RECOGNITION OF TRAFFIC SIGN SYMBOLS BY A GENERATIVE LEARNING METHOD

Hiroyuki Ishida, Tomokazu Takahashi, Ichiro Ide, Yoshito Mekada, Hiroshi Murase

Graduate School of Information Science, Nagoya University.
Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8603 Japan

ABSTRACT

We present a novel training method for recognizing traffic sign symbols. The symbol images captured by a car-mounted camera suffer from various forms of image degradation. To cope with degradations, similarly degraded images should be used as training data. Our method artificially generates such training data from original templates of traffic sign symbols. Experimental results show the effectiveness of the proposed method for traffic sign symbol recognition.

1. INTRODUCTION

Technologies for supporting drivers with car-mounted cameras have gained considerable industrial interest in recent years. Traffic sign recognition is one of the important tasks. Various attempts have been carried out on the detection of traffic signs: On the other hand, relatively few studies have been conducted on the category classification of extracted signs. Furthermore, most are mainly oriented toward high-quality images. This paper focuses on the classification of degraded traffic sign symbols. We propose a method that automatically generates training data in accordance with actual degradation characteristics.

2. TRAINING BY GENERATIVE LEARNING

This method consists of two steps. The first is the parameter estimation step. The second is the generation step of training data. Training data are generated from an original image by three degradation models: rotation, blurring, and segmentation error. These models are defined with generation parameters, as shown in Fig. 1.

Distribution of generation parameters is estimated from actual images, which are used to simulate degradations. The parameters for each image need to be estimated to calculate distribution. Parameter vector p consisting of the generation parameters is defined as:

$$p = (\theta_x, \theta_y, \theta_z, \gamma, \quad x, \quad y, w, h). \tag{1}$$

This vector is used to generate a degraded image from an original image. Figure 2 illustrates this estimation step.

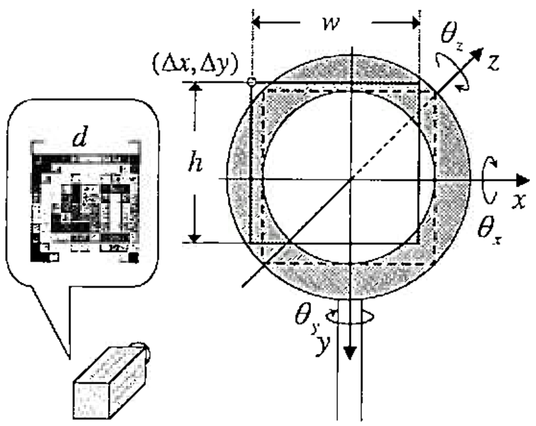

Fig. 1. Parameters for degradation models

Parameter distribution is estimated from multiple parameter vectors $\hat{p}$ computed from captured images. This method assumes a normal distribution. Average vector μ and covariance matrix Σ are then obtained from the multiple vectors $\hat{p}$ by

$$\mu = \mathcal{E}[\hat{p}], \tag{2}$$

$$\Sigma = \mathcal{E} \ (\hat{p} - \mu)(\hat{p} - \mu)^t \ . \tag{3}$$

Once parameter distribution is estimated, parameter vector g, which follows the estimated distribution, is produced from μ and Σ by the following parameter-producing function:

$$g = \Sigma^{\frac{1}{2}} r + \mu, \tag{4}$$

where r denotes a vector composed of standard normal random numbers and $\Sigma^{\frac{1}{2}}$ denotes the Cholesky decomposition of Σ. Figure 3 illustrates this generation step. Various parameter vectors are produced, and correspondingly, various training data of all categories and sizes are generated.

3. RECOGNITION METHOD

The subspace method is used in the recognition step. Let $x^{(c)}_{\{n,d\}}$ be a vector consisting of $d \times d$ pixels of category c's n-th training data whose size is d. Eigenvectors are constructed

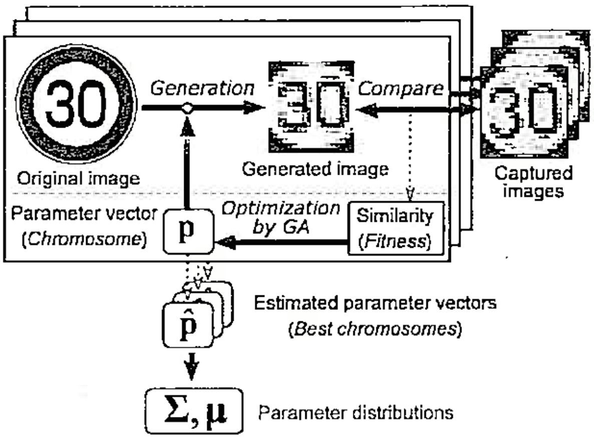

Fig. 2. Parameter estimation step

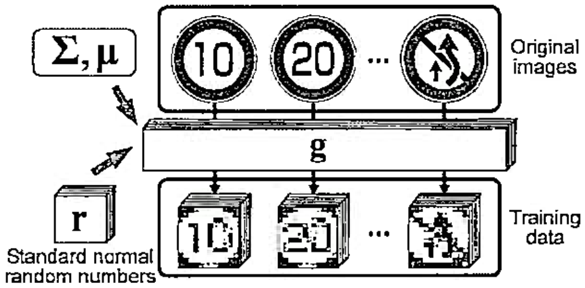

Fig. 3. Generation of training data

for each category c and for each size d. Initially matrix $X_d^{(c)}$ is constructed from N training data ($n = 1, \cdots, N$) by

$$X_d^{(c)} = \left[x_{\{1,d\}}^{(c)} \quad \cdots \quad x_{\{N,d\}}^{(c)} \right]. \tag{5}$$

Auto-correlation matrix $Q_d^{(c)}$ is computed by

$$Q_d^{(c)} = X_d^{(c)} \ {X_d^{(c)}}^{t}. \tag{6}$$

Eigenvectors are derived from $Q_d^{(c)}$, of which $e_{\{l,d\}}^{(c)}$ ($l = 1, \cdots, L$) with the largest L ($< N$) eigenvalues are used for recognition.

A given image is classified to category c that maximizes similarity, which is defined as the sum of the squared inner product between the given image and the eigenvectors. Given M image frames of the same target, let z_m be a vector normalized identically as the training data from the m-th image ($m = 1, \cdots, M$) whose size is d_m; the recognition result is obtained by

$$\hat{c} = \arg\max_{c} \sum_{m=1}^{M} \sum_{l=1}^{L} \left\langle e_{\{l,d_m\}}^{(c)}, z_m \right\rangle^2. \tag{7}$$

4. EXPERIMENT

An experiment was performed using video data captured by a car-mounted camera on a sunny morning. The video data contained fifteen traffic signs. Instead of Eq. (4), training data were generated using a parameter producing function in which $\Sigma^{\frac{1}{2}}$ was weighed on as

$$g = k\Sigma^{\frac{1}{2}} r + \mu. \tag{8}$$

The number of the generated training data was 200 ($N = 200$). Recognition rates in six cases ($k = 0, 1/4, 1/2, 1, 2, 4$) were compared. The case of $k = 1$ was identical to the proposed method, since Eq. (8) equals Eq. (4). In the recognition step, ten successive frames were integrated ($M = 10$), and ten eigenvectors were used ($L = 10$).

Recognition rates are presented in Table 1

Table 1. Recognition rates.

k	0	1/4	1/2	1	2	4
Rates (%)	57.4	89.2	91.7	92.9	91.4	91.2

Compared with the case of $k = 0$, in which similarity to an average pattern was evaluated, the recognition rates were drastically improved. Although the other cases of k ($k = 1/4, 1/2, 2, 4$) also exhibited high recognition rates, the case of $k = 1$ was the most effective. Note that the case using estimated distributions ($k = 1$) was the most appropriate for recognizing signs captured in similar conditions. This result indicates that GA-based parameter estimation successfully worked and also exhibited the superiority of the proposed method over the other ks.

5. CONCLUSION

In this paper, a method for recognizing traffic sign symbols was proposed. Degradation parameters were defined to simulate actual degradations. Based on the defined model, degradation characteristics were estimated from a small number of captured images for training, and training data for all categories were generated. The usefulness of our method for degraded traffic sign images was experimentally demonstrated.

The proposed method is applicable for any sign by combining it with other traffic sign detection methods. As for future research, the effectiveness of the method should be evaluated under various weather conditions and at various times of day.