about 10 ms/frame. Therefore the BCT is slower than a color segmentation approach but it is much more efficient, with about the same amount of false positive.

The BCT approach dispenses with the bright/dark contrast assumptions made in the CT by enabling the object to be detected to have dark/bright or bright/dark contrast. We show that the unsigned version of the BCT performs better than the signed BCT, because it can handle the case of road signs with mixed contrast (dark/bright and bright/dark). We also investigate the impact of image gradient estimation in the *normalized RGB* colorspace. Gradient orientation is estimated using the red channel. Using a mixture of normalized red and blue channels to define the gradient magnitude brings better results than using only the red channel, as the blue signs edges are given more weight when casting their votes.

The VBT (Belaroussi *et al.*, 2009a) is a new transformation for angle vertex and bisector detection. It takes the image gradient as input and outputs two arrays, accumulating evidence of respectively angle vertex and angle bisector. A geometric model of the gradient orientation is implemented using a pair-wise voting scheme: normal vectors of two adjacent sides of a triangle have a specific relationship depending on the corresponding vertex angle. We have illustrated the results of the

VBT on some synthetic images, with the two output accumulators. For instance, in figure 2, two cases of triangles are shown with light/dark or dark/light contrast. In both cases, we can see that the VBT is almost invariant to in-plane rotation. The VBT is also quite robust to vertex occlusion and missing edges as shown in figure 2. Indeed, set of voters are taken from the sides of a triangle so if a consistent portion of the triangle edges is still visible, its vertices can be found even if they are occluded.

The BCT and VBT approaches are compared to a monovariate cumulative method (radial symmetry transform and Hough transform) in (Belaroussi *et al.*, 2010a). We estimate the performances of these detectors on three different databases of respectively 48 images 360×270 , 89 images 640×480 and 847 images 480×540 . They are evaluated in terms of detection rate, false positive per image and Dice coefficient. On a test set of 847 images¹, the BCT is able to detect 95% of the 21 circular or rectangular road signs for 5% of False Positive Per Image (FPPI). On the same set containing 10 triangular signs, the VBT has a correct detection rate of 100% for a FPPI of 1%. These results are consistent with the state of the art in this field. Furthermore, these two pose free detectors can be used in conjunction with a recognition stage especially in the case of videos, which typically requires a detection rate higher than 90%. The BCT can process efficiently a 480×540 image in an about 30 ms. The VBT is able to accurately detect vertices and bisectors of triangular road signs in a 480×540 image in about 67 ms. On a 2.3GHz Pentium IV with 3Go RAM, the average processing time of a 480×540 image is 30 ms for the BCT whereas the VBT is able to accurately detect vertices and bisectors of triangular road signs in about 67 ms.

The influence of a degradation of the image acquisition is studied on a test set of 40 images. The tolerance to an additive white Gaussian noise with a level of up to -32 db (gaussian standard deviation ranging from 0 to 40) is investigated. To analyze the influence of blurring on the performances of the detectors, each image is convolved with a Gaussian smoothing filter with a standard deviation ranging from 0 to 5 (31×31 filter size). Both algorithms obtain more than 80% of correct detection rate at a noise level of -20 dB ($\sigma = 10$). The BCT detection rate is almost invariant to the level of blurring, whereas the VBT keeps more than 90% of correct detection rate for Gaussian filters size of up to 11×11 .

1. Introduction

Pour l'identification des panneaux routiers, la plupart des méthodes par traitement d'image se composent de deux étapes : la détection des panneaux dans la scène routière et la reconnaissance de leur type. Dans le cas d'un flux vidéo, un algorithme de poursuite de cible est généralement ajouté. Les systèmes d'identification des panneaux de signalisation routière sont généralement évalués globalement. Cette démarche ne permet malheureusement pas d'analyser finement les performances de chaque étape. Il est donc difficile de savoir si c'est l'étape de

1. This image database is publicly available at <http://www.itowns.fr/benchmarking> with its hand labeled ground truth.

détection ou l'étape de reconnaissance qu'il convient d'améliorer afin d'obtenir un système plus performant. Nous nous focalisons ici sur la détection rapide et efficace des panneaux routiers, ceux de police (IISR-R-1-6, 2008) en particulier. En effet, la détection est souvent une tâche plus difficile que la reconnaissance. Par ailleurs, la qualité de la détection conditionne la qualité des résultats de reconnaissance. Notons que les travaux présentés dans cet article ne prennent pas en considération le cas des panneaux de direction, dont la détection doit combiner détection de textes et de panneaux.

1.1. Applications de la détection de la signalisation

De nombreuses applications nécessitent une localisation précise des panneaux, particulièrement en milieu urbain. Des millions de kilomètres étant à traiter, et la signalisation n'étant pas constante dans le temps, une automatisation de cette tâche est nécessaire pour les applications suivantes :

- relevé cartographique, notamment pour les outils multimédias de navigation 3D immersive (Géoportail IGN, Google StreetView...),
- diagnostic et gestion de l'infrastructure,
- aide à la navigation en complément du GPS,
- assistance à la conduite en lien avec la robotique.

En particulier pour les systèmes de transports intelligents de nouvelle génération, où une communication véhicule-infrastructure ou véhicule-véhicule est présente, l'usager ainsi que le gestionnaire de la route ont le besoin de connaître l'évolution de l'état de la signalisation en temps réel.

1.2. Les méthodes de détection de panneaux

Les algorithmes de détection de panneaux dans une image fixe peuvent être rangés selon les trois catégories suivantes :

- modélisation colorimétrique : une segmentation en composantes connexes fondée sur un modèle de couleur est opérée. Les régions d'intérêt sont ensuite validées par un algorithme de reconnaissance ou un modèle d'apparence. Ces méthodes sont les plus rapides mais aussi les moins robustes aux variations des conditions d'éclairage ;
- modélisation géométrique : les contours de l'image sont analysés par une approche structurelle ou globale. Ces méthodes sont généralement plus robustes que celles photométriques parce qu'elles traitent le gradient de l'image, et peuvent traiter des images en niveaux de gris ;
- méthodes avec apprentissage : un classifieur (cascade, SVM, réseaux de neurones) est entraîné sur une base d'exemples. Il est appliqué sur une fenêtre glissante qui parcourt l'image à plusieurs échelles. Ces méthodes combinent géométrie et photométrie mais peuvent être une étape coûteuse en temps de calcul. Elles nécessitent la constitution d'une base d'apprentissage par type de panneaux, étape fastidieuse lorsque le nombre d'objets à détecter est grand.

Un grand nombre d'approches utilisent une segmentation s'appuyant sur un modèle de la couleur principale du panneau (de la Escalera *et al.*, 2003), les candidats étant ensuite filtrés par une étape de reconnaissance (algorithmes génétiques, ACP, *template matching* à différentes positions et échelles) ou un algorithme de poursuite dans le temps. Malheureusement cette approche est sujette à d'éventuelles sur-segmentations et sous-segmentations qui induisent des cibles manquées. Ces méthodes exploitent un seuillage dans l'espace HSV et plus particulièrement sur le canal *Hue*, la teinte étant supposée invariante aux changements de luminance (de la Escalera *et al.*, 2003), (Shaposhnikov *et al.*, 2002), (Paclik *et al.*, 2002). Alternativement, le canal rouge $r = R/(R + G + B)$ de l'espace rgb normalisé (Dutilleul *et al.*, 2007) ou les rapports R/B et G/B peuvent être préférés (Soheilian *et al.*, 2010) (de la Escalera *et al.*, 1997). Dans (Broggi *et al.*, 2007) le choix s'est porté sur l'espace YUV après une correction chromatique préalable compte tenu des valeurs des pixels de la chaussée théoriquement gris ($R = G = B$). D'autres espaces tels que Lab (Reina *et al.*, 2006) ou un mélange d'espaces (Timofte *et al.*, 2009) sont utilisés mais le plus souvent la détermination du seuil est empirique. On trouve cependant des méthodes de seuillage adaptatif dans (Paulo *et al.*, 2007) ou (Vitabile *et al.*, 2001). La classification de la couleur d'un pixel par un réseau de neurones est aussi proposée dans (Ohara *et al.*, 2002) et par un SVM dans (Zhu *et al.*, 2006). L'approche que nous proposons n'opère aucune sélection sur la couleur : nous traitons simplement les canaux rouge et bleu normalisés pour détecter le centre des panneaux et évaluer leur taille.

Un exemple de détecteur photométrique est proposé dans (Timofte *et al.*, 2009). Un SVM est utilisé pour valider les régions d'intérêt provenant d'une sélection colorimétrique suivie d'une transformée de Hough pour les cercles. Le taux de détection obtenu est de 95 % pour 20 % de faux positifs par image. Une approche plus robuste est proposée par (Bahlmann *et al.*, 2005) : un détecteur à sélection automatique de filtres spatiaux (Adaboost) opérant sur différents espaces colorimétriques est utilisé, suivi d'une étape de *template matching*, puis d'un filtre temporel pour la détection et la reconnaissance de panneaux de limitation de vitesse dans une séquence d'images. Les taux de faux positifs et de faux négatifs est présenté sur une base de test d'images (1 700 exemples positifs et 40 000 négatifs), rendant difficile la comparaison avec d'autres approches. De plus, seuls les panneaux circulaires sont traités, et une base d'apprentissage de 4 000 exemples positifs est nécessaire. La cadence de traitement (détection, reconnaissance et poursuite) est de 100 ms/img, pour une image PAL 384×288 parcourue par 5 détecteurs (traitant des fenêtres glissantes carrées respectivement de taille 14, 20, 28, et 54 pixels).

Un exemple récent de détecteur fondé sur l'apprentissage est constitué d'une cascade attentionnelle de classifieurs sélectionnés par Adaboost (Baro *et al.*, 2009). La taille de la fenêtre d'entrée d'un détecteur est de 30×30 (taille minimale de panneaux détectables). Cela permet de détecter des panneaux circulaires et triangulaires : quatre détecteurs sont construits, modélisant respectivement les panneaux de Danger, Céder le passage, Interdiction et Obligation. Cette approche obtient les résultats suivants :

- cercle *Interdiction* : 70 % de détections pour 3.65 % de faux positifs par image,
- cercle *Obligation* : 60 % de détections pour 0.95 % de faux positifs par image,
- triangle *Danger* : 65 % de détections pour 2.25 % de faux positifs par image,

– triangle *Céder le passage* : 75 % de détections pour 2.8 % de faux positifs par image.

La cadence de traitement n'est pas donnée car l'algorithme est exploité hors ligne. Le cas des panneaux rectangulaires n'est pas abordé. L'approche que nous proposons obtient de meilleurs résultats en termes de détections, bien que ce soit sur des bases de test différentes.

Notons que dans (Baro *et al.*, 2009), l'étape suivante de reconnaissance de panneaux nécessite au préalable les paramètres de leur transformation affine. Une transformation en symétrie radiale est alors opérée pour les panneaux circulaires, tandis qu'une transformée de Hough associée à une détection de coins est utilisée pour déterminer les bords des triangles. Les performances en généralisation des modèles d'apparence sont limitées par la variété des exemples de la base d'apprentissage. Ainsi, (Ruta *et al.*, 2009b) comparent la détection de panneaux circulaires par transformée de Hough à celle obtenue par une cascade de classifieurs boostés et sélectionnent la transformée de Hough pour sa flexibilité et sa rapidité.

Dans le cas d'une approche géométrique, la détection de panneaux se fait à partir des contours de l'image. Ainsi dans (Fang *et al.*, 2003) des réseaux de neurones auto-associateurs, dont les poids sont fixés de manière à correspondre à la forme recherchée (cercle, triangle, octogone), servent de filtres non linéaires convoluant l'image en niveau de gris et le canal *Hue*. Un système d'inférences floues exploite les deux cartes obtenues pour détecter les panneaux. Ce traitement traite une image en 1 à 2 secondes, et est intégré dans un système de poursuite et de reconnaissance, mais les résultats ne sont présentés que sous la forme d'illustrations. La détection de limitation de vitesse est aussi assurée par une transformée de Hough par (Miura *et al.*, 2000), les panneaux rectangulaires d'information étant détectés par un seuillage colorimétrique dans YUV suivi d'une projection horizontale et verticale du gradient et d'une reconnaissance du Kanji. Les performances de cette approche sont seulement illustrées par quelques exemples.

(Garcia-Garrido *et al.*, 2006) utilisent une transformée de Hough pour détecter les cercles (limitation de vitesse ou Stop) ou les lignes pour les triangles. Leur approche sélectionne les contours fermés dans une certaine gamme de rapport largeur/hauteur, ce qui réduit la robustesse aux rotations hors-plan. Un triangle est détecté si un contour est formé de trois droites se coupant avec un angle de 60° (*Céder le passage* et *Danger*). La cadence de ces traitements varie de 20 ms/img à 200 ms/img en fonction du nombre de contours détectés. Les candidats détectés sont vérifiés par une reconnaissance de panneaux par réseaux de neurones. Ceux qui sont validés sont ensuite poursuivis dans le temps à l'aide d'un filtre de Kalman. Cette approche testée sur des vidéos permet de détecter, au moins une fois, 97 % des 435 panneaux de limitation de vitesse et 94 % des 312 panneaux de *Danger*, mais le nombre de faux positifs et le nombre total d'images ne sont pas révélés par les auteurs. Pour la signalisation triangulaire, dans (Piccioli *et al.*, 1996), des régions d'intérêt sont sélectionnées en s'appuyant sur un seuillage dans le plan *Hue-Saturation* de HSV. Des polygones sont ajustés aux chaînes de contours linéaires, et sous l'hypothèse que les triangles ne sont ni inclinés, ni distordus par la perspective, des triangles équilatéraux avec un côté horizontal et deux côtés de pente égale à $\pm\pi/3$ radians sont recherchés. Cet algorithme dépend largement de la segmentation par la couleur, et il est fortement contraint en termes d'orientation (il se restreint aux panneaux vus sous une perspective fronto-parallèle).

Dans le cas de séquences d'images en niveaux de gris, (Barnes *et al.*, 2003) utilisent la transformation en symétrie radiale TSR avec les mêmes étapes de validation pour la détection de panneaux de limitation de vitesse 40 et 60 miles/h. (Caraffi *et al.*, 2008) proposent une détection des signes de fin de limitation de vitesse dans des séquences d'images en niveau de gris. Leur approche consiste à détecter les transitions clair/sombre/clair : une transformée de Hough pour les cercles valide les candidats retenus, et un filtrage temporel est requis pour éliminer les nombreuses fausses alarmes. Ces étapes de filtrage ne sont pas nécessaires dans l'approche que nous proposons.

D'après notre analyse, la TSR est une transformation monovariée où chaque point de contour vote dans plusieurs accumulateurs (un par échelle) indépendamment de son voisinage. Cela induit un nombre relativement grand de faux positifs. Dans la transformation en symétrie généralisée TSG, introduite par Reisfeld (Reisfeld *et al.*, 1995), chaque vote est issu d'une paire de points. Cette transformation est particulièrement adaptée à la mise en évidence des symétries radiales et axiales. Elle agit comme une transformée de Hough bivariée, ce qui a pour conséquence de réduire le nombre de fausses alarmes et de la rendre moins sensible au bruit qu'une transformation monovariée.

La TSG n'est pas suffisamment sélective et il arrive encore que des paires de gradients sans symétrie particulière contribuent de manière significative aux votes en un point donné de l'accumulateur.

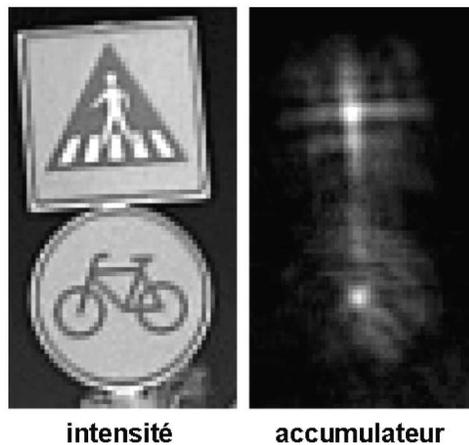


Figure 1. TC standard : image source et accumulateur

La transformation chinoise TC introduite par (Milgram *et al.*, 2008) focalise le processus de vote sur les paires de gradients ayant une symétrie centrale. Cette spécialisation corrige les inconvénients de la TSG comme démontré par (Belaroussi *et al.*, 2008). Elle est capable de détecter les axes et centres de symétrie des panneaux circulaires et rectangulaires comme le montre la figure 1. Nous proposons ici une variante de la TC appelée transformation chinoise bilatérale TCB (Belaroussi *et al.*, 2009b) pour la détection de panneaux circulaires et polygonaux (non triangulaires).

Contrairement à la TC standard, la TCB permet de détecter des objets sans hypothèse *a priori* sur leur contraste clair/sombre ou sombre/clair. Elle est invariante à la position et l'orientation de l'objet à détecter, et est robuste à des petites rotations hors du plan de l'image

Loy et Barnes (Loy *et al.*, 2004) proposent une adaptation de la transformation en symétrie radiale TSR aux panneaux triangulaires : chaque point de contour vote pour un segment perpendiculaire à son gradient et à une distance donnée. Le processus de vote est réalisé pour un ensemble fini de rayons et un accumulateur est finalement obtenu en sommant les contributions sur tous les rayons. Une étape de reconnaissance par corrélation croisée valide ou élimine les candidats obtenus. Leur approche est testée sur 15 images contenant 15 panneaux triangulaires : ils obtiennent un taux de détection de 100 % pour 10 faux positifs (66 % de faux positifs par image). Une TSR pour les octogones et une autre pour les carrés sont proposées. Sur 30 images contenant 15 octogones et 19 carrés, le taux de détection est de 95 % pour 15 faux positifs (10 % de faux positifs par image). Le temps de traitement est 50 ms/img pour une image 320×240 .

Une approche similaire est testée plus largement par (Foucher *et al.*, 2009) pour une application de prédétection de panneaux. Une segmentation fondée sur un seuillage colorimétrique produit un ensemble de candidats qui sont validés par une TSR. L'inconvénient majeur de la TSR pour les triangles est qu'elle peut seulement trouver la taille et la position des formes : elle ne peut distinguer un panneau de *Céder le passage* d'un panneau d'*intersection*. Notre approche est capable de détecter précisément l'orientation d'un panneau triangulaire car ses trois sommets sont détectés, avec moins de faux positifs que la TSR puisque cette dernière est une fonction monovariée du gradient.

L'approche proposée par (Cyganek, 2008) requiert, contrairement à notre méthode, une segmentation s'appuyant sur un seuillage dans le plan H-S de HSI. Un détecteur de coins, basé sur les caractéristiques locales binaires, traite les composantes connexes trouvées. La dernière étape consiste à filtrer, parmi tous les triplets de points saillants, ceux qui vérifient certaines contraintes de taille et

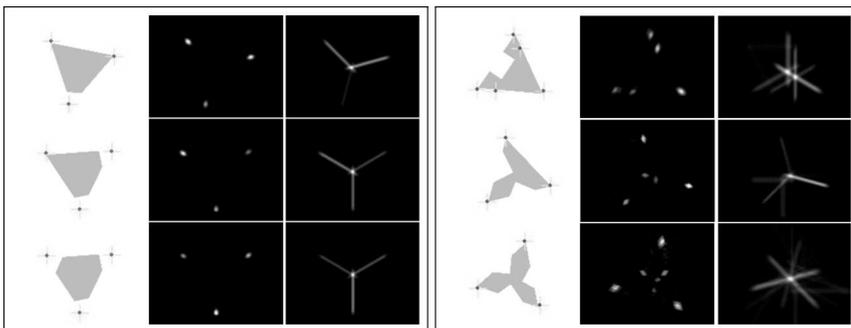


Figure 2. Traitement par la TSB de triangles (gris sur blanc) avec un, deux et trois sommets occultés, puis un, deux et trois côtés occultés. De gauche à droite : image originale avec les sommets détectés, l'accumulateur des sommets S_{accu} , et l'accumulateur des bissectrices B_{accu}

d'orientation. Un inconvénient de cette approche est que lorsqu'un sommet de triangle est occulté, le panneau ne peut être détecté. La figure 2 illustre quelques cas d'occultations traités par notre algorithme nommé transformation en sommet et bissectrice TSB : les sommets détectés par la TSB sont représentés par une croix jaune dessinée sur l'image en intensité. On peut voir que la TSB est robuste, dans une large mesure, aux occultations même lorsque les trois sommets sont manquants.

(Choi *et al.*, 2004) ont récemment proposé la transformation en symétrie pour les coins TSC utilisant un modèle paramétrique bivarié de l'orientation du gradient de l'image, inspiré de la transformation en symétrie généralisée TSG de (Reisfeld *et al.*, 1995). Deux points du contour votent pour leur point milieu, la contribution dépendant de leur appartenance à un angle donné. Cette fonction d'appartenance est fondée sur la relation entre l'orientation du gradient de ces points (P_i, P_j) : plus la vraisemblance qu'ils appartiennent aux deux côtés d'un angle est grande, plus leur vote pour le milieu $P = (P_i + P_j)/2$ est important. Cette méthode induit un biais dans la localisation du sommet du coin. Dans notre approche, un modèle de coins différent est proposé : les points sont groupés par paires mais leur vote est accumulé au niveau du sommet de l'angle auquel ils appartiennent. De plus, la bissectrice de cet angle est aussi modélisée.

Nous avons appliqué le second algorithme proposé à la détection de panneaux de signalisation triangulaires. Plus qu'un détecteur de coins, notre approche est capable de détecter trois sommets et leur bissectrice, et en déduit le centre du cercle inscrit au triangle ainsi que la taille et l'orientation du triangle. Aucun filtre de couleur n'est nécessaire, et notre approche est valide pour des triangles ayant un contraste sombre/clair ou clair/sombre. De plus, elle est invariante aux rotations dans le plan de l'image et robuste aux petites rotations hors du plan de l'image.

La détection de panneaux fondée sur la TCB et la TSB a été comparée à des approches fondées sur les transformées TSR et transformée de Hough dans (Belaroussi *et al.*, 2010a). Il est ressorti de cette étude que l'accumulation bi-variée est plus efficace qu'une transformation mono-variée lorsque la taille des panneaux est suffisante, typiquement supérieure à 10 % de la taille de l'image.

1.3. Organisation de l'article

La section 2 rappelle la transformation en symétrie généralisée TSG et introduit les transformations chinoise standard TC et bilatérale TCB. On y expose une analyse comparée de ces opérateurs pour la tâche de détection des panneaux de signalisation circulaires et rectangulaires.

La section 3 introduit la transformation en symétrie de coins TSC, et la transformation en sommet et bissectrice TSB. La différence entre ces opérateurs est discutée et l'originalité de notre approche est discutée.

Les résultats expérimentaux, sur des images synthétiques et des images de route en milieu urbain, sont discutés à la section 4. On montre notamment que la TCB et la TSB ne sont pas dépendantes de l'orientation globale du panneau, ne dépendent pas non plus du contraste clair/sombre ou sombre/clair, et sont rapides et efficaces en environnement extérieur. La supériorité de notre approche par rapport à une approche fondée sur la couleur est démontrée. La robustesse aux conditions d'éclairage est

vérifiée sur des bases de test différentes. Le comportement en présence d’occultations ou de distorsions est discuté. La robustesse aux bruits et aux flous est étudiée.

2. Détection de panneaux circulaires et rectangulaires

2.1. Transformation en symétrie généralisée TSG

La transformation en symétrie généralisée (TSG) (Reisfeld *et al.*, 1995) est un opérateur sans contrainte d’orientation absolue, qui dans son principe cherche à détecter des portions de contours symétriques. La TSG a inspirée différents algorithmes, dédiés à un type de symétrie ou à un type de forme à détecter.

La TSG calcule un indice de symétrie à partir des orientations des gradients de deux pixels (P_i, P_j) en votant au point milieu P du segment $[P_i P_j]$. Pour chaque point P d’une image, un ensemble de votants est défini par :

$$\Gamma(P) = \{(P_i, P_j) \mid \frac{P_i + P_j}{2} = P\} \quad [1]$$

Chaque pixel P_i a un vecteur gradient \mathbf{n}_i faisant un angle θ_i par rapport à l’horizontale. Chaque paire de vecteurs gradients $(\mathbf{n}_i, \mathbf{n}_j)$ est plus ou moins proche d’une certaine symétrie axiale et centrale selon les orientations (θ_i, θ_j) . Cette proximité est estimée par une fonction à deux termes explicités dans la figure 3 et nommée fonction de phase. Cette fonction est ensuite pondérée par une fonction de la distance $\|\mathbf{P}_i \mathbf{P}_j\|$, et par la norme des gradients $r_i = \|\mathbf{n}_i\|$ et $r_j = \|\mathbf{n}_j\|$. La fonction $D(i, j)$ utilisée pour la distance permet d’atténuer les votes de paires éloignées, donc

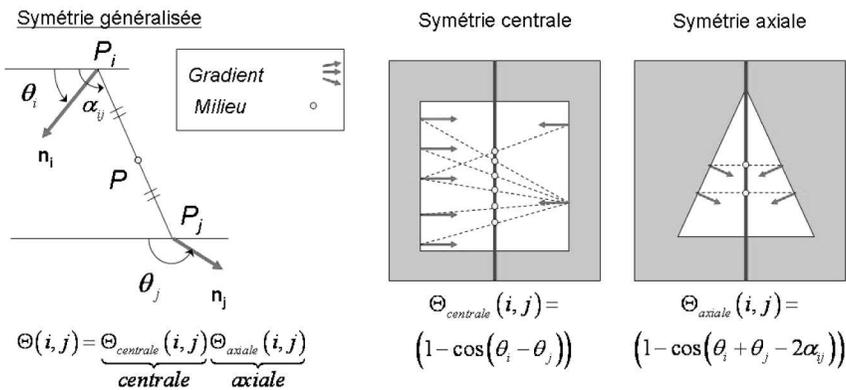


Figure 3. La symétrie centrale et la symétrie axiale entre deux vecteurs gradient \mathbf{n}_i et \mathbf{n}_j peuvent être modélisées par des contraintes sur les paramètres $(\theta_i, \theta_j, \alpha_{ij})$.

Dans la Transformation en Symétrie Généralisée TSG, l’indice de symétrie est le produit de deux termes : $\Theta_{centrale}$ et Θ_{axiale} pour les symétries centrale et axiale respectivement

de rendre la détection locale à une certaine échelle. L'indice de symétrie en un point P est la somme sur l'ensemble Γ_P de toutes ces contributions :

$$S(P) = \sum_{(i,j) \in \Gamma_P} \underbrace{D(i,j)}_{\text{distance}} \underbrace{\Theta_{axiale}(i,j)}_{\text{Symétrie axiale}} \underbrace{\Theta_{centrale}(i,j)}_{\text{Symétrie centrale}} \underbrace{r_i r_j}_{\text{Module gradients}} \quad [2]$$

En pratique, r_i peut être une fonction logarithme du module du gradient : $r_i = \log(\|\mathbf{n}_i\|)$ ou une autre fonction croissante. Au contraire, $D(i,j)$ est une fonction décroissante avec la distance $\|P_i P_j\|$, typiquement :

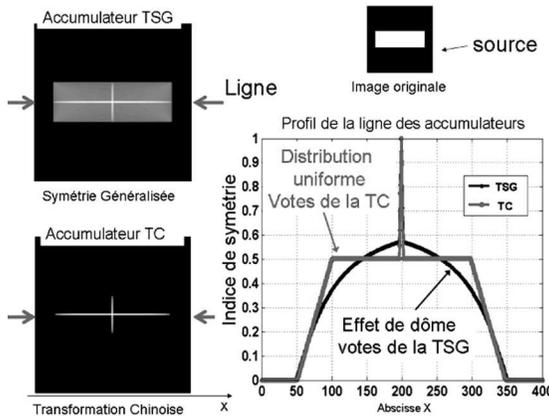


Figure 4. Accumulateurs de la TSG et de la TC pour une forme rectangulaire : les deux axes de symétrie (horizontal et vertical) sont mis en valeur

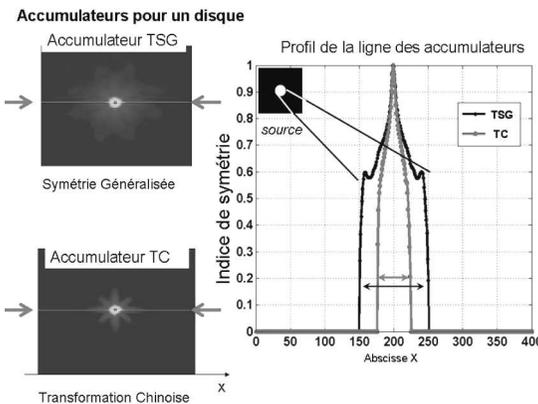


Figure 5. Accumulateurs de la TC et de la TSG dans le cas d'une forme circulaire. La localisation du centre par la TC est plus précise et le pic de la TSG est deux fois plus large que celui de la TC

$$D(i, j) = \frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{\|P_i - P_j\|}{2\sigma}} \quad [3]$$

où le facteur d'échelle σ dépend de la taille des objets recherchés. La formule de $\Theta(i, j)$ est indiquée sur la figure 3. Elle combine un indice de symétrie centrale et un indice de symétrie axiale: $\Theta(i, j) = \Theta_{refl}(i, j) \times \Theta_{rad}(i, j)$.

2.2. Transformation chinoise TC

Pour modéliser la symétrie axiale, une autre approche a récemment été proposée, la TC (Milgram *et al.*, 2008), et appliquée à la détection des yeux sur des images de visages.

Comme pour la TSG, l'espace d'accumulation de la TC est celui de l'image, et une paire de points de contour vote pour leur milieu. La différence entre ces deux transformations est la fonction de phase $\Theta(i, j)$, c'est-à-dire les conditions requises pour former une paire. La transformation chinoise ne fait appel qu'à un terme de symétrie centrale, mais avec une condition sur l'alignement qui permet d'englober aussi le cas de la symétrie axiale :

$$\Theta(i, j) = \underbrace{W_\beta(\theta_i - \alpha_{ij})}_{\text{Alignement restreint}} \underbrace{W_\delta(|\theta_i - \theta_j| - \pi)}_{\text{Symétrie centrale}} \quad [4]$$

où W_R est un noyau uniforme de rayon R :

$$W_R(x) = 1 \quad \text{si } |x| < R, \quad 0 \quad \text{sinon} \quad [5]$$

Le premier terme de la fonction de phase est une contrainte sur l'alignement entre $\mathbf{P}_i\mathbf{P}_j$ et \mathbf{n}_i . La valeur de δ est petite pour assurer que la deuxième contrainte $|\theta_i - \theta_j| = \pi$ est vérifiée numériquement, de sorte que seuls les gradients ayant une direction opposée sont inclus dans l'ensemble des votants. Comme pour la TSG, un facteur d'échelle σ définit la région d'influence de chaque pixel par le biais de $D(i, j)$.

La TC peut sembler proche de la TSG mais elle est assez différente. En effet, bien qu'elle ne modélise pas explicitement la symétrie axiale, elle est aussi capable de détecter des axes de symétrie (cas de la détection des doigts de la main (Belaroussi *et al.*, 2008)). Autre exemple, les accumulateurs obtenus avec ces transformées sont comparés dans la figure 4, dans le cas d'une image avec un rectangle blanc sur fond noir. Avec la TC, la majeure partie des votes est accumulée près des axes, tandis que pour la TSG, la contribution des paires de gradients se diffuse loin des axes. Les deux profils obtenus sur les deux accumulateurs au même endroit (indiqué par des flèches) montrent la distribution horizontale de l'indice de symétrie pour chacune des transformations. Le pic central est le point d'intersection des axes. On observe que la TC résulte en une distribution quasi uniforme des votes tandis que les votes de la TSG forment un dôme : la TC est donc plus sélective que la TSG. La figure 5 illustre

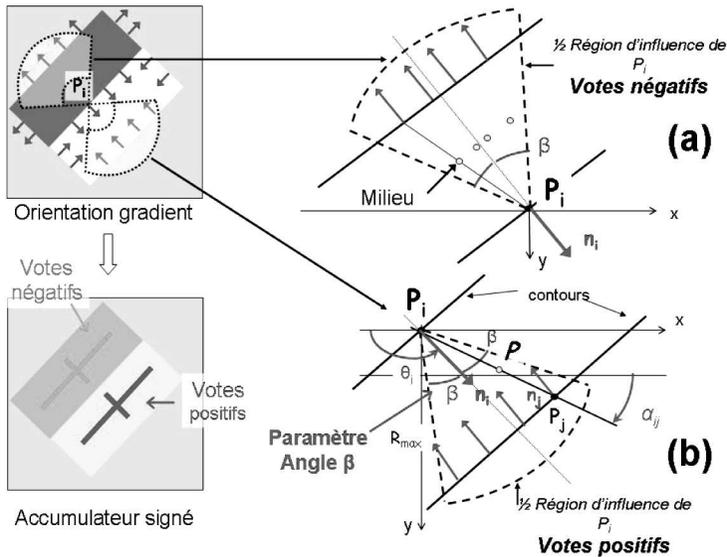


Figure 6. Transformation chinoise bilatérale signée : (a) Dans le cas d'un contraste sombre/clair, une paire de points de contour avec des gradients d'orientations opposées (et divergents) $(\mathbf{n}_i, \mathbf{n}_j)$ résulte en une contribution négative dans l'accumulateur. (b) Une paire de points de contour avec des gradients d'orientations opposées $(\mathbf{n}_i, \mathbf{n}_j)$ mais pointant l'un vers l'autre contribue positivement à l'accumulateur (rectangle clair/sombre)

le cas d'un disque blanc sur fond noir : la TC est aussi plus précise que la TSG. La figure 1 montre l'accumulateur obtenu par la TC sur des panneaux circulaires et polygonaux.

2.3. Transformation chinoise bilatérale TCB

La TC fait toutefois l'hypothèse que l'objet est clair sur un fond sombre dans l'image, de sorte que les paires de points à considérer sont réduites aux couples (P_1, P_2) ayant des gradients convergents (voir figure 6b). Cette hypothèse est généralement vérifiée dans le cas d'un iris ou d'un oeil, mais pas dans le cas de la détection de panneaux. Elle peut être assurée par une segmentation couleur, notamment si on cherche à détecter des panneaux d'une certaine couleur (Foucher *et al.*, 2009). Pour prendre en compte les cas de gradients d'orientations divergents, nous proposons plutôt de modifier le terme de phase de la TC. Nous avons testé deux alternatives : dans la première, les gradients divergents ont une contribution négative (c'est la TCB signée, illustrée par les figures 6 et 7), et dans la seconde, ils ont la même contribution positive (c'est la TCB non signée) :

$$\Theta(i, j) = \left(\underbrace{W_\beta(\theta_i - \alpha_{ij})}_{\text{clair/sombre}} \underbrace{\pm}_{\text{signe}} \underbrace{W_\beta(\theta_i - \alpha_{ij} - \pi)}_{\text{sombre/clair}} \right) \times \underbrace{W_\delta(|\theta_i - \theta_j| - \pi)}_{\text{Symétrie centrale}} \quad [6]$$

où δ est la précision avec laquelle l’assertion $|\theta_i - \theta_j| = \pi$ est vérifiée numériquement ($\delta = \frac{2\pi}{N}$ si on quantifie l’orientation du gradient sur N valeurs). Comme précédemment, α_{ij} est l’angle que fait le vecteur $\mathbf{P}_i\mathbf{P}_j$ avec l’axe horizontal, voir la convention en figure 3. β est la tolérance angulaire entre les vecteurs \mathbf{n}_i et $\mathbf{P}_i\mathbf{P}_j$ et sur l’alignement entre les gradients \mathbf{n}_i et \mathbf{n}_j . Ce paramètre délimite la zone d’influence du point P_i comme l’indique la figure 6. Une valeur petite de β est adaptée à une forme strictement circulaire tandis que la valeur $\beta = \pi/2$ peut être utilisée pour des formes plus variées. Néanmoins $\beta = \pi/2$ augmente substantiellement le nombre de votants et par voie de conséquence le temps de calcul, ainsi qu’en général le nombre de faux positifs. Concernant la pondération $D(i, j)$ par la distance, il est plus rapide d’utiliser un seuillage sur la distance entre deux points, plutôt qu’une gaussienne :

$$D(i, j) = W_{R_{max}}(\|P_i - P_j\|) - W_{R_{min}}(\|P_i - P_j\|) \tag{7}$$

où R_{min} et R_{max} sont respectivement la taille minimale et maximale d’un panneau. Les panneaux sont donc recherchés dans la gamme de taille $[R_{min} R_{max}]$. L’accumulateur est incrémenté d’une valeur proportionnelle au produit de $\Theta(i, j)$ avec les fonctions du module des gradients $r_i = \log(1 + \|\mathbf{n}_i\|)$, soit :

$$Accu_{TCB}(P) = \sum_{(i, j) \in \Gamma(P)} D(i, j)\Theta(i, j) r_i r_j \tag{8}$$

Dans le cas où l’on cherche à détecter des formes polygonales ayant quatre côtés ou plus, il faut utiliser une valeur de $\beta = \frac{\pi}{8}$ au moins. Cela peut se faire en quantifiant l’orientation du gradient sur 8 directions comme dans le code de Freeman.

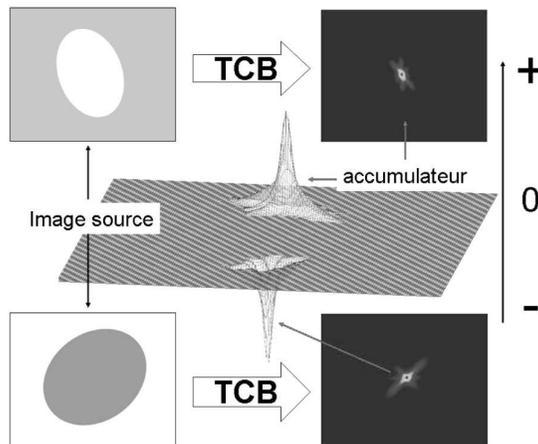


Figure 7. Accumulateur de TCB signée dans le cas d’ellipses clair/sombre et sombre/clair : la TCB est invariante aux rotations dans le plan image

Comme le montre la figure 10, la TCB a pour sortie deux accumulateurs, un pour les centres de symétrie des objets, et un autre dont est déduit leur rayon. Le centre de symétrie d'un objet peut être extrait de l'accumulateur [8]. Par contre, pour déterminer l'étendue spatiale de l'objet détecté, un autre accumulateur est nécessaire pour enregistrer la somme des distances entre votants $\|P_i - P_j\|$, ce qui permet de calculer leur moyenne :

$$Rayon(P) = \frac{1}{N} \sum_{(i,j) \in \Gamma(P)} \|P_i - P_j\| / 2 \quad [9]$$

où $N = Card(\Gamma(P))$. L'estimation du rayon de l'objet détecté est avantageuse, en termes de taille mémoire, vitesse et précision, comparée aux méthodes multi-échelles telles que celles présentées dans (Barnes *et al.*, 2003) et (Loy *et al.*, 2004) qui utilisent un accumulateur par rayon.

Dans la TCB signée, les objets sombre/clair ont une contribution négative tandis que ceux de contraste clair/sombre ont un vote positif. La figure 7 montre l'accumulateur dans le cas de formes elliptiques avec des contrastes et des orientations différents : on remarque que la TCB est invariante aux rotations dans le plan de l'image. La robustesse à de petites rotations hors du plan de l'image est illustrée avec des polygones sur la figure 8. Comme l'expliquent (Park *et al.*, 2006), l'emploi d'une méthode de votes signées renforce la sélectivité de la transformation. Cela dit l'application de la TCB signée est problématique dans le cas de panneaux non plein (indication et interdiction). Les deux types de contraste sont alors présents et les votes négatifs diminuent la contribution des votes positifs pour le centre du panneau. Les résultats expérimentaux montrent ainsi que la TCB non signée améliore notablement les performances en détection par rapport à la TCB signée.

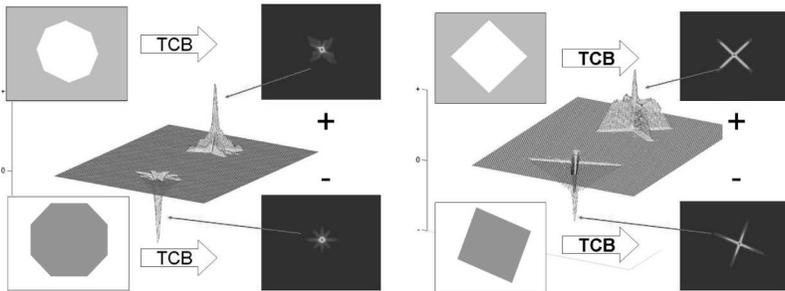


Figure 8. TCB signée dans le cas de formes polygonales : image en niveaux de gris et accumulateur correspondant. Les contours des objets de contraste clair/sombre (en haut) ont une contribution positive dans l'accumulateur ; ceux de contraste sombre/clair (en bas) contribuent avec un vote négatif

2.4. Combinaison avec la couleur

Les transformées précédentes font une sélection sur la forme, mais les panneaux de signalisation routière sont identifiables aussi par leurs couleurs. Il y a plusieurs

façons de combiner la sélection sur la forme et les couleurs. Nous avons donc testé plusieurs possibilités.

Nous avons ainsi testé une approche TC + Camshift s'appuyant sur un modèle de couleur, dont les étapes sont détaillées par le diagramme de la figure 9. Un histogramme 2D de la couleur rouge, dans le plan *Hue-Saturation* de HSV, est construit à partir d'une base différente de la base de test. Cette base d'apprentissage est constituée de 5 950 pixels extraits manuellement de panneaux rouges contenus dans les 48 images de la base *traffic sign image database* (Grigorescu *et al.*, 2003). L'histogramme 2D est rétroprojeté dans l'image avec pour résultat la carte des probabilités de la couleur rouge. Cette carte est seuillée afin de trouver des régions d'intérêt. Le champ de gradient de ces régions est estimé et mis en entrée de la transformation chinoise. La taille des objets ayant un indice de symétrie suffisant est déterminée en opérant une segmentation par l'algorithme du Camshift (Bradsky, 1998). La TC sert à déterminer les centres de symétrie et à initialiser la fenêtre du Camshift.

Une deuxième approche a été testée : une vue d'ensemble du traitement par la TC ou TCB (signée ou non) sur les contours du canal rouge normalisé est donnée en figure 10. La taille des objets ayant un indice de symétrie suffisant est déterminée à partir d'un second accumulateur détaillé à la section précédente (équation 9).

Plutôt que d'estimer le module du gradient en un point I par $|\mathbf{n}_I| = |\nabla r|$, il est aussi intéressant de l'estimer par $|\mathbf{n}_I| = |\nabla r| + |\nabla b|$ pour faire ressortir les contours des panneaux bleus. Le canal rouge normalisé est calculé par $r = \frac{R}{R+G+B}$, le bleu par $b = \frac{G}{R+G+B}$. Cette variante est aussi testée dans l'évaluation expérimentale.

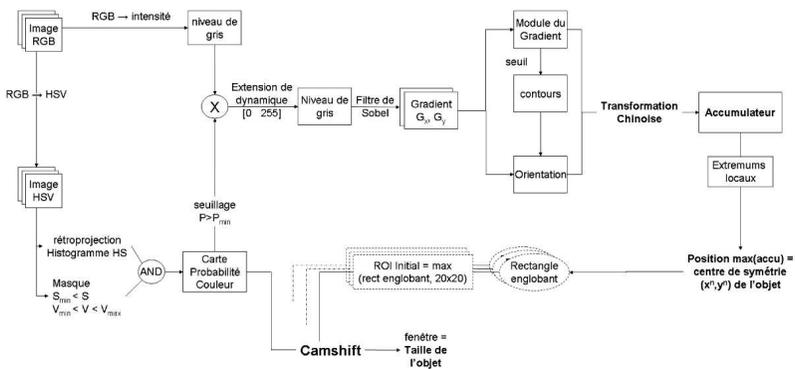


Figure 9. Diagramme du traitement par TC+Camshift. Un modèle de couleur rouge est construit dans le plan Hue-HueSaturation de HSV. Il est rétroprojeté dans l'image HSV : les régions vraisemblablement rouges sont traitées par TC pour trouver les centres de symétrie, et une segmentation par Camshift est opérée à partir de chaque centre

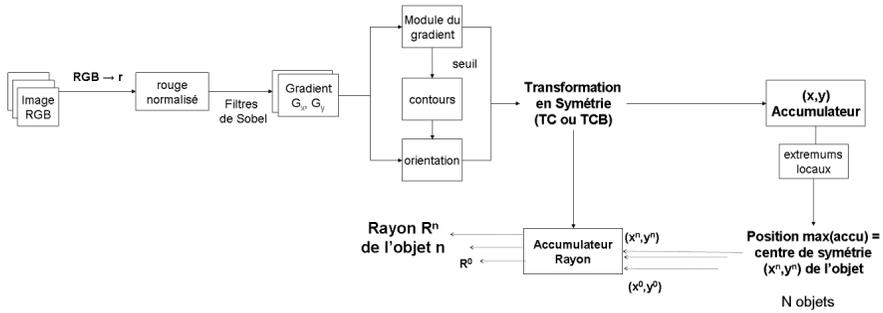


Figure 10. Vue d'ensemble de la détection de symétrie (par TC ou TCB). L'image couleur est convertie dans l'espace colorimétrique rgb normalisé. La composante rouge normalisé est traitée : on estime l'orientation et le module de son gradient. La transformation en symétrie (TC ou TCB) détecte les objets symétriques (centre et taille)

3. La détection de panneaux triangulaires

Le cas des panneaux triangulaires ne peut pas être traité par les transformées précédentes. Nous proposons donc une transformation spécifique pour les triangles qui combine détection des sommets et du centre du triangle.

3.1. De la TSG à la détection de coins

La Transformation en Symétrie de Coins (TSC) introduite dans (Choi *et al.*, 2004), adapte la TSG aux caractéristiques de symétrie d'un angle γ de valeur angulaire donnée. Ainsi, la contribution d'une paire de points (P_i, P_j) à la fonction de phase $\Theta(i, j)$ est constituée de deux termes : le premier modélise la symétrie axiale et le second estime le degré d'appartenance à la classe des coins d'angle γ . Le vote du couple (P_i, P_j) est aussi accumulé en leur milieu P :

$$\Theta_\gamma(i, j) = \underbrace{\Theta_{axiale}(i, j)}_{\substack{\text{Symétrie} \\ \text{axiale}}} \exp\left(-\frac{(\beta_{ij} (|\theta_i - \theta_j| - \pi - \gamma))^2}{2\sigma_\gamma}\right) \quad [10]$$

Le terme de symétrie axiale est maximal lorsque les gradients sont symétriques par rapport à la médiatrice du segment $[P_i P_j]$. Le second terme a une valeur maximale lorsque le coin d'angle $|\theta_i - \theta_j| - \pi$ est égal à γ . σ_γ est un paramètre contrôlant l'adéquation à γ . $\beta_{ij} = \left[\pi \sqrt{1 - \cos^2(\theta_i - \theta_j)}\right]^{-1}$ est une fonction de rejet

assurant que P_i et P_j ne sont pas alignés sur le même contour. Enfin, l'indice de symétrie de la TSC en P , pour un angle γ donné, est la somme des contributions des paires de points de Γ_P :

$$S_\gamma(P) = \sum_{(i,j) \in \Gamma_P} D(i,j) \Theta_\gamma(i,j) r_i r_j \tag{11}$$

À notre avis, la pertinence du terme de symétrie axiale est discutable : deux points appartenant à des côtés adjacents d'un angle ne vérifient pas nécessairement cette condition de symétrie comme le montre la figure 11. Il semble préférable d'exclure ce critère de symétrie par rapport à la médiatrice. De plus, dans la TSC le vote d'une paire de points est accumulé en leur milieu, ce qui ne met pas en valeur le coin mais un nuage de points situés près du coin : les résultats de la TSC sont donc biaisés.

3.2. Transformation en sommet et bissectrice TSB

Dans l'approche que nous proposons, une paire de points (P_i, P_j) vote pour le sommet du coin A auquel ils appartiennent :

$$\Gamma(A) = \{(P_i, P_j) \mid \mathbf{AP}_i \cdot \mathbf{n}_i = 0, \mathbf{AP}_j \cdot \mathbf{n}_j = 0\} \tag{12}$$

$\mathbf{n}_i = \left[\frac{\partial I}{\partial x} \quad \frac{\partial I}{\partial y} \right]^T$ étant le gradient de l'image I au point $P_i = (x_i, y_i)$, l'équation de la perpendiculaire au gradient en P_i est :

$$y = a_i (x - x_i) + y_i \quad \text{où} \quad a_i = -\frac{\frac{\partial I}{\partial x}}{\frac{\partial I}{\partial y}} \tag{13}$$

En conséquence, le sommet A de l'angle formé par les droites perpendiculaires aux gradients en P_i et P_j a pour coordonnées :

$$\begin{cases} x_A = \frac{1}{a_i - a_j} (y_j - a_j x_j - y_i + a_i x_i) \\ y_A = a_i (x_A - x_i) + y_i \end{cases} \tag{14}$$

On peut alors construire un tableau accumulant la présence d'un sommet en A comme dans le cas de la TSG :

$$S_{accu}(A) = \sum_{(i,j) \in \Gamma(A)} D(i,j) \Theta_\gamma(i,j) r_i r_j \tag{15}$$

avec

$$D(i,j) = W_{L_{max}}(\|\mathbf{P}_i \mathbf{P}_j\|) \tag{16}$$

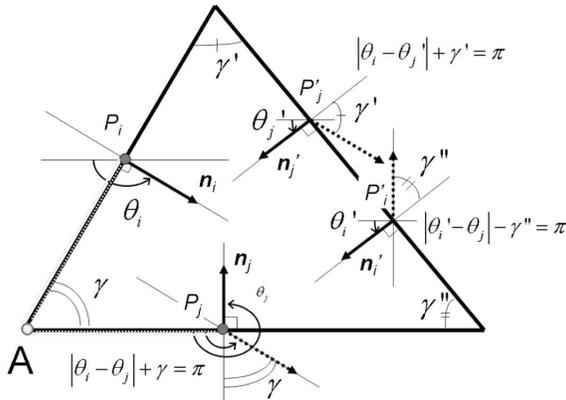


Figure 11. Orientation du gradient des côtés d'un triangle. La relation entre les orientations θ_i et θ_j de deux côtés adjacents dépend de l'angle au sommet γ , car chaque vecteur gradient \mathbf{n}_i et \mathbf{n}_j est orthogonal à son côté

où $W_{L_{max}}$ est un noyau uniforme de rayon L_{max} . De cette manière, les pixels trop éloignés ne sont pas appariés. On choisit pour la pondération par rapport aux modules du gradient une fonction logarithme r_i et r_j la même fonction que pour la TCB :

$$r_i = \log(1 + \|\mathbf{n}_i\|) \quad [17]$$

Notons, qu'il est aussi possible de ne pas spécifier d'angle γ , la fonction de phase étant alors définie par :

$$\Theta(i, j) = \Theta_{//}(i, j) = 1 - \delta(\cos(\theta_i - \alpha_{ij})) \quad [18]$$

où $\alpha_{ij} = \angle(O\mathbf{x}, \mathbf{P}_i\mathbf{P}_j)$ est l'angle de $\mathbf{P}_i\mathbf{P}_j$ par rapport à l'horizontale (voir figure 3), et δ est le symbole de Kronecker. La fonction $\Theta_{//}(i, j)$ rejette les cas où $|\theta_i - \alpha_{ij}| = \pi/2$ modulo π , P_i et P_j étant alors sur des contours parallèles. Les maxima locaux du tableau S_{accu} , après seuillage permettent de détecter les sommets d'angles.

Pour détecter le centre du panneau, la paire de points (P_i, P_j) vote aussi pour une portion $[AB]$ de la bissectrice de l'angle au sommet A, donc d'angle $\angle(P_iAB) = \angle(P_iAP_j)/2$:

$$\Gamma(AB) = \{(P_i, P_j) \mid \mathbf{AP}_i \cdot \mathbf{n}_i = 0, \mathbf{AP}_j \cdot \mathbf{n}_j = 0, \angle(P_iAB) = \angle(P_iAP_j)/2\} \quad [19]$$

Le segment $[AB]$ est construit de façon à ce que sa longueur soit la taille maximale L_{max} du triangle à détecter dans l'image. Un indicateur de la localisation des bissectrices est donc simultanément construit en accumulant les votes selon les segments $[AB]$:

$$B_{accu}([AB]) = \sum_{(i,j) \in \Gamma(AB)} D(i,j) \Theta_{\gamma}(i,j) r_i r_j \tag{20}$$

Par construction, chaque bissectrice d'un angle apparaît sous la forme d'un segment dans l'accumulateur B_{accu} . Si une forme triangulaire est présente dans l'image, son centre est donc un maximum local de B_{accu} du fait que les trois bissectrices se coupent en un même point.

3.3. Détection de signalisation triangulaire

Dans le cas des panneaux triangulaires, des contraintes supplémentaires nous amènent à définir une fonction de phase plus spécifique $\Theta(i, j)$. Lorsque le triangle est approximativement équilatéral : l'angle de base est $\gamma_0 = \pi/3$ à une précision $\Delta\gamma$ donnée. Deux points de contour de gradient \mathbf{n}_i et \mathbf{n}_j , ayant pour orientation respectivement θ_i et θ_j , appartiennent à deux côtés adjacents du triangle équilatéral si l'une des deux conditions suivantes est vérifiée:

$$||\theta_i - \theta_j| - \pi - \gamma_0| \leq \Delta\gamma \tag{21}$$

ou

$$||\theta_i - \theta_j| - \pi + \gamma_0| \leq \Delta\gamma \tag{22}$$

Le choix de $\Delta\gamma$ détermine la précision avec laquelle l'égalité $|\theta_i - \theta_j| - \pi = \gamma_0$ ou $-\gamma_0$ doit être vérifiée numériquement : ces deux cas sont illustrés sur la figure 11.

De plus, il ne faut pas être trop sévère sur le choix de $\Delta\gamma$. En effet, la projection d'un panneau de signalisation dans le plan image dépend de son orientation dans le repère caméra : son image est en toute généralité un triangle quelconque, mais l'intervalle $\gamma \in]\frac{\pi}{6}, \frac{\pi}{2}[$, correspondant à $\Delta\gamma = \pi/6$, regroupe un large spectre de points de vue habituels de la caméra. On peut envisager une précision plus fine si on a une connaissance *a priori* des angles à détecter. Nous avons ainsi testé les cas $\Delta\gamma = \pi/6$, $\Delta\gamma = \pi/12$ et $\Delta\gamma = \pi/18$. Les meilleurs résultats sont obtenus pour une précision intermédiaire $\Delta\gamma = \pi/12$ (Belaroussi *et al.*, 2010b).

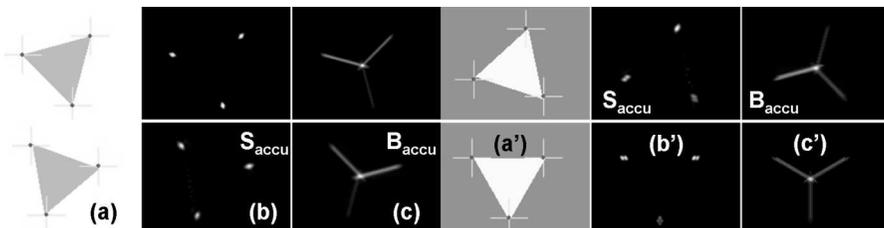


Figure 12. Cas de triangles clair/sombre ou sombre/clair avec différentes orientations. Colonne (a) : image source (niveau de gris) avec les sommets détectés (croix). Colonne (b) : accumulateur des sommets S_{accu} . Colonne (c) : accumulateur pour les bissectrices B_{accu} . La TSB est relativement invariante aux rotations dans le plan

Afin de modéliser plus spécifiquement l'angle $\gamma = \pi/3$, la fonction de phase est conçue pour n'apparier que les points appartenant aux côtés adjacents d'un angle de $\gamma_0 = \pi/3$:

$$\Theta(i, j) = \Theta_{//}(i, j) \left(\underbrace{W_{\frac{2\pi}{N}} (|\theta_i - \theta_j| - \pi - \gamma)}_{\text{angle aigu}} + \underbrace{W_{\frac{2\pi}{N}} (|\theta_i - \theta_j| - \pi + \gamma)}_{\text{angle obtus}} \right) \quad [23]$$

La largeur du noyau $W_{\Delta\gamma}$ est la précision $\Delta\gamma$. On la calcule à partir du choix d'une constante N qui est le nombre de bissectrice considéré : $\Delta\gamma = \frac{2\pi}{N}$.

Comme expliqué précédemment, on calcule deux accumulateurs, celui des sommets S_{accu} et celui des centres B_{accu} . Notre stratégie pour détecter un triangle est de commencer par détecter les centres de triangles en utilisant B_{accu} , puis de regrouper les sommets correspondants par triplet en exploitant S_{accu} . Un centre est détecté en recherchant un maximum local de l'accumulateur B_{accu} supérieur à un certain seuil τ_B . Pour chaque détection, on cherche dans S_{accu} les trois maxima locaux les plus proches supérieurs à un seuil τ_S . Si un point de l'accumulateur B_{accu} est associé à trois sommets, ils sont définis respectivement comme centre et sommets d'un triangle.

3.4. Exemples sur des images synthétiques

Sur la figure 12, des triangles blancs sur fond gris ou gris sur fond blanc sont traités. Dans les deux cas, on observe que la TSB est pratiquement invariante aux rotations dans le plan de l'image. La figure 2 montre que la transformation est aussi robuste à l'occultation d'un ou plusieurs sommets du triangle, ainsi qu'à la rupture partielle d'un ou plusieurs côtés. En effet, les points votant appartiennent à des côtés adjacents du triangle. Tant qu'une portion consistante de ces contours est visible, ses sommets peuvent être retrouvés même s'ils sont cachés.

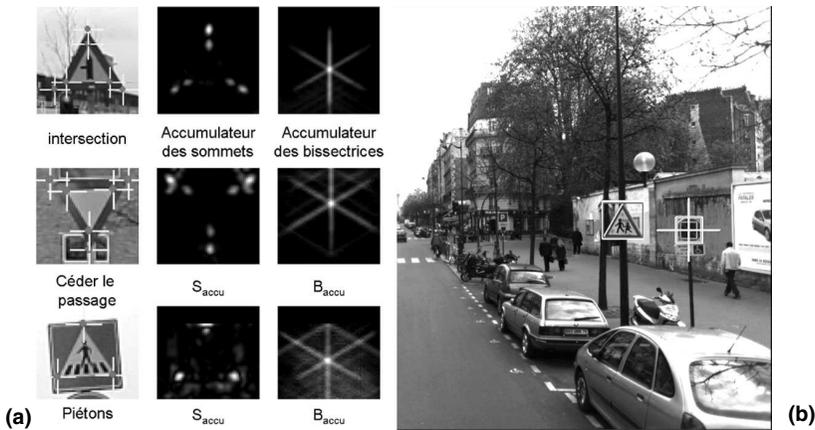


Figure 13. (a) Les trois types de formes triangulaires présentes dans la base de test. (b) Exemple de détections sur une image de la base Stéréopolis

4. Évaluations expérimentales

4.1. Les bases de test

Les performances des algorithmes et les variantes sont évaluées sur 3 bases d'images différentes, dans lesquelles les panneaux pris en compte sont ceux dont la taille est comprise entre 32×32 et 70×70 :

- Base *traffic sign image database* (Grigorescu *et al.*, 2003). Elle est constituée de 48 images de taille 360×270 pixels contenant 38 panneaux bleus (passage piétons, vélos) et 18 panneaux rouges (intersection, céder le passage). Parmi ces 55 panneaux routiers, 17 sont circulaires, 21 sont rectangulaires (uniquement des panneaux C20) et 40 contiennent un triangle. Ces derniers se divisent en trois catégories (figure 13a) :

- passage piétons C20 : un triangle blanc dans un panneau rectangulaire bleu (ce sont aussi les 21 panneaux rectangulaires).
- avertissement intersection : un triangle blanc, contenant un pictogramme, bordé par un triangle rouge, dont un sommet est situé au dessus de la base horizontale,
- avertissement céder le passage: un triangle blanc bordé par un triangle rouge avec un sommet situé en dessous de la base horizontale.

- Base *Paris 15* : 89 images, de taille 640×480 , prises dans les rues de Paris 15° (voir figure 14), contenant 81 panneaux circulaires rouges et un total de 92 panneaux de signalisation (11 bleus ou jaunes, rectangulaires ou circulaires).

- Base *Stéréopolis* (Belaroussi *et al.*, 2010a) : 847 images de dimensions 480×540 contenant 10 panneaux triangulaires et 21 panneaux circulaires ou rectangulaires (voir figure 13b). C'est une base publique : <http://www.itowns.fr/benchmarking>.

Les performances de la TCB (panneaux circulaires et rectangulaires) ont été mesurées sur la base *traffic sign image database*, la base *Paris 15* et la base *Stéréopolis*. Les performances de la TSB ont été estimées sur la base *traffic sign image database* et la base *Stéréopolis*.

4.2. Performances en détection de panneaux circulaires et rectangulaires

4.2.1. Comparaison entre TCB, TC standard et une approche couleur

Les algorithmes de la TC et de la TCB sont testés avec les mêmes étapes de prétraitement illustrées par le diagramme de la figure 10. La comparaison est réalisée sur la base *Paris 15* : la figure 14 illustre quelques exemples de détections correctes et de faux positifs pour la TCB. Comparé à la TCB, l'algorithme de la TC est moins efficace comme le montre les courbes de la figure 15. Les courbes ROC de la figure 15a montrent le Taux de Détections Correctes TDC en fonction du taux de Faux Positifs Par Image (FPPI) (Maji *et al.*, 2009) pour chaque algorithme. La figure 15b montre le coefficient nommé Dice pour la TC et la TCB en fonction du seuil sur l'accumulateur avec :



Figure 14. Exemples de segmentation de panneaux par TCB sur une scène urbaine (base de test Paris 15°)

$$\text{TDC} = \frac{VP}{P} \quad \text{FPPI} = \frac{FP}{\text{Nb Img}} \quad \text{Dice} = \frac{2VP}{VP + FP + P}$$

où VP est le nombre de vrais positifs, FP le nombre de faux positifs, P est le nombre total de panneaux dans la base de test, et Nb Img le nombre total d'images. Le tableau 1 indique les performances de ces deux algorithmes pour un point donné des courbes ROC : la TCB est capable de détecter 79 panneaux sur 92 avec 25 faux positifs, tandis que la TC standard n'en détecte que 69 pour 24 faux positifs. En effet,

Tableau 1. Performances en détection de panneaux sur la base Paris 15 de 89 images 640×480 en milieu urbain, contenant 92 panneaux dont 81 rouges et 11 bleus ou jaunes

Méthode	Nombre de cibles	Vrais Positifs	Faux Positifs	Cadence
TCB Bilaterale	92	79 (86 %)	25	30 ms/img
TC Standard	92	69 (75 %)	24	30 ms/img
TC + Camshift (rouge)	81	61 (75 %)	22	10 ms/img

comme on traite le canal rouge normalisé (voir diagramme de la figure 10), les panneaux rouges ont de grandes chances d'avoir un contraste clair/sombre et auquel cas ils peuvent être détectés par la TC et la TCB. Les panneaux ayant une autre couleur, en particulier les signalisations bleus, ont plus de chances d'apparaître avec un contraste sombre/clair : la plupart d'entre eux sont détectés par la TCB mais pas par la TC standard. Le coefficient Dice atteint une valeur maximale de 82 % pour la TCB contre 77 % seulement pour la TC. Avec une différence de 5 %, la transformation chinoise bilatérale améliore significativement la transformation chinoise standard.

L'approche TC+Camshift est plus rapide que l'algorithme de la TCB mais manque plus de panneaux (voir tableau 1). En effet, sur les 81 panneaux rouges présents dans la base de test, 75 % seulement sont détectés pour 22 faux positifs, à une cadence moyenne de 10 ms/image (contre 30 ms/image pour la TCB) sur un processeur standard Pentium IV@1,2GHZ. Cela illustre le manque de généralité des approches fondées sur un modèle de couleur. Ce modèle étant construit sur une caméra différente de celle de la base de test, un grand nombre de panneaux rouges n'ont pas la même signature dans le plan H-S que celle de l'histogramme des couleurs mémorisées. Ces panneaux ne peuvent donc pas être détectés par TC+Camshift.

Par ailleurs, la TCB contourne le problème de constance des couleurs en traitant le canal rouge normalisé : $r = \frac{R}{R+G+B}$. Ce canal est relativement invariant aux conditions d'éclairage de la scène. De plus, il permet de détecter des panneaux de couleurs quelconques : le bleu apparaît le plus souvent avec un contraste sombre/clair à l'inverse du rouge, mais aussi les panneaux jaunes ou verts sont généralement assez contrastés pour être détectés dans cette composante chromatique.

On notera toutefois que les taux de FPPI du tableau 1 sont de l'ordre de 30 %, ce qui est relativement important. Par exemple, les performances obtenues par les TSR proposées par (Loy *et al.*, 2004) sur 30 images contenant 15 octogones et 19 carrés,

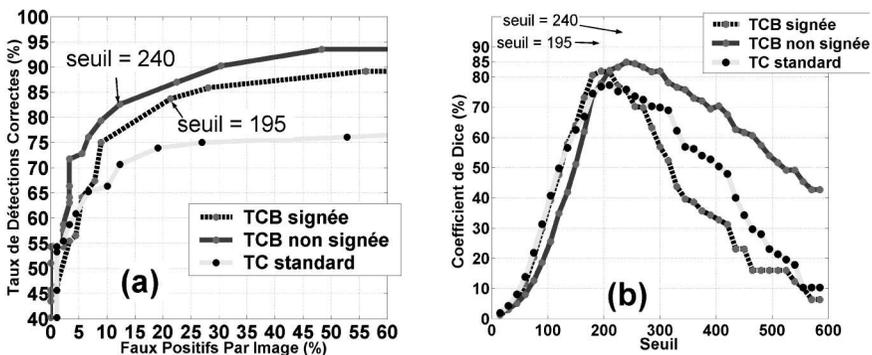


Figure 15. Comparaison des performances de la TC bilatérale (noir) et de la TC standard (clair), sur une base de test de 89 images 640×480 contenant 92 panneaux. (a) Courbes ROC : taux de détections correctes TDC (%) en fonction du taux de faux positifs par image FPPI (%). (b) Coefficient Dice (%) en fonction du seuil sur l'accumulateur

sont de 95 % de taux de détection 10 % de faux positifs par image (mais ils ajoutent une étape de reconnaissance). Cela vient du fait que nous ne traitons que le canal rouge normalisé pour estimer le gradient de l'image. Comme nous le montrons à la section suivante, une représentation plus judicieuse du gradient des couleurs améliore sensiblement les résultats de la TCB.

4.2.2. Comparaison entre TCB signée et non signée

La figure 15 compare aussi les performances des TCB signée et non signée. Le coefficient Dice atteint une valeur maximale de 85 % pour la TCB non signée contre 82 % pour la TCB signée. La supériorité des performances de la TCB non signée sur la TCB signée est aussi vérifiée sur la base *Stéréopolis*.

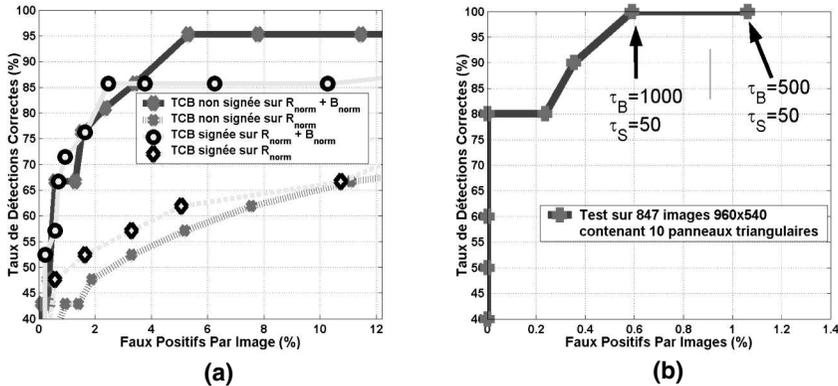


Figure 16. (a) Courbes ROC de la TC bilatérale sur la base *Stéréopolis* composée de 847 images 480×540 contenant 21 panneaux non triangulaires. Plusieurs variantes sont représentées : TCB signée et non signée, module du gradient estimé sur le canal rouge normalisé (R_{norm}) ou calculé par la somme de celui du canal rouge normalisé et de celui du canal bleu normalisé ($R_{norm} + B_{norm}$). (b) Courbe ROC de la TSB construite en faisant varier le seuil sur l'accumulateur B_{accu} , le seuil $\tau_S = 50$ étant fixe. La base de test *Stéréopolis* contient 10 panneaux triangulaires

La figure 16a montre que les résultats sont améliorés si le module du gradient privilégie ceux rouges et ceux bleus. Ainsi, quand on traite une combinaison des gradients des canaux rouges et bleus normalisés, on augmente le contraste des panneaux bleus par rapport à ceux dans le canal du rouge normalisé seul. Les meilleures performances sont obtenues en combinant le rouge et le bleu, et en détectant les symétries par transformée bilatérale non signée.

REMARQUE. – Les images de la base *Stéréopolis* sont de dimensions 1920×1080 et des panneaux y apparaissent avec une grande taille. Afin d'accélérer les calculs, nous réduisons chaque image d'un facteur 2 : les images traitées ont alors pour dimensions 960×540 . De plus les panneaux ne sont cherchés que sur le côté droit de l'image, la partie traitée a alors la taille 480×540 . Enfin, on se limite aux panneaux dans l'intervalle $[R_{min}, R_{max}] = [32, 70]$ dans l'image traitée. En particulier, ils ne sont pas détectés lorsqu'ils sont trop lointains ou trop proches du capteur.

4.3. Détection de panneaux de signalisation routière triangulaires

Notre algorithme a été évalué en premier sur la base *traffic sign image database*. L'algorithme est capable de détecter correctement 38 panneaux sur 40 avec seulement 2 faux positifs, comme le résume tableau 2.

Tableau 2. Performances en détection de panneaux de signalisation routière triangulaires sur la base *traffic sign image database* de 48 images 360 × 270

Images	Panneaux (triangles)	Vrais Positifs	Faux Positifs	Seuil τ_B	Seuil τ_S
48	40	38 95 %	2	500	50

Les tests sur la base *Stéréopolis* montrent une sensibilité du résultat de la TSB pas trop importante par rapport à la valeur du seuil τ_B . La courbe ROC de la figure 16b des performances de la TSB sur la base *Stéréopolis* donne le point de fonctionnement utilisé pour calculer les valeurs du tableau 2 : $\tau_B = 500$ et $\tau_S = 50$.

4.4. Analyse en sensibilité

Nous avons étudié l'influence d'une dégradation du système d'acquisition de l'image sur la base *traffic sign image database* constituée de 48 images contenant 38 cercles ou rectangles et 40 triangles. Nous nous sommes intéressés aux cas d'un bruit additif et d'un lissage gaussien. La figure 18 donne des exemples des transformations opérées sur la base de test. La TCB non signée est utilisée avec un seuil $\tau_{TCB} = 270$ et la TSB avec les seuils $\tau_B = 500$ et $\tau_S = 50$, quel que soit le niveau de bruit ou de lissage.

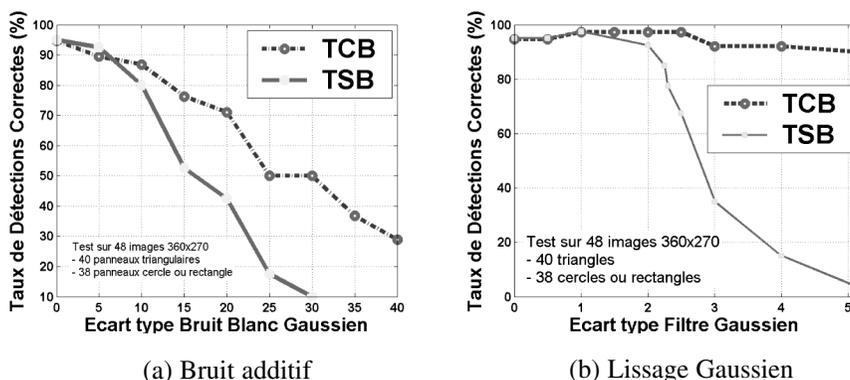


Figure 17. Variation du taux de détection de la TCB non signée et de la TSB en fonction (a) de l'ajout d'un bruit blanc Gaussien, et (b) d'un floutage de l'image. La base de test est la base *traffic sign image database*

La tolérance à un bruit additif est étudiée en ajoutant à chaque pixel de chaque image un bruit blanc gaussien d'écart type variant de 0 à 40 (– 32dB). La figure 17a représente les taux de détection de la TCB et de la TSB en fonction du niveau de bruit. On peut voir que les taux ont une décroissance lente quand l'écart type augmente, en particulier dans le cas de la TCB. La TSB reste compétitive avec un taux de détection supérieur à 80 % pour un niveau de bruit de – 20 décibels ($\sigma = 10$).

Pour analyser l'influence d'un flou dans l'image sur la détection de panneaux, on convolue chaque image avec un filtre Gaussien ayant un écart type variant de 0 à 5. L'écart type σ est proportionnel à la taille du filtre : $\sigma = 2$ correspond à un lissage sur un voisinage 11×11 . La figure 17b présente le taux de détections corrects pour la TCB (graphe en pointillés) et la TSB. On peut voir que le TDC de la TCB reste stable même après un lissage très fort ($\sigma = 5 \Leftrightarrow$ lissage sur un voisinage 31×31 , voir figure 18f). La TSB est aussi robuste à un niveau de flou inférieur à $\sigma = 2$ (figure 18e), avec un taux de détection supérieur à 92 %.

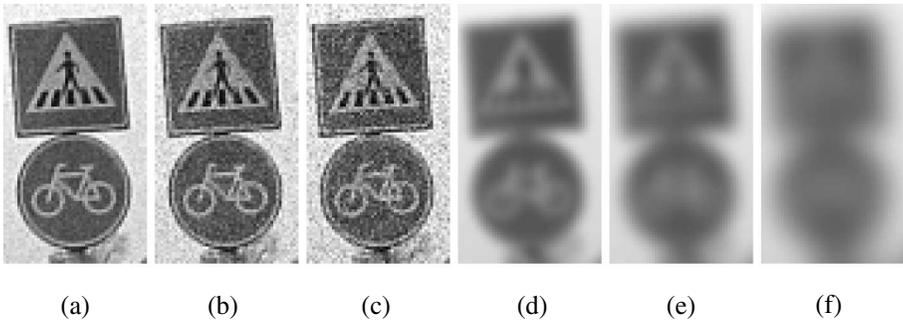


Figure 18. Exemples de dégradations testées : bruit additif d'écart type (a) $\sigma = 10$, (b) $\sigma = 20$, (c) $\sigma = 30$; filtre Gaussien d'écart type (d) $\sigma = 2$, (e) $\sigma = 3$, (f) $\sigma = 5$

4.5. Discussion

Concernant la détection de panneaux circulaires et polygonaux (hors triangle), la TCB est capable de détecter 95 % de panneaux pour 5 % de FPPI, sur la base *Stéréopolis*. Ces résultats améliorent ceux de la TSR présentés dans (Loy *et al.*, 2004) avec 95 % de taux de détection pour 10 % de faux positifs par image, et ceux obtenus par l'approche Adaboost de (Baro *et al.*, 2009) pour les cercles (60 à 70 % de réussite pour 1 à 4 % de FPPI).

La TSB est capable de détecter 100 % des panneaux pour 1 % de FPPI (et moins si on considère le seuil $\tau_B = 1000$), sur la base *Stéréopolis*. Ces résultats sont meilleurs que ceux obtenus par la TSR (Loy *et al.*, 2004) qui atteint un taux de détection de 100 % pour 66 % de faux positifs par image sur une base de 15 images contenant 10 triangles. Ils améliorent aussi les performances obtenues par les cascades de classifieurs boostés (Baro *et al.*, 2009) (65 % à 75 % des panneaux triangulaires détectés pour 2.25 % à 2.8 % de FPPI, nombre d'images de test non

communiqué). L'approche proposée par (Timofte *et al.*, 2009) obtient 95 % de réussite pour 20 % de FPPI tous panneaux confondus.

Ces performances sont suffisantes pour être couplées à une étape de reconnaissance des panneaux. Ainsi (Baro *et al.*, 2009) arrivent à système atteignant 90 % de détections correctes (au plus) en utilisant l'information provenant de paires d'images stéréo. Avec ce taux de détection (par image fixe), ils arrivent à poursuivre et reconnaître correctement plus de 90 % des panneaux dans le cas de vidéos. En effet, dans une séquence, un même panneau apparaît dans plusieurs images consécutives, ce qui augmente les chances de le détecter (et le reconnaître) au moins une fois.

La cadence moyenne de traitement de la TSB sur cette base de test est de 67 ms/img, celle de la TCB est de 30 ms/img, sur des images 480×540 . Ces temps ne tiennent pas compte des prétraitements de l'image, tels que le calcul du gradient et de son orientation, mais tiennent compte de la recherche de maxima locaux dans les accumulateurs. L'approche par TSR de (Loy *et al.*, 2004) met 50 ms/img pour des images 320×240 . En comparaison, l'approche par cascade attentionnelle de (Bahlmann *et al.*, 2005) traite une image 384×288 en 100 ms.

Ces temps de calcul sont très rapides pour des applications *hors ligne* telles que l'annotation/enrichissement de cartographie. Pour des systèmes d'aide à la navigation où l'on cherche à informer le conducteur de son environnement, ces cadences sont amplement suffisantes. Pour un système de conduite assistée en revanche, ces temps de calcul sont encore loin des cadences atteintes par exemple par un ESP (>25 Hz). Toutefois si la disposition de la caméra par rapport au véhicule est connue, on peut délimiter une région de recherche dans l'image et diminuer grandement le nombre de points de contour à prendre en compte (la complexité des algorithmes bivariés est en $\mathcal{O}(n^2)$ où n est le nombre de points de contour).

5. Conclusion et perspectives

Nous avons présenté une méthode de détection des panneaux circulaires et rectangulaires utilisant un détecteur de symétrie efficace, la transformation chinoise bilatérale. Elle peut traiter une image 640×480 en 30 ms avec un taux de détection élevé : 86 % pour 25 faux positifs sur un ensemble de 89 images contenant 92 panneaux de différentes couleurs et formes (à l'exception des panneaux triangulaires). Elle améliore les résultats de la transformation chinoise standard qui ne détecte que 75 % des panneaux pour 24 faux positifs sur cette base de test. De plus, la TCB est plus facile à généraliser que la TSR car elle peut détecter de la même manière cercles, carrés, rectangles, losanges et polygones. Elle est aussi plus précise que la TSG car elle est spécialisée sur les paires de points ayant des gradients symétriques par rapport à leur milieu. Cette méthode a fait l'objet d'une publication dans les actes de la conférence ISVC 2009 (Belaroussi *et al.*, 2009b).

Pour traiter le cas des panneaux triangulaires, nous avons introduit une nouvelle transformation pour la détection du sommet d'un angle et de sa bissectrice simultanément, même en cas d'occultations partielles. Elle peut traiter efficacement une image 360×270 en moins de 50 ms pour détecter des panneaux de signalisation triangulaires. Elle atteint un taux de détection élevé avec 38 panneaux sur 40

correctement détectés pour 2 faux positifs, sur une base de 48 images 360×270 , et 10 panneaux sur 10 sur une base de test de 847 images 480×540 pour 9 faux positifs. Cette approche a été présentée aux conférences WACV 2009 (Belaroussi *et al.*, 2009a) et RFIA 2010 (Belaroussi *et al.*, 2010b).

La méthode présentée a été comparée à deux autres algorithmes issus de l'état de l'art dans (Belaroussi *et al.*, 2010a). Nous montrons que les approches cumulatives bivariées sont plus robustes que les approches monovariées (transformées TSR et transformée de Hough). La base de test *Stéréopolis* et sa vérité terrain est disponible publiquement à l'adresse suivante : <http://www.itowns.fr/benchmarking>.

L'étape suivante dans ce travail est d'ajouter un module de reconnaissance pour éliminer les fausses alarmes constituées par des objets symétriques tels que les phares de voitures, les fenêtres, enseignes et logos ou encore certains piétons. Enfin, ces algorithmes peuvent encore être optimisés pour application en temps réels à bord de véhicules.

Remerciements

Les travaux présentés dans cet article ont été financés par l'Agence Nationale de Recherche, projet MDCO-iTowns : Image-based Town On-line Web Navigation and Search engine <http://www.itowns.eu/>. Les auteurs remercient Eric Dumont pour ses nombreuses remarques et corrections.

Bibliographie

- Bahlmann C., Zhu Y., Ramesh V., Pellkofer M., Koehler T. (2005). « A system for traffic sign detection, tracking, and recognition using color, shape, and motion information », *IEEE Intelligent Vehicles Symposium*, IEEE, p. 255-260.
- Barnes N., Zelinsky A., Fletcher L. (2003). « Traffic sign recognition and analysis for intelligent vehicles », *Image and Vision Computing*, vol. 21, n° 3, p. 247-258.
- Baro X., Escalera S., Vitria J., Pujol O., P. Radeva (2009). « Traffic Sign Recognition Using Evolutionary Adaboost Detection and Forest-ECOC Classification », *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, p. 113-126.
- Belaroussi R., Foucher P., Tarel J.-P., Soheilian B., Charbonnier P., Paparoditis N. (2010a). « Road Sign Detection in Images : A Case Study », *to appear in Proceedings of International Conference on Pattern Recognition (ICPR'10)*, Istanbul, Turkey.
- Belaroussi R., Milgram M. (2008). « A Real Time Traffic Sign Detection by Symmetry Transform Using a Two Cameras System », *4th International Symposium on Visual Computing ISVC'08*, Lecture Notes in Computer Science, Springer, Las Vegas, Nevada, USA, p. 703-712, décembre.
- Belaroussi R., Tarel J.-P. (2009a). « Angle Vertex and Bisector Geometric Model for Triangular Road Sign Detection », *IEEE Workshop on Applications of Computer Vision WACV'09*, Institute of Electrical and Electronics Engineers, Snowbird, Utah, USA, p. 577-583, 7-9 décembre.

- Belaroussi R., Tarel J.-P. (2009b). « A Real-Time Road Sign Detection using Bilateral Chinese Trans-form », *5th International Symposium on Visual Computing ISVC'09*, Lecture Notes in Computer Science, Springer, Las Vegas, Nevada, USA, p. 1161-1170, 30 novembre-2 décembre.
- Belaroussi R., Tarel J.-P. (2010b). « Modèle géométrique du sommet et de la bissectrice d'un angle pour la détection de panneaux triangulaires », *Reconnaissance des Formes et Intelligence Artificielle RFIA'10*, AFRIF-AFIA, Caen, 19-22 décembre.
- Bradsky G. (1998). « Computer Vision Face Tracking For Use in a Perceptual User Interface », *Intel Technology Journal*.
- Broggi A., Cerri P., Medici P., Porta P.-P., Ghisio G. (2007). « Real Time Road Signs Recognition », *Proceedings of IEEE Intelligent Vehicles Symposium (IV'07)*, Istanbul, Turkey, p. 981-986, juin.
- Caraffi C., Cardarelli E., Medici P., Porta P., Ghisio G., Monchiero G. (2008). « An algorithm for Italian de-restriction signs detection », *IEEE Intelligent Vehicles Symposium*, IEEE, p. 834-840.
- Chen S.-Y., HSieh J.-W. (2008). « Boosted road sign detection and recognition », *Proceedings of IEEE Int. Conference on Machine Learning and Cybernetics*, vol. 7, p. 823-826.
- Choi I., Chien S.-I. (2004). « A generalized symmetry transform with selective attention capability for specific corner angles », *Signal Processing Letters*, vol. 11, n° 2, p. 255-257.
- Cyganek B. (2008). « Real-Time Detection of the Triangular and Rectangular Shape Road Signs », *Advanced Concepts for Intelligent Vision Systems ACIVS'07*, Lecture Notes in Computer Science, Springer, p. 744-755.
- de la Escalera A., Armingol J., Mata M. (2003). « Traffic sign recognition and analysis for intelligent vehicles », *Image and Vision Computing*, vol. 21, n° 3, p. 247-258.
- de la Escalera A., Armingol J., Mata M. (1997). « Road Traffic Sign Detection and Classification », *IEEE Transactions on Industrial Electronics*, vol. 44, n° 6, p. 848-859.
- Dutilleux G., Charbonnier P. (2007). « Métaheuristiques biologiques pour la détection de la signalisation routière », in P. Siarry (ed.), *Optimisation en traitement du signal et de l'image*, Traité IC2, série traitement du signal et de l'image, Hermes, chapter 10, p. 271-294, février.
- Fang C., Chen S., Fuh C. (2003). « Road sign detection and tracking », *IEEE Transactions on Vehicular Technology*, vol. 52, n° 5, p. 1329-1341.
- Foucher P., Charbonnier P., Kebbous H. (2009). « Evaluation of a Road Sign Pre-detection System by Image Analysis », *International Conference on Computer Vision Theory and Applications VISAPP'09*, Lecture Notes in Computer Science, Springer, Lisbonne, Portugal, p. 362- 367, février.
- Garcia-Garrido M., Sotelo M., Martin-Gorostiza E. (2006). « Fast traffic sign detection and recognition under changing lighting conditions », *Proceedings of IEEE Intelligent Transportation Systems Conference (ITSC'06)*, Toronto, Canada, p. 811-816.
- Grigorescu C., Petkov N. (2003). « Distance sets for shape filters and shape recognition », *Transactions on Image Processing*, vol. 12, n° 10, p. 1274-1286.
- IISR-R-1-6 (2008). Instruction interministérielle sur la signalisation routière, Régulation n° 1 à 6, Ministère de l'Écologie, de l'Énergie, du Développement durable et de la Mer.
- Kuo W.-J., Lin C.-C. (2007). « Two-stage road sign detection and recognition », *Proceedings of the IEEE International Conference on Multimedia and Expo ICME'07*, p. 1427-1430, juillet.

- Loy G., Barnes N. (2004). « Fast shape-based road sign detection for a driver assistance system », *International Conference on Intelligent Robots and Systems IROS'04*, IEEE/RSJ, Sendai, Japan, p. 70-75, septembre.
- Maji S., Malik J. (2009). « Object Detection Using a Max-Margin Hough Transform », *Proceedings of Computer Vision and Pattern Recognition CVPR*, p. 1038-1045.
- Maldonado-Bascon S., Lafuente-Arroyo S., Gil-Jimenez P., Gomez-Moreno H., F.Lopez-Ferreras (2007). « Road-Sign Detection and Recognition Based on Support Vector Machines », *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, n° 2, p. 264-278, juin.
- Milgram M., Belaroussi R., Prevost L. (2008). « Multi-stage Combination of Geometric and Colorimetric Detectors for Eyes Localization », *13th International Conference Image Analysis and Processing ICIAP'05*, Lecture Notes in Computer Science, Springer, p. 1010-1017.
- Miura J., Tsuyoshi K., Shirai Y. (2000). « An Active Vision System for Real-Time Traffic Sign Recognition », *Proceedings of IEEE Intelligent Transportation Systems Conference (ITSC'00)*, p. 52-57.
- Ohara H., Nishikawa I., Miki S., Yabuki N. (2002). « Detection and recognition of road signs using simple layered neural networks », *Proceedings of the 9th International Conference on Neural Information Processing ICONIP'02*, vol. 2, p. 626-630, novembre.
- Paclik P., Novovicova J., Pudil P., Somol P. (2000). « Road Sign Classification using the Laplace Kernel Classifier », *Pattern Recognition Letters*, vol. 21, p. 1165-1173.
- Park C.-J., Seob K.-S., Choib H.-M. (2006). « Symmetric polarity in generalized symmetry transformation », *Pattern Recognition Letters*, vol. 27, n° 7, p. 854-857.
- Paulo C., Correia P. (2007). « Automatic detection and classification of traffic signs », *Proceedings of 8th Int. Workshop on Image analysis for Multimedia Interactive Services (WIAMIS'07)*, p. 11-14.
- Piccioli G., Micheli E. D., Parodi P., Campani M. (1996). « Robust method for road sign detection and recognition », *Image and Vision Computing*, vol. 14, n° 3, p. 209-223.
- Reina R. V., Sastre R. L., Arroyo S. L., Jimenez P. G. (2006). « Adaptive traffic road sign panels text extraction », *Proceedings of the 5th WSEAS International Conference on Signal Processing, Robotics and Automation ISPRA'06*, p. 295-300, février.
- Reisfeld D., Wolfson H., Yeshurun Y. (1995). « Context Free Attentional Operators : the Generalized Symmetry Transform », *International Journal of Computer Vision*, vol. 14, n° 2, p. 119-130.
- Ruta A., Li Y., Liu X. (2009a). « Real-time traffic sign recognition from video by class-specific discriminative features », *Pattern Recognition*, vol. 43, p. 416-430.
- Ruta A., Porikli F., Watanabe S., Li Y. (2009b). « In-vehicle camera traffic sign detection and recognition », *Machine Vision and Applications*, décembre.
- Shaposhnikov D., Lubov N., Podladchikova L., Golovan A., Shevtsova N., Hong K., Gao X. (2002). « Road Sign Recognition by Single Positioning of Space-Variant Sensor Window », *Proceedings of 15th International Conference on Vision Interface*, Calgary, Canada, p. 213-217.
- Soheilian B., Arlicot A., Pappadimitris N. (2010). « Extraction de panneaux de signalisation routière dans les images couleurs », *Actes du 17ème congrès francophone en Reconnaissance de Formes et Intelligence Artificielle, RFIA'10*, Caen, France, p. 743-750, Janvier.

- Timofte R., Zimmermann K., Gool L. V. (2009). « Multi-View Traffic Sign Detection, Recognition, and 3D Localisation », *IEEE Workshop on Applications of Computer Vision WACV'09*, Institute of Electrical and Electronics Engineers, Snowbird, Utah, USA, p. 69-76, 7-9 décembre.
- Vitabile S., G.Pollacia, G.Pilato, F.Sorbello (2001). « Road signs recognition usign a dynamical pixel aggregation technique in the HSV color space », *Proceedings of 11th Int. conf. On Image Analysis and Processing (ICIAP 2001)*, p. 572-577.
- Zhu S., Liu L. (2006). « Traffic sign recognition based on color standardization », *Proceedings of the IEEE International Conference on Information Acquisition*, p. 951-955, août.

Article reçu le 22/02/2010
Version révisée le 11/08/2010



Rachid Belaroussi : ingénieur physicien de l'ESPCI-ParisTech en 2002 et Docteur de l'Université Pierre et Marie Curie (Paris VI), il a soutenu sa thèse en décembre 2006 sur la détection et la poursuite de visage dans des images et vidéos. Il est chercheur contractuel à l'Inrets-LCPC depuis février 2009.



Jean-Philippe Tarel : ingénieur de l'École Nationale des Ponts et Chaussées (1991) et Docteur de l'Université Paris IX-Dauphine (1996), il a effectué sa thèse au sein de l'Institut National de Recherche en Informatique et Automatique (INRIA). Depuis 1999, il est chargé de recherches au LCPC, après un postdoc à Brown University, USA, en 1997-98. À l'heure actuelle, il s'intéresse particulièrement à la détection d'objets, à l'estimation robuste, et à l'apprentissage statistique.