A novel unsupervised approach to discovering regions of interest in traffic images

Zhenyu An [a], Zhenwei Shi [a,*], Ying Wu [b], Changshui Zhang [c]

[a] *Image Processing Center, School of Astronautics, Beihang University, Beijing 100191, P.R. China*

[b] *Department of Electrical Engineering and Computer Science, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA*

[c] *State Key Laboratory on Intelligent Technology and Systems, Tsinghua National Laboratory for Information Science and Technology (TNList), Department of Automation, Tsinghua University, Beijing 100084, P.R. China*

Abstract

Analyzing image of traffic scenes plays a major role in intelligent transportation systems. Regions of interest, including traffic signs, vehicles or some other man-made objects, largely attract drivers' attention. With different prior knowledge, conventional approaches generally define and build dedicated detectors to each class of such regions. In contrast, this paper focuses on explaining what regions in traffic images can be of interest, which is a critical problem yet neglected before. Instead of pre-defining the detectors, a computational model based on an unsupervised way is proposed. The core idea is to simulate an image with multiple bands from the given traffic image by stacking the spatial information. Our study shows that the distribution of such data can be captured by a simplex in a linear subspace, and each data point can be represented by a linear reconstruction over the set of vertices of the simplex. An effective method to identify the simplex vertices is proposed. These simplex vertices actually comprise the core elements in the regions of interest, as

physically they correspond to regions with saturated colors. Comparisons of the proposed approach and conventional methods on computational complexity and practical extensive experiments are implemented. The results validate and show the efficacy of the proposed approach.

Key words: Image of traffic scene (ITS); Regions of interest; Simplex vertex; Matrix factorization.

1 Introduction

Recently, intelligent transportation system attracts more and more attention for its wide applications. Generally, the entire ITS processing chain contains four parts, including segmentation [1–3], detection [4], recognition [5] and tracking [6]. Various methods have been proposed for different areas like sign detection and recognition [7–9], driver distraction detection [10,11] and occupant pose inference [12,13].

Among the different methods, thresholding approaches [1,2] are usually employed for segmentation purpose, in which visual experience always plays an important role. For the purpose of detection and classification, different dedicated descriptors, like edges [6], Histogram of Oriented Gradients (HOG) [9], Local Binary Pattern (LBP) [14] and their extensive forms, are carefully designed and applied. Canny edge detection or the similar methods are always applied as the only feature in [15], where Ruta et al. reported an 100% detection rate, but it is "the case for one of sign types" [16]. Haar-like features are introduced to be used with edges and also achieves effective results [17]. Some local features also attracted attentions. HOG is one of the represen-

* Corresponding author. Tel.: +86-10-823-39-520; Fax: +86-10-823-38-798.
Email address: shizhenwei@buaa.edu.cn (Zhenwei Shi).

tative feature which is based on creating histograms of gradient orientations on patches of the image, In [9] and [18], HOG is respectively combined with boosting methods and support vector machine (SVM) to accomplish detection and recognition tasks. Vehicles and pedestrians are also crucial targets as they convey traffic information. Negri et al. aimed at detecting pedestrians in surveillance video sequences [19], by designing a family of oriented histogram descriptors and a cascade of boosted classifiers. Duric et al. used Darboux motion model to estimate relative vehicle motions in traffic scenes [20]. Jia et al. discussed the detection of vehicles in front-view static images with frequent occlusions [21], they constructed Bayesian problem's formulations and designed Markov chain to detect vehicles.

However, two crucial problems still remain unsolved despite the above developments. First, an essential problem in ITS processing — "what are regions of interests (also denoted as ROIs)" — is neglected. Researchers paid much attention to classifying various features of interesting regions by applying training methods. These features are specifically designed and they exploit the spatial or texture information, but no deep researches or mathematical descriptions on the universal characteristics of ROIs are provided. The other problem is the neglect of color information application in ITS. Note that, although various segmentation methods could accomplish the task via color thresholding, they have not been deeply exploited. For instance, in order to use those different features, gray images are usually first obtained by removing colors of ITS [22]. Therefore, besides spatial information, how to use the color information in ITS should also be taken into consideration.

To solve the problems, two main issues will be discussed in the paper: Mathematical description and definition of ROI; Physical explanation of the definition. In the paper, we will first give a mathematical definition of ROI. To

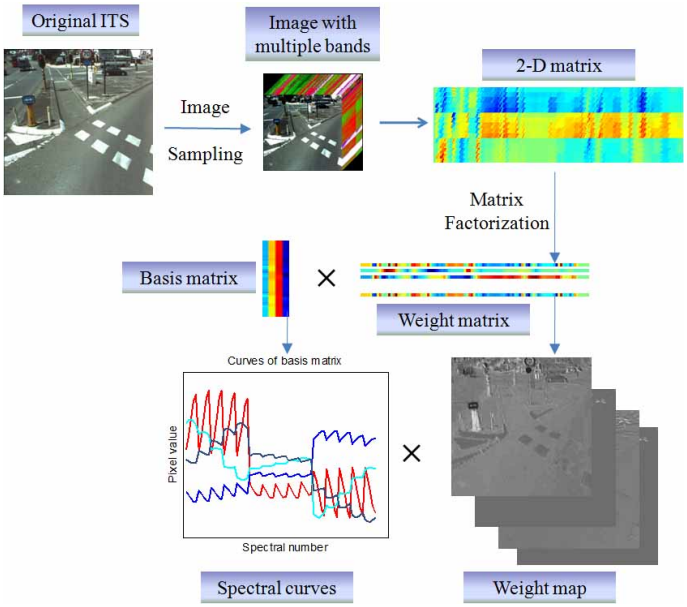

Fig. 1. Proposed processing chain for ITS with some critical steps, including image sampling, vertex extracting and matrix factorization. The factorized basis matrix and weight matrix has their definite physical meanings.

confirm its rationality, a computational model is then provided as illustrated in Fig. 1, with its core idea is to exploit the geometrical characteristic of ITS. In Fig. 1, the original ITS with colors red, green and blue is transformed into a 3-D image with its color number is larger than 3. In such circumstance, each color plane will be named as "band" or "spectrum" instead of "color" because their broader meanings. In the following discussion, the image with multiple bands will be named IMB for short. In the paper, one can see that, such data can be mathematically captured by a simplex in a proper linear subspace, and each data point can be represented by a linear reconstruction over the set of simplex vertices. Exploiting the characteristics, we will discuss the essence of ROI and its physical meaning.

The rest of the paper is organized as follows: In Section 2, the definition of ROI is provided. In Section 3, a novel processing chain for handling ITS is discussed. We propose an IMB simulating approach to obtain IMB. An effective vertex extracting algorithms is also proposed to obtain simplex vertices. In Section

4, experiments on the popular databases are implemented. Finally, the paper comes to the conclusion in Section 5.

2 Statement on regions of interest in ITS

Compared with images of other scenes, ITS has specifical characteristics for its vast man-made scenes. Since vehicles usually move on the roads, different objects in the ITS like traffic signs, buildings on either side of the roads, and even the other cars or pedestrians ahead of the camera could be important for drivers. Among them, traffic signs and lights are the most representative objects that inform drivers about the road conditions and restrictions. They are designed to be outstanding from the backgrounds for visual convenience. Therefore, these ROIs have relatively simpler but saturated colors. Generally, three characteristics of the ROIs in ITS could be summarized as follows:

i. An ITS may have several ROIs, and each region is inside connected.
ii. Only one color exists inside each region and it does not change much between the pixels in the region.
iii. The pixel colors in each region always have the same or similar values as the extremal values in one band of the ITS. For instance, if we have a ROI of blue traffic sign, then it has similar value as the extremal value of ITS in the blue band (sometimes they are the same).

To mathematically describe the ROIs, we first give the maximum and minimum values of band k in the ITS $\mathbf{F}^{I\times J\times K}$ as follows:

$1'$: $max\_bandk = max\{f(x,y,k), x \subseteq \{1,2,\ldots,I\}, y \subseteq \{1,2,\ldots,J\}, k \subseteq \{1,2,\ldots,K\}\}$

$2'$: $min\_bandk = min\{f(x,y,k), x \subseteq \{1,2,\ldots,I\}, y \subseteq \{1,2,\ldots,J\}, k \subseteq$

$\{1, 2, \ldots, K\}\}$

where I, J and K respectively denote the image width, height and band number, and $f(x, y, k)$ denotes the pixel value of ITS in location (x, y) and band k. Essentially, ROIs of ITS are the sets of coordinates and usually comprised by several subsets. For an ITS, its ROIs could be represented as: $C : \{C_1, C_2, ..., C_i, ..., C_M\}$, where M is the number of ROIs, and C_i is one of the ROIs which could be described in Definition 1.

Definition 1 (ROI in ITS) *Given a coordinate set C_i in ITS $\mathbf{F}^{I\times J\times K}$, it is one ROI of ITS if for any element $(x_0, y_0) \in C_i$ and two small thresholds $\delta_1 > 0$, $\delta_2 > 0$, there exists a small m-neighborhood area set $\Omega^m \subset C_i$ and a band $k \in \{1, 2, ..., K\}$. For every pixel $(x_t, y_t) \in \Omega^m$, the following four equations hold simultaneously:*

$$|x_0 - x_t| < m \tag{1}$$
$$|y_0 - y_t| < m \tag{2}$$
$$|f(x_0, y_0, k) - f(x_t, y_t, k)| < \delta_1 \tag{3}$$
$$min\{|max\_bandk - f(x_t, y_t, k)|, |min\_bandk - f(x_t, y_t, k)|\} < \delta_2 \tag{4}$$

Note that the definition is concluded from the process of traffic sign and signal designing, so it is more likely a descriptive concept, and it needs confirmation from other point of view. In the next section, a novel approach to study the intrinsic relationship between these ROIs of ITS and simplex vertices will be proposed.

3 Methodology

In this section, a novel approach for handling ITS will be discussed. The approach will be further used to confirm the rationality of the above definition

of ROI from the geometrical view. To begin with, we first give pre-knowledge on simplex as they will be exploited in the paper. The relationship between simplex and ITS is also discussed.

3.1 Pre-knowledge of simplex and ITS

Simplex is a widely applied fundamental concept [23,24]. For a simplex C_S with vertices $\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, ..., \mathbf{m}_k]$, each element $\mathbf{r} \in C_S$ could be reconstructed by:

$$\mathbf{r} = \mathbf{M}\alpha + n \tag{5}$$

where n is the noise, $\alpha = [\alpha_1, \alpha_2, ..., \alpha_k]^T$. In practical application, the weight vector α is usually nonnegative and sum-to-one, which means $\{\alpha \subset \mathbb{R}^k : \mathbf{1}^T\alpha = 1, 0 \preceq \alpha\}$, where $\mathbf{1}$ is a $k\times$ 1 vector of ones.

If we add the sum-to-one limitation to the pixels in original ITS, then the original RGB-color module could be normalized using the following equations:

$$N_R = \frac{R}{R+G+B}; N_G = \frac{G}{R+G+B}; N_B = \frac{B}{R+G+B} \tag{6}$$

where R, G, and B are the pixel values of bands red, green and blue, respectively. N_R, N_G and N_B stand for the projected value of three bands, respectively. This new color module, which is usually named normalized RGB (N-RGB) color module, has been widely applied in traffic image processing. The values in the N-RGB color module satisfy the sum-to-one restriction because for each pixel, equation $N_R + N_G + N_B = 1$ always holds.

Fig. 2 illustrates the scattered points of ITS in RGB and N-RGB color space. Compared with points in the original RGB color space as shown in Figs. 2(b) and (c), the scattered data points that have been projected to the N-RGB

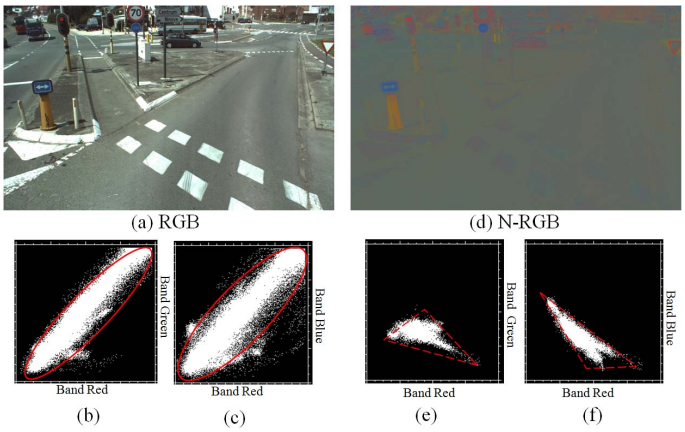

Fig. 2. Scattered points of ITS in RGB and N-RGB color module. (a) and (d) are the ITS in RGB and N-RGB color module, respectively. Figures (b), (c) and (e), (f) are the scattered points of figures (a) and (d), respectively.

space form a more standard simplex (Figs. 2(e) and (f)). Meanwhile, vertices are crucial elements for a simplex because every data point inside the simplex could be linearly represented by the vertices. This characteristic of simplex will be deeply exploited to explain the rationality of the definition on ITS. In the next section, we focus on finding the vertices in a proper subspace and exploring the relationship between them and ROI.

3.2 Three problems in designing methodology

To explore the intrinsic relationship between data simplex and ROIs, three problems should be taken into consideration and solved.

a. The first problem is how to increase the number of spectral dimension of ITS. For the original ITS with three colors, we could obtain only 3 vertices, and it is not enough for the complicated traffic scenes. Therefore, the first problem is how to increase the dimension of the data set and obtain more available vertices to represent the simplex of ITS. For the purpose, the adjacent pixels of the original ITS will be sampled and stacked as the data in spectral dimension of ITS.

b. The second problem is how to reduce redundancy and confirm the proper projected subspace. The sampling process introduces redundancy, which makes some points be outside of simple. Hence, how to remove the redundant information and obtain a simplex with clear boundaries is another crucial problem.
c. The third problem is how to extract vertices of the data simplex. In the paper, an extracting method based on the maximum volume is proposed. It is an effective method in vertex extracting and will be discussed in the following section.

3.3 Simulating image with multiple bands

The possible approach for the first question is to gain the band number of ITS. In this section, an available process for simulating image with multiple bands (IMB) is proposed, and it is based on sampling spatial information in the original ITS.

An example for simulating IMB with 12 bands is illustrated in Fig. 3. For a given ITS $\mathbf{F}^{I\times J\times K}$, where I, J, and K are respectively the width, height and band number, we use a sliding window with the size $w \times h$ to go through it in both horizontal and vertical directions. Three steps are then involved in the simulating process.

1) Sampling: We sample pixels using the above window. Different patches will be obtained and each patch has the size $2 \times 2 \times 3$ because ITS usually has three color planes. Therefore, each patch contains three subpatches and each subpatch has the size 2×2.
2) Vectorization: For each color (i.e. for each plane) of the sampled patch, the 2×2 pixels are regrouped as a 4×1 vector. The regrouping approach

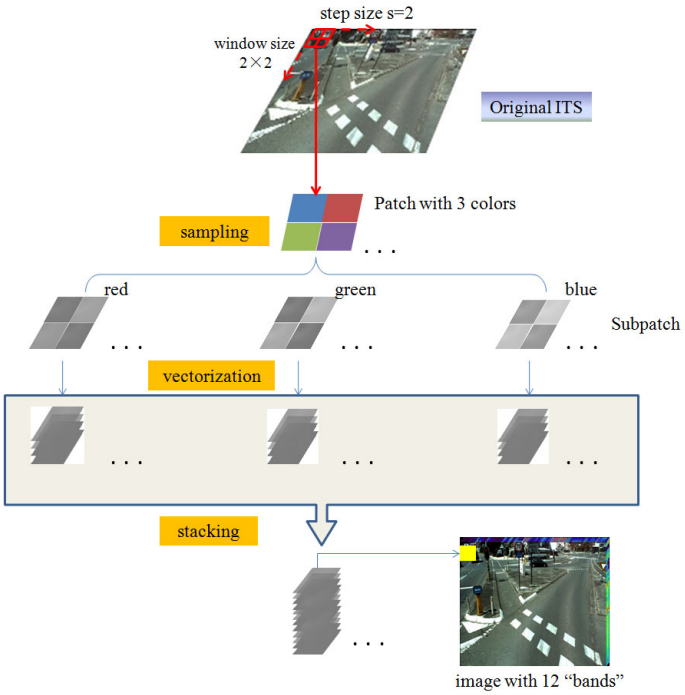

Fig. 3. Process for simulating image with multiple bands. This is an example for simulating IMB with 12 bands, as the window size is 2 × 2 and the step size is 2.

is just a simple vectorization process. For every patch, we could obtain three vectors as shown in Fig. 3.

3) Stacking: For each patch, we simply stack the above three vectors and obtain a vector with size 12×1. We thus obtain one pixel in the simulated data (as denoted as the yellow square in the right bottom image). Since the sampling process is implemented through the whole image, amount of patches and vectors are obtained. They will be further stacked, and we finally obtain the new data—IMB.

In the process, two main parameters affect the size of the obtained IMB: one is the sliding window size $w \times h$, the other is the sliding step size s. The spatial resolution of the simulated IMB will be the same with the original ITS if the step size is 1. On the contrary, the spatial resolution of IMB will degrade if the step size is larger than 1. The band number of IMB relies on the size of sliding window and it is three times of sliding window width and height,

namely $w \times h \times 3$. This process is a simple yet effective approach for increasing the dimension of the original ITS, and the simulated IMB would be used in the subsequent processing.

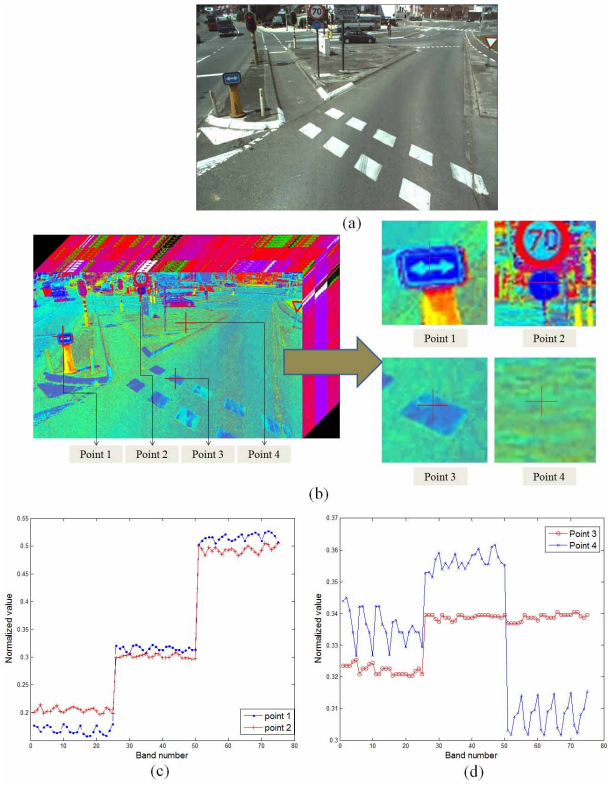

Fig. 4. Simulated IMB result and its spectra. (a) Original ITS. (b) Simulated IMB with 3-D structure and four points in the IMB. (c) The similar spectral curves of points 1 and 2. (d) The different spectral curves of points 3 and 4.

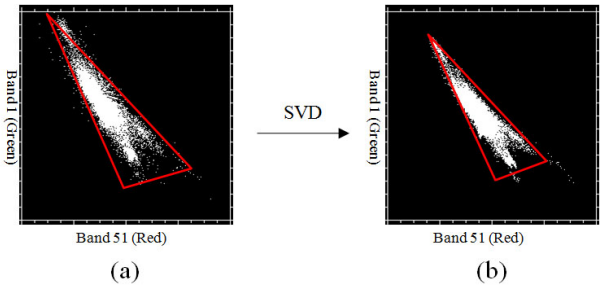

Fig. 5. Examples of scattered points of IMB before and after using SVD. (a)-(b) show the scattered points of band 1 with respect to band 51 before and after using SVD, respectively.

Fig. 4 illustrates an simulated IMB and its spectral curves. An IMB with 75 bands is obtained by implementing the proposed simulating process. Fig. 4(a) is the original ITS. The left figure in Fig. 4(b) is the simulated IMB with de-

graded spatial resolution because the step size is set to $s = 2$. It is shown with false color for visual convenience. Obviously, the gained bands are essentially the vectorization of pixels in a sliding window in different color planes of ITS. Besides the increasing bands, the following critical characteristics could also be concluded from the simulated results: 1. Objects with similar colors have similar spectral curve trends. In practical application, spectral curves of points 1 and 2 (white color in the original ITS) have similar shapes as displayed in Figs. 4(b) and (c). 2. Objects with different colors usually contain different spectral information. As shown in Figs. 4(b) and (d), the curves of two points in the background (points 3 and 4) have significant differences because of the different neighboring pixels in ITS.

3.4 SVD and dimension identification

The simulating process could effectively help to obtain a new image with more band numbers. However, if the window size is too large, then the IMB will have too many bands. Besides, the window size is usually set to be larger than step size. Although this could avoid abandoning some useful information, it may introduce redundancy because some pixels may be used too many times. For instance, In Fig. 5(a), although the scattered points in band 1 and 51 of IMB (originally the red and blue bands in ITS, respectively) also approximately form the simplex, some data points are outside the simplex, which could be regarded as the redundancy of the data. Therefore, before finding a proper subspace where simplex could perfectly captures the data sets, we should reduce the redundancy. In the paper, singular value decomposition (SVD) [25,26] will be applied to accomplish the task.

For a given matrix $\mathbf{A}^{L\times N}$, SVD finds three matrices $\mathbf{U}^{L\times L}$, $\Lambda^{L\times N}$ and $\mathbf{V}^{N\times N}$

which satisfy $\mathbf{A} = \mathbf{U}\Lambda\mathbf{V}^T$ where $\mathbf{V}^T$ is the transpose of $\mathbf{V}$. The columns of $\mathbf{U}$ are the eigenvectors of the matrix $\mathbf{A}\mathbf{A}^T$ and the columns of $\mathbf{V}$ are the eigenvectors of the matrix $\mathbf{A}^T\mathbf{A}$. If the eigenvalues of $\mathbf{A}\mathbf{A}^T$ are $\lambda_1, ..., \lambda_r$ (r is the rank of $\mathbf{A}\mathbf{A}^T$) and $\lambda_1 \geq \lambda_2 \geq ..., \geq \lambda_r$, then the matrix Λ are composed by setting $\Lambda_{ii} = \sqrt{\lambda_i}$ for $1 \leq i \leq r$, and zero otherwise. Usually, Λ_{ii} is named as the singular value. Λ_{ii} is ranked from large to small, only the largest several singular values could reconstruct the original data with negligible error. Projecting the data to the space that spanned by the eigenvectors which corresponds to the several largest singular values, we could obtain the projected data that represents the essential structure of the original data. Meanwhile, the number of the largest singular values could be regarded as the vertex number of simplex. Therefore, in the proposed method, SVD could be applied as a preprocessing step to denoise data and project the original data to a proper space. The whole steps are shown as in the Algorithm, where the threshold δ is an experimental parameter and it has effects on the reducing dimension. In the practical application, 0.999 is good for ITS.

Algorithm 1 Algorithm for denoising and reducing redundance using SVD.

(1) Transform the simulated 3-D IMB into a 2-D data $\mathbf{A}^{L\times N}$, where L is the spectral number and N is the pixel number.

(2) Implement SVD (the function *svd* in software MATLAB) on the data $\mathbf{A}\mathbf{A}^T$, we obtain matrices $\mathbf{U}^{L\times L}$, $\Lambda^{L\times N}$ and $\mathbf{V}^{N\times N}$ that satisfy $\mathbf{U}\times\Lambda\times\mathbf{V} = \mathbf{A}\mathbf{A}^T$.

(3) Initialize $p = 1$, and increase p until the condition $\frac{\sum_{i=1}^{p}\Lambda_{ii}}{\sum_{i=1}^{r}\Lambda_{ii}} > \delta$ is first meet, then p is the vertex number. Herein, δ is a proper threshold.

(4) Implement SVD (the function *svds* in MATLAB) on the data $\mathbf{A}$ with p largest values, then the parameter of function *svds* is set to p and we obtain $\mathbf{U}_p$, $\mathbf{S}_p$ and $\mathbf{V}_p$ that satisfy $\mathbf{U}_p\times\mathbf{S}_p\times\mathbf{V}_p \approx \mathbf{A}$. The approximately equal symbol is adopted because we have neglected small eigenvalues via the function *svds*. Finally, we could calculate the projected matrix $\mathbf{A}_p$ of original data $\mathbf{A}$ by using $\mathbf{A}_p = \mathbf{U}_p^T \times \mathbf{A}$.

Results of the denoised IMB via SVD are illustrated in Figs. 5(c) and (d). The 3-D image and its scattered figures of band 1 (originally red band) with

respect to bands 51 (originally blue band) are displayed. Compared with Figs. 5(a) and (b), most interference points in the original data are eliminated. Meanwhile, we see that the points in the red triangle of Fig. 5(d) become much denser than those in Fig. 5(b), which implies that we have effectively denoised IMB while preserving the simplex property by applying SVD. Next, we focus on how to extract vertices from the simplex.

3.5 Extracting simplex vertices with maximum volume

A simplex vertex extracting method based on maximizing volume is proposed by synthesizing the methods N-FINDR [27] and vertex component analysis (VCA) [28], which respectively face the problem of computational amount and insistent results. As shown in Fig. 6, the scattered points form a simple convex set, and it could be captured by different simplexes (the green dash line and the red solid line). The different simplexes have different areas or volumes. Obviously, the simplex with the maximum volume (red solid line) is the better one because it suits the data set and the data points are all inside the simplex. To obtain such simplex, we first give the formula for calculating the volume of simplex. For a simplex with vertices $\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, ..., \mathbf{m}_p]$, then the volume of the simplex could be mathematically written as:

$$V(\mathbf{M}) = \frac{1}{(p-1)!}|det([\begin{matrix}\mathbf{M} \\ \mathbf{1}^T\end{matrix}])| \tag{7}$$

where $p!$ is factorial operation, $\mathbf{1}$ is a column vector with all its values are 1. $|det(\cdot)|$ calculate the absolute value of determinant.

The extracting method for vertices is proposed with its core principle as illustrated in Fig. 7. In this 2-D convex set Cs, each point $\mathbf{r}_i \in Cs$ is first projected

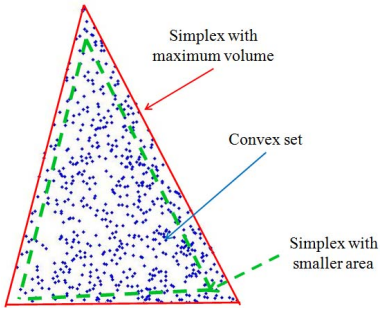

Fig. 6. Illustration of maximum volume for convex set. The triangle with red solid lines is more appropriate compared with the green dashed lines, as the former has larger volume while surrounding the data set.

onto a randomly initialized hyperplane $\mathbf{f}_1$ with $v_i = \mathbf{r}_i^T * \mathbf{f}_1$, where v_i is the projected data. Therefore, we could obtain a vector that comprised by these data $\mathbf{v} = v_1, v_2, ..., v_N$, where N is the total number of the set Cs. Hence, the point $m_\mathbf{A} \in C_s$, with the maximum value in the projected $\mathbf{v}$ that could be calculated by $< m_\mathbf{A}^T, \mathbf{f}_1 >= max(\mathbf{v}(j), j = 1, ..., N)$, is exactly the first vertex in the figure. After that, with the equation $\mathbf{f}_2 = \mathbf{I} - \mathbf{B}\mathbf{B}^+$, we obtain the second hyperplane $\mathbf{f}_2$ that is orthogonal to the subspace spanned by $\mathbf{f}_1$. Herein, $\mathbf{I}$ is a vector with all its values are one. $\mathbf{B}$ is the space spanned by $\mathbf{f}_1$ and $\mathbf{B}^+$ is the pseudo inverse of $\mathbf{B}$. Projecting the original convex set points to $\mathbf{f}_2$, and we obtain the second vertex $m_\mathbf{B}$ that is corresponding to the maximum projected value. Similarly, the process is iterated until all the vertices are obtained. And we could obtain all the vertices of the data simplex and they form the basis matrix $\mathbf{M}$.

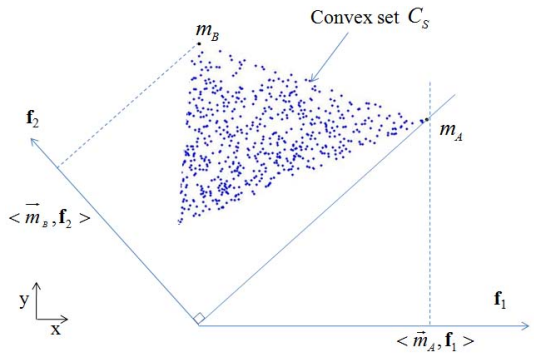

Fig. 7. Two-dimensional example for extracting vertices of data simplex.

An imperceptible problem in the above extracting step is the uncertainty of the extracting results, which is caused by the random initialization of $\mathbf{f}_1$. To obtain a relatively more stable result, we repeat the searching process and obtain different simplexes with different volumes. As discussed above, the best simplex contains the maximum volume. Thus, the simplex with the maximum volume is exactly the result we expect.

Each fractional weight vector α in equation (5) could be calculated after obtain the vertex vectors. Assume that each column of the matrix $\mathbf{H}$ is comprised by the weight vector, then the above factorization could be written as $\mathbf{A} = \mathbf{M} \times \mathbf{H}$, where $\mathbf{A}$ is the observation data (the IMB in our problem) and each column of $\mathbf{M}$ is a vertex vector. As the data set and basis matrix have been obtained, this problem could be simplified into a least squares problem. Different algorithms could be used for it. Least square method (LSM) [24], a widely used method, will be applied in our approach. Mathematically, it could be calculated as $\mathbf{H} = \mathbf{M}^{+}\mathbf{A}$, where $\mathbf{M}^{+}$ is the pseudo inverse. Note that, each row of $\mathbf{H}$ represents the distribution of the object corresponding to a specifical vertex. It could be reshaped into a 2-D map for visual analysis, as shown in the right bottom of Fig. 2. Such map could be quite useful in other image processing areas, and it will be further discussed in the experiments.

The pseudocode for the entire processing chain is shown in Algorithm 2. In the algorithm, operations $\mathbf{A}(:, k)$ and $\mathbf{A}(k, :)$ respectively represent the k-th column and k-th row elements of matrix $\mathbf{A}$. $\mathbf{v}(k)$ is stands for the k-th element of vector $\mathbf{v}$. Operation $\mathbf{B} = \{\mathbf{B}, \mathbf{C}\}$ is used to store $\mathbf{C}$ into $\mathbf{B}$, and $\mathbf{B}\{k\}$ represents the k-th element (matrix or set) of $\mathbf{B}$.

Algorithm 2 Proposed processing chain for ITS.

(1) Initialization: given an ITS $\mathbf{F}^{I\times J\times K}$, sampling window size $w \times h$, step size s, maximum iteration times $maxiter$, threshold δ, $\mathbf{I}\text{ndex}_{all} = \emptyset$ and matrix set $\mathbf{M}_{all} = \emptyset$ (They are used to respectively store iteration number and the calculated basis matrix in the iteration).

(2) Transform the ITS in RGB color module into a new version in N-RGB color module.

(3) Implement the image simulating process with the sampling parameters w, h and s. Obtain the IMB $\mathbf{G}^{W\times H\times L}$ where L is the band number and $L = w \times h \times 3$.

(4) Vectorizing each band of $\mathbf{G}$ into a row vector that has the size $1 \times N$ where $N = W \times H$, and we could obtain the 2-D data $\mathbf{A}^{L\times N}$.

(5) Obtain the subspace dimension p and projected data A_p by $\mathbf{A}_p = \mathbf{U}_p^T\mathbf{A}$ as stated in the Section 3.4.

(6) **for** $kk = 1$ to $maxiter$

(a) Randomly initialize the $p \times p$ auxiliary matrix $\mathbf{B}$;

(b) **for** $ii = 1$ to p

(i) Implementing the equation: $\mathbf{f} = ((\mathbf{I}-\mathbf{B}\mathbf{B}^+))/\|(\mathbf{I}-\mathbf{B}\mathbf{B}^+)\|_F$, we obtain the vector $\mathbf{f}$ that is orthonormal to the subspace spanned by $\mathbf{B}$. $\|\cdot\|_F$ is the Frobenius norm. $\mathbf{B}^+$ could be calculated with the function $pinv$ in the MATLAB.

(ii) Project the data $\mathbf{A}_p$ into the vector $\mathbf{f}$ by $\mathbf{v} = \mathbf{f}^T\mathbf{A}_p$.

(iii) Calculate the extreme point of $\mathbf{v}$ and update the matrix $\mathbf{B}$.

$$k = argmax_{j=1,\ldots,N}|\mathbf{v}(j)| \tag{8}$$

$$\mathbf{B}(:, ii) = \mathbf{A}_p(:, k) \tag{9}$$

(iv) Update the expected vertex matrix:

$$\mathbf{M}(:, ii) = \mathbf{A}_p(:, k) \tag{10}$$

$$index(:, ii) = k; \tag{11}$$

(c) **end for**

(d) Store the indices of simplex and calculate the volume of simplex (Vos) spanned by $\mathbf{M}$ via implementing equation (7)

$$\mathbf{I}\text{ndex}_{all} = \{\mathbf{I}\text{ndex}_{all}, index\} \tag{12}$$

$$\mathbf{M}_{all} = \{\mathbf{M}_{all}, \mathbf{M}\} \tag{13}$$

$$Vos(kk) = V(\mathbf{M}_{all}(kk)) \tag{14}$$

(7) **end for**

(8) Obtain the simplex with maximum volume.
$ind = argmin_{j=1,\ldots maxiter}(Vos(j))$,
$\mathbf{M}_{final} = \mathbf{U}_p\mathbf{A}_p(:, \mathbf{I}\text{ndex}_{all}\{ind\})$.

(9) Calculate the weight matrix $\mathbf{H}$ via the LSM, namely, the equation $\mathbf{H} = (\mathbf{M}_{final}^T\mathbf{M}_{final})^{-1}\mathbf{M}_{final}^T\mathbf{A}$.

3.6 Complexity analysis

The computational complexity of the proposed method contains two main parts, SVD and vertex extraction. In the former part, the decomposition of $\mathbf{A}\mathbf{A}^T$ costs $2NL^2+L^3$ flops, and the projection $\mathbf{A}_p = \mathbf{U}_p^T \times \mathbf{A}$ costs $2NLp$ flops. In the latter part, vertex extraction step costs $2Np^2$ flops (the computation of $\mathbf{f}^T\mathbf{A}_p$) in each iteration, and the calculation for weight matrix $\mathbf{H}$ cost about $2NLp$ flops. The whole computational complexity is $2NL^2 + L^3 + 4NLp + 2Np^2 \times maxiter$. Note that, the relation $N \gg L > p$ holds, so the computational cost could be simplified as $Comp1 \approx 2NL^2 + 4NLp + 2Np^2 \times maxiter$.

3.7 Relationship between the vertices and ROIs

In this section, analyses on the relationship of the simplex vertices in ITS and its ROIs will be provided. For simplicity, we take an ITS $\mathbf{F}^{I\times J\times K}$ with two bands (namely, $K = 2$) as an example. Under the circumstance, its vertex extracting process is exactly the same as shown in Fig. 7.

For this image with two bands, we have its maximum values of two bands which could be denoted as max_bandx and max_bandy in x and y directions, respectively. (m_{Ax}, m_{Ay}) and (m_{Bx}, m_{By}) are used to represent the coordinates of points m_A and m_B. Meanwhile, $f(m_{Ax})$ and $f(m_{By})$ respectively denote the pixel value in x coordinate of vertex m_A and pixel value of vertex m_B in y coordinate. As shown in Fig. 7, vertices m_A and m_B respectively have the maximum values in x and y directions, namely, the following two equations hold:

$$f(m_{Ax}) = max\_bandx; f(m_{By}) = max\_bandy$$

Note that when applying SVD to project the IMB data into a proper space

and obtain dimension number of data in the extracting process, we also have accomplished image denoising step. Thus, the data points with maximum values would not be emerged 'alone', because the pixel would be regarded as the noise and eliminated if it has few adjacent pixels. Therefore, in terms of the extracted vertex m_A, there exists a small m-nearest neighborhood area Ω_1, and for every pixel $(x_1, y_1) \in \Omega_1$, it has the similar pixel value as the m_A, thereby leading the following equations:

$$|m_{Ax} - x_1| < m \tag{15}$$
$$|m_{Ay} - y_1| < m \tag{16}$$
$$|f(m_{Ax}, m_{Ay}) - f(x_1, y_1)| < \delta_1 \tag{17}$$

where δ_1 denotes a small value. Note that the inequality (21) is just a special form of (8), so after combining the equation (17) with the above three equations (21)-(23), we will see that these four expressions completely satisfy the definition of ROI as in Definition 3. Hence, we could come to the critical conclusion that the extracted vertex m_A is just in the ROIs of ITS. Obviously, the same conclusion on the vertex m_B could also be obtained and demonstrated in the similar way. This conclusion is easily comprehended in the intuitive sense. For a simplex, the most important elements are its vertices or the data points near the vertices as they always attract people's attention at first glance. So they could be regarded as the interesting points in data simplex. Experiments on the real data will confirm the efficacy of the proposed methods.

4 Experiments

In this section, we report results of experiments, aiming at discussing the relationship between the simplex vertices and the ROIs in practice. Compar-

isons between the proposed method and some conventional color thresholding methods will also be implemented.

Two publicly available ITS databases will be used as shown in Fig. 8: 1) The German Traffic Sign Detection Benchmark (GTSDB) [1]; 2) KUL Belgium Traffic Signs data set (KUL Data set) [2] [29]. Experiments will be implemented on those data sets.

(a) GTSDB (c) KUL

Fig. 8. Examples of two data sets. (a)-(b) are respectively images of GTSDB and KUL data sets.

4.1 Experiments on the GTSDB data set

In the experiments on GTSDB data set, parameters are empirically set to $w \times h = 3 \times 3$ and $s = 3$. Fig. 9 illustrates the results using the proposed method. Green crosses in Fig. 9(a) mark the location of the vertices. Spectral curves of the vertices are shown in Fig. 9(b). Figs. 9(c)-(f) illustrate the weight maps.

The vertices are usually located in the ROIs of ITS. In Fig. 9(a), points 1 and 3 are respectively located in the blue traffic sign and red car. Point 2 is located in the green grasses and point 4 is located in the edges of the white car. Generally, most of the data points in the ROIs that convey crucial traffic

[1] url: `http://benchmark.ini.rub.de/`
[2] url: `http://www.esat.kuleuven.be/psi/visics/software-and-datasets`

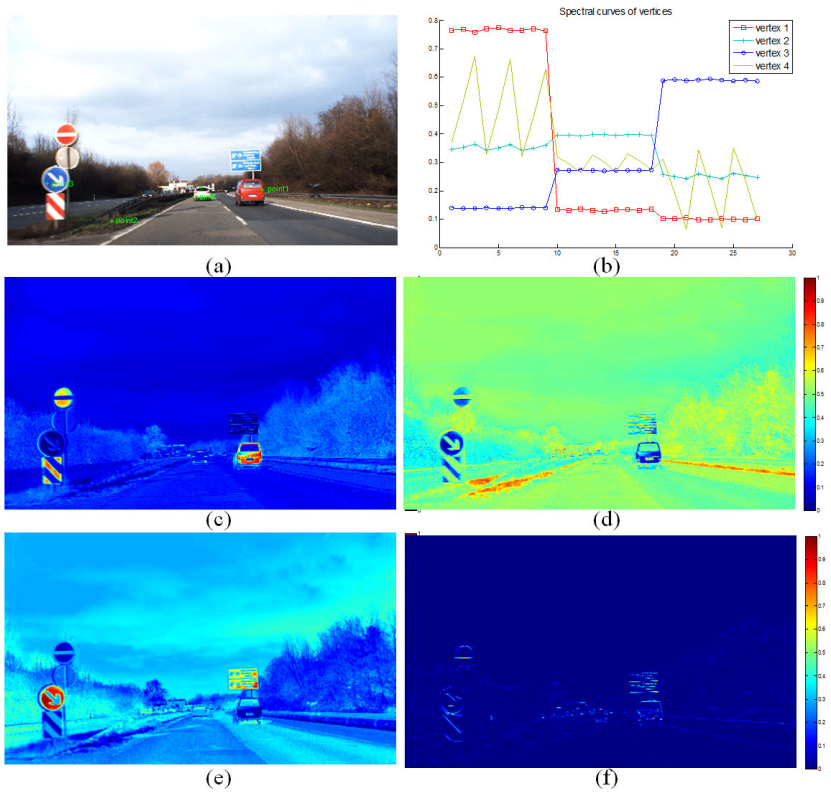

Fig. 9. Extracted vertices using the proposed method. (a) is the original ITS and the green crosses mark the location of the vertices. (b) show the spectral curves of the vertices. (c)-(f) illustrate the weight maps.

information have similar saturated colors. Therefore, the obtained vertices of data simplex are usually located in the ROIs with dominant color. The conclusion that the simplex vertices of ITS are the elements in ROIs could be then obtained, as it is theoretically proved in section 3.7. Fig. 9(b) shows the curves of the vertex vectors. Curves 1 to 3 respectively represent the colors red, green and blue in the original ITS. Curve of vertex 4 represent a mixed color, so it has a more irregular shapes compared with the other three curves. It informs us that the different objects have different curve shapes.

Each row vector of $\mathbf{H}$ could be transformed into the same spatial size of IMB. Therefore, weight matrix $\mathbf{H}$ stands for distribution of the corresponding vertex and it will facilitate the subsequent image processing. Fig.s 9(c)-(f) shows all the weight maps with false color for visual convenience. The "red" pixel

represents the larger weight of the corresponding vertex, while the "blue" pixel represents the less components of the corresponding vertex. Figs. 9(c) and (e) respectively show the red traffic signs and car, which are visually prominent. The green belts along the road are outstanding in Fig. 9(d), and the scene in Fig. 9(f) represents the edges of the original ITS. Weight maps not only provide us with information on components of the basis, but also have apparent physical meaning. The scene corresponding to the vertices will be outstanding in the map. Using simple thresholding method, we could easily obtain a segmenting result. This characteristic would be quite helpful for the subsequent segmentation and detection. In the next section, we will quantitatively compare the proposed method and some conventional methods based on assessing the segmentation results.

4.2 Comparison experiments with other different methods

Since ROIs have relatively higher pixel values than backgrounds in the maps, one natural application of the weight map is ROI segmentation. Using simple thresholding method, we could segment the ROIs from the image. To quantitatively assess the proposed method, we will compare the segmenting results of the proposed method and methods in different color space, including RGB [30], HSV [31], and Ohta [32]. Besides, some popular automatic image segmentation methods including K-means [33], Mean-shift [34] and Chan-Vese [35] are also used to evaluate the performance because of the their close relationships. Before experiments, traffic signs of images from the databases are manually labeled. Although the proposed method could obtain ROIs with different colors simultaneously, we only compare the segmenting results of the blue traffic signs for fairness. Therefore, all the class numbers in the automatic segmentation methods are set to 2. The metric error rate (ER) is used

to quantitatively assess the segmentation results of different methods. ER is defined as:

$$ER = \frac{BT + TB}{TN} \tag{18}$$

where BT represents the number of background pixels identified as target pixels, TB represents the number of target pixels identified as background pixels, and TN denotes the total number of pixels of the input image. Obviously, its reference value is 0.

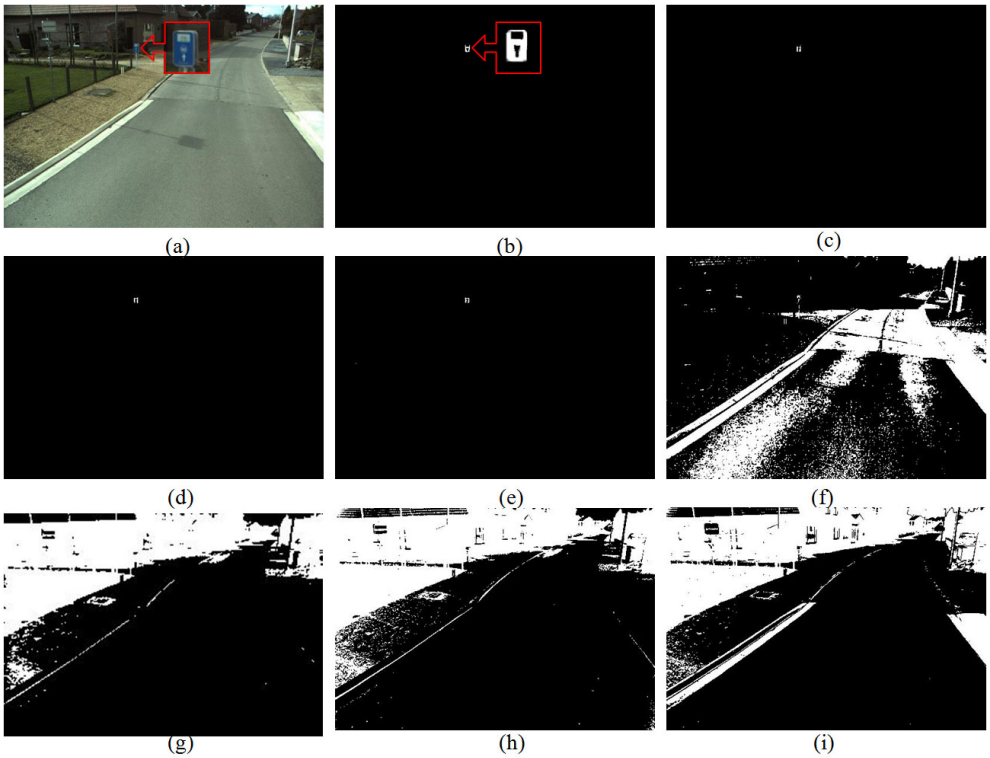

Fig. 10. Illustration of comparison results using different methods. (a) Original ITS. (b) Labeled image. (c)-(f) Segmentation results using the proposed methods, methods in Ohta, HSV and RGB color space. (g)-(i) Segmentation results using the methods Chan-Vese, K-means and Mean-shift

Fig. 10 shows the results using the above different methods as an example. Compared with the results in Figs. 10(f)-(i), the results of Figs. 10(c)-(e) are much better because they are much closer to the labeled image in Fig. 10(b). It implies that the proposed method and methods in HSV and Ohta obtains behave much better than other methods. ER of the different methods

are presented in tabular form in Table 1. The metric in bold type are the best of all. Again, the proposed method obtains the lowest error rate, while the conventional segmentation methods including Chan-Vese, K-means and Mean-shift are not approximate for handling ITS. Nevertheless, it should be clear that the segmentation experiment is just one nature application of the proposed method, the more critical destination of the method is to provide us a novel view on the relations between ROIs and the data simplex vertices.

Table 1
Calculated ER using different methods.

	Our method	Ohta	HSV	RGB	Chan-Vese	K-means	Mean-shift
ER	**0.0006**	0.0009	0.0011	0.0021	0.3807	0.4816	0.4067

5 Conclusion and prospective work

In the paper, a novel unsupervised approach to discovering the ROIs of ITS based on exploring spectral domain is proposed. In the approach, an IMB is transformed from the original ITS, and it forms a convex set which could be captured by a simplex. To obtain the best simplex with maximum volume, an effective vertex extracting method is designed. Then the linear combination model in equation (5) and the LSM are applied to calculate the weight matrix $\mathbf{H}$. In addition, mathematical definition of ROIs is provided based on its characteristics. By comparing the definition and the vertex extracting method, we conclude that the simplex vertices are in the ROIs of ITS, which is also demonstrated by the experiments on different databases. Therefore, the proposed approach provides us a novel view on ROIs of ITS, as it exploits the spectral information and geometrical characteristics of ITS data.

One natural application of the proposed method derives from the physical meaning of the weight matrix $\mathbf{H}$, as it provides important information on the

object distributions. Each row of **H** could be transformed into a 2-D image with the same spatial size as the IMB, then different objects corresponding to the vertices will be salient in different maps. Hence, these maps facilitate to extracting crucial traffic information and are important for the subsequent image segmentation or object detection, which also demonstrates the efficacy of the proposed method.

6 Acknowledgment

The work was supported by the National Natural Science Foundation of China under the Grants 61273245 and 91120301, the 973 Program under the Grant 2010CB327904, the Program for New Century Excellent Talents in University of Ministry of Education of China under the Grant NCET-11-0775, the funding project of State Key Laboratory of Virtual Reality Technology and Systems, Beihang University under the Grant VR-2014-ZZ-02, and the Fundamental Research Funds for the Central Universities under the Grant YWF-14-YHXY-028.

References

[1] H. Gomez-Moreno, S. Maldonado-Bascon, P. Gil-Jimenez, and S. Lafuente-Arroyo, Goal evaluation of segmentation algorithms for traffic sign recognition, IEEE Transactions on Intelligent Transportation Systems, 11(4) (2010) 917-930.

[2] H. Fleyeh, Shadow and highlight invariant colour segmentation algorithm for traffic signs, in Proc. IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2006, pp. 1-7.

[3] S. Chabrier, B. Emile, H. Laurent, C. Rosenberger, and P. Marche, Unsupervised

evaluation of image segmentation application to multi-spectral images, in Proc. 17th International Conference on Pattern Recognition (ICPR), 2004, vol. 1, pp. 576-579.

[4] M. Garcia-Garrido, M. Sotelo, and E. Martin-Gorostiza, Fast traffic sign detection and recognition under changing lighting conditions, in: Proc. IEEE Conference on Intelligent Transportation Systems, M. Sotelo, Ed., 2006, pp. 811-816.

[5] W. Ritter, Traffic sign recognition in color image sequences, in: Proc. IEEE Conference Intelligent Vehicles Symposium, Detroit, MI, 1992, pp. 12-17.

[6] M. Meuter, C. Nunn, S. M. Gormer, S. Muller-Schneiders, and A. Kummert, A decision fusion and reasoning module for a traffic sign recognition system, IEEE Transactions on Intelligent Transportation Systems, 12(4) (2011) 1126-1134.

[7] J. Greenhalgh and M. Mirmehdi, Real-Time detection and recognition of road traffic signs, IEEE Transactions on Intelligent Transportation Systems, 13(4) (2012) 1498-1506.

[8] F. Zaklouta and B. Stanciulescu, Real-time traffic-sign recognition using tree classifiers, IEEE Transactions on Intelligent Transportation Systems, 13(4) (2012) 1507-1514.

[9] G. Overett and L. Petersson, Large scale sign detection using HOG feature variants, in: Proc. IEEE Intelligent Vehicles Symposium, 2011, pp. 326-331.

[10] M. Trivedi and S. Cheng, Holistic sensing and active displays for intelligent driver support systems, Computer, 40(5) (2007) 60-68.

[11] C. Tran and M. M. Trivedi, Vision for driver assistance: Looking at people in a vehicle, Visual Analysis of Humans, (2004) 597-614.

[12] A. Doshi, S. Cheng, and M. Trivedi, A novel active heads-up display for driver assistance, IEEE Transactions on Systems, Man, and Cybernetics-Part B, 39(1) (2009) 85-93.

[13] A. Doshi and M. Trivedi, Attention estimation by simultaneous observation of viewer and view, in: Proc. IEEE Computer Vision and Pattern Recognition Workshops (CVPRW) 2010, pp. 21-27.

[14] S. H. Lee, J. Y. Choi, Y. M. Ro, and K. N. Plataniotis, Local color vector binary patterns from multichannel face images for face recognition, IEEE Transactions on Image Processing, 21(4) (2012) 2347-2353.

[15] A. Ruta, Y. Li, and X. Liu, Real-time traffic sign recognition from video by class-specific discriminative features, Pattern Recognition, 43(1) (2010) 416-430.

[16] A. Mogelmose, M. M. Trivedi, and T. B. Moeslund, Vision-based traffic sign detection and analysis for intelligent driver assistance systems: perspectives and survey, IEEE Transactions on Intelligent Transportation Systems, 13(4) (2012) 1484-1497.

[17] V. Prisacariu, R. Timofte, K. Zimmermann, I. Reid, and L. Van Gool, Integrating object detection with 3D tracking towards a better driver assistance system, in: Proc. 20th International Conference on Pattern Recognition (ICPR), Aug. 2010, pp. 3344-3347.

[18] Y. Xie, L.-F. Liu, C.-H. Li, and Y.-Y. Qu, Unifying visual saliency with HOG feature learning for traffic sign detection, in: Proc. IEEE Intelligent Vehicles Symp., Jun. 2009, pp. 24-29.

[19] P. Negri, N. Goussies, P. Lotito, Detecting pedestrians on a Movement Feature Space, Pattern Recognition, 47 (2014) 56-71.

[20] Z. Duric, R. Goldenberg, E. Rivlin, and A. Rosenfeld, Estimating relative vehicle motions in traffic scenes, Pattern Recognition, 35(6) (2002) 1339-1353.

[21] Y. Q. Jia, C. H. Zhang, Front-view vehicle detection by Markov chain Monte Carlo method, Pattern Recognition, 42 (2009) 313-321.

[22] M. L. Song, D. C. Tao, C. Chen, X. L. Li and C. W. Chen, Color to gray: visual cue preservation, IEEE Transactions on Pattern Analysis and Machine

Intelligence, 32(9) (2010) 1537-1552.

[23] N. Jorge and W. Stephen, Numerical optimization, Springer. 2nd ed., 2006.

[24] S. Boyd and L. Vandenberghe. Convex optimization, Cambridge University Press. 2004.

[25] S. Deerwester, S. Dumais, T. Landauer, G. Furnas and R. Harshman. Indexing by latent semantic analysis, Journal of the American Society of Information Science, 41(6) (1990) 391-407.

[26] J. Demmel and W. Kahan, Accurate singular values of bidiagonal matrices, SIAM Journal on Scientific and Statistical Computing, 11(5) (1990) 873-912.

[27] M. E. Winter, N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data, in: Proc. SPIE Image Spectrometry V, 1999, vol. 3753, pp. 266-277.

[28] J. M. P. Nascimento and J. M. B. Dias, Vertex component analysis: A fast algorithm to unmix hyperspectral data, IEEE Transactions on Geoscience and Remote Sensing, 43(4) (2005) 898-910.

[29] R. Timofte, K. Zimmermann, and L. Van Gool, Multi-view traffic sign detection, recognition, and 3D localisation, in: 2009 Workshop on Applications of Computer Vision (WACV), New York: Springer-Verlag, Dec. 2011, pp. 1-15.

[30] V. Prisacariu, R. Timofte, K. Zimmermann, I. Reid, and L. Van Gool, Integrating object detection with 3D tracking towards a better driver assistance system, in: Proc. 20th ICPR, Aug. 2010, pp. 3344-3347.

[31] S. Lafuente-Arroyo, P. Garcia-Diaz, F. J. Acevedo-Rodriguez, P. Gil-Jiménez, and S. Maldonado-Bascon, Traffic sign classification invariant to rotations using support vector machines, in: Proc. Advanced Concepts for Intelligent Vision Systems, Aug./Sep. 2004, pp. 256-262.

[32] Y. Ohta, T. Kanade, and T. Sakai, Color information for region segmentation, Computer Graph and Image Processing, 13(3) (1980) 222-241.

[33] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, An efficient k-means clustering algorithm: Analysis and implementation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(7) (2002) 881-892.

[34] D. Comaniciu and P. Meer, Robust analysis of feature spaces: color image segmentation, in: Computer Vision and Pattern Recognition, Proceedings., 1997 IEEE Computer Society Conference on, 1997, pp. 750-755.

[35] T. F. Chan, B. Y. Sandberg, and L. A. Vese, Active contours without edges for vector-valued images, Journal of Visual Communication and Image Representation, 11(2) (2000) 130-141.