Road Sign Detection and Location Using PTIGHT Based on BP Neural Network

Jia-Jyun Yan  
National Taipei University of Technology Graduate Institute of Computer and Communication Engineering  
gary.wolf1@msa.hinet.net

Rong-Chin Lo  
National Taipei University of Technology Graduate Institute of Computer and Communication Engineering  
rclo@ntut.edu.tw

In the paper, we propose a new automatic road sign detection and location system to assist path planning and location of an autonomous land vehicle (ALV). In the system, first the Back-Propagation Neural Networks (BPNN) is used to search the road sign of interest (RSOI), then opening of morphological is employed to eliminate the noises of the RSOI. In the real environment road signs usually affected by affine distortion. Therefore, we use perspective-transformation-invariant generalized Hough transformation (PTIGHT) to detect and locate the road signs accurate. Several experimental results are also shown the practicability of the process system.

Keywords: Stereo Vision · Neural networks · Road Sign detection · Perspective · Generalized Hough transform (GHT).

1. Introduction

For the past few years, road sign detection and recognition are respected in Intelligent Transportation System (ITS) area. Road signs are used to regulate traffic, warn drivers, and provide useful information to help make driving safe and convenient.

Recently, many techniques have been developed to detect road signs [4]–[7]. However, most deal cases with single image and simple background, and most models of road signs were first described in terms of color histograms. Identifying a road sign extracted from an image was accomplished by simply comparing its color histogram. However, their method failed to account for the structural evidence of road signs. Some people combined colors and shapes to detect road signs. First, they used a color threshold to segment the input image. They then extracted special color regions, such as red and blue, from the image. Second, specific angle corners, including 60° and 90°, were used to convolve with these special color regions. Finally, to examine which corners constituted a fixed shape road sign, the shapes of road signs were used. Unfortunately, they did not consider the cover of the road sign and perspective of the road sign.

Due to the use of the point spread function to express the perspective reference table [3], the required dimensionality of the Hough counting space (HCS) for the PTIGHT is reduced to 2D. The PTIGHT is based on the generalized Hough transform (GHT) using to a new perspective reference table (PR-table). Intrinsically, the PR-table is a 2D array of cells that preserve all perspective information. In order to build a PR-table from a given planar shape, we apply perspective transformation on the given planar shape to derive all possible perspective shape images from all viewing positions and orientations. In this paper, how to determine parameters (angle, size, distance, etc.) of the road sign accurate is important. We propose a new approach as a contribution of two parts. First, we perform BPNN and morphological opening on a scene with cluttered background to detect the RSOI of sign and eliminate the noise. The other part, we use the PR-table, which is created from the real road sign with all perspective mappings, to extract the exact parameters.

2. Outline of road sign detection and location

The color space i.e. RGB usually changes with light intensities, fading and shadows. So, in the study we use the HSI model for road sign detection. The BPNN consists the dimension of input is three neurons, and number of output neuron is one.

![System framework diagram.](image-url)

The sigmoid transfer function is adopted for all, whose output range from 0 to 1. After recognition of BPNN,
the output of the neuron is 1 to make sure the input block is road sign and other cases are 0. Figure 1 shows the system-framework.

3. Forward and Inverse Perspective Transformations

There is much work [3] described the formulas for general forward and inverse perspective transformations between two planes. In this paper, we consider the perspective transformation relationship between the plane \( \pi_0 \) and its image plane \( \pi \) of a perspective planar shape as shown in Fig. 2. The framework of the transformation can be described with two coordinate systems. One is the camera coordinate system (CCS) denoted as X-Y-Z. The other is the natural coordinate system (NCS) denoted as \( X_0, Y_0 \) on perspective shape plane \( \pi_0 \). The observer’s viewpoint is located at the origin \( O = (0, 0, 0) \) of the CCS and the visual axis coincides with the Z-axis. Point \( F = (0, 0, C_0) \) in plane \( \pi_0 \) is called the fixation point, where the parameter \( C_0 \) represents the distance between the observer’s viewpoint and the fixation point.

![Perspective transformation relationship between a perspective shape plane and its image plane](image)

From the geometric relations shown in Fig. 2, the transformations between perspective shape plane \( \pi_0 \) and image plane \( \pi \) can be derived according to the following steps:

1. The equation of image plane \( \pi \) in the CCS is \( Z = f \), where \( f \) is the distance between the observer’s viewpoint (or the lens center of the camera) and the image plane.

2. The normal to perspective shape plane \( \pi_0 \) in the CCS is
   \[
   n = (\sin \tau \cos \sigma, -\sin \sigma, \cos \tau \cos \sigma)
   \]

3. The equation of perspective shape plane \( \pi_0 \) in the CCS is
   \[
   n \cdot [X \ Y \ Z] = C_2 \cos \tau \cos \sigma
   \]
   which can be transformed into
   \[
   Z = A X' + B Y' + C_2
   \]
   where \( A = -\tan \tau, B = \tan \sigma / \cos \tau \).

4. Let the viewpoint \( O \), point \( P \) in \( \pi \) with CCS coordinates \((X', Y', Z')\), and point \( Q \) in \( \pi_0 \) with CCS coordinates \((X'', Y'', Z'')\) be collinear. Then \( P \) is the image of \( Q \). This implies the following equations:
   \[
   \frac{X'}{X} = \frac{Y'}{Y} = \frac{Z'}{f}
   \]

Then, we can solve \( X', Y', \) and \( Z' \) from Eqs. (2) and (3) to be

\[
X' = \frac{C_2 X}{f - AX - BY}, \quad Y' = \frac{C_2 Y}{f - AX - BY},
\]

\[
Z' = \frac{f C_2}{f - AX - BY}
\]

5. Now, we can derive the inverse perspective transformation from image plane \( \pi \) to perspective shape plane \( \pi_0 \). First, we transform the CCS coordinates \((X', Y', Z')\) of \( Q \) in \( \pi_0 \) into its NCS coordinates \((X_0, Y_0)\) as follows:

\[
\begin{bmatrix}
X_0 \\
Y_0 \\
0 \\
1
\end{bmatrix} = \text{Tilt} (\sigma) \cdot \text{Pan}(\tau) \cdot \text{Translation} (-C_2) \cdot \begin{bmatrix}
X' \\
Y' \\
Z'
\end{bmatrix}
\]

where \( \text{Tilt}(\sigma) = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos \sigma & \sin \sigma & 0 \\
0 & -\sin \sigma & \cos \sigma & 0 \\
0 & 0 & 0 & 1
\end{bmatrix} \)

\( \text{Pan}(\tau) = \begin{bmatrix}
\cos \tau & 0 & -\sin \tau & 0 \\
0 & 1 & 0 & 0 \\
\sin \tau & 0 & \cos \tau & 0 \\
0 & 0 & 0 & 1
\end{bmatrix} \)
Then, from Eqs. (4) and (5), we obtain the inverse
perspective transformation, which maps the CCS
coordinates \((X, Y)\) of \(P\) to the NCS
coordinates \((X', Y')\) of \(Q\), as follows:

\[
\begin{bmatrix}
X' \\
Y'
\end{bmatrix} = \begin{bmatrix} K & L \\ f & 1 \end{bmatrix} \begin{bmatrix} X \\
Y \\
1 \end{bmatrix}
\]

where

\[
K = \frac{f}{\sin \sigma \tan \sigma + \cos \sigma (C_2 - X_0 \sin \sigma)} + Y_0 \cos \tau \tan \sigma
\]

\[
L = \begin{bmatrix}
\cos \sigma \cos \sigma & \sin \sigma \cos \sigma & \sin \tau \tan \sigma \\
0 & 1
\end{bmatrix}
\]

4. Describe of Conventional Generalized
Hough Transform

To use the conventional GHT to detect an
arbitrary template planar shape, it is necessary to set
up a 2D HCS. Each cell in the HCS has a value
specifying the possibility that the reference point of
the template shape to be detected is located at the cell.
Before performing GHT, the R-table for the template
shape is built up in advance by the following steps: (1)
select a suitable pixel \(R\) in the given template shape
as the reference point; (2) rotate the shape \(180^\circ\)
with respect to \(R\); (3) trace all the template shape
points and construct the R-table consisting of the
displacement vectors between all the shape points
and \(R\).

In the GHT process, all the displacement vectors
of the R-table are superimposed on each shape point
in the input image. The value of each corresponding
cell in the HCS pointed to by a displacement vector is
incremented by one. If there exists any cell with its
value exceeding a preselected threshold value and
being the maximum in the HCS, then it is determined
that the template shape is detected at the location of
the cell. See Figs. 3(a), 3(b), and 3(c) for an
illustrative example.
illustrated in Figs. 3(d) and 3(e). Sometimes, several candidate locations may be verified, and the one with the largest counter value is selected, which will be called the optimal candidate.

5. PR-Table Construction

Given an image $T$ of a planar shape which we surveyed before, we can construct the PR-table from it using the relevant formulas derived previously. As an illustrative example, let the given template shape image $T$ be shown in Fig.4 which includes a planar arrow shape. The PR-table are constructed from image $T$ by the following major steps.

![Figure 4. Three perspective images (denoted as a, b, and c) derived by applying perspective transformations with different parameters.](image)

(1) Survey the parameters of the image $T$ which is perpendicular to the ground, including size, distance, edge, etc.

(2) Compute the locations of the image points of all possible perspective shapes in the image plane $T$ with different perspective parameters. For examples, three possible perspective shape images denoted as a, b and c are shown in Fig.4.


In actual applications, given a template image of a planar shape and an input perspective image containing a perspectively-transformed version of the planar shape, how do we verify that the shape in the input image is a perspectively-transformed version of the planar shape in the template image. It is mentioned previously that the conventional GHT is unsuitable for this problem. So the PTIGHT is proposed. The basic steps are similar to those of the conventional GHT, including construction of a PR-table. However, an additional process, called inverse PTIGHT, is proposed for verification of candidate shape locations. The algorithm of PTIGHT as follows:

**Input:** the RSOI.

**Steps:**

1. perform the perspective transformation on the known road sign shape with Eqs. (1) through (7) to create the following template-set acted as PR-table:
   - Template-set = \{ $S_{\tau,\sigma,a}$ | $k \in \{1$ (rectangle), $2$ (triangle), $3$ (circle)\}, and $-40^\circ \leq \tau,\sigma \leq 40^\circ$, step $10^\circ$, and $5m \leq d \leq 9m$, step $1m$ \}, where $\tau = \text{pan}$, $\sigma = \text{tilt}$, $d = \text{distance}$.

2. Find the road sign from RSOI by voting with the PR-table (or Template-set) if the maximum value of HCS is larger than a threshold.

**Output:** Output the parameters of extracted road sign including $\tau$, $\sigma$ and distance.

7. Experimental results

In the system, we translated the road sign in different parameters including pan, tilt and distance to create the corresponding template. For convenience, the template is indexed by the number $m$. The number $m$ is defined as follows.

$$m = (\tau + 40) \times 10 + (\sigma + 40) + \frac{d}{100}$$
Figure 5. Show the experimental results with different parameters. (a) pan = 0°, tilt = 0°, d = 5m (a1) show the result of BPNN. (a2) show the result of morphological. (a3) show the edge. (b) pan = 40°, tilt = 0°, d = 5m, (c) pan = 40°, tilt = 0°, d = 9m.

All the experiments is implemented on a Intel T2450(2.0 GHz). The images of road signs are in size 260 × 300 pixels. The PR-Table contains 405 templates for each road sign shape.

8. Reference

[1] 塗 alguém, 葉永, 『多目的傳遞圖形分及位置』, 國立台北科技大學機電整合所, 2008。