Towards an automatic system for road lane markings extraction in large-scale aerial images acquired over rural areas by hierarchical image analysis and Gabor filter

HANG JIN*†, YANMING FENG† and MAOXUN LI‡

† Faculty of Science and Technology, Queensland University of Technology, Brisbane, Australia
‡ College of Urban Economics and Public Administration, Capital University of Economics and Business, Beijing, P. R. China

An automatic approach for road lane markings extraction from high-resolution aerial images is proposed, which can automatically detect the road surfaces in rural areas based on the hierarchical image analysis, and the procedure is facilitated by the road centerlines obtained from the low-resolution images. The lane markings are further extracted on the generated road surfaces with 2D Gabor filters. The proposed approach is performed on the aerial images of the Bruce Highway around Gympie, Queensland. Evaluation of the generated road surfaces and lane markings using four representative test fields has validated the proposed method.

1. Introduction

Accurate and detailed road models are required in many applications, such as traffic monitoring, and advanced driver assistance systems, which are designed to achieve active safety purposes through lane departure warning, lateral collision prevention, and speed limit etc. The next generation car navigation systems tend to be cooperative navigation systems using vehicle to vehicle and vehicle to infrastructure communications, requiring the detailed road surfaces and roadside safety features, e.g. the number of lanes, the locations of lanes, symbol markings, various traffic and road signs. This paper focuses on the automatic extraction of road surfaces and lane details.

The popular method for lane information acquirement is through vehicle-based Mobile Mapping System (MMS), where the road lane markings can be detected and reconstructed in the field using laser scanner or close range photogrammetric imagery. Although accurate lane features can be obtained through this method, it is costly and tedious to acquire the accurate road details especially over large rural areas.

Another approach to acquire lane data is through feature extraction from remotely sensed images, which has been a long-standing research topic within photogrammetry and remote sensing community. However, due to the limitation of the ground resolution of images, the majority of existing approaches have only focused on the detection of road centerline rather than the lane details. Research efforts have been centralized in a number of institutions, resulting in various approaches to the problem, including multi-scale approaches (Baumgartner *et al*. 1999), knowledge-based extraction (Trinder and Wang 1998) and context cues (Hinz and Baumgartner 2000).

∗ Corresponding author. Email: jinhang3695@gmail.com

Only a few approaches involved the detection of lane markings in the road extraction. For instance, Steger *et al*. (1997) extracted the collinear road markings as bright objects in large scale photographs when the roadsides exhibit no visible edges. Hinz and Baumgartner (2003) utilized road marking features, detected as thin bright lines with symmetric contrast, as the evidence for the presence of a road. Another approach of road extraction with pavement markings detection was presented by Zhang (2004), where the road markings and zebra crossing are segmented based on coloristic and geometric characteristics. In a word, road pavement markings were only regarded as a clue to reconstruct the road network for most of the existing methods. Therefore, the quality requirements (Tournaire and Paparoditis 2009), such as robustness, quality, completeness, are far below the lane level applications. In a more recent work, Kim *et al*. (2006) built a system to extract pavement information relied on simple image processing algorithms. Tournaire and Paparoditis (2009) proposed a specific approach for dashed lines and zebra crossing reconstruction, which relied on external knowledge introduced in the detection and reconstruction process based on primitives that extracted in the images. With the event of rapid development of sensor technologies, the spatial resolution have being greatly increased, which makes it feasible to accurately detect the lane features from aerial images.

In the reminder of this paper, the hierarchical image analysis combined with Gabor filtering is introduced into a new framework for road lane markings extraction from geo-referenced aerial imagery.

2. Methodology

In order to detect the road surfaces and lane markings appropriately, it is beneficial to understand the physical characteristics of road features, which can directly influence their visual characteristics.

2.1 *Road characteristics*

The properties of roads can basically be divided into two types: spectral and geometric (Gruen and Li 1995, Vosselman and Knecht 1995). The spectral properties of roads refer to different materials of road surfaces, e.g. asphalt or concrete, which appear as particular color signature on the images. The geometric properties refer to the constant widths, continuous curvature changes, and homogeneous local orientation distributions etc.

The characteristics of road pavement markings as specified by (Queensland Department of Main Roads 2001) are: (i) their shapes and sizes are constrained by strict specifications; (ii) they constitute high contrasted objects. The geometric specifications of road pavement markings are given in figure 1. The color of pavement marking is generally white. Marking lines, which may be unbroken (continuous) or broken, are categorized as longitudinal, transverse and parking control. In this paper, only the extraction of longitudinal markings is concentrated on, as they comprise the overwhelming majority of road pavement markings.

Figure 1. Geometric specifications of the pavement markings in a rural arterial road environment (Queensland Department of Main Roads 2001).

2.2 *Hierarchical image analysis*

The hierarchical approach is developed on the facts that different characteristics of road features can be optimally detected in different scales. Indeed road lane details can only be detected in high-resolution images. The problem is that many local

disturbances such as shadows or vehicle can also greatly degrade the results. On the other hand, the roads centerlines can be extracted in a stable manner from lower resolution images even in the presence of background objects such as trees or cars, where they are basically treated as homogeneous bands with different lengths and orientations. Therefore, it is beneficial to extract the desired road features from images of different resolutions and subsequently combine the individual results into a single refined output.

2.3 *Gabor filtering*

2D Gabor filter, which was extended from 1D Gabor by Daugman (1985), have been successfully applied to a variety of image processing and pattern recognition problems, such as image enhancement (Lindenbaum *et al*. 1994), image compression (Daugman 1988), texture analysis (Bovik *et al*. 1990), object detection (Weber and Casasent 2001), and image segmentation (Soares *et al*. 2006). 2D Gabor filter can be used to extract the road lane markings thanks to its following properties: (i) tuneable to specific orientations, (ii) adjustable orientation bandwidth, and (iii) robust to noise. Furthermore, it has optimal joint localization in both spatial and frequency domains. Therefore, Gabor filter can be considered as orientation and scale tunable edge and line (bar) detectors (Manjunath and Ma 1998), which makes it a superior tool to detect the geometrically restricted linear features, such as the road pavement markings.

2.3.1 The Gabor functions. The general functionality of the 2D Gabor filter family can be represented as a Gaussian function modulated by a complex sinusoidal signal. Specifically, the 2D Gabor filter can be defined in both the spatial domain $g(x, y)$ and the frequency domain $G(u, v)$. The 2D Gabor function in spatial domain can be formulated as (Cai and Liu 2000):

$$g(x, y) = \exp\left\{-\pi\left(\frac{x_r^2}{\sigma_x^2} + \frac{y_r^2}{\sigma_y^2}\right)\right\}\exp\{j2\pi(u_0 x + v_0 y)\} \tag{1}$$

Its 2D Fourier transform is expressed as

$$G(u, v) = \exp\left\{-\pi\left[(u - u_0)_r^2\sigma_x^2 + (v - v_0)_r^2\sigma_y^2\right]\right\} \tag{2}$$

where $j = \sqrt{-1}$; σ_x and σ_y are the scaling parameters of the Gaussian envelope for the x- and y-axis, respectively; (u_0, v_0) presents the spatial frequencies of the sinusoid carrier in Cartesian coordinates, which can also be expressed in polar coordinates as (f, ϕ), where $f = \sqrt{u_0^2 + v_0^2}$, $\phi = \arctan(v_0/u_0)$, and the subscript 'r' stands for a rotation operation as follows:

$$x_r = x\cos\theta + x\sin\theta$$

$$y_r = -y\sin\theta + y\cos\theta$$

where θ is the rotation angle of the Gaussian envelope.

2.3.2 Determination of Gabor filter parameters. The correct determination of Gabor filter parameters is the central issue for the lane pavement markings extraction process. In order to effectively and accurately extract road lane markings with different sizes and thicknesses from aerial images using Gabor filters, we proposed an efficient method to determine the Gabor filter parameters.

1) Determination of θ

θ stands for the orientation of the span-limited sinusoidal grating. The rough direction φ ($\varphi \in [0, \pi)$) of the road centreline is already obtained from the low

resolution image, thus the orientation θ ($\theta \in [0, \pi)$) of Gaussian envelope is given as perpendicular to φ by:

$$\theta = (\varphi + \pi/2)\%(\pi)$$

where % is the modulo operator.

2) Determination of f

f is the frequency of the sinusoid, which determines the 2D spectral centroid positions of the Gabor filter. This parameter is derived with respect to the width of road lane markings. If f is set too large, the Gabor filter will produce two peaks at the edges of road markings. Therefore, in order to produce a single peak for the given line, the output of the Gabor filter must satisfy the following conditions:

$$Out(t_0) \geq Out(t_1)$$

where W is the line width, and $W/2 \geq t_1 > t_0 \geq 0$. From the above condition, the frequency can be obtained within the following bounding:

$$0 < f \leq 1/W_m$$

where W_m is the width of the road marking in pixel. The details of the proofing process can be referred to (Liu *et al*. 2003). Other ground objects such as white vehicles may also be detected if the frequency is set too small. Thus, in order to only produce peaks on the road markings, the parameter f also should satisfy:

$$f > 1/W'$$

where W' is the width of other white features, and $W' > W_m$. Therefore, the frequency f is given by:

$$1/W' < f \leq 1/W_m$$

In our experiments, we set $f = 1/W_m$, which will produce only a single peak in the output of the filter on road markings regardless of the values of σ_x and σ_y.

3) Determination of σ_x and σ_y

The parameters σ_x and σ_y determine the spread of the Gabor filter in ϕ and θ directions respectively. According to (Liu *et al*. 2003), σ_x and σ_y have the following parameter constraint:

$$\sigma_y = k\sigma_x$$

where k is a constant. As the road lane markings have strict orientation and enough distance between adjacent lanes, we set $k = 1$ to simplify the calculation.

The relationship between the orientation bandwidth $\Delta\theta$ and the frequency f within the frequency domain is illustrated in figure 2, which can be given by:

$$\Delta\theta = 2\arctan\left(\frac{l}{f}\right)$$

where $\Delta\theta$ is the orientation bandwidth, and both h and l are radius. It gives

$$l = f\tan(\Delta\theta/2)$$

Figure 2. The uncertainty relation in frequency domains. $\sigma_x = \sigma_y$, $\phi = 90°$, $\Delta\theta$ is the orientation bandwidth, and both h and l are radius (only the positive frequencies are shown).

Applying the 3dB frequency bandwidth in V direction when $\phi = 90°$ to equation (2), we have

$$G(u_0, h)|_{\phi=90°} = \exp[-\pi(h\sigma_x)^2] = \sqrt{2}/2$$

It gives

$$h = \frac{\sqrt{\frac{\ln 2}{2\pi}}}{\sigma_x}$$

As $h = l$, σ_x can be further given as

$$\sigma_x = \frac{\sqrt{\frac{\ln 2}{2\pi}}}{f\tan(\Delta\theta/2)}$$

According to orientation bandwidths of cat cortical simple cells (Liu *et al.* 2003), the mean angle covers a range from 26° to 39°. After examining the line extraction results over the above range, we find it appropriate to set $\Delta\theta = 30°$. Then σ_x and σ_y can be further obtained by:

$$\sigma_x = \sigma_y = 0.58/f$$

3. Road feature extraction scenario

The workflow of our road lane markings extraction algorithm is given in figure 3. We can see that the algorithm can be basically divided into three stages: (i) image preprocessing, (ii) road centerline extraction in low resolution images, and (iii) road surfaces and lane markings detection in high resolution images. The preprocessing is designed to geometrically correct the raw aerial image and improve its quality. Road centerline extraction in the low resolution image can further provide the orientation and location of the road surfaces, and road markings are finally extracted on the road surfaces using Gabor filter. Besides, a shifting window process, which transfers the original large image into sub-images along the road centerline, is utilized to deal with large aerial images.

Figure 3. Detailed flowchart of the proposed system for road markings extraction.

The geometric correction, which can be accomplished with the commercial digital photogrammetric software, is employed to compensate for the distortions introduced by factors such as variations of the sensor platform, and relief displacement. After the image ortho-rectification, the geometric specification of the lane markings can be utilized in the road marking extraction. The hierarchical image analysis is then employed to detect the road centerlines and surfaces in low and high resolution images, respectively.

3.1 *Shifting window process*

A single frame of digital aerial image acquired by Vexcel UltraCam-D has a dimension of 11500×7500 pixels with three channels (Red, Green and Blue), which occupies 253 M bytes, and a mosaic image could be even larger. It is not practical to directly input such a large image into the road marking extraction process. Therefore, a "shifting widow" image process is proposed and implemented to optimize the road lane feature extraction from such large images.

After the processing of low resolution images, the generated centerlines of the road surfaces are then mapped to the high resolution image, and a chain structure is required to present the road centerline, given by:

$$\sum_{i=0}^{n-1} P_i,$$

where P_i is the ith pixel on the centerline with coordinates (x_i, y_i), and n is the total number of pixels.

Based on the road centerline obtained from the low resolution image, the high resolution image can be transformed into sub-images along the road centerline. This sub-image with $N \times N$ pixels is called a shifting window, where a new sub-image is introduced based on the location of the previous sub-image as well as the road centerline. The shifting window process is illustrated in figure 4, where the square boxes represent the shifting windows, the dashed line is the road centerline, and the two solid lines indicate the road sides.

<div align="center">Figure 4. The illustration of the shifting window process.</div>

To optimally locate the shifting windows within the high resolution image, the adjacent shifting windows must not overlap, and they also must cover the whole road surface without omission. To deal with this problem, an automatic sub-image generation method, which is developed from the bisection theorem, is developed in this paper.

The original bisection method is a root-finding algorithm which repeatedly bisects an interval and then selects a subinterval in which a root must lie for further processing. Here, we are intended to find the centers of the shifting windows by repeatedly bisecting an interval and then selecting a subinterval in which the optimal center pixels must lie.

The steps to find the optimal center pixel of the shifting window are:

(1) Select the first pixel $P_0 = (x_0, y_0)$ on the road centerline $\sum_{i=0}^{n-1} P_i$, and set $x_{\text{start}} = x_0, x_{\text{end}} = x_0 + w$, where x_{start} is the minimum x coordinate of the shifting window, x_{end} is the maximum x coordinate of the shifting window, and w is the width of the shifting window.

(2) Calculate the maximum y coordinate y_{max} and the minimum y coordinate y_{min} within the range $[x_{\text{start}}, x_{\text{end}}]$; and then compare the value of $y_{\text{max}} - y_{\text{min}}$ with h (height of the shifting window):

    (a) if $y_{\text{max}} - y_{\text{min}} < h$, then save $(x_{\text{center}}, y_{\text{center}})$ to the center pixel list C, where $x_{\text{center}} = (x_{\text{start}} + x_{\text{end}})/2$, and $y_{\text{center}} = (y_{\text{max}} + x_{\text{min}})/2$. Go to step 5.

    (b) if $y_{\text{max}} - y_{\text{min}} > h$, then set $x_{\text{temp}} = (x_{\text{end}} + x_{\text{start}})/2$, where x_{temp} is the temporary variable in the calculation procedure.

(3) Calculate the maximum y coordinate y_{max} and the minimum y coordinate y_{min} within the range $[x_{\text{start}}, x_{\text{temp}}]$; and then compare the value of $y_{\text{max}} - y_{\text{min}}$ with h again:

if $y_{\text{max}} - y_{\text{min}} < h$, then set $x_{\text{temp}} = (x_{\text{end}} + x_{\text{temp}})/2$.

if $y_{\text{max}} - y_{\text{min}} > h$, then set $x_{\text{end}} = x_{\text{temp}}$, and $x_{\text{temp}} = (x_{\text{start}} + x_{\text{temp}})/2$.

(4) Continue step 3 until $|x_{\text{temp}} - x_{\text{end}}| < \varepsilon$, where ε is the threshold to stop the finding loop. ε is defined as 1 pixel to accurate locate the shifting window in

this paper. Then set $x_{\text{end}} = x_{\text{temp}}$, and save $(x_{\text{center}}, y_{\text{center}})$ to the center pixel list C, where

    (a) $y_{\text{center}} = (y_{\text{max}} + x_{\text{min}})/2$;

    (b) The value of x_{center} is as followings:

        (i) $x_{\text{center}} = w/2$ if $x_{\text{end}} < w$;

        (ii) $x_{\text{center}} = x_{\text{end}} - w/2$ if $x_{\text{end}} > w$ or $x_{\text{start}} < W - w$, where W is the width of the high resolution image;

        (iii) $x_{\text{center}} = W - w/2$ if $x_{\text{start}} > W - w$; End the finding loop.

(5) Set $x_{\text{start}} = x_{\text{k}}$, $x_{\text{end}} = x_{\text{start}} + w$, where x_{k} is the last pixel within the former shifting window, then go to step 2.

After applying the above steps, we obtained the center pixel set $C = \sum_{i=0}^{k-1} P_i$ of all shifting windows, where P_i is the location of window center for the i_{th} shifting, and k is the total number of shifting windows. A shifting window is determined by its center pixel as well as its width and height.

3.2 Road centerline detection in low-resolution image

The original image is firstly decomposed to the desired resolution using the 2D discrete wavelet transform (DWT). In the image decomposition via wavelet transform, properly choosing wavelet is an important issue. The Bior(9-7) filter bank is utilized for the image decomposition, as it has almost the best performance in maintaining good visible quality within over 4300 candidate filter banks tested in (Villlasenor *et al.* 1995). The road can then be detected in the obtained low resolution image with fewer disturbances such as shadows or vehicle.

To ensure that only necessary features would involve in the separation of road class, color space transform is employed to select the appropriate image chromatic information. Features that contribute little to the image classification should be discarded. Therefore, the principal component analysis (PCA) is utilized to reduce the dimension while preserve the essential information.

The road surfaces appear relatively white compared with other features, which means it has relatively low saturation values. The saturation is given by:

$$S = 1 - \frac{\min(R,G,B)}{I} \tag{3}$$

where R, G, and B indicate the Red, Green and Blue band value, and I is the intensity, defined as $I = \frac{R+G+B}{3}$

As the vegetation areas have relatively low value of blue component in RGB, the C_b component in YC_rC_b color space (a color space used as the color pipeline in video and digital photography systems) is selected to distinguish the road surfaces from these vegetation features. C_b is given by:

$$C_b = -0.148R - 0.291G + 0.439B + 128 \tag{4}$$

The 1st PCA component, the saturation band, and the C_b component are stretched by histogram equalization, and they are further fused using 1st PCA component as R band, saturation component as G band, and C_b component as B band.

After the data preparation, the image segmentation approach is used to classify the road surfaces from other ground objects. The ISODATA method is used in this work to segment the aerial image due to the fact that an unsupervised clustering method is usually superior to a supervised clustering approach as no training set is needed.

The connected component analysis is utilized to remove small noises that are misclassified into road class. A modified Wang-Zhang thinning algorithm

(Krishnapuram and Chen 1993) is further chosen to extract the road skeletons thanks to its property of structure-preserving.

Significant variations in road surfaces reflectance and shadows may disconnect the extracted road segments. In (Amini *et al.* 2002), a perceptual grouping algorithm was used to restore the longer straight road sides segments from the fragmented line segments. Mainly based on the geometric properties, a similar approach is used to link the disconnected road centerline segments here. Basically, two segments are linked if the following constraints are fulfilled:

1) Two centerline segments are adjacent and their endpoints are within 10 pixels.
2) Two centerlines are approximately collinear, which means that the difference of their orientations τ_o and the lateral distance τ_d are less than a given threshold, here $\tau_o < 5°$ and $\tau_d < 2$ pixels.
3) The overlapped part of the two adjacent centerlines is less than 10% of the length of either segment.

The road centerlines should have homogeneous local orientation distributions. However, the generated road centerline is actually zigzag due to the disturbances from shadows or variation of road width. The linear regression is an appropriate method to fit a linear relation between the extracted road centerline points.

3.3 *Road surfaces detection in high-resolution image*

The road centreline has been obtained from the low-resolution image in the previous step. Meanwhile, the orientation and the rough position of the road surfaces within high-resolution image also can be easily acquired. The orientation φ is calculated using the two endpoints $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ of the centreline segment (See figure 5), which can be derived by

$$\varphi = \arctan\left(\frac{y_2 - y_1}{x_2 - x_1}\right)$$

To indicate the rough position of the road surfaces in high-resolution images, a buffer is constructed using the acquired information. The generated road centreline may not be exactly located in the centre of the road surfaces, thus the buffer width should be large enough to contain the whole road surfaces. In our experiment, the buffer width W_b is defined to be the average width of the road surfaces, which can be easily calculated by (Liu *et al.* 2003):

$$W_b = \frac{2S_{um}}{S_{um} - S'_{um}}$$

where S_{um} is the total sum of pixels on the extracted road surfaces, and S'_{um} is the sum of pixels for the road surfaces after deleting its edge points.

The buffer zone is as shown in figure 5 between the two dashed lines in parallel with the solid black line. The road surfaces should be within the buffer zone, and thus the majority of other features outside the buffer can be removed, which will greatly reduce the complexity of image classification. As the vegetation regions have relatively low value of C_b component, only the C_b band is utilized to extract the road surfaces in high resolution image.

Figure 5. The representation of orientation φ and the buffer zone, *w* is the average width of road surface, P_1 and P_2 are the two endpoints of road centreline.

As one of the widely used techniques for monochrome image segmentation, histogram thresholding is utilized here to segment the aerial image. The threshold is then automatically determined by the Otsu's algorithm (Otsu 1979), which finds the optimal threshold T by maximizing

$$V(T) = \frac{\left(\bar{\mu}\omega(T) - \mu(T)\right)^2}{\omega(T)\mu(T)}$$

where $\omega(T) = \sum_{i=0}^{T} p_i$, $\mu(T) = \sum_{i=T+1}^{255} p_i$, $\bar{\mu} = \sum_{i=0}^{255} i p_i$, and p_i is the probability of pixels with grey level i in the image. Segments are selected as road features if the following two criteria are satisfied: (i) the segment is in the buffer zone, (ii) the length to width ratio γ is larger than a preset threshold, here $\gamma > 2$. Besides, the connected component analysis is utilized here to remove small noises that are misclassified into the road class.

The road sides should have homogeneous local orientation distributions. Thus, the linear regression is an appropriate method to fit a linear relation between the extracted road edge points.

3.4 *Lane marking extraction*

To reduce the computational load, only the 1st PCA component of the original color image is selected. The parameters of Gabor filter are determined as given in section 2.3.2.

After the implementation of Gabor filtering, the Otsu thresholding algorithm is employed to segment the road markings, and the binary morphological opening and closing along the road direction is further utilized to smooth the generated road markings and remove other misclassified objects. The morphological opening and closing operators are presented as $\otimes$ and $\odot$, which are given by (Zhu *et al.* 2005):

$$A \otimes B = (A \ominus B) \oplus B$$
$$A \odot B = (A \oplus B) \ominus B$$

where A is a binary image, B is the binary element structure, $\oplus$ and $\ominus$ are morphological dilation and erosion operators, which are represented as:

$$A \oplus B = \{a + b | a \in A, b \in B\}$$
$$A \ominus B = \{x | b \in B, b + x \in A\}$$

Connected component analysis (CCA) is also applied on the filtered image to remove smaller noise. CCA groups the pixels into connected components based on pixel connectivity and calculate the area for each component, while small regions whose area is under 50 pixels are deleted.

The lane segments may be corrupted by many facts: occlusion, e.g. trees above the road surfaces; worn-out painting of lane lines; dirty markings on the road surfaces. For instance, even if road markings are painted with strict specifications, they are often partially eroded due to weather and traffic conditions. Besides, cars often occlude the road markings, which may lead to the marking detection failure.

The vehicle can be eliminated from the road markings by utilizing the following indicators: (i) elongation – the ratio of the major axis to the minor axis of the polygon, and (ii) lengths of the major and minor axis. The elongation measure of vehicle is smaller than the road lane markings, and the length of the major and minor axis of vehicle are within certain ranges. In our experiment, the major axis length of the vehicle is set to be within 2 to 10 m, while minor axis is set to be between 1.5 and 3 m.

In the filtered image, the pavement markings have much higher brightness values than the road pavement surfaces. The geometric properties and spatial relationships are further utilized to select the lane marking candidates.

1) The geometric properties of the lane markings were calculated: the lengths, widths, and the areas are examined by the road construction manuals (Queensland Department of Main Roads 2001). Only regions satisfying the rules specified in the manual were selected as suitable candidates.

2) The distance and orientation differences between neighboring lane marking candidates are calculated and analyzed. Only the candidates with similar orientation and particular distances would be preserved. The orientation differences must be no more than 5°, and the distances between parallel markings should be approximate 3.5 m.

After the road surfaces and pavement markings detection in sub-images, the generated results are further spliced to obtain the final road model. As the road sides are mostly acquired by linear regression or parallel line approximation, the road sides in adjacent shifting window images may not match very well, as the example shown in figure 6(*a*). Therefore, the relevant endpoints of road sides in adjacent sub-images are replaced by their mean, and the road sides are further determined by their new endpoints (figure 6(*b*)).

Figure 6. Endpoints matching of road edges, (a) Two adjacent sub-images, (b) Endpoints-matched road edges.

3.5 *Result evaluation*

Quantitative evaluation of the experimental results is of great importance for the practical applications of the road features extraction algorithms. However, so far only a few approaches on assessment of road axes extraction can be found in the literatures. Heipke *et al.* (1997) and Wiedemann *et al.* (1998) compared the generated road centerline with a buffer around the reference data, and computed the quality measures principally include completeness, correctness and quality. This evaluation method using vector-based buffer analysis may be used to assess the quality of road marking extraction, but it is not suitable for the evaluation of road surfaces extraction.

To define an error, the 'reference data' or 'more accurate data' is commonly used as the 'true' value in an experimental study (Shi *et al.* 2003). The basis is to compare the automated (derived) results against a manually compiled, high quality reference model. The evaluation of the detected road surfaces is given in figure 7. The detection errors primarily consist of commission and omission errors. The commission error is the area of the falsely detected road surfaces, while omission error is the area of the omitted road surfaces.

Figure 7. Evaluation of the extracted road surfaces: solid area is the referenced road surface; dashed region indicates the detected road surface.

Following the concept of error matrix, the evaluation matrices for the accuracy assessment of road surfaces detection can be defined at the pixel level as follows:

1) Detection rate

$$d = \frac{TP}{TP + FN}$$

2) False alarm rate

$$f = \frac{FP}{TP + FP}$$

In the above equation, TP (true positive) is the number of road surface pixels correctly identified, FN (false negative) is the number of road surface pixels identified as other objects, FP (false positive) is the number of non-road pixels identified as road surfaces.

The detection rate indicates how well pixels of road surfaces that have been correctly classified; the false alarm rate denotes the probabilities of pixels been uncorrected classified into actual category of road surfaces.

The road marking accuracy evaluation is carried out by comparing the extracted pavement markings with manually plotted road markings used as reference data as presented in (Wiedemann *et al.* 1998), and both data sets are given in vector representation. The buffer width is predefined to be the average width of the road markings, and we set it to be 15 cm in our experiment. Then the accuracy measures are carried out by detection rate, false alarm rate and RMS (root mean square) difference, given as:

1) Detection rate

$$d = \frac{\text{length of matched reference}}{\text{length of reference}}$$

2) False alarm rate

$$f = \frac{\text{length of unmatched extraction}}{\text{length of extraction}}$$

3) RMS difference

$$RMS = \sqrt{\frac{\sum_{i=1}^{n}(d(e_i;r)^2)}{l}}$$

where n is the number of pieces of matched extraction, $d(e_i;r)$ indicate the shortest distance between the *i*th piece of the matched extraction and the reference network.

4. Experiment and discussion

The experiments were carried out on the aerial images from the Bruce Highway Gympie part, Queensland. The objective of the experiment was to determine the performance of the proposed road feature extraction approach quantitatively over the study area.

4.1 *Data set*

Thanks to Queensland Department of Transport and Main Roads, a set of high-resolution aerial panchromatic images taken on 28 November of 2008 using the UltraCam-D digital camera is made available for this study. The camera was mounted on a Cessna Conquest 441 aircraft flying at an altitude of 1100 m. The image scale is 1:11081, and pixel size is 9 μm (about 10 cm GSD). The test sites are parts of the Bruce highway, located in Gympie, Queensland. The width of traffic lane is 3.5 m on average. The original scale of the photogrammetric images is 7500×11500 pixels, with three color bands (Red, Green, and Blue). To test the efficiency of large image process, image mosaic is utilized to expand the image scale.

4.2 *Experiment and analysis*

The properly determination the size of the shifting window is of great importance for the road features detection in large images. A smaller window may not cover the

whole road surfaces; on the other hand, a larger window may not guarantee the local homogeneous direction, which is required by linear regression and Gabor filter, and it would also increase the computing complexity. The average width of the road surfaces is about 15 m (including the road shoulders), which is 150 pixels in line with the 10 cm GSD of the testing image dataset. Thus the window size is defined to be 512×512 pixels based on our experiments.

The parameters of Gabor filter are determined as given in section 2.3.2. The orientation of the markings is obtained by the detected centerline from the low resolution image. The average width of road markings is 4 pixels, thus the frequency f is set to be 0.25, while the axis scaling parameters σ_x and σ_y of the Gaussian function is set to be 2.32.

The road feature extraction result for the first experiment is shown in figure 8. The test aerial mosaic image has a size of 16384×8192 pixels, where the road in the image is relatively straight, and with less obstructions from shadows and vehicles. It can be seen that the road surfaces are almost perfectly extracted and the lane markings are also well detected thanks to our strategy, and only very few false alarms exists caused by vehicles.

To assess the extraction results quantitatively, the manual plotted road surface and lane markings are employed as the true data, and the reference lane markings are illustrated in figure 9. The buffer width is set to be 2 pixels – half the average width of pavement markings. As it is reviewed in section 1 that within all the road feature extraction methods as we know, only two strategies, which are proposed by Kim *et al.* (2006) as well as Tournaire and Paparoditis (2009), are specially designed to extraction lane information from remotely sensed images. Kim's lane marking extraction method is chosen here for the comparison because Tournaire's method focuses on only particular dashed markings. We apply Kim's method in each shifting window image to extract lane markings and further filter out misclassified objects using the extracted road surface.

Figure 8. The road surface and lane marking extraction result in test aerial mosaic image I: (a) original aerial image; (b) extracted road surface and lane markings; (c) region A in a larger scale; (d) region B in a larger scale.

Figure 9. The manual plotted lane marking used as reference in test site I.

The accuracies evaluation for our strategy and Kim's method upon image I is given in table 1. The overall detection rates of lane markings for both methods are over 99%, and the false alarm rates are below 3%, which indicates that both strategies are capable of dealing with images with less effect of noise on the road surface.

Table 1. Result evaluation for road feature extraction in test image I.

The results for test image II and III are given in figure 10 and figure 12, respectively. The manual plotted lane marking are illustrated in figure 11 and figure 13. The difference between test image II and III is that the road in test image II has big sinuous curves while the road in test image III has several traffic islands on its surface. And there are several vehicles and shadow on the road surfaces for both test scenes.

Figure 10. The road surface and lane marking extraction result in test aerial mosaic image II: (a) original aerial image; (b) extracted road surface and lane markings; (c) region A in a larger scale; (d) region B in a larger scale.

Figure 11. The manual plotted lane marking used as reference in test site II.

Figure 12. The road surface and lane marking extraction result in test aerial mosaic image III: (a) original aerial image; (b) extracted road surface and lane markings; (c) region A in a larger scale; (d) region B in a larger scale.

Figure 13. The manual plotted lane marking used as reference in test site III.

Table 2 and table 3 present the quantitative assessment of the results for both approaches. The results for both test images have detection rate of about 96% and false alarm rate below 6.5%. Kim's method has a larger completeness rate (approximate 1%), but at the meantime, its misclassification is significant higher than our algorithm (about 4%). The RMS differences of both methods are about 3 cm.

Table 2. Result evaluation for road feature extraction in test image II.

Table 3. Result evaluation for road feature extraction in test image III.

In test image IV (as shown in figure 14), the road surfaces are greatly obstructed by the shadows casted by trees, which increase the difficulty of road marking extraction. The extraction of lane segments may be corrupted by many facts, such as occlusions, worn-out painting of lane lines, and dirty marks on the road surface. An example is given in figure 14 (image A), the bottom lane marking is partly worn-out and has less contrast with the road surface, is not entirely detected.

Figure 14. The road surface and lane marking extraction result in test aerial mosaic image IV: (a) original aerial image; (b) extracted road surface and lane markings; (c) region A in a larger scale; (d) region B in a larger scale.

Figure 15. The manual plotted lane marking used as reference in test site IV.

In heavily shadowed regions, Gabor filter will leave out lane segments that are completely obstructed by shadows due to their poor contrast with road surfaces, and enhance the lane markings in non-shadow and partly-shadowed areas. Top-hat based strategy can remove the affection from shadow to some extent, it also enhance many other ground objects in the meantime. The statistics in table 4 has confirmed the above conclusion.

Table 4. Result evaluation for road feature extraction in test image IV.

Both road lane feature extraction approaches based on top-hat and Gabor filtering achieves satisfactory results in non-shadowed regions as it is shown in table 1, 2 and 3. Similar conclusions can be drawn from the visual analysis of figure 16, which is an example of road marking extraction in a non-shadowed area. All the three long lane markings are correctly extracted using both methods, while the Gabor filter is superior to top-hat on enhancing the lane features as well as minimizing the affections from other ground objects, e.g. trees.

Figure 16. An example of road lane marking extraction in non-shadowed region: (*a*) the original aerial image; (*b*) and (*c*) are the top-hat filtered result and the extracted lane marking; (*d*) and (*e*) are the Gabor filtered image and the generated pavement markings, respectively.

Gabor filter has a relatively low detection rate but with low false alarm rate compared with top-hat based method. An example is as given in figure 17 that several shadow features on the road surface are misclassified into lane marking when using Kim's method. In contract, although some lane segments in the shadow regions are

not correctly detected using our proposed approach, there is no misclassification within the generated lane segments.

Figure 17. An example of road lane marking extraction on shadowed road surface: (*a*) the original aerial image; (*b*) and (*c*) are the top-hat filtered result and the extracted lane marking; (*d*) and (*e*) are the Gabor filtered image and the generated pavement markings, respectively.

5.   Concluding remarks

In this paper, we have presented a new approach to detect road surfaces and pavement markings in high resolution aerial images over rural areas, based on hierarchical image analysis and Gabor filtering. The experimental results on four aerial mosaic images have indicated that over 95% of road surfaces and near 95% of the pavement markings had been correctly extracted, preliminarily proving that the presented strategy is promising. The omission of road surfaces and road markings resulted from occlusions, the poor condition of pavement markings, and partial shadows over the markings.

Overall, the findings of this research can be summarized as follows:

1) Road surfaces and pavement markings can be simultaneously extracted from high resolution remote sensed images. This will significantly facilitate the generation of highly accurate digital road maps.
2) Uses of Gabor filter can enhance the pavement marking as well as constraint the affections from other ground objects, which greatly reduces the difficulty of pavement markings extraction.
3) Shifting window process is an efficient method to deal with large images, by splitting the original images into sub-images along the road centerline detected from a low resolution image.

As part of the future research, the proposed framework will be further improved so that more seriously obstructed road scenes can be dealt. For instance, an automatic vehicle detection approach may be introduced to reduce the effects of vehicles. GPS Real-time kinematic positioning solutions from a probe vehicle could be a good choice to recover the lane markings in areas where there are large obstructions, for example, large numbers of skyscrapers or trees would greatly deteriorate the extraction result in urban or forest areas.

References

AMINI, J., SARADJIAN, M. R., BLAIS, J. A. R., LUCAS, C., and AZIZI, A., 2002, Automatic road-side extraction from large scale imagemaps. *International Journal of Applied Earth Observation and Geoinformation*, **4**, pp. 95-107.

BAUMGARTNER, A., STEGER, C., MAYER, H., ECKSTEIN, W., and EBNER, H., 1999, Automatic road extraction based on multi-scale, grouping, and context. *Photogrammetric Engineering & Remote Sensing*, **65**, pp. 777-785.

BOVIK, A. C., CLARK, M., and GEISLER, W. S., 1990, Multichannel texture analysis using localized spatial filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**, pp. 55-73

CAI, J., and LIU, Z., 2000, Off-line unconstrained handwritten word recognition. *International Journal of Pattern Recognition and Artificial Intelligence*, **14**, pp. 259-280.

DAUGMAN, J. G., 1985, Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of Optical Society of America, A: Optics and Image Science*, **2**, pp. 1160-1169.

DAUGMAN, J. G., 1988, Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **36**, pp. 1169-1179.

GRUEN, A., and LI, H., 1995, Road extraction from aerial and satellite images by dynamic programming. *ISPRS Journal of Photogrammetry and Remote Sensing*, **50**, pp. 11-20.

HEIPKE, C., MAYER, H., WIEDEMANN, C., and JAMET, O., 1997, Evaluation of automatic road extraction. *International Archives of Photogrammetry and Remote Sensing*, **32**, pp. 47-56.

HINZ, S., and BAUMGARTNER, A., 2000, Road extraction in urban areas supported by context objects. *International Archives of Photogrammetry and Remote Sensing*, **33**, pp. 405-412.

HINZ, S., and BAUMGARTNER, A., 2003, Automatic extraction of urban road networks from multi-view aerial imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, **58**, pp. 83-98.

JIN, X., and DAVIS, C. H., 2005, An integrated system for automatic road mapping from high-resolution multi-spectral satellite imagery by information fusion. *Information Fusion*, **6**, pp. 257-273

KIM, J. G., HAN, D. Y., YU, K. Y., KIM, Y. I., and RHEE, S. M., 2006, Efficient extraction of road information for car navigation applications using road pavement markings obtained from aerial images. *Canadian Journal of Civil Engineering*, **33**, pp. 1320-1331.

KRISHNAPURAM, R., and CHEN, L.-F., 1993, Implementation of parallel thinning algorithms using recurrent neural networks. *IEEE Transactions on Neural Networks*, **4**, pp. 142-147.

LINDENBAUM, M., FISCHER, M., and BRUCKSTEIN, A., 1994, On Gabor's contribution to image enhancement. *Pattern Recognition*, **27**, pp. 1-8.

LIU, Z., CAI, J., and BUSE, R., 2003, *Handwriting recognition: soft computing and probabilistic approaches*, pp. 145-172 (Berlin, Germany: Springer Verlag).

MANJUNATH, B. S., and MA, W. Y., 1998, Texture features for browing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **18**, pp. 837-842.

OTSU, N., 1979, A threshold selection method from gray level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, **9**, pp. 62-66.

QUEENSLAND DEPARTMENT OF MAIN ROADS, 2001, *Guide to pavement markings*, pp. 9-58 (Brisbane, Queensland: Department of Main Road).

SHI, W., CHEUNG, C. K., and ZHU, C., 2003, Modelling error propagation in vector-based buffer analysis. *International Journal of Geographical Information Science*, **17**, pp. 251-271.

SOARES, J. V. B., LEANDRO, J. J. G., CESAR, R. M., JELINEK, H. F., and CREE, M. J., 2006, Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification. *IEEE Transactions on Medical Imaging*, **25**, pp. 1214-1222

STEGER, C., MAYER, H., and RADIG, B., 1997, The role of grouping for road extraction. In *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, A. Gruen, E. P. Baltsavias, and O. Henricsson (Ed.), pp. 1931-1952 (Basel, Switzerland: Birkhäuser).

TOURNAIRE, O., and PAPARODITIS, N., 2009, A geometric stochastic approach based on marked point processes for road mark detection from high resolution aerial images. *ISPRS Journal of Photogrammetry and Remote Sensing*, **64**, pp. 621-631.

TRINDER, J. C., and WANG, Y., 1998, Automatic road extraction from aerial images. *Digital Signal Processing*, **8**, pp. 125-224.

VILLLASENOR, J. D., BELZER, B., and LIAO, J., 1995, Wavelet filter evaluation for image compression. *IEEE Transactions on Image Processing*, **4**, pp. 1053-1060.

VOSSELMAN, G., and KNECHT, J. D., 1995, Road tracing by profile matching and Kalman filtering. In *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, A. Gruen, O. Kuebler, and P. Agouris (Ed.), pp. 265-274 (Basel, Switzerland: Birkhaeuser Verlag).

WEBER, D. M., and CASASENT, D. P., 2001, Quadratic Gabor filters for object detection. *IEEE Transactions on Image Processing*, **10**, pp. 218-230.

WIEDEMANN, C., HEIPKE, C., MAYER, H., and JAMET, O., 1998, Empirical evaluation of automatically extracted road axes. In *Empirical Evaluation Methods in Computer Vision*, J. B. Kevin, and P. J. Phillips (Ed.), pp. 172-187 (Silver Spring, MD: IEEE Computer Society Press).

ZHANG, C., 2004, Towards an operational system for automated updating of road databases by integration of imagery and geodata. *ISPRS Journal of Photogrammetry and Remote Sensing*, **58**, pp. 166-186.

ZHU, C., SHI, W., PERARESI, M., LIU, L., CHEN, X., and KING, B., 2005, The recognition of road network from high-resolution satellite remotely sensed data using image morphological characteristics. *International Journal of Remote Sensing*, **26**, pp. 5493-5508.

FIGURES

1. Separation lines, broken lane

3 m   9 m

100 mm

2. Separation lines

100 mm

3. Unbroken lane lines, edge lines

150 mm
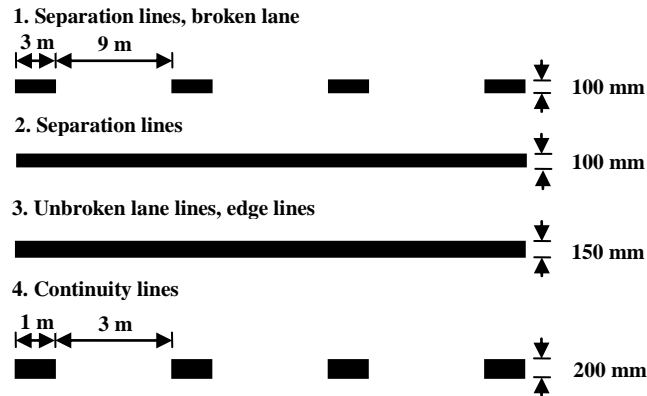
4. Continuity lines

1 m   3 m

200 mm

Figure 1. Geometric specifications of the pavement markings in a rural arterial road environment (Queensland Department of Main Roads 2001).
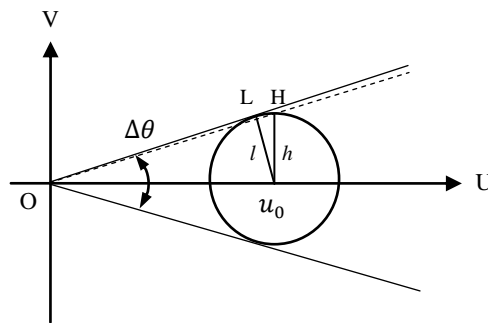
Figure 2. The uncertainty relation in frequency domains. $\sigma_x = \sigma_y$, $\phi = 90°$, $\Delta\theta$ is the orientation bandwidth, and both h and l are radius (only the positive frequencies are shown).
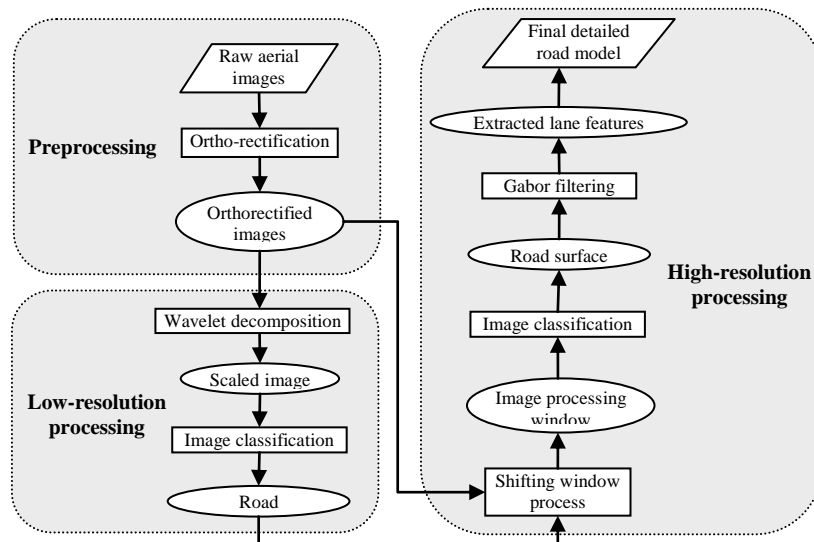
Figure 3. Detailed flowchart of the proposed system for road marking extraction.
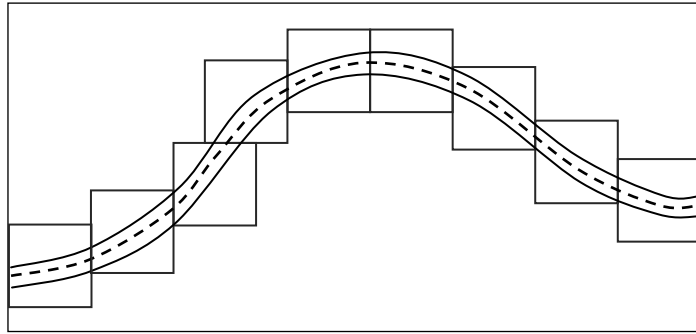
Figure 4. The illustration of the shifting window process.
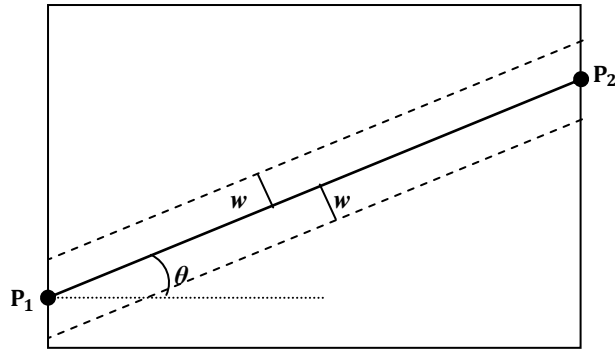
Figure 5. The representation of orientation φ and the buffer zone, w is the average half width of road surface, P_1 and P_2 are the two endpoints of road centreline.
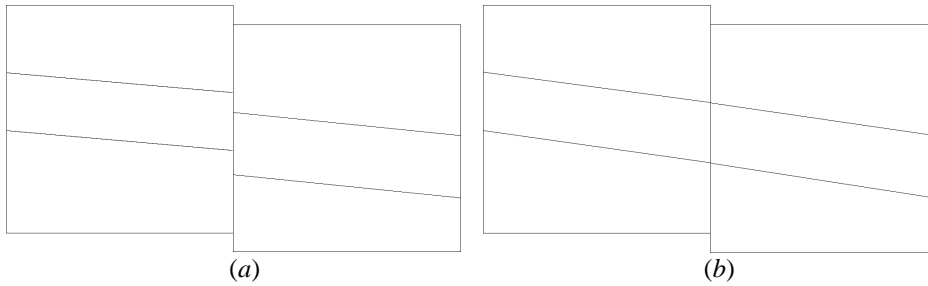
(*a*)   (*b*)

Figure 6. Endpoints matching of road edges, (a) Two adjacent sub-images, (b) Endpoints-matched road edges.

Correct detection (TP)

Commission error (FP)

Omission error (FN)
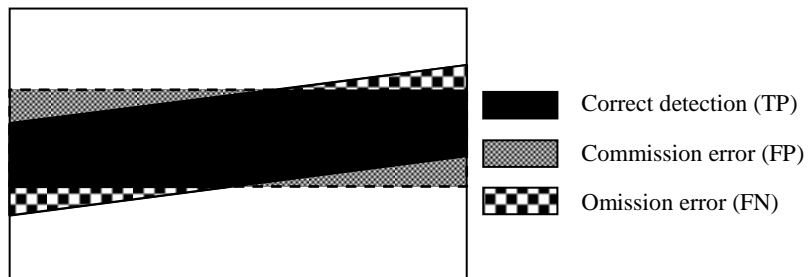
Figure 7. Evaluation of the extracted road surface: solid area is the referenced road surface; dashed region indicates the detected road surface.
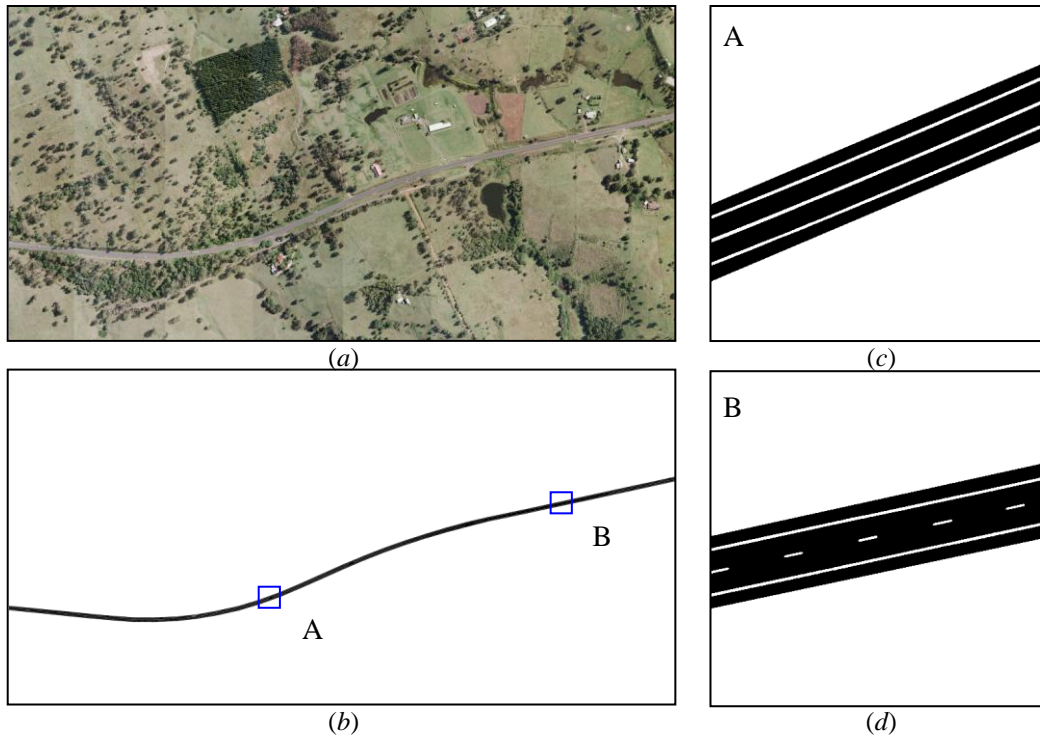
Figure 8. The road surface and lane marking extraction result in test aerial mosaic image I: (a) original aerial image; (b) extracted road surface and lane markings; (c) region A in a larger scale; (d) region B in a larger scale.
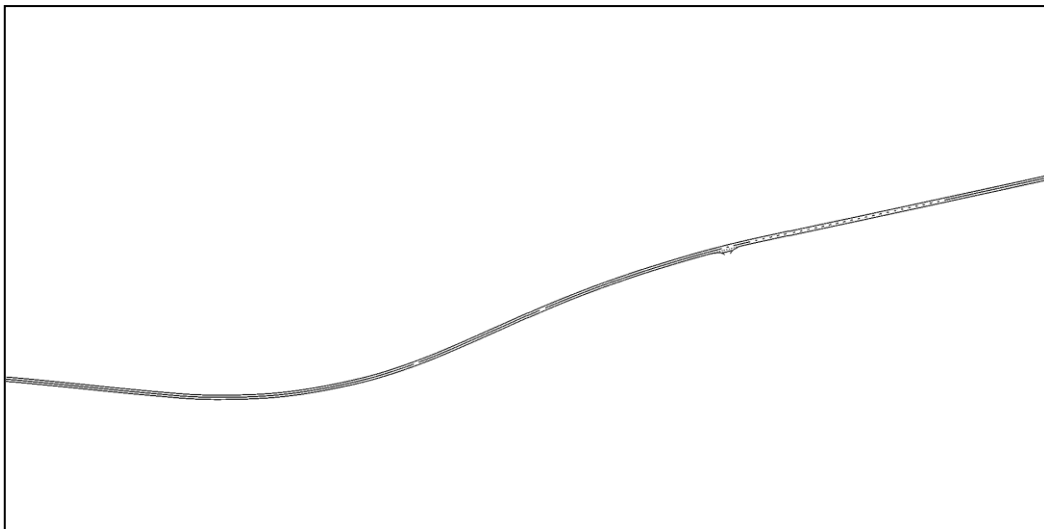
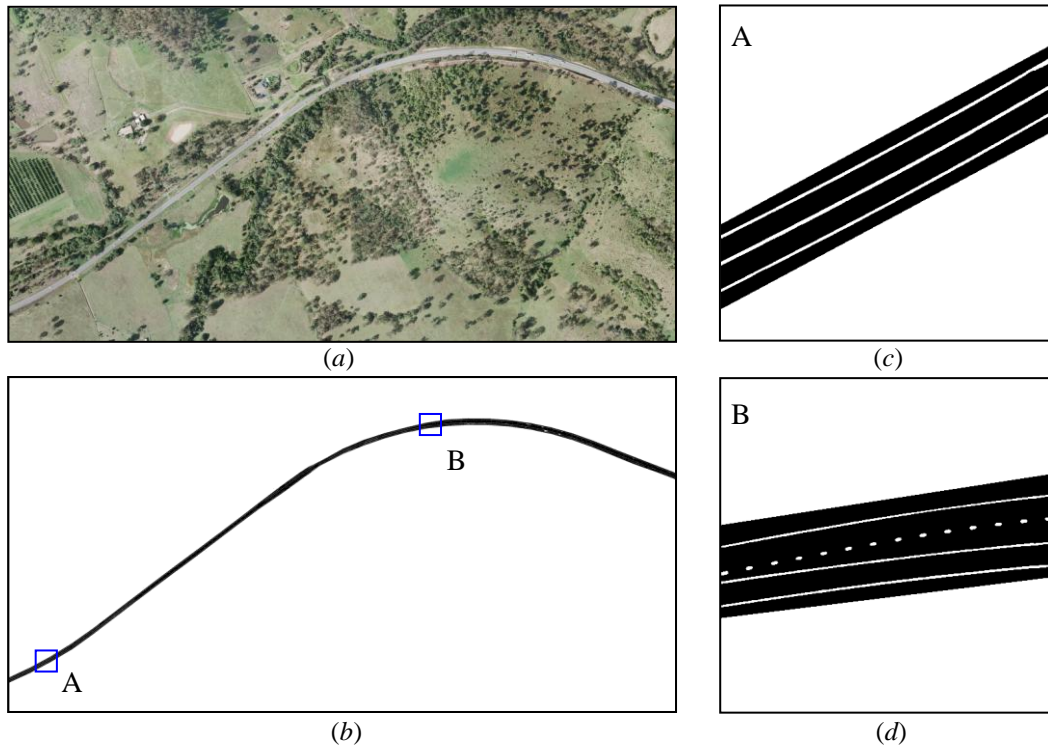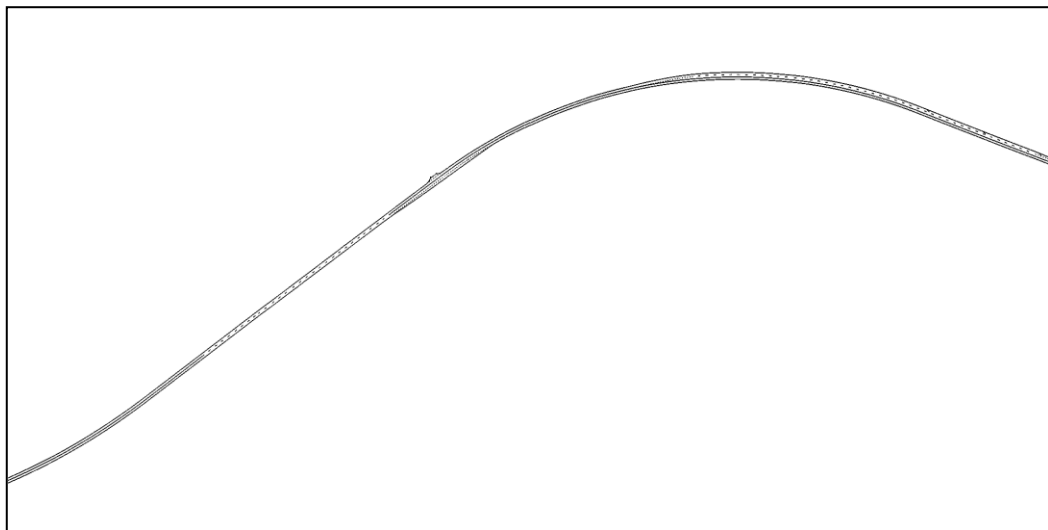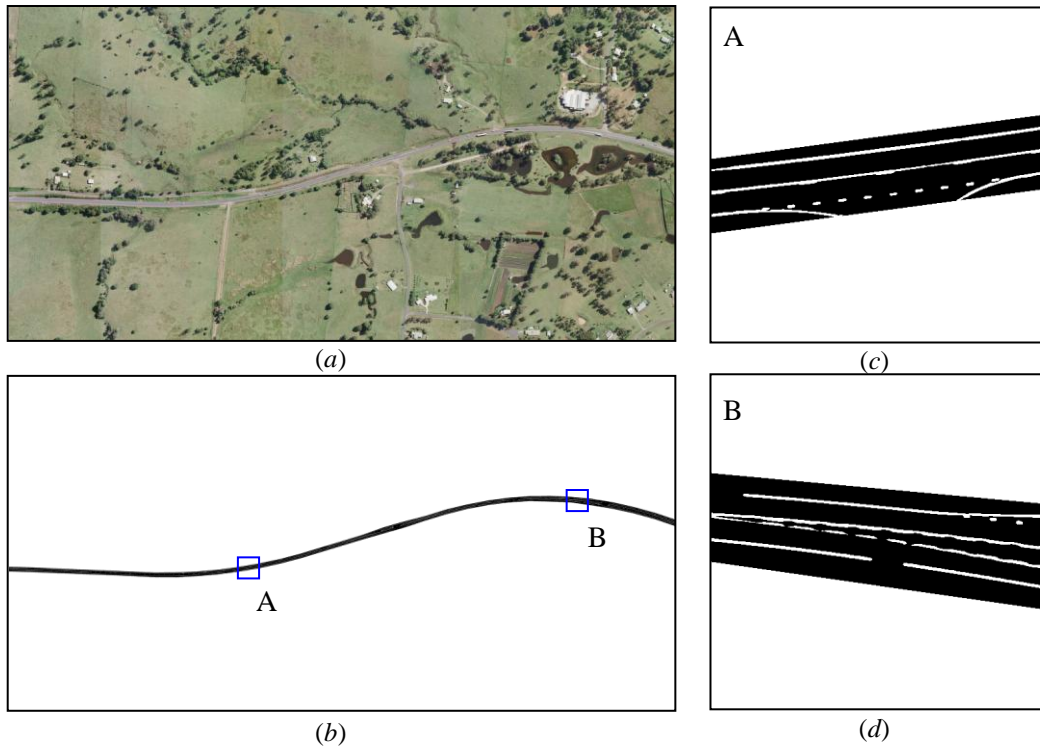Figure 9. The manual plotted lane marking used as reference in test image I.

Figure 10. The road surface and lane marking extraction result in test aerial mosaic image II: (a) original aerial image; (b) extracted road surface and lane markings; (c) region A in a larger scale; (d) region B in a larger scale.

Figure 11. The manual plotted lane marking used as reference in test image II.

Figure 12. The road surface and lane marking extraction result in test aerial mosaic image III: (a) original aerial image; (b) extracted road surface and lane markings; (c) region A in a larger scale; (d) region B in a larger scale.
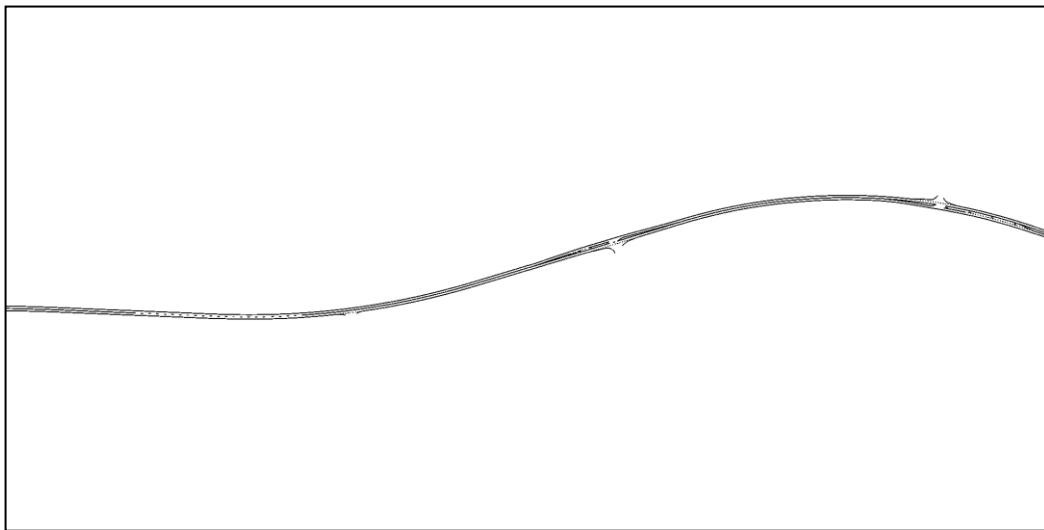

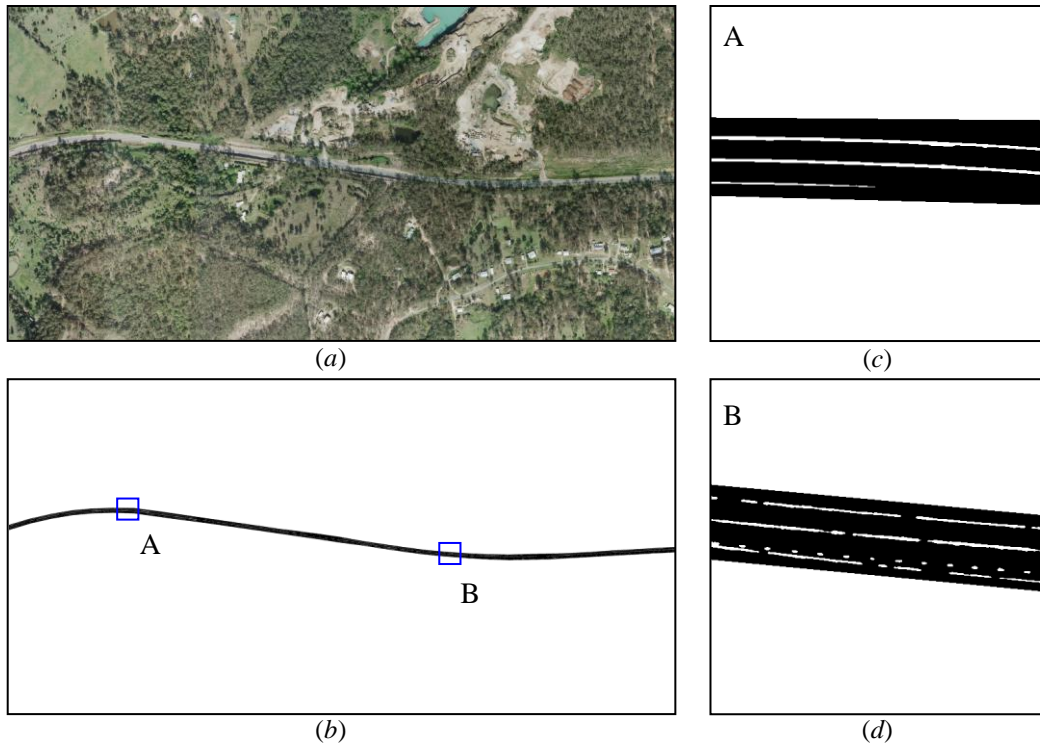Figure 13. The manual plotted lane marking used as reference in test site III.

Figure 14. The road surface and lane marking extraction result in test aerial mosaic image IV: (a) original aerial image; (b) extracted road surface and lane markings; (c) region A in a larger scale; (d) region B in a larger scale.
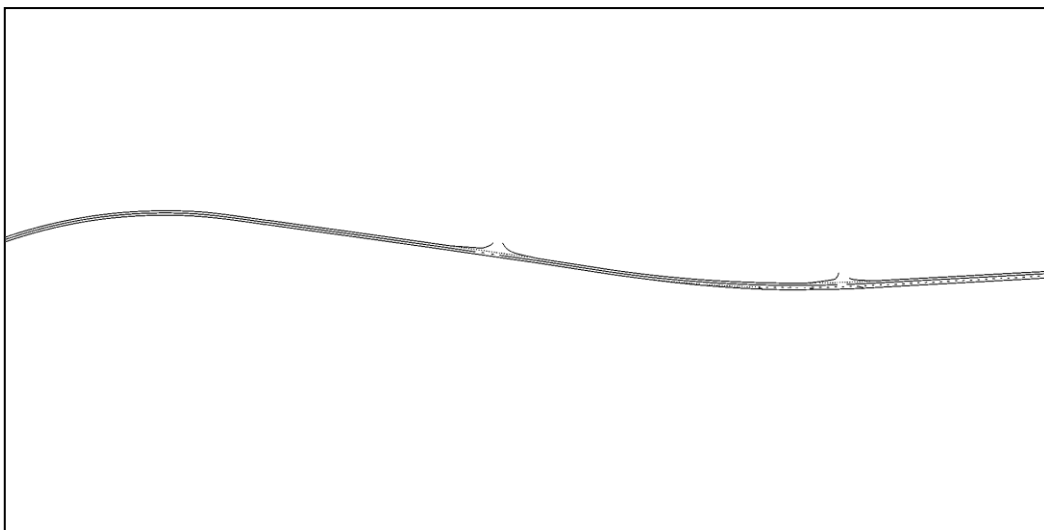

Figure 15. The manual plotted lane marking used as reference in test site IV.
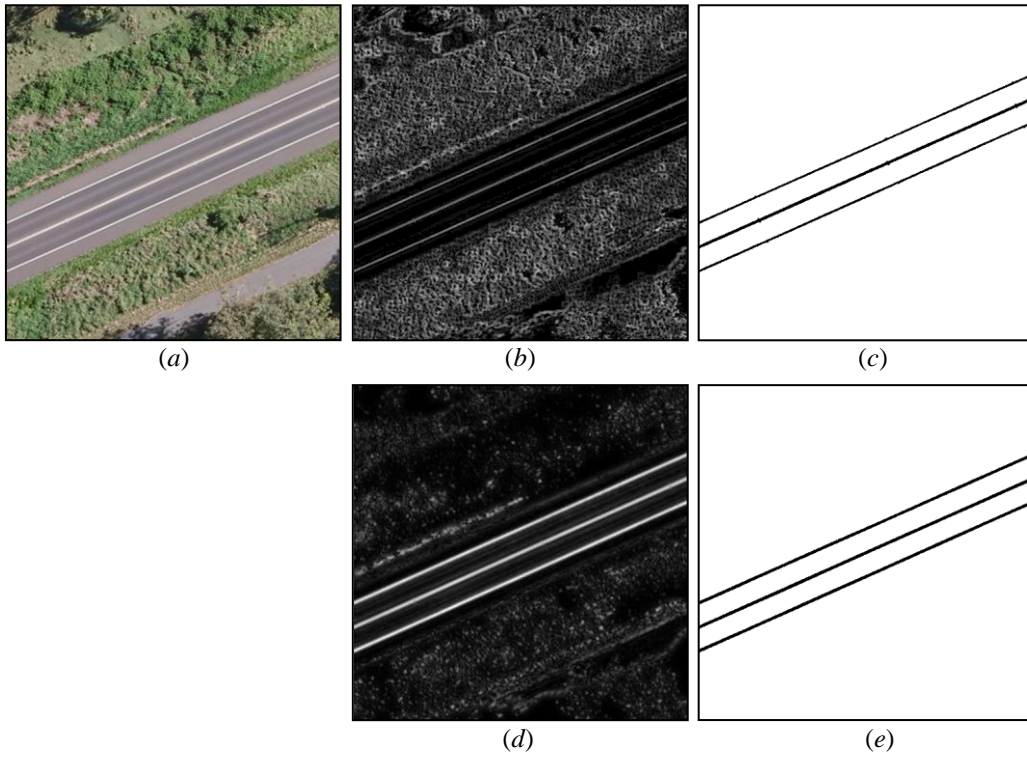
Figure 16. An example of road lane marking extraction in non-shadowed region: (a) the original aerial image; (b) and (c) are the top-hat filtered result and the extracted lane marking; (d) and (e) are the Gabor filtered image and the generated pavement markings, respectively.
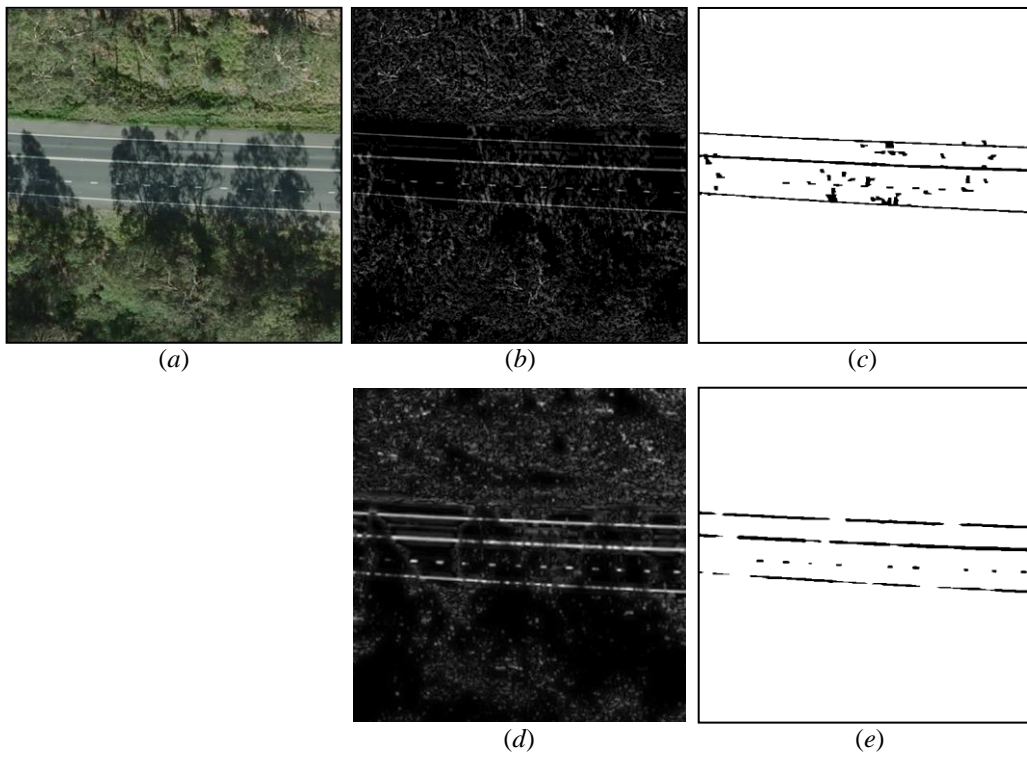

Figure 17. An example of road lane marking extraction on shadowed road surface: (a) the original aerial image; (b) and (c) are the top-hat filtered result and the extracted

lane marking; (d) and (e) are the Gabor filtered image and the generated pavement markings, respectively.

TABLES

Table 1. Result evaluation for road feature extraction in test image I.

	Road surface extraction		Lane marking detection		
	Detection rate (%)	False alarm rate (%)	Detection rate (%)	False alarm rate (%)	RMS difference (cm)
Our approach	99.8	2.9	99.3	0.6	2.4
Kim's method	\	\	99.5	2.5	2.2

Table 2. Result evaluation for road feature extraction in test image II.

	Road surface extraction		Lane marking detection		
	Detection rate (%)	False alarm rate (%)	Detection rate (%)	False alarm rate (%)	RMS difference (cm)
Our approach	98.2	3.2	94.5	2.7	2.7
Kim's method	\	\	95.3	5.8	2.8

Table 3. Result evaluation for road feature extraction in test image III.

	Road surface extraction		Lane marking detection		
	Detection rate (%)	False alarm rate (%)	Detection rate (%)	False alarm rate (%)	RMS difference (cm)
Our approach	99.2	3.2	94.7	2.0	3.2
Kim's method	\	\	96.2	6.5	3.6

Table 4. Result evaluation for road feature extraction in test image IV.

	Road surface extraction		Lane marking detection		
	Detection rate (%)	False alarm rate (%)	Detection rate (%)	False alarm rate (%)	RMS difference (cm)
Our approach	95.8	2.2	82.5	1.0	3.8
Kim's method	\	\	86.9	13.5	3.5