Pothole detection on asphalt pavements from 2D-colour pothole images using fuzzy *c*-means clustering and morphological reconstruction

Yashon O. Ouma[a],[*], M. Hahn[b]

[a] Department of Civil and Structural Engineering, Moi University, 30100 Eldoret, Kenya
[b] Department of Geomatics, Stuttgart University of Applied Sciences, Schellingstr., 70174 Stuttgart, Germany

ARTICLE INFO

ABSTRACT

In general, potholes on asphalt pavements can be detected and represented in 2D and 3D. However, pothole detections through 3D imaging and image reconstructions have proven to be expensive in terms of acquisition equipment and the computational and processing requirements and time. For potholes at incipient formations, their detection, representation and quantification in terms of the surface-area are important for timely maintenance and repairs. By casting pavement image segmentation for pothole detection as a problem of clustering multivariate features within mixed pixels (*mixels*), this study presents a low-cost 2D vision image-based approach for the detection of potholes on asphalt road pavements in urban areas. The approach in this study is based on the *a priori* integration of multiscale texture-based image filtering for *textons* representation using wavelet transform, into the superpixel clustering of the pavement defects and non-defects using fuzzy *c*-means (FCM) algorithm. For the extraction of the defects extrema (minima and maxima) in the hybrid wavelet-FCM clustering results, fine segmentation based on morphological reconstruction is adopted to further smoothen and recognize the contour of the detected potholes. The methodology is implemented in a MATLAB prototype, tested and validated using 75 experimental image datasets. With a mean CPU run-time of 95 seconds, the average detection accuracies by comparing the study results and the manually segmented ground-truth data were determined using the Dice coefficient of similarity, Jaccard Index and sensitivity metric as 87.5%, 77.7% and 97.6% respectively. The average magnitudes of the mean and standard deviation of the percentage errors in pothole size extractions were detected as 8.5% and 4.9% respectively. The results of the study show that with well-planned road condition surveys, the proposed algorithm is suitable for the detection and extraction of incipient potholes from 2D vision images acquired using low-cost consumer-grade imaging sensors.

1. Introduction

In the recent past, a great deal of research has been dedicated to the development of innovative methods and algorithms to improve on the widely used manual-based "walk-and-record" road condition surveys. Despite the fact that automated approaches have been proposed, the manual-based pavement condition surveys are still predominantly being used [1,2], for the identification, categorization and quantification of the types and degrees of pavement defects [3,4]. In retrospect, the American Transport Research Bureau (ATRB) and the UK's Transport Research Laboratory (TRL) among other road maintenance agencies have recommended that the manual interventions be eliminated in order to reduce road survey costs, and also highlighted the drawbacks of the current automated road survey systems [2,5].

Road pavement defects can broadly be classified as surface defects and elevation-oriented defects. Potholes, patching and bleeding are surface-elevation related defects which can be categorized according to area-depth, and are mostly caused by loss of road layer aggregates as summarized in Table 1. For these types of distresses, the surface-area and depth or elevation information are an important geometric factor in determining the extent and severity of the defect.

For maintenance and repairs (M & R), potholes should be detected at the incipient stages before interfering with the road subbase or base as illustrated in Fig. 1. In such cases, the surface-areas of the detected potholes [7], become more significant than their depth d_i (Fig. 1). If early and regular road maintenance is neglected, the incipient or minor cracks on the road surface eventually cause substantial structural damage to the road, which results in part into potholes.

Notably, the detection of potholes from 2D or 3D imagery should be understood from the point of view of the pavement engineer and the contractor, when it comes to road maintenance and repairs. That is, the incipient pothole treatment as applied in partial, full-depth or injection

Table 1

Categorization of surface elevation-based defects on flexible asphalt pavements.

Road distress category	Distress characterization and distress type image		
Area-depth or surface-elevation based road pavement distresses	(i) Pothole	(ii) Patching	(iii) Excess asphalt (Bleeding)

patching is often carried out within a surface-area extent. Thus the pothole detection system is supposed to ascertain the existence of the potholes and estimate of the areal extents. This means that both the depth and diameter of the pothole defects are particularly crucial when the degree of severity of the pothole defect is of interest in M & R for purposes of ranking and prioritization.

Therefore in this study, pothole depth is not considered since the aim is to detect potholes at the initial stages before the depth exceeds the design asphalt layer depths d_1 or d_2 (Fig. 1). This implies that of significance is the detected pothole surface-area A_{d_i} for the corresponding design depth d_i. In Fig. 1, A_{Rd_i} refers to the estimated surface-area during the actual pothole treatment or repair. In scenarios where the potholes are considered to be less severe, A_{Rd_i} can be considered to be equivalent to A_{d_i}. It is also worth noting that during M & R implementations, the volume of interest is normally not the exact depth/volume of the detected pothole area, but the volume of cut as approximated by the area *abcd* in Fig. 1.

The automated pavement condition survey methods such as: 2D aerial imageries and terrestrial photogrammetric data; 2D stereo-imagery and laser point clouds, can be categorized as: (i) vibration-based accelerometers sensors [8]; (ii) 3D-stereovision techniques [9]; (iii) three-dimensional based laser scanning [10], and (iv) two-dimensional vision image based techniques [11]. Although laser-scanning systems provide highly accurate geometrical data of the pavement profile and distress detection, the cost of the sensors is still relatively high, which limits their application for routine pavement assessments [11–13]. Furthermore, the related computational approaches are expensive in terms of data processing, and the approach cannot be applied over a wide area for fast pothole detection as reported in Jo and Ryu [14] and Koch and Brilakis [11]. In contrast, the use of 2D image data for road condition surveys is widely used and has also yielded better results [4,6,84]. In general, the detection of elevation-based defects such as potholes is hindered by the costs involved in the data acquisitions and processing [4].

For this study, 2D vision images are considered appropriate for rapid and road condition survey, and for the early detection of potholes on asphalt pavements. However, 2D image processing for pavement surface assessments is a challenging task due to the inherent complexities in the image characteristics and processing for pavement defect

detections [6,11,15–18].

In 2D images, pavement image segmentation is the most widely used approach for pothole detection (e.g. [7,11]). The primary segmentation methods used include thresholding, clustering and edge detection and region extraction [19,85]. These conventional methods tend to restrict each point of the data set to exclusively two clusters of data and no-data. Furthermore, hard segmentation methods cannot maintain much information which makes them suspect to noise and imaging artifacts. As a consequence, the segmentation results are often crisp, meaning that each pixel of the image belongs to exactly just one class. Most pavement surfaces images are ambiguous and have indistinguishable histograms. In such images, it is not easy for classical thresholding techniques, such as Otsu [20] and Kapur et al. [21], to find suitable criterion of similarity or closeness for thresholding, since they only work well when two consecutive gray-levels of the image are distinct. In 2D images therefore, issues such as limited spatial resolution, poor contrast, overlapping intensities, noise and intensity inhomogeneity makes hard segmentation a difficult task and often results in low accuracies in defect detection.

Statistically, image segmentation is an ambiguous problem because of the following reasons [86]. First, the statistical characteristics of local features comprising of colour, texture, edge, and contour do not usually show the same degree of homogeneity or saliency at the same spatial or quantization scale. As such segmentation results are not expected to be unique, and instead should prefer a hierarchy of segmentations at multiple scales [22]. Secondly, even after accounting for variations due to the scale, the different spectral or textural regions may still contain some intrinsic complexities, making it a difficult statistical problem to determine the correct number of segments and their dimensions. As such, a good segmentation algorithm for the detection of features should be able to group similar image pixels into regions whose statistical characteristics comprising of colour and or texture are homogeneous or stationary, and whose boundaries are simple and spatially accurate [23].

There is therefore the need to improve on the algorithms for detection of pothole geometrics including surface-area, compactness, shape, orientation and location. This study seeks to extend on the pothole detection process by using multiclass feature clustering approach, based on fuzzy *c*-means (FCM). FCM clustering is a soft
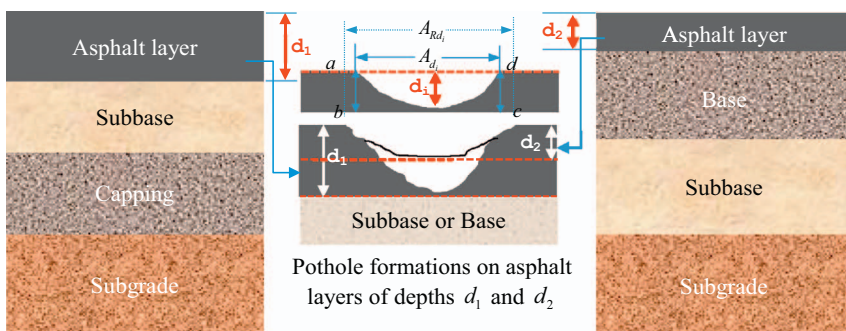

Fig. 1. One-dimensional depiction of cross-sections of flexible pavement profiles [6] and possible incipient pothole structural failures of depth d_1 and d_2 with the corresponding surface-areas A_{d_i}.

segmentation method that has been used extensively to improve on the compactness of the regions with its cluster validity and simple implementation [24]. It is an unsupervised clustering technique that is based on the idea of uncertainty of belonging, which is described by a membership grading, and has the ability to employ more information from the image in comparison with the crisp or hard segmentation methods.

To boost the performance of the FCM, the original image is filtered and smoothened *a priori*, by using a multiscale discrete wavelet transform (DWT). The filtering process groups pixels into regions of coherent texture for *texton* representation so as to obtain superpixels. The DWT-FCM approach is able to perform soft segmentation while compensating for intensity inhomogeneities on the superpixels. The DWT is preferred to the undecimated-wavelet decomposition since the latter requires higher computational load and yields high redundancy in the computed wavelet coefficients, making the subsequent computational processing such as edge detection and classification expensive [25]. Notably, despite the fact that the incorporation of filters such as the Gaussian and bilateral filters may enhance the fuzzy clustering performance, these filters apply only on noise of limited frequency bandwidth, thus leaving noise of certain frequency untreated [26]. The low-frequency components of wavelets are capable of detecting the candidate defect pixels, hence characterizing the coarse structure of the data and also identifying the long-term trends in the original data. However, due to the anisotropic properties of wavelets, the wavelets transform is not able to effectively and automatically extract pavement defects with non-directional patterns and high textural variations such as potholes. Furthermore, while wavelets are able to effectively represent discontinuities in at least one-dimensional signals, the tensor-product construction may not be flexible in reproducing the continuity and shape representation in 2D images [27]. As such, a spectral aggregation approach is necessary in order to cluster and differentiate defects from non-defects.

FCM depends on the Euclidean distance between pixels based on the assumption that each feature is of equal importance. However, in most real-world applications, features are not considered equally important. Thus, this assumption may seriously affect clustering performance. To address this drawback in FCM, this study introduces the mathematical morphological reconstruction to reduce the noise of both high and low-frequencies especially in the pothole edges by referring to the original image to successfully segment the potholes. That is, the result should be the production of segmentations that allow regions or classes to overlap, as opposed to standard hard segmentations that enforce a binary decision on whether a pixel is inside or outside the object. The use of the morphological reconstruction is based on the fact that it is suitable for: the extraction of the minima and maxima extremum; hierarchical image construction, and has the advantage of speed and noise resistance [28]. As such it can be able to compensate for the non-pothole defect artifacts such as stones, pebbles and leaves which remain after clustering.

The integrated hybrid wavelet-FCM clustering and morphological reconstruction is proposed due to the following advantages over other techniques: (i) non-parametric -in the sense that no noise distribution needs to be assumed in advance; (ii) unsupervised - since the dataset with missing or incomplete class labels can be appropriately handled, and (iii) robust - the performance of pattern clustering using FCM should not be affected significantly by small deviations or it should not deteriorate drastically due to noise or outliers.

1.1. Empirical approach for pothole detection in 2D imagery

Fig. 2(a) shows a typical 2D-image frame of a pavement, whereby Fig. 2(b)(i) is a tile-image cropped from an image frame. Visually, the potholes in Fig. 2(b)(i) have different spectral reflectances and spatial textures. In Fig. 2(b)(ii) however, the gray scale values of the potholes A and B and the linear crack C are all observed to be less than those of the

surrounding unweathered pavement surfaces. In pavement distress detection, the objective should be to detect and categorize the different distress types. However, by using DWT-SMF method proposed by Ouma and Hahn [17], the results in Fig. 2(b)(iii) and (b)(iv) are obtained, whereby the darker pothole B and the linear distress segment C are well detected as compared to pothole A. This illustrates the difficulty in pavement distress detection in which the differentiations between the unweathered material and the weathered pavements is not a straightforward image segmentation task.

From the results in Fig. 2, it can be argued that vision-based methods operating on 2D images may be limited in detecting the surface-elevation defect. While 3D reconstruction and laser scanning methods may address this limitation, they are not only costly, but are only capable of detecting some defects and not all of them simultaneously [4]. The illustration in Fig. 2(b) points to the fact that it is important to develop pavement-defect-specific models. This is because most of the techniques reported in pavement distress detections suffer from dependence on multispectral data, classification, and local or global registration and variations. In order to address the above phenomenon in 2D image based pothole detection, the pothole detection system should be based on the fuzzy intensity and morphological properties of pavement distress features and structures such that the two-dimensional pothole geometric properties are taken into account.

Given a 2D image, the objective in feature extraction is to accurately locate, for every edge pixel, the following feature cues: orientation, intensity differences on edge sides, subpixel positions and curvature. A major difficulty in pavement images is the intensity inhomogeneity artifact, which causes a shading effect to appear over the image. The artifacts can significantly degrade the performance of segmentation methods that assume that the intensity value of a defect class is constant over the image. This forms the hypothesis in this study that in 2D vision-based pothole detection, robust smoothing at a suitable scale is a prerequisite for good defect-detection results and is empirically illustrated in Fig. 3.

Fig. 3(a)(i) shows an image with similar pixel values within and outside the candidate pothole region. The observed phenomenon in Fig. 3(a)(i) is attributed to the effect in 2D imaging due to limitations of the imaging systems in terms of the resolution and illumination conditions. Physically, the presence of foreign objects such as stones and pebbles within the pothole need to be compensated. This in part causes "spill-in" and "spill-out" effects between the defect and non-defect regions as depicted in Fig. 3(a)(iii), such that the desired detection or segmentation in Fig. 3(a)(ii) may not be achieved. The second reason for the observations in Fig. 3(a) can be attributed to the pothole pixels comprising of different types and fractions of features or *mixels*, which should be modeled and segmented independently.

Fig. 3(b) presents a section of a pavement image with the corresponding histogram plot. The edge pixels labeled X between regions A and B in Fig. 3(b) can be classified to belong to either of the two regions due to the *mixel* problem. However after filtering and smoothing (Fig. 3(c)), the pixels within region X have distinct grayscale values and the edge gradient is much larger, such that the difference between areas A and B can be estimated and mapped as in Fig. 3(d), hence minimizing the "spilling" effects. Arguably therefore, it is only after efficient smoothing that it is possible to cluster potentially homogeneous regions and estimate the edge of the pothole. From the analysis in Fig. 3, it can be concluded that potholes are defined by randomized shapes and structures that cannot easily fit any analytical descriptions, hence the reliance on intensity and textural models is required.

To achieve the filtering and smoothing in pothole detection, the wavelets transform is used in this study. This is because the multiscale-wavelet smoothing uses dual filtering, such that while approximating the original image, it harmonizes the spectral reflectance within a given neighborhood in (x,y) directions, and effectively minimizes or eliminates noise and imaging artifacts within the image, such that noise is reduced preferentially over image signal. Because wavelet transform
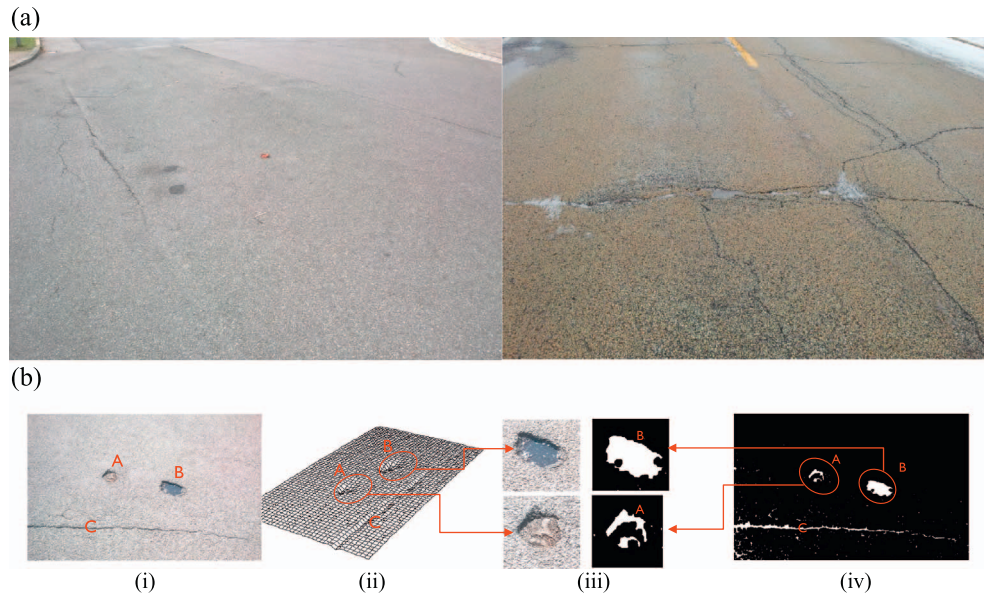
(a)

(b)

(i) (ii) (iii) (iv)

Fig. 2. (a) Sample wide area 2D-image frames acquired using smartphone. (b) Selected image-tile visualization and detection of potholes and linear distresses on asphalt pavements.

concentrates the signal power, if a constant background is subtracted from the wavelet transform of an image, then the signal should suffer less loss than the noise. This results in the minimization of over-segmentation during clustering, especially in regions with low feature contrasts in the RGB imagery.

2. Literature review on pothole detection

2.1. Imaging and sensor methods for pothole defect detection

As already stated in the introduction, the use of sensors for automated pothole surveys on pavement surfaces can be categorized into vibration-based, laser-scanning and vision-based techniques. Vibration-based techniques employ the gradient variations from accelerometer data. According to Yu and Yu [8], the accuracy of detection of potholes

using accelerometers is lower than that achieved with the other sensors such as cameras and lasers, since the potholes are only detected when a vehicle's wheels traverse a pothole, and consequently false detections are likely to occur while vehicles drive over *e.g.* manhole covers, grates and speed bumps [14]. Further studies on vibration-based methods for pothole detections can be found in: Cong et al. [29], Jang et al. [30]; Madli et al. [31], Wang et al. [32] and Chen et al. [33].

Because of the 3D geometry of potholes defects, several research studies have adopted depth information obtained from either 2D stereovision or 3D laser scanning. Hou et al. [9] and Wang [34] used stereovision imaging technique to create 3D surface models for pavement condition assessment. The study results showed that there were problems with complete 3D reconstruction using stereovision methods attributed to the complexities in feature matching. To improve on the results from Hou et al. [9], Salari and Bao [12] combined stereovision
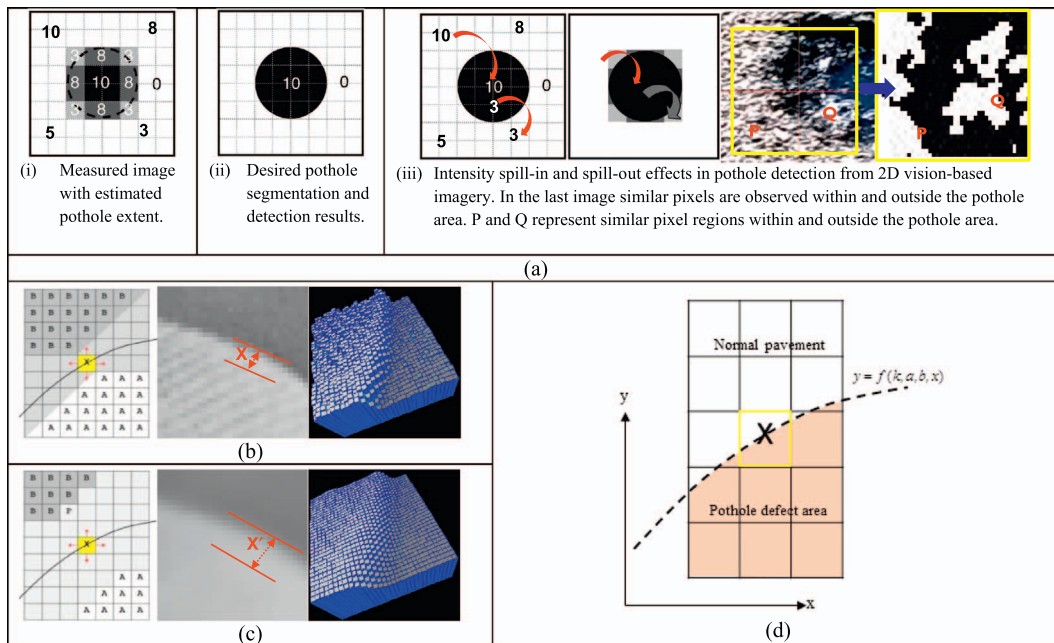

(i) Measured image with estimated pothole extent.

(ii) Desired pothole segmentation and detection results.

(iii) Intensity spill-in and spill-out effects in pothole detection from 2D vision-based imagery. In the last image similar pixels are observed within and outside the pothole area. P and Q represent similar pixel regions within and outside the pothole area.

(a)

(b)

(c)

(d)

Fig. 3. (a) Conceptualization of the influence of imaging processing on pothole detection with the "spill-in" and "spill-out" effects. (b) Section of original image-tile with corresponding histogram. (c) Smoothened or filtered section of original image showing intra- and inter-spectral and textural homogeneity. (d) Edge detection and estimation illustration in filtered imagery.

data with colour segmentation methods in pavement distresses inspection by using a neural network classifier and applying a probabilistic relaxation thresholding. The results were effective for the detection of different pavement surfaces.

Bursanescu et al. [35] proposed a 3D vision system based on laser sensors to obtain road-surface profiles. In NCHRP [36] and Wang [34], laser sensors were used for road-surface-roughness and rutting measurements. Li et al. [10] proposed a 3D transverse scanning technique based on IR laser sensors and developed a system for 3Dsurface profile generation for the detection of pavement distortions such as rutting and shoving. Yu and Salari [37] presented laser-imaging techniques to collect road information, and used artificial neural network algorithm for pothole detection and severity estimation based on index of severity.

As argued in Salari and Bao [18]; Koch and Brilakis [11]; NCHRP [36], and Zhou et al. [38], the currently available imagery and laser scanner based methods are expensive technologies and systems and therefore in most cases the equipment cost is a barrier in adopting the automated methods. From the different sensor techniques for road condition surveys, it is observed that most studies are based on image-processing techniques due in part to the fact that 2D images yield better results in road surface distress detections [39,84]. However, environmental conditions such as light and shadow conditions, different background textures, and non-crack patterns compromises the assessment outcomes [17]. As such improvements on 2D colour image segmentation techniques for pothole detections are necessary [40,41].

2.2. On pothole detection from 2D colour imagery

Despite the fact that the limitations and drawbacks in manual pavement-assessment methods have resulted into the development of automated defect assessments in the recent past, a majority of the commercial pavement-assessment tools are based on image-processing approaches for defect detection [39]. Zhou et al. [42] presented a method for detection of potholes and patches in 2D images using DWT. The method was however limited by the fact that it was restricted to the level of presence [4]. Karuppuswamy et al. [43] modeled potholes as circular obstacles with a white colour, and of 2-ft in diameter. By using a standard imaging board, potholes were detected through blob matching. Lin and Liu [44] proposed a pothole detection system from 2D images, in which partial differential equations were used for initial segmentation to isolate distress regions. A non-linear support vector machine was then applied to differentiate between potholes and other types of distresses.

For distress detection on asphalt pavements, Nejad and Zakeri [45] presented a comparative evaluation of the discriminating power of several multiresolution texture analysis techniques using wavelet, ridgelet, and curvelet-based texture descriptors. They concluded that for pothole detection, the curvelet-based signatures outperformed the other multiresolution techniques, yielding an accuracy of 97.9%. By segmenting a road surface image into defect and non-defect regions using a histogram shape-based thresholding algorithm, Koch and Brilakis [11] detected pothole defects using morphological thinning and elliptic regression with an accuracy of 86%. They further improved on the method through the use of automated video-based pothole detection, which was able to incrementally update the texture signature for non-patches areas, and to track the detected potholes [15].

For detecting a variety of road-related objects such as lanes, road signs and potholes, Danti et al. [46] presented a system of image processing and classification. To detect potholes, black and white threshold was applied to the image that highlights the pothole area. The results showed that the algorithm was unable to deliver desirable result, as it segmented many undesirable areas within an image. They recommended that a more effective filtering method should be adopted in order to improve on the accuracy of the approach. Jog et al. [41] presented an approach based on visual 2D recognition and 3D reconstruction for detecting and measuring the width, quantity, and

depth of potholes using a monocular camera mounted on a mobile vehicle.

Through clustering, Buza et al. [7] presented a pothole detection system based on the method by Koch and Brilakis [11]; however the thresholding algorithm was modified to Otsu's image thresholding, so as to automatically determine the correct threshold value necessary per frame. The approach was based on unsupervised vision-based method and does not require expensive data collection equipment, filtering and training. The method was able to roughly estimate the pothole surface-areas with accuracy of approximately 81%. A vision approach was also employed by Murthy and Varaprasad [47], with images obtained from a camera mounted on top of a vehicle, and a custom MATLAB code used to detect potholes.

In Ryu et al. [48], a 2D vision-based method for detection of potholes on asphalt and concrete pavements was presented. By using a histogram shape-based thresholding and maximum entropy, potholes were detected from the surroundings. A median filter was further used to remove noise such as cracks, and the actual potholes extracted by comparing the candidate regions with background regions through histogram and standard deviation features. Their study results yielded an overall accuracy of 73.3%. Radopoulou and Brilakis [49] and Radopoulou et al. [50] detected potholes from 2D video frames by respectively using the semantic *texton* forests (SFT) algorithm and through classification of distress severity from vision images and elevation signals.

From the literature survey, 2D vision-images it is seen to suffice the detection of potholes on asphalt pavements, especially if the defects are at the incipient stages of formations as discussed in Fig. 1. Methodologically, most of the reported studies relied on the use of thresholding, which as stated in the introduction has limitations in autonomous image segmentations and as such there is need for improved algorithmic accuracy as presented in reviews by Radopoulou and Brilakis [4] and Schnebele et al. [6]. It is conclusive that because pavement images have high levels of intensity variation and texture content, the detection of distresses is generally an ill-conditioned problem.

In general, despite the fact that the performances of image-segmentation algorithms has improved significantly over the years, these improvements have come partly at the price of ever more sophisticated feature selection processes, more complex statistical models, and more costly optimization techniques. The drawbacks of some of the widely used segmentation algorithms, as presented above, for pavement pothole detections, can be summarized as follows [51]:

(i) Clustering algorithms such as *k*-means and expectation-maximization (EM) are influenced by initial-value estimations, and therefore have relatively low stability. The methods also are more sensitive to noise and intensity inhomogeneities, resulting into the inaccurate segmentation results [52].

(ii) Thresholding is an iterative procedure, which typically does not take into account the spatial characteristics of an image and cannot be applied to multichannel images. This makes it more sensitive to noise and intensity inhomogeneities, which occur in pavement images. Region-growing like thresholding is another segmentation approach that is rarely used alone but usually within a set of image-processing operations, and has the disadvantage that it requires manual interaction to obtain the seed point(s). The main limitation in using thresholding is that, in its simplest form, it is only generates two classes and it is not possible to apply to multichannel images [51].

(iii) Supervised classifiers in general require tedious and robust training data that are usually manually segmented, and then used as references for the automatic segmenting new data. This makes them subjective and influenced by the training data templates, and do not perform spatial modelling.

(iv) The Markov Random Field (MRF) statistical model has the difficulty in the selection of proper parameters for controlling the

strength of spatial interactions. For example, too high a setting can result in an excessively smooth segmentation and loss of important structural details.

(v) Artificial neural networks (ANN) are inherently parallel networks, and their processing are usually simulated on a standard serial computer, thus reducing computational advantage. Because of the many interconnections used in a neural network, spatial information can be easily incorporated into its classification procedures. However, ANNs require good and large training samples, making them subjective and also require significant computational time.

From the literature review, it is conclusive that to improve on pothole detection, the task should be divided into two steps comprising of pattern extraction, to achieve clean information comprising of candidate pixels and pattern classification, to provide failure or defect information. This forms the basis for the use of wavelets for noise filtering and region smoothing [46], and the extension of the detection process to the clustering of the scene multivariate features using FCM and finally fine segmentation of wavelet-FCM results using morphological reconstruction approach in order to resolve the "spill-in" and "spill-out" phenomenon and to extract the extrema for pothole defect detection.

3. Methodology

Because of the inherent spectral and spatial complexities on pavement surfaces as illustrated in Figs. 2 and 3, a non-crisp clustering approach is preferred. There exist several clustering algorithms proposed for dealing with different problems such as partitioning clustering, hierarchical clustering, neural network based clustering, mixture model based clustering and kernel based clustering [53]. The most widely used clustering algorithm is the partitioning clustering, such as *k*-means, fuzzy *c*-means, and their variations. FCM is better than *k*-means because of its non-crisp characteristics. However, FCM tends to perform poorly under noisy environment and therefore modified fuzzy clustering algorithms have been proposed [54,55]. This drawback is overcome in this study by using the hybrid approach of scale-based spectral and textural filtering using wavelets, *a priori* to FCM clustering. From the clustered information into *n*-classes, the extraction of the minima and maxima extrema is necessary in order to segment the pothole defect pixels. Morphological reconstruction is an operator provided by mathematical morphology, which is suitable for hierarchical segmentation and is effective in extrema extractions as well as hierarchical image construction [28]. Further, morphological reconstruction has the advantage of providing a good symbolic description of the image and has the ability to extract geometrical information in signals through appropriate transformations by referring to the original image.

3.1. 2D image filtering using multiscale wavelet transform

Since image textures may often contain both statistical and structural properties, a spatial analysis method should be able to represent both types of properties in order to completely describe the texture [56]. Various methods for texture feature analysis have been proposed during the last decades [57]. However the image texture analysis problem remains difficult and still subject to intensive research. A major class of feature extractors relies on the assumption that texture can be defined by the local statistical properties of pixel gray-levels.

From an image histogram, first-order statistics can be derived and used as texture features. However, first-order statistics do not suffice for adequate texture description, and therefore second-order statistics are necessary as efficiently reflected in features computed from the gray-level co-occurrence matrices [58]. Second-order statistics have also been found to be inadequate [59], as such other texture analysis schemes such as Markov Random Fields [60]; fractal models [61] and Wold decomposition [62] have been proposed.

Several multichannel texture analysis systems have been developed [63–65]. However in the last decade, wavelet theory has emerged and has become a mathematical framework which provides a more formal, solid and unified framework for multiscale image analysis [66,67]. The main advantage of wavelet analysis is that it has the ability to perform a local analysis by analyzing a localized area of a larger signal. Furthermore, wavelets analysis is translation invariant process which has the desirable property for the accurate localization of region boundaries in terms of spectral and textural feature properties [68]. This is attributed to the hypothesis that wavelets, through dual filtering is capable of revealing some aspects of data that other signal analysis techniques often miss, such as trends, breakdown points, and discontinuities in higher derivatives.

In summary, 2D-DWT is computed by applying a separable filter bank to the image according to Eqs. (1)–(4) [69].

$$L_n(b_i, b_j) = [H_x * [H_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \tag{1}$$

$$D_{n1}(b_i, b_j) = [H_x * [G_y * I_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \tag{2}$$

$$D_{n2}(b_i, b_j) = [G_x * [H_y * I_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \tag{3}$$

$$D_{n3}(b_i, b_j) = [G_x * [G_y * I_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \tag{4}$$

where $*$ denotes the convolution operator; $\downarrow 2,1 (\downarrow 1,2)$ is the subsampling along the rows and columns and $I_0 = I(\vec{x})$ is the original image. H and G are respectively the low and high bandpass filters. I_n is obtained by dual low-pass filtering and is therefore referred to as the low resolution image at scale n. The detail images D_{ni} are obtained by bandpass filtering in a specific direction and contain directional detail information at level or scale n. The original image I_0 is thus represented by a set of subimages at several scales $\{I_d, D_{ni}\}_{i = 1, 2, 3, n = 1, \dots, d}$, which is a multiscale representation of depth d or level of decomposition.

In practice, the multiscale wavelet filtering and denoising can be divided into the following three steps according to Donoho and Johnstone [70]:

(i) decompose the noisy signal into the time-frequency domain by a selected set of orthonormal wavelet basis.
(ii) threshold the wavelet coefficients by suppressing coefficients smaller than a specific value and preserving other larger coefficients, and
(iii) reconstruct the thresholded coefficients into the original time domain.

Typical wavelet types found in literature are the Haar, Morlet, Gaussian, Meyer, Daubechies and Symlet wavelets. Obviously, the selection of wavelet basis and defining a proper threshold to discard partial detail coefficients are imperative for a successful denoising. The most widely used wavelet basis in the context of image analysis, and also adopted in this study, is the Daubechies wavelet [66]. While the initial Haar basis function, also termed as first-order Daubechie's basis, is best suited in representing step signals or piecewise constant signals, the Daubechies basis function is preferred for smoother signals [71].

Thresholding in wavelet analysis has been proposed through either hard or soft methods [70]. While hard thresholding discards wavelet coefficients below a threshold τ and maintains the same level for wavelet coefficients that exceed the threshold value, soft thresholding discards wavelet coefficients below the threshold τ, but also shrinks larger coefficients towards zero by the value of the threshold. Thus, hard thresholding can result in better reproduction of peak discontinuities, while soft thresholding with a larger bias can give better visual quality and fewer artifacts. The implementation of the 2D-DWT in this work follows from Ouma and Hahn [17], whereby optimal wavelet level is obtained through empirical analysis and thresholding of the optimal-level wavelet is implemented by using soft multi-thresholding so as to obtain at least three classes, and not the crisp two-

classes.

3.2. Fuzzy c-means clustering

FCM is a soft extension of the hard k-means, and is a popular clustering algorithm based on the minimization of the distance-based objective function [72]. The technique attempts to partition every image pixel into a collection of M-fuzzy cluster centers with respect to some given criteria [73]. Fuzzy cluster analysis allows gradual memberships of data points to clusters measured as degrees in [0, 1]. This gives it the flexibility to express the fact that data points can belong to more than one cluster. Furthermore, these membership degrees offer a much finer degree of detail of the data model. Aside from assigning a data point to clusters in shares, membership degrees can also express how ambiguous or definite a data point should belong to a cluster [74]. The mathematical representation and implementation of the FCM is briefly explained below.

By letting $X = (x_1, x_2, \ldots, x_N)$ denote an image with N pixels to be partitioned in c clusters, where x_i represents multispectral (features) data, the FCM algorithm is an iterative optimization that minimizes the cost function J_m defined according to Eq. 5. The objective function is based on the minimization of the Euclidean distance between the input data x_k and the cluster center c_i, subject to the following probability constraints: $\Sigma_{i=1}{}^{C} u_{ik} = 1$, $1 \le k \le n$ and $0 < \Sigma_{k=1}{}^{n} u_{ik} < n$, $1 \le i \le c$.

$$J_m(U, C) = \sum_{i=1}^{N} \sum_{k=1}^{C} (u_{ik})^m (d_{ik})^2 \tag{5}$$

where u_{ik} is the degree of membership of pixel x_k in the ith cluster center; $d_{ik} = \|x_k - c_i\|^2$ is the Euclidean distance between any measured feature and the center; c_i is the d-dimension clustering center and $m \in [1, +\infty]$ is the weighted matrix defining the fuzziness factor which is used to control the fuzzy degrees of the cluster results.

If $U_i = (u_{1i}, u_{2i}, \ldots, u_{Mi})^T$ is the set of membership degree of the ith pixel associated with each cluster center; x_i is the ith pixel in the image, and c_i is the ith cluster center, then $U = (U_1, U_2, \ldots, U_N)$ is the membership degree matrix and $C = (c_1, c_2, \ldots, c_M)$ is the set of cluster centers.

The objective function J_m of the FCM clustering technique reveals the clustering quality of the output images in terms of the degree of compactness and uniformity of the cluster centers. Specifically, a smaller value of J_m indicates a more compact and uniform cluster center set that leads to more desirable clustering results. However, a closed-form solution for calculating the minimum value of J_m does not exist because different types of input images consist of different pixel distributions. Hence, different expressions of the objective function J_m are produced. Consequently, a formula that may be specifically used to calculate the minimum value of J_m for all types of images during the FCM clustering process does not exist [75]. To achieve minimization of the objective function J_m, the alternative strategy is to carry out the FCM clustering technique in an iterative manner. Fuzzy partitioning is carried out through an iterative optimization of the objective function, with the update of membership u_{ik} and the cluster centers c_i according to Eqs. 7 and 8, and the iteration stops when Eq. 9 is satisfied. In order to get the minimum value of $J_m(U, C)$ let:

$$\begin{cases} \frac{\partial J_m(U,C)}{\partial u_{ik}} = 0 \\ \frac{\partial J_m(U,C)}{\partial c_i} = 0 \end{cases} \tag{6}$$

After an alternative optimization, the membership function u_{ik} and the cluster center c_i are respectively updated as in Eq. (7), and the update of the cluster centers by c_i as in Eq. (8):

$$u_{ik} = \frac{1}{\sum_{j=1}^{C} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} = \frac{d_{ik}^{-2/(m-1)}}{\sum_{j=1}^{C} d_{jk}^{-2/(m-1)}} \tag{7}$$

$$c_i = \frac{\sum_{k=1}^{N} (u_{ik})^m \cdot x_k}{\sum_{k=1}^{N} (u_{ik})^m} \tag{8}$$

where $m > 1$ is known as the fuzzifier with $m = 2$ as set by Bezdek [72]; C is the number of clusters and N is the number of map units, and the iteration stops when:

$$\max_{ik} \|u_{ik}^{(t+1)} - u_{ik}^{(t)}\| < \varepsilon \tag{9}$$

where ε is a termination criterion between 0 and 1, whereas t is the iteration number.

This procedure converges to a local minimum or a saddle point of J_m, and therefore multiple runs may produce different results as the minimization process is significantly sensitive to the initial randomly selected cluster centers. Starting with an initial guess for each cluster center, the FCM converges to a solution for c_i representing the local minimum or a saddle point of the cost function. Convergence can be detected by comparing the changes in the membership function or the cluster center at two successive iteration steps.

From the above formulation, the cost function is minimized when pixels close to the centroid of their clusters are assigned high-membership values, and low-membership values are assigned to pixels with data that are far from the centroid. The membership function represents the probability that a pixel belongs to a specific cluster. In the FCM algorithm, the probability is dependent solely on the distance between the pixel and each individual. The output from FCM clustering is a list of cluster centers and n membership-grades for each pixel, where n is a number of desired clusters. A pixel will be assigned to the cluster with highest membership-grade. Further details on the FCM clustering algorithm can be found in [74].

From the above representation, FCM is not robust to tolerate noise or outliers because of assigning relatively high membership values to outliers across c-clusters [55]. This is because for an outlier, all the ratios d_{ik}/d_{jk} will be close to unity, leading to all membership values of an outlier close to $1/c$. These rather high membership values may cause an improper estimation of cluster centers since outliers will attract the centroids of clusters. Therefore if a small weight (membership) can be assigned to the "noisy" data points and a large weight to the "clean" data points, the estimation of cluster centers will be more accurate. This conceptualization inspires the proposed denoising using DWT, in order to boost the performance of the conventional FCM clustering.

3.3. Morphological reconstruction

Morphological reconstruction is based on two input images, a marker image as the starting point and a mask image as the constraint, and a structuring element which is used to define connectivity [76]. Morphological reconstruction processes the marker image, based on the characteristics of the mask image. The high-points or peaks, in the marker image specify where processing begins, and the processing continues until the image values stop changing. Conceptually, morphological reconstruction can be thought of as repeated dilations of the marker image until the contour of the marker image fits under the mask image. In this way, the peaks in the marker image are spread out or dilated. Reconstruction by dilation, that is the geodesic transformations of a marker image F and a mask image G, can be represented in Eqs. (10)–(11) [77].

$$R_G(f) = \delta_G^{(i)}(f) \text{ with } \delta_G^{(i)}(f) = \delta_G^{(i+1)}(f) \tag{10}$$

$$\delta_G^{(n)}(f) = \delta_G^{(n-1)}(f) \text{ with } \delta_G^{(1)}(f) = \delta^{(B)}(f) \wedge g \tag{11}$$

where $\delta_G^{(n)}$ is a geodesic dilation and B is a at disc-shaped structuring element with a suitable radius.

For the two grayscale images F and G defined on the same domain D, taking their values in the discrete set $\{0, 1, \ldots, N-1\}$ and such that $F \le 1$ (i.e. for each pixel $p \in D_I, J(p) \le I(p)$), the reconstruction $\rho_I(F)$ of
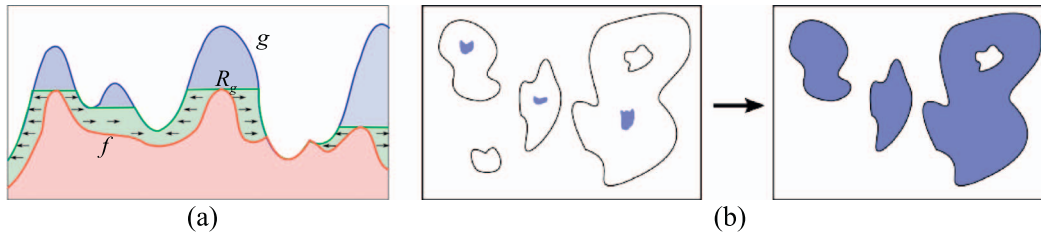
Fig. 4. (a) One-dimensional illustration of morphological reconstruction of marker image *F* through repeated dilations under the mask image *G*. (b) Markers inside the objects in the left image and the image on the right show the reconstructed objects and the final binary image.
(Figure adopted from [79]).

*G*from *F* is given as in Eq. 12, which can be modified such that the grayscale reconstruction $\rho_I(F)$ of *G* from *F* is obtained by iterating grayscale geodesic dilations of *F* under *G* until stability is reached according to Eq. 13 [78].

$$\rho_I(F)(p) = \max\{k \in [0, N-1] \mid p \in \rho_{T_k(I)}(T_k(J))\} \tag{12}$$

$$\rho_I(F) = \lim_{x \to +\infty} \delta_I^{(n)}(F) = \lim_{x \to +\infty} \underbrace{\delta_I^{(1)} \circ \delta_I^{(1)} \circ \dots \circ \delta_I^{(1)}(F)}_{n-\text{times}} \tag{13}$$

As illustrated in Fig. 4(a), the marker image intensity profile is represented as the red line defined by *f*, and the mask image intensity profile is represented as the blue line *g*. The final image intensity profile is represented as the green line R_g [79]. The arrows show the directions of propagation from the marker intensity profile to the mask intensity profile. The green regions between *f* and *g* shows the changes introduced by the morphological reconstruction process. When applied to binary images, morphological reconstruction pulls out the connected components of an image identified by a marker image as illustrated in Fig. 4(b), where the dark patches inside three objects in the mask image *G*, on the left correspond to the marker image *F*. The result of the morphological reconstruction is shown on the right of Fig. 4(b). The reconstruction $\rho_I(F)$ of marker *F* from mask *G* is the union of the connected components of *F* which contain at least a pixel of *G* [79]. In Fig. 4, the marker intensity profile is propagated spatially but is bounded by the mask image's intensity profile.

In this study, 8-connected neighborhood is chosen so that at any given dilation, the objective is to have one object. Just like binary reconstruction extracts the marked components of the mask, grayscale reconstruction extracts the peaks (maxima) of the mask which are marked by the marker image. The primary difference between binary and grayscale morphological reconstruction algorithms is that in binary reconstruction, any pixel value change is necessarily the final value change, whereas a value update in grayscale reconstruction may later be replaced by another value update.

3.4. Pothole detection performance measures

To objectively measure the quality of the detected and extracted potholes, it is important to use measures which provide information about the segmented regions and boundaries. This, as already stated above, is because segmentation is an ill-defined problem, as there is no single manual ground-truth segmentation against which the output of an algorithm may be compared. Rather the comparison should be made against the set of possible perceptually consistent interpretations of the image, of which only a minuscule fraction is usually available [80].

In this study the Dice coefficient of similarity and the Jaccard Index are used to determine the validity of the study results, and is implemented as presented in Ouma and Hahn [17]. These measures compares the results obtained from the algorithm to a set of manually segmented images, and determines the degrees of over- and under-segmentations by determining the fractions of spatial overlaps between the segmentation results and the ground-truth data. The advantage of these measures is that they count the fractions of pairs of pixels whose

labels are consistent between the computed segmentation and the ground-truth, averaging across multiple ground-truth segmentations.

If *S* is the segmentation provided by the proposed algorithm and *G* is the manually segmented image, then the Dice coefficient d_C is defined according to Eq. 14, for two similarly labeled regions *r* in *S* and *G*. Dice coefficient can also be expressed in terms of the Jaccard index j_I as expressed in Eq. 13, where $\overline{G}$ is the average ground-truth data. The ground-truth set is generally defined as $G = \{g_1, g_2, \dots, g_N\}$, and $\overline{G} = \sum_1^N g_i/n$ where *n* is the number of manually segmented images. In this study, $n = 3$ is adopted such that at least two matching ground-truth datasets are merged [17].

$$j_I(S, G_n) = \frac{\sum F \cdot \overline{G}}{\sum F + \sum \overline{G} - \sum F \cdot \overline{G}} = \frac{|F \cap \overline{G}|}{|F| + |\overline{G}| - |F \cap \overline{G}|} \tag{13}$$

$$d_C(S, G_n) = \frac{2^* \sum F \cdot \overline{G}}{\sum F + \sum \overline{G}} = \frac{2^* |F \cap \overline{G}|}{|F| + |\overline{G}|} = \frac{2^* j_I}{(1 + j_I)} \tag{14}$$

The Dice coefficient like the Jaccard Index is in the range [0, 1], where high values indicate a large similarity between the segmented image and the ground-truth reference data. Because of their formulations, d_C is always larger than j_I, except at 0 and 1 when both are the same. In this study, the results of the Dice and Jaccard accuracy measures are compared with the sensitivity metric, alongside the pothole surface-area detection error rates. Sensitivity or recall refers to the proportion of images containing potholes that have been correctly classified or segmented.

To further analyze the accuracy of the results, a volume overlap between the 2D binary images is also presented, in terms of the surface-area, so as to enable the visualization of the difference between the variability of average manual segmentation and the results from the proposed methodology. From the surface-area differences between the ground-truth and segmentation results, the mean and standard deviation of error percentages in pothole size and shape extractions are also determined for the experimental results. The surface-area of the detected pothole is approximated according to the Eq. 15, and can be defined by an optimal rectangular bounding box.

$$\begin{aligned} A_{d_i} &= l_p \times w_p \times \sum_x \sum_y I_p(x, y) \\ &= l_p^2 \cdot \sum_x \sum_y I_p(x, y) \end{aligned} \tag{15}$$

where A_{d_i} is the total pothole area as detected within an imaging frame; $l_p = w_p$ is the size of a defect or pothole pixel I_P; (x, y) is the coordinate or position of the defect pixel I_P, and therefore $\Sigma_x \Sigma_y I_p(x, y)$ is the total defect pixel counts which is based on the total number of defect pixels in the image and a 2×2 averaging window.

The implementation of the proposed methodology and the results validation procedure is schematically presented in Fig. 5. Details on the processing steps are as presented in our initial study as reported in Ouma and Hahn [17]. The prototype is implemented in MATLAB version 7.11 (R2010b) and supported by the MATLAB Image Processing Toolbox.
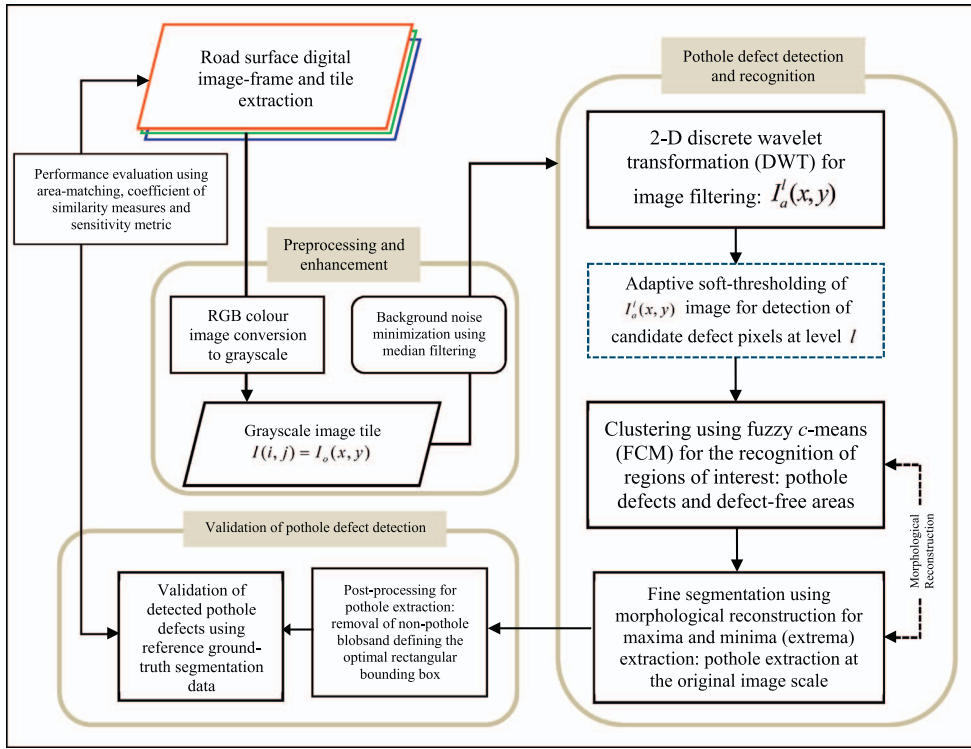
Fig. 5. Schematic workflow of the approach for pothole detection on asphalt pavements using the proposed hybrid wavelet-FCM and morphological reconstruction from 2D imagery. $I_a^l(x,y)$ is the filtered image at DWT optimal-level *l*.

4. Image data acquisition and test data

Currently, monitoring of the road conditions is carried out either *via* the following techniques or a combination thereof [81]: (i) manual inspections where raters visually survey and judge the road condition [3]; (ii) specialized road condition survey vehicles comprising of different motion and image sensors, and (iii) in some cases citizens who call in their observations [13]. Due to the volume of roads that needs to be inspected, manual surveys are limited in terms of productivity and more so given the fact that the capacity of every road in a given network is often utilized to full capacity due to high traffic volumes, leading to an ever faster degradation that necessitates more frequent inspections. The requirement for higher inspection frequency cannot be tackled manually due to subjectivity and accuracy in data collection, time, cost and human resource constraints. The second method, also known as automated road condition survey, is generally expensive in terms of equipment costs, processing techniques involved and investments in skilled human personnel. The third method is carried out voluntarily through reports or by citizens' complaints, and is most of the times about severe pavement distresses, like large potholes on main roads [13].

For urban roads, it is desirable to have a low-cost automated system that can be used to monitor the roads on a continuous basis with minimal human intervention. This study exploits the pervasive and 'smart' nature of consumer-grade smartphone devices to collect the road condition data through smartphone colour camera imaging. By using such devices, no dedicated and expensive platforms and drivers are needed for automated data collection. In this study, the main device for the data collection is the car-charged and GPS enabled Samsung Galaxy S5 camera, with a resolution of 1080×1920 pixels. By mounting the camera on the windshield of a Toyota Hiace van, using a universal car windshield mount holder, a camera shutter delay of 5-frames per second in automated mode was used in order to accommodate the variable vehicle speeds. The dash-cam images were captured in terms of wide frames, from which the tiles, referred to as the test datasets were extracted for implementing the algorithm.

Most of the test images used in this study were taken when the weather was overcast or cloudy, during high-shadow or after light rains, since these conditions offered good lighting and imaging conditions, hence minimizing the artifacts encountered in during to natural and environmental conditions. From the captured data and for further processing, experimental distresses images with minimal blur [82] were selected at 10-m chainages. The smartphone-GPS was used to locate the chainages of the observed pavement failures for the case study of 3-km Nandi road in Eldoret Town, Kenya [3]. The experimental data sets were varied in terms of pothole sizes and shapes, different imaging conditions such as noise, background features and illumination and shadow conditions and also taking into consideration the various pavement conditions such as pavement discolorations and other pavement defects like linear cracking. With these variations, a total of 75-test image datasets were selected for testing and validating the proposed FCM-DWT approach. All the selected test images contained potholes, and the proposed method only detects and extracts potholes in a single image-tile, but does not automatically identify pothole-images from a range of different pavement images. The characteristics of the test data sets comprising of pothole defects and the non-pothole defects are summarized in Table 2.

5. Results and analysis

5.1. Pothole detection and extraction results

Representative results for the 2D vision-based detection of pothole defects are presented in Fig. 6. Tiles of the original images are presented in Fig. 6(a), and Fig. 6(b) and (c) respectively shows the detections of candidate pothole defect areas using DWT, before and after soft-thresholding, at level-3 of Daubechies-4 (db4). Despite the fact that thresholding of the wavelets transform smoothened image results (Fig. 6(c)) is able to show the presence of not only potholes, but also linear cracks, there are also observed artifacts within the image which are attributed to noise and background information. Through clustering using FCM, the results in Fig. 6(d) indicate that three-clusters are determined. These clusters are characterized by pothole distresses, other distresses like linear cracks, non-distress features and no-data regions.

Table 2
Characteristics of the test datasets depicting pothole defects and non-pothole defects.

Test pothole image data characteristics	Sample image data
– Illumination and light intensity variations – Background asphalt background variations – Cracks – Oil stains – Patches – Pebbles – Shadows – Other noise or artifacts	

By using morphological reconstruction between the FCM image and the filtered original image, as depicted in Fig. 5, the final pothole segmentation and detection results are obtained as presented in Fig. 6(e). In order to determine the minimal indicator for the extent of pothole surface-area, especially during maintenance and repairs, the results for the first two processing output in Fig. 6(e) are automatically defined within a minimal rectangular bounding box. Fig. 6(f) presents the corresponding manually delineated ground-truth data, which is the average of at least two or three manual pothole area segmentations, as explained in Section 3.4.

In some of the resulting morphologically reconstructed images in Fig. 6(e), small blobs representing non-potholes are observed. These blobs characterize either linear distresses segments or dark asphalt areas which are aggregated during the clustering process and are not completely eliminated by using the morphological reconstruction filtering. The blobs are intuitively eliminated on the basis of their size and proximity to the pothole. That is, the smaller and further the blobs are from the pothole, the higher the probability that they are neither parts of linear distresses nor potholes. The final results presented in Fig. 6(e) are thus those cleaned of the non-pothole regions. However, in
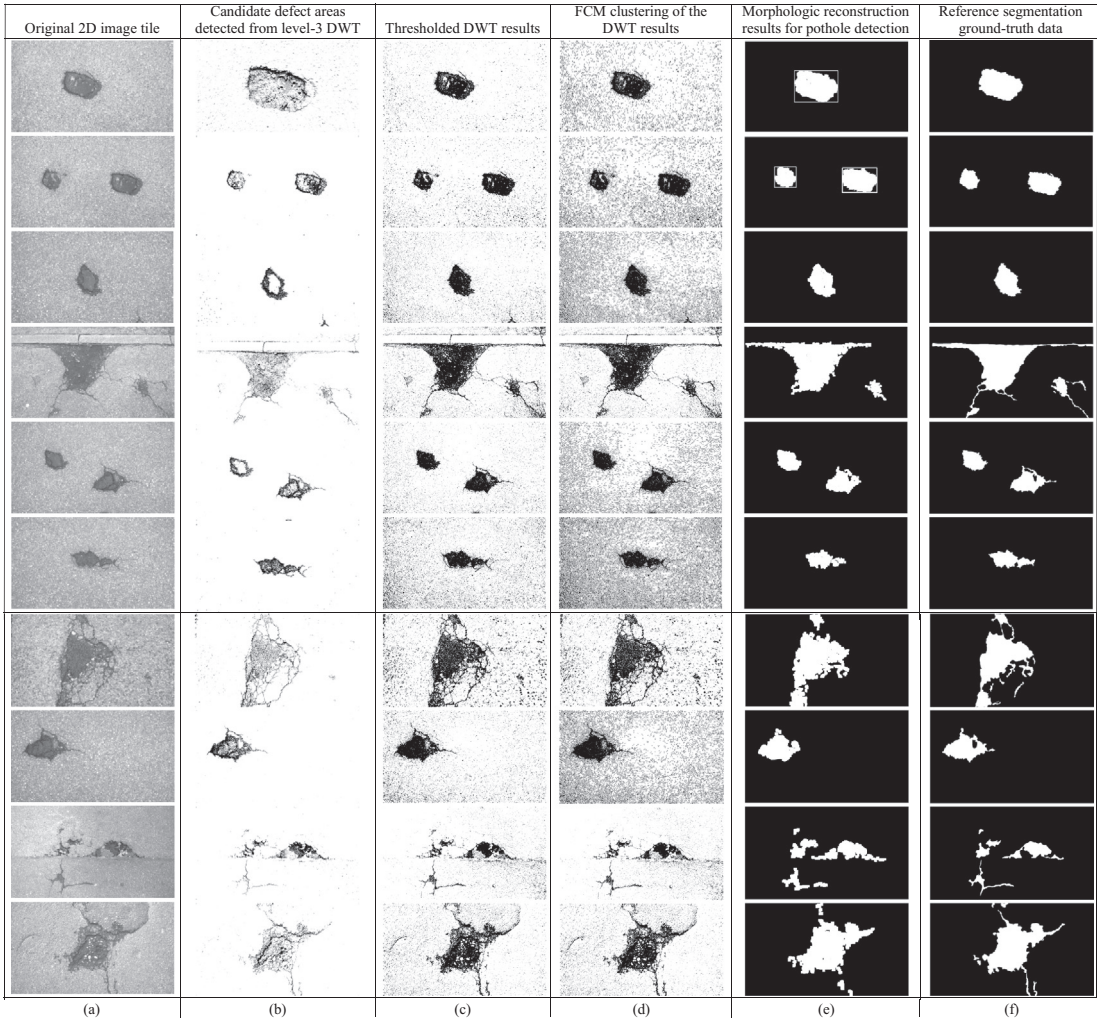

Fig. 6. Pothole detection results: (a) original image; (b)–(d) intermediate results; (e) final results from morphological reconstruction, and (f) ground-truth reference image.

computing the pothole detection accuracy, the results before clean-up are used as presented and discussed in Section 5.2 below.

For the image preprocessing, clustering and morphological reconstruction, the average CPU run-time for the proposed approach was 95 s. The overall results show that the proposed approach is able to detect the pavement distresses, by extracting the potholes in terms of their shapes and surface-area extents. From the study hypothesis, it is observed that the presented algorithmic approach is more biased in detecting round shaped features than the linear cracks. That is, despite the fact that linear features are also characterized by dark pixels as potholes, their spatial sizes and textural differences inhibit their detections using the proposed algorithm and are only detectable when their sizes exceed a certain width.

The results in Fig. 6 shows that the empirically set out conceptualization in Figs. 2 and 3, that the pothole detection system should be independent of other distresses present within the scene as well as noise and background information, is achieved to a high degree. Comparatively, the run-time for the proposed approach is on average 4-times faster than the clustering of the original image. The final results of the study show the ability of the proposed approach to detect the presence of the potholes in shape, size and compactness.

5.2. Quantitative verification of pothole detection from 2D-vision imagery

To validate the results of the detection and extraction of the pothole defects, the ground-truth reference data from manual image segmentation are used. As presented in Ouma and Hahn [17], the manual segmentation results from three experts, were compared and merged based on overlap and neighborhood rules such that: (i) a pixel marked as a crack by two or more experts is considered as a crack pixel, and (ii) every pixel marked as a crack and next to a pixel kept by step (i) or (ii) is also considered as a crack. The reliability of the reference segmentation images was evaluated by comparing the percentage area overlap and the mean distance between each pixel detected by only one expert and not kept in the reference image, and the reference segmentation. For the 75 distress test images, the reference segmentation ground-truth data was accepted when overlaps were more than 90% and the mean distances are less than 10% between any two reference segmentation images.

Using the above derived ground-truth data, the accuracy of detection using the Dice and Jaccard based similarity measures were determined with the maximum similarity being observed as 96.2% and 92.7% respectively. In cases where the degrees of overlaps between the ground-truth and the results were more than 95%, few outliers characterized by lines and or clustered blobs were detected on the pavement surface. Notably, the image results which registered high similarities with the ground-truths were taken either before or in the afternoons, *i.e.* when there is maximum shadow effect. In similar cases, some of the images were acquired after light rainfall. On the other hand, the pothole detections with least similarities to the ground-truth were mostly observed for images which were captured when the shadow effects are at a minimum. For these sets of test images, which accounted for 9% of the test data, the least corresponding Dice and Jaccard indices were determined as 81.8% and 68.2% respectively. This observation points to the fact that road condition surveys should be planned, in terms of the time of day, weather condition and season for data capture, so as to obtain the best results [13].

From the volume overlap analysis presented in Fig. 7, it is observed that the regularly formed and near-elliptical shaped potholes are easier to detect, hence indicating maximum-overlaps with the ground-truths (Fig. 7(a)). This is in part due to the fact that their shadow formations are regular, making easier to segment. Similar results are observed in Fig. 7(c), in which more than one pothole is detected within the same image frame with an accuracy of more than 90%. On the contrary, in Fig. 7(b) and (d), where there are other pavement distresses characterized by multiple linear cracks, slightly larger overlaps between the

detected potholes and the ground segmentations are observed. Notably the potholes with the multiple pavement deformations are mostly formed around the road kerbs.

In all the results for the 75 test images, it is observed that the shape and size of the ground-truth partially coincides with that of the segmented potholes, or in most cases, the pothole ground-truth is slightly smaller in size than the detected potholes. This implies that the proposed approach is accurate to a high-degree in detecting the pothole edges, however in a few results there is an observed marginal over-clustering around the edges of the potholes. For the 75 test datasets, the overall average accuracy using the Dice coefficient of similarity and the Jaccard index were determined as 87.5% and 77.7% respectively. The relative magnitudes of the mean and standard deviation of error percentage in pothole size and shape extraction for the experimental datasets were respectively determined as 8.5% and 4.9%.Since the objective as indicated in the introduction section is on the detection of the presence and extent of the pothole, the results in this study are accurate and acceptable for purposes of maintenance and repairs of incipient potholes. Most importantly, the algorithm is able to estimate the size and shapes of the potholes, with minimal deviations as depicted in volume analysis in Fig. 7. The results presented in Figs. 6 and 7 are particularly important in scenarios where the shapes of the potholes are complex and cannot be mathematically approximated.

For a user, the relationship between the accuracy of the detected potholes should be correlated to the accuracy of extraction of the size and shape of the potholes. This relationship is presented in Fig. 8, with the results showing that high detection accuracy does not automatically translate into an accurate extraction of the shape and surface-area of the detected pothole. The sample results in Fig. 8 represent results from the highest, the median and least detections for the surface-area differences and the accuracy measures. The results in Fig. 8 show that the errors in the surface-areas were not more than 10% in magnitude.

Comparatively from Fig. 8, for the test images #2 and #17, it is observed that the relative magnitudes of the errors of the detected surface-areas of the potholes are nearly similar. However, the degree of similarity to the ground-truths is not exactly the same as determined using the Jaccard index. On the other hand, by using the Dice coefficient, the detection accuracies are seen to be more correlated to the surface-area error magnitudes. On the contrary, by comparing test images #19 and #20, it is observed that while the pothole surface-area detection errors are nearly equal, the degrees of similarities differ. Similar results are observed for test images #10 and #13, and 14 and 15. While it can be argued that these relationships are magnitude dependent, it can be concluded that a user should always check the performance and validity of pothole detection in terms of both the surface-area error and the detection accuracy.

By comparing the Dice coefficient and the degree of sensitivity for pothole detection, it is observed that sensitivity, which is the counterpart of specificity, shows that all the pothole detections results are at above 94% with an average sensitivity of 97.6%. This observation affirms our earlier argument in Ouma and Hahn [17] that sensitivity and specificity metrics tend to overestimate the actual accuracy. From the analysis, the Dice similarity measure is considered to be more optimal for pothole detection accuracy assessment and can be used as an indicator for correlating the accuracy of pothole surface-area extraction. From the validation results, it is evident that the performance of the proposed algorithm is acceptable with an average accuracy of 87.5% using the Dice similarity coefficient.

In general, the pothole defect regions which are most difficult to detect are those that are very shallow, hence present insignificant shadows and are formed without regular or near-regular shapes. Those that are not very shallow and are near-elliptical in shape are detected with accuracies of more than 90%. The misdetections were observed to be concentrated particularly around the edges of the potholes and the image frame or scene. This observation implies that edge detection for potholes is in some scenarios a complicated 2D-image processing task,
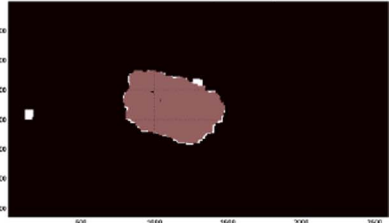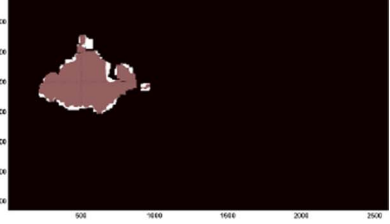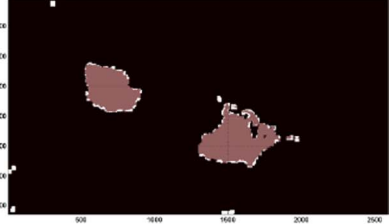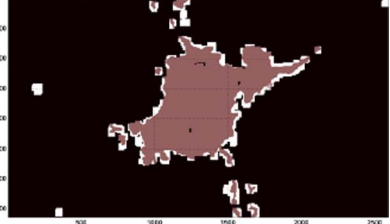
Experimental image-tile reference	Original image-tile	Magnitude of the difference between the detection results and ground-truth before removal of non-pothole blobs
(a) Test image #12		
(b) Test image #15		
(c) Test image #16		
(d) Test image #19		

Fig. 7. Evaluation of pothole detection results comparing original image and the magnitude of the surface-area differences between the detected potholes and the ground-truth.
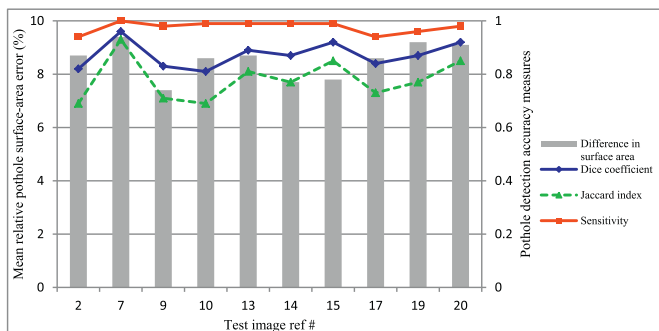

Fig. 8. Relationship between pothole size and shape detection and the corresponding detection accuracies.

which may not be completely resolved by using 2D-RGB imagery alone. This is attributed to the fact that at the pothole edges, there is a high degree of multispectral and multispatial variations which cannot be resolved without a spatial constraint such as additional texture and or depth information.

5.3. Further discussions and analysis

In order to further illustrate and verify the performance of the proposed algorithm, the result of an image frame-tile with pothole is presented in Fig. 9. In Fig. 9(a) the original image in colour and grayscale with the regions marked A, B, C and D for illustrations is presented. Fig. 9(b) shows the results of the multiscale filtering, which is derived from level-3 of the approximation image. As presented and compared in Fig. 9(f), it is empirically evident that at a specific scale the proposed filtering using multiscale wavelets is effective in region smoothing and edge enhancement. For example, in terms of edge preservation and enhancement, the results in Fig. 9(f)(i) depict an intra-smoothing within regions A and B resulting in the preservation of the edge region between the two regions. This is what is envisaged in the empirical hypothesis and is necessary in order to boost the feature identification and clustering using FCM as illustrated in Fig. 3.

Further in Fig. 9f(ii), there is observed intra- and inter-region smoothing which constitutes to the aggregation of features. As evident in Fig. 9(f)(ii), the regions C and D in the original image are constituted by heterogeneous features. However at level-3 of wavelets filtering, the
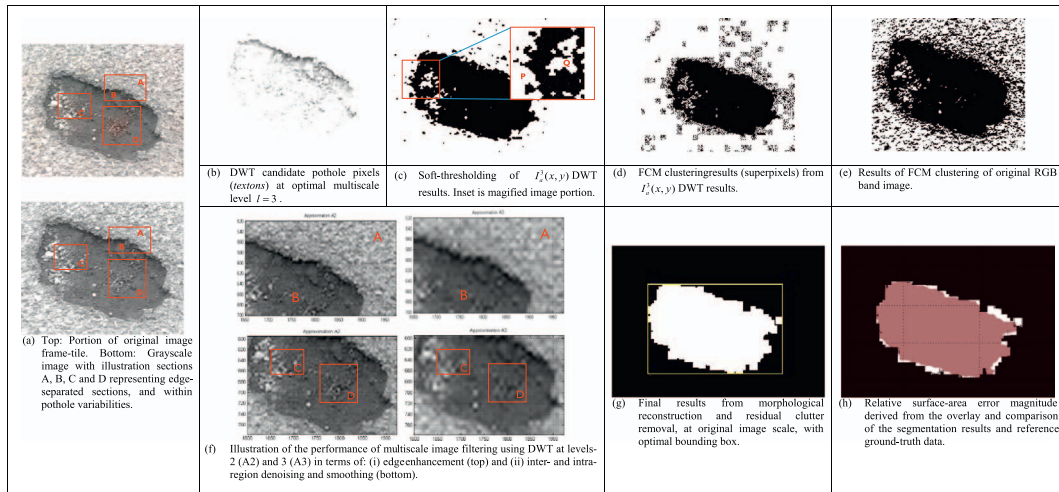
Fig. 9. Illustration of the processing sequence and performance of pothole detection algorithm using the wavelet-FCM clustering and morphological reconstruction.

results in image A3 shows that the regions C and D tend to be smoother and more homogeneous as compared to the original image and the preceding wavelet levels. And as expected, the results at scales above level-3 are much coarser and tend to over-aggregate and oversmoothen the features, resulting into the blurring of information within the regions.

By applying soft-thresholding to the optimal wavelet filtering results, the result in Fig. 9(c) is obtained. The results in Fig. 9(c) comprise of unclustered but indicative information on the presence potholes, with "spill-in" and "spill-out" effects as illustrated in the inset image in regions P and Q. The subsequent results of the hybrid wavelet-FCM clustering are presented in Fig. 9(d). As compared to the three-band FCM clustering in Fig. 9(e), it is observed that the conventional FCM not only has a higher processing run-time, but also presents low quality clustering results as compared to the wavelet-FCM approach. The final pothole segmentation output is presented in Fig. 9(g), which is the product of the morphological reconstruction for extrema extractions from the DWT-FCM results. An optimal bounding box is used to automatically highlight the possible minimum or maximum rectangular area in the displayed frame-tile, so that an end-user can easily estimate the area A_{d_i} for M & R purposes.

A comparative overlay of the ground-truth reference image and the segmentation results (Fig. 9(h)) shows that the pothole is detected with Dice similarity and Jaccard indices of 0.95 and 0.91 respectively, and a sensitivity metric value of 0.99. The results in Fig. 9(h) indicate that there are observed minor errors in the magnitude of surface-areas and shape detections of the pothole, especially around the edges. These as discussed below are attributed to by the "spill-in" and "spill-out" effects, which are contributed to by either over-clustering due to low-spectral resolution and or the lack of spatial information. In overall, the empirical illustrations in Fig. 9 show that the proposed algorithm performs quite well in preserving the two-dimensional geometric properties of pothole such as edges, area and shape and localization in space and position.

In further assessment of the performance of the algorithmic approach, it is important to investigate the results in which the algorithm has failed to produce good results. While the results shows that 2D vision-based imaging is suitable for imaging pothole defects, one of the constraints is that the data capture must be planned and carried out at specific times of the day, weather conditions and or seasons. Poor illuminations without shadow effects will result into lack of sufficient cues for pothole detection as illustrated in Fig. 10. This, coupled with low-spectral resolution camera sensors may result into poor detections and extractions of the actual size and extent of pothole defect.

As shown in Fig. 10(a), the presence of shadow improves the

detection of pothole distress, in such a way that part of the pothole region marked A is detected, while parts of region B are not detected. Empirically, the poor detection of region B is attributed to the fact that the region is only slightly deteriorated as the presence of part of the asphalt layer is still visible. Secondly, the region B is shallower in depth than region A, and hence the shadow effect is minimal. To estimate the optimal area for reconstruction purposes, a bounding box is constructed on the final image as shown in the right image. From Fig. 10(b), it is observed that 2D image data is suitable in detecting the centroid $P_{25}(x,y)$ of the pothole, and can thus be used as a detector and an indicator of presence of a pothole for maintenance and repairs. Thus the results in Fig. 10(a) and (b), though are of slightly lower accuracy, can still serve as indicators in the automated pothole detection process.

With improved illumination and planned data acquisition, the performance of the methodology improves as illustrated in Fig. 10(c), where two potholes are accurately detected in shape and size with Dice similarity and Jaccard index coefficients of 0.89 and 0.76 respectively. The results in Fig. 10(c) not only illustrate and emphasize the significance of shadow based cues in pothole detection in 2D-vision images, but also that combined improvements can be made in terms of edge detection.

The initial processing results of the study, as presented and discussed in Figs. 6–10, shows that the approach of 2D pavement image filtering using wavelets is effective in reducing noise without over-blurring the edges in the image. The difficult part of the implementation is on the setting of the level of signal noise in the wavelet transform by choosing the optimal scale of smoothing. In the acquired pavement images there are almost always areas at the edge of the image that contain no signal, but only noise. These areas can be used to characterize the noise in the image, such that a user chosen percentage of the maximum noise can be subtracted from the wavelet transform of the image. This process may reduce the noise more than it does the signal and may not blur the edges significantly. However, there may be also some reductions in the signal power which can reduce smaller structures in the filtered image and may also add some local artifacts around edges. As such the amount of signal subtracted should be limited.

The observations in Figs. 6–7, where non-pothole areas are also detected after morphological reconstruction can be attributed to the drawbacks in the FCM clustering as discussed in the results section, in which case the marker image, which is the product of FCM clustering, should be contain only true object pixels so that no other regions are reconstructed. Therefore the use of more aggressive thresholds and spatial information may be applied to improve on the accuracy of the results as suggested by Xu et al. [83]. This is because the conventional FCM method clusters pixels only based on the illuminations and does
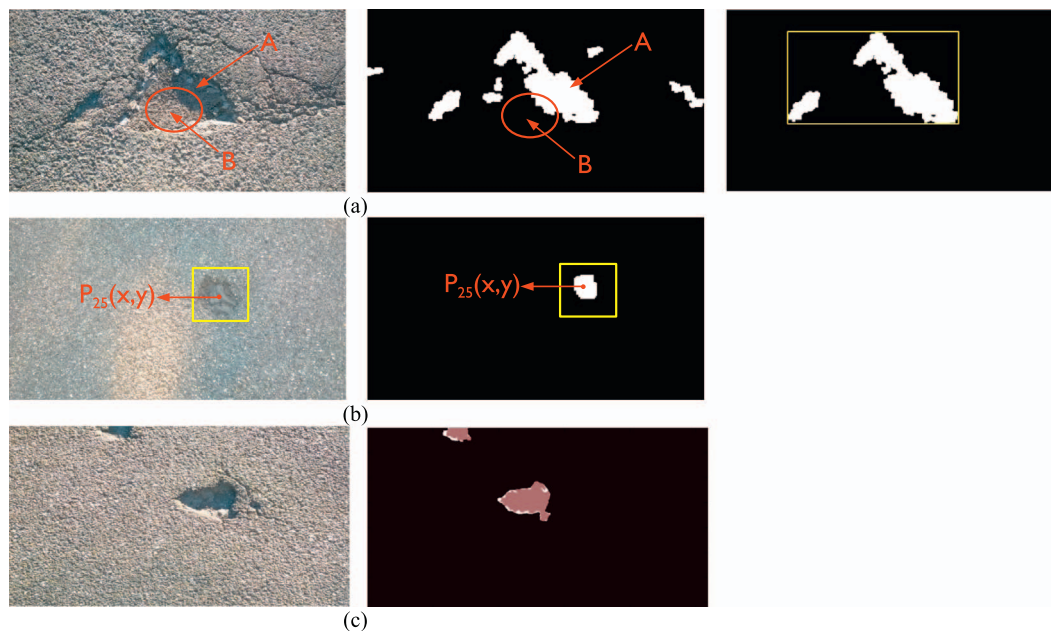
Fig. 10. Pothole detection results under different environmental and pavement surface and conditions and variations such as: shadow, background and illumination and pothole surface spectral and textural differences.

not pay attention to the locations of pixel. This makes the FCM method become highly sensitive to any kind of additive noise, and implies that if there is a noisy pixel in a homogeneous and heterogeneous level, this single pixel is attributed to a different pixel of its adjacent pixels and an unconventional output is realized in the segmentation stage. This observation could have contributed to the results in the study where there is some observed degree of misclassifications, even where the shadow influence was optimal.

While the proposed algorithm performs well in detecting potholes in general, the clustering phase can be further improved by introducing spatial neighborhood weighting especially for edge detection in the FCM algorithm. Furthermore, the image pixels in the immediate neighbors possess relatively similar feature data, therefore, the probability that adjacent pixels belongs to the same cluster will be high as illustrated in Figs. 2 and 3. Since the conventional FCM algorithm does not fully utilize this information, a noisy pixel can be wrongly classified because of its abnormal feature data. This fronts for, besides the *a priori* filtering, the incorporation of spatial information into the fuzzy *c*-means in order to improve on the clustering accuracy results in autonomous pothole detection.

6. Conclusions

The maintenance, restoration and improvement of urban road infrastructure is one the grand global challenges as outlined by the US National Academy of Engineering, the UK Royal Academy of Engineering and the Chinese Academy of Engineering. This study presents an autonomous approach for the improvement of pothole detection based on 2D image processing. The method sequentially comprises of multiscale wavelet transform filtering with the objective of background noise minimization and image smoothing, followed by unsupervised fuzzy *c*-means clustering, and morphological reconstruction for fine tuning the pothole detection and segmentation results in order to extract the pothole defects from the non-defects. Results from 75 experimental datasets show that the approach is much faster than the conventional FCM and accurately estimates the shapes and sizes of the potholes with an average Dice coefficient of similarity of 87.5%. The mean and standard deviations for the errors in the shape and size detections between the ground-truth and the detected potholes were

respectively determined as 8.5% and 4.9%. In overall, the proposed method is considered to be an optimal compromise between accuracy, cost and applicability in pothole detection.

With the high costs of 3D imaging and or reconstruction systems, this study recommends the use of 2D vision-based imaging system for rapid detections of incipient pothole defects on asphalt pavement surfaces. Due to the visual nature of the approach, it is evident that the solution is dependent on lighting conditions, obstructions in the line of view, rain and any other factors that visually enhance or impair the ability to see the target potholes on the pavement surfaces. As such improvements in the algorithm by incorporating spectrometric and spatial information are recommended. This will enable the gathering of detailed spectral information regarding the chemical and mineral properties of the asphalt surface at a higher spatial resolution. There is also need to modify the FCM clustering algorithm so as to take into account additional spatial information. In addition, considerations of low-cost 3D imaging systems needs to be investigated and integrated into the low-cost pothole detection systems, especially in studies where accuracies in depth and volume are necessary such as in distress severity and prioritization analysis for road maintenance and repairs. Finally, for an autonomous pothole detection and extraction system, a method that detects potholes in an arbitrary sequence of pavement imagery should be introduced, *a priori* to the actual detection, verification and quantification of the pothole geometric elements

Acknowledgements

References

[1] R. Medina, J. Gmez-Garca-Bermejo, E. Zalam, Automated visual inspection of road surface cracks, Proceeding of the 27th ISARC, IIT Madras, 25–27 June 2010, Chennai, India, 2010, pp. 155–164.
[2] A. Wimsatt, S. Hurlebaus, T. Scullion, E. Fernando, Promising existing and

emerging technologies and techniques, Proceedings of the International Symposium on Nondestructive Testing for Design Evaluation and Construction Inspection, Washington, DC, USA, 18 January 2008, 2008.

[3] Y.O. Ouma, J. Opudo, S. Nyambenya, Comparison of fuzzy AHP and fuzzy TOPSIS for road pavement maintenance prioritization: methodological exposition and case study, Adv. Civ. Eng. 2015 (2015) 1–17 http://dx.doi.org/10.1155/2015/140189(140189).

[4] S.C. Radopoulou, I. Brilakis, Improving road asset condition monitoring, Transp. Res. Procedia 14 (2016) (2016) 3004–3012 http://dx.doi.org/10.1016/j.trpro.2016.05.436.

[5] S. McRobbie, A. Wright, Crack Detection on Local Roads. TTS Research-Phase I, Transport Research Laboratory, Washington, DC, USA, 2005.

[6] E. Schnebele, B.F. Tanyu, G. Cervone, N. Waters, Review of remote sensing methodologies for pavement management and assessment, Eur. Transp. Res. Rev. 7 (7) (2015) 1–19 http://dx.doi.org/10.1007/s12544-015-0156-6.

[7] E. Buza, S. Omanovic, A. Huseinnovic, Pothole detection with image processing and spectral clustering, Proceedings of the 2nd International Conference on Information Technology and Computer Networks, Antalya, Turkey, 8–10 October 2013, 2013, pp. 48–53.

[8] B.X. Yu, X. Yu, Vibration-Based System for Pavement Condition Evaluation, Proceedings of the 9th International Conference on Applications of Advanced Technology in Transportation Engineering, Chicago, IL, USA, 13–15 August 2006, 2006, pp. 183–189 http://dx.doi.org/10.1061/40799(213)31.

[9] Z. Hou, K. Wang, W. Gong, Experimentation of 3D pavement imaging through stereovision, International Conference on Transportation Engineering, ASCE, Reston, VA, 2007, pp. 376–381 http://dx.doi.org/10.1061/40932(246)62.

[10] Q. Li, M. Yao, X. Yao, B. Xu, A real-time 3D scanning system for pavement distortion inspection, Meas. Sci. Technol. 21 (1) (2010) 015702 http://dx.doi.org/10.1088/0957-0233/21/1/015702.

[11] C. Koch, I. Brilakis, Pothole detection in asphalt pavement images, Adv. Eng. Inform. 25 (3) (2011) 507–515 http://dx.doi.org/10.1016/j.aei.2011.01.002.

[12] E. Salari, G. Bao, Automated Pavement Distress Inspection Based on 2D and 3D Information, IEEE International Conference on Electro/Information Technology (EIT), IEEE, Piscataway, NJ, 2011, pp. 1–4 http://dx.doi.org/10.1109/eit.2011.5978575.

[13] S. Varadharajan, S. Jose, K. Sharma, L. Wander, C. Mertz, Vision for road inspection, IEEE Winter Conference on Applications of Computer Vision (WACV), 2014 2014, pp. 115–122 http://dx.doi.org/10.1109/wacv.2014.6836111.

[14] Y. Jo, S. Ryu, Pothole detection system using a black-box camera, Sensors 15 (2015) 29316–29331 http://dx.doi.org/10.3390/s151129316.

[15] C. Koch, G.M. Jog, I. Brilakis, Automated pothole distress assessment using asphalt pavement video data, J. Comput. Civ. Eng. 27 (4) (2013) 370–378 http://dx.doi.org/10.1061/(ASCE)CP.1943-5487.0000232.

[16] T.S. Nguyen, S. Begot, F. Duculty, J.-C. Bardet, M. Avila, Pavement cracking detection using an anisotropy measurement, IASTED International Conference on Computer Graphics and Imaging (CGIM), Acta Press, Calgary, AB, 2010, pp. 80–87 http://dx.doi.org/10.2316/p.2010.679-079.

[17] Y.O. Ouma, M. Hahn, Wavelet-morphology based detection of incipient linear cracks in asphalt pavements from RGB camera imagery and classification using circular radon transform, Adv. Eng. Inform. 30 (3) (2016) 481–499 http://dx.doi.org/10.1016/j.aei.2016.06.003.

[18] E. Salari, G. Bao, Pavement distress detection and severity analysis, Proc. SPIE 7877 (2011) 78770C http://dx.doi.org/10.1117/12.876724.

[19] K.S. Fu, J.K. Mui, A survey on image segmentation, Pattern Recogn. 13 (1) (1981) 3–16 http://dx.doi.org/10.1016/0031-3203(81)90028-5.

[20] N. Otsu, A threshold selection method from gray-level histogram, IEEE Trans. Syst. Man Cybern. 9 (1) (1979) 62–66 http://dx.doi.org/10.1109/tsmc.1979.4310076.

[21] J.N. Kapur, P.K. Sahoo, A.K.C. Wong, A new method for gray-level picture thresholding using the entropy of the histogram, Comput. Vis. Graphics Image Process. 29 (3) (1985) 273–285 http://dx.doi.org/10.1016/0734-189x(85)90125-2.

[22] Z. Tu, S. Zhu, Image segmentation by data-driven Markov Chain Monte Carlo, IEEE Trans. Pattern Anal. Mach. Intell. 24 (5) (2002) 657–673 http://dx.doi.org/10.1109/34.1000239.

[23] R. Haralick, L. Shapiro, Image segmentation techniques, Computer Vision, Graphics, and Image Processing, 29(1) 1985, pp. 100–132 http://dx.doi.org/10.1016/s0734-189x(5)90153-7.

[24] J. Nayak, B. Naik, H.S. Behera, Fuzzy c-means (FCM) clustering algorithm: a decade review from 2000 to 2014. Computational intelligence in data mining, Spring 2 (2015) 133–149 2015 http://dx.doi.org/10.1007/978-81-322-2208-8_14.

[25] N. Kingsbury, 1999. Image processing with complex wavelets. Philosophical Transactions of the Royal Society A: Mathematical, Phys. Eng. Sci. 357 (1760) 2543–2560 http://dx.doi.org/10.1098/rsta.1999.0447.

[26] S.-L. Jui, L. Chao, G. Haibing, A. Ajith, E.H. Aboul, X. Kai, Fuzzy c-means with wavelet filtration for MR image segmentation, Nature and Biologically Inspired Computing (NaBIC), 2014 Sixth World Congress on, 2014 IEEE, 2014, pp. 12–16 http://dx.doi.org/10.1109/nabic.2014.6921884.

[27] M. Do, M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, IEEE Trans. Image Process. 14 (12) (2005) 2091–2106 http://dx.doi.org/10.1109/tip.2005.859376.

[28] J.C. Nunes, Y. Bouaoune, E. Delechelle, O. Niang, P. Bunel, Image analysis by bidimensional empirical mode decomposition, Image Vis. Comput. 21 (12) (2003) 1019–1026 http://dx.doi.org/10.1016/s0262-8856(03)00094-5.

[29] F. Cong, H. Hautakangas, J. Nieminen, O. Mazhelis, M. Perttunen, J. Riekki, T. Ristaniemi, Applying wavelet packet decomposition and one-class support vector machine on vehicle acceleration traces for road anomaly detection, Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 7951 LNCS (PART 1), 2013, pp. 291–299 http://dx.doi.org/10.1007/978-3-642-39065-4_36.

[30] J. Jang, A. Smyth, Y. Yang, D. Cavalcanti, Road surface condition monitoring via multiple sensor-equipped vehicles, IEEE Conference on Computer Communications Workshops, 2015 2015, pp. 43–44 http://dx.doi.org/10.1109/infcomw.2015.7179334.

[31] R. Madli, S. Hebbar, P. Pattar, V. Golla, Automatic detection and notification of potholes and humps on roads to aid drivers, IEEE Sensors J. 15 (8) (2015) 4313–4318 http://dx.doi.org/10.1109/jsen.2015.2417579.

[32] H.-W. Wang, C.-H. Chen, D.-Y. Cheng, C.-H. Lin, C.-C. Lo, A real-time pothole detection approach for intelligent transportation system, Math. Probl. Eng. 2015 (2015) (2015) 869627(7 pp.), http://dx.doi.org/10.1155/2015/869627.

[33] K. Chen, G. Tan, M. Lu, J. Wu, CRSM: a practical crowdsourcing-based road surface monitoring system, Wirel. Netw 22 (3) (2016) 765–779 http://dx.doi.org/10.1007/s11276-015-0996-y.

[34] K.C.P. Wang, Challenges and feasibility for comprehensive automated survey of pavement conditions, Proceedings of the International Conference on Applications of Advanced Technologies in Transportation Engineering, ASCE, Beijing, China, 2004, pp. 531–536 http://dx.doi.org/10.1061/40730(144)99.

[35] L. Bursanescu, M. Bursanescu, M. Hamdi, A. Lardigue, D. Paiement, Three-dimensional infrared laser vision system for road surface features analysis, Proc. SPIE 4430 (1) (2001) 801–808 http://dx.doi.org/10.1117/12.432808.

[36] National Cooperative Highway Research Program (NCHRP), Automated Pavement Distress Collection Techniques, Transport Research Board of the National Academies, Washington, DC, 2004 https://doi.org/10.17226/23348.

[37] X. Yu, E. Salari, Pavement pothole detection and severity measurement using laser imaging, In Proceedings of the IEEE International Conference on Electro/Information Technology, Mankato, MN, USA, 15–17 May 2011, 2011, pp. 1–5 http://dx.doi.org/10.1109/eit.2011.5978573.

[38] J. Zhou, P.S. Huang, F.-P. Chiang, Wavelet-aided pavement distress image processing, Proc. SPIE 5207 (1) (2003) 728–739 http://dx.doi.org/10.1117/12.509670.

[39] M. Jahanshahi, F. Jazizadeh, S. Masri, B. Becerik-Gerber, Unsupervised approach for autonomous pavement-defect detection and quantification using an inexpensive depth sensor, J. Comput. Civ. Eng. 27 (6) (2013) 743–754 http://dx.doi.org/10.1061/(asce)cp.1943-5487.0000245.

[40] M. Gavilan, D. Balcones, O. Marcos, D.F. Llorca, M.A. Sotelo, I. Parra, M. Ocaña, P. Aliseda, P. Yarza, A. Amirola, Adaptive road crack detection system by pavement classification, Sensors 11 (10) (2011) 9628–9657 http://dx.doi.org/10.3390/s111009628.

[41] G.M. Jog, C. Koch, M. Golparvar-Fard, I. Brilakis, Pothole properties measurement through visual 2D recognition and 3D reconstruction, Proceedings of the ASCE International Conference on Computing in Civil Engineering, Clearwater Beach, FL, USA, 17–20 June 2012, 2012, pp. 553–560 http://dx.doi.org/10.1061/9780784412343.0070.

[42] J. Zhou, P.S. Huang, F.-P. Chiang, Wavelet-based pavement distress detection and evaluation, Opt. Eng. 45 (2) (2006) 027007 http://dx.doi.org/10.1117/1.2172917.

[43] J. Karuppuswamy, V. Selvaraj, M. Ganesh, E. Hall, Detection and avoidance of simulated potholes in autonomous vehicle navigation in an unstructured environment. Intelligent Robots and Computer Vision XIX: Algorithms, Techniques, and Active Vision, Proc. SPIE 4197 (2000) 70–80 http://dx.doi.org/10.1117/12.403788.

[44] J. Lin, Y. Liu, Potholes detection based on SVM in the pavement distress image, 9th International Symposium on Distributed Computing and Applications to Business, Engineering and Science, 2010 IEEE, 2010, pp. 544–547 http://dx.doi.org/10.1109/dcabes.2010.115.

[45] F. Nejad, H. Zakeri, A comparison of multi-resolution methods for detection and isolation of pavement distress, Expert Syst. Appl. 38 (3) (2011) 2857–2872 http://dx.doi.org/10.1016/j.eswa.2010.08.079.

[46] A. Danti, J. Kulkarni, P. Hiremath, An image processing approach to detect lanes, pot holes and recognize road signs in Indian roads, Int. J. Model. Optim. 2 (6) (2012) 658–662 http://dx.doi.org/10.7763/ijmo.2012.v2.204.

[47] S. Murthy, G. Varaprasad, Detection of potholes in autonomous vehicle, IET Intell. Transp. Syst. 8 (6) (2014) 543–549 http://dx.doi.org/10.1049/iet-its.2013.0138.

[48] S.-K. Ryu, T. Kim, Y.-R. Kim, Image-based pothole detection system for ITS service and road management system, Math. Probl. Eng. 2015 (2015) 968361(10 pp.), http://dx.doi.org/10.1155/2015/968361.

[49] S. Radopoulou, I. Brilakis, Detection of multiple road defects for pavement condition assessment, EG-ICE 2015 - 22nd Workshop of the European Group of Intelligent Computing in Engineering, July 13–16, 2015, Eindhoven, The Netherlands, 2015.

[50] S. Radopoulou, I. Brilakis, K. Doycheva, C. Koch, A framework for automated pavement condition monitoring, Construction Research Congress 2016: Old and New Construction Technologies Converge in Historic San Juan - Proceedings of the 2016 Construction Research Congress, CRC 2016, 2016, pp. 770–779 http://dx.doi.org/10.1061/9780784479827.078.

[51] D.L. Pham, C. Xu, J.L. Prince, Current methods in medical image segmentation, Annu. Rev. Biomed. Eng. 2 (2000) (2000) 315–337 http://dx.doi.org/10.1146/annurev.bioeng.2.1.315.

[52] Y.Q. Zhao, Y.L. Zan, X.F. Wang, G.Y. Li, Fuzzy c-means clustering-based multilayer perception neural network for Liver CT images automatic segmentation, Control and Decision Conference (CCDC), Xuzhou, May 2010, 2010, pp. 3423–3427 http://dx.doi.org/10.1109/ccdc.2010.5498558.

[53] R. Xu, D. Wunsch, Survey of clustering algorithms, IEEE Trans. Neural Netw. 16 (3) (2005) 645–678 http://dx.doi.org/10.1109/tnn.2005.845141.

[54] R.N. Dave, R. Krishnapuram, Robust clustering methods: a unified review, IEEE Trans. Fuzzy Syst. 5 (2) (1997) 270–293 http://dx.doi.org/10.1109/91.580801.

[55] K.L. Wu, M.S. Yang, Alternative c-means clustering algorithms, Pattern Recogn. 35 (10) (2002) 2267–2278 http://dx.doi.org/10.1016/s0031-3203(01)00197-2.

[56] Y. Huang, K. Chan, Texture decomposition by harmonics extraction from higher order statistics, IEEE Trans. Image Process. 13 (1) (2004) 1–14 http://dx.doi.org/10.1109/tip.2003.819432.

[57] T.R. Reed, J.M.H. du Buf, A review of recent texture segmentation and feature extraction techniques, CVGIP 57 (3) (1993) 359–372 http://dx.doi.org/10.1006/ciun.1993.1024.

[58] R.M. Haralick, K. Shanmugan, I. Dinstein, Texture features for image classification, IEEE Trans. Syst. Man Cybern. 3 (3) (1973) 610–621 http://dx.doi.org/10.1109/tsmc.1973.4309314.

[59] A. Gagalowicz, A new method for texture field synthesis: some applications to the study of human vision, IEEE Trans. Pattern Anal. Mach. Intell. 3 (5) (1982) 520–533 http://dx.doi.org/10.1109/tpami.1981.4767145.

[60] G.C. Cross, A.K. Jain, Markov random field texture models, IEEE Trans. Pattern Anal. Mach. Intell. 5 (1) (1983) 25–39 http://dx.doi.org/10.1109/tpami.1983.4767341.

[61] A.P. Pentland, Fractal-based description of natural scenes, IEEE Trans. Pattern Anal. Mach. Intell. 6 (6) (1984) 661–674 http://dx.doi.org/10.1109/tpami.1984.4767591.

[62] F. Liu, R.W. Picard, Periodicity, directionality, and randomness: Wold features for image modeling and retrieval, IEEE Trans. Pattern Anal. Mach. Intell. 18 (7) (1996) 722–733 http://dx.doi.org/10.1109/34.506794.

[63] J. Beck, A. Sutter, R. Ivry, Spatial frequency channels and perceptual grouping in texture segregation, Comput. Vis. Graphics Image Process. 37 (2) (1987) 299–325 http://dx.doi.org/10.1016/S0734-189X(87)80006-3.

[64] A.K. Jain, Learning texture discrimination masks, IEEE Trans. Pattern Anal. Mach. Intell. 18 (2) (1996) 195–205 http://dx.doi.org/10.1109/34.481543.

[65] M. Unser, M. Eden, Multiresolution feature extraction and selection for texture segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 11 (7) (1989) 717–728 http://dx.doi.org/10.1109/34.192466.

[66] I. Daubechies, Ten Lectures on Wavelets, 1992 Capital City Press, Montpellier, Vermont, 1992http://dx.doi.org/10.1137/1.9781611970104.

[67] S. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Anal. Mach. Intell. 11 (7) (1989) 674–693 http://dx.doi.org/10.1109/34.192463.

[68] S. Liapis, E. Sifakis, G. Tziritas, Colour and texture segmentation using wavelet frame analysis, deterministic relaxation, and fast marching algorithms, J. Vis. Commun. Image Represent. 15 (1) (2004) 1–26 http://dx.doi.org/10.1016/s1047-3203(03)00025-7.

[69] S. Mallat, A Wavelet Tour of Signal Processing, 2nd ed., Academic Press, New York, 1999http://dx.doi.org/10.1016/b978-0-12-374370-1.x0001-8.

[70] D.L. Donoho, I.M. Johnstone, Adapting to unknown smoothness via wavelet shrinkage, J. Am. Stat. Assoc. 90 (432) (1995) 1200–1224 http://dx.doi.org/10.1080/01621459.1995.10476626.

[71] R. Ganesan, T.K. Das, V. Venkataraman, Wavelet-based multiscale statistical process monitoring: a literature review, IIE Trans. 36 (9) (2004) 787–806 http://dx.doi.org/10.1080/07408170490473060.

[72] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, 1981 Kluwer Academic Publishers, 1981, http://dx.doi.org/10.1007/978-1-4757-0450-1.

[73] S.S. Reddi, S.F. Rudin, H.R. Keshavan, An optimal multiple threshold scheme for image segmentation, IEEE Trans. Syst. Man Cybern. (1984) 661–665 SMC-14 http://dx.doi.org/10.1109/tsmc.1984.6313341.

[74] V.J. De Oliveira, W. Pedrycz, Advances in Fuzzy Clustering and its Applications, John Wiley & Sons, Ltd, 2007, http://dx.doi.org/10.1002/9780470061190.

[75] K.S. Tan, N.A. Mat Isa, W.H. Lim, Color image segmentation using adaptive unsupervised clustering approach, Appl. Soft Comput. 13 (4) (2013) 2017–2036 http://dx.doi.org/10.1016/j.asoc.2012.11.038.

[76] L. Vincent, Morphological grayscale reconstruction in image analysis: applications and efficient algorithms, IEEE Trans. Image Process. 2 (2) (1993) 176–201 http://dx.doi.org/10.1109/83.217222.

[77] A. Sopharak, B. Uyyanonvara, S. Barman, S. Vongkittirux, N. Wongkamchang, Fine exudate detection using morphological reconstruction enhancement, Int. J. Appl. Biomed. Eng. 3 (1) (2010) 45–50.

[78] L. Vincent, Morphological grayscale reconstruction: definition, efficient algorithm and applications in image analysis, Proc. IEEE Conf. on Comp. Vision and Pattern Recognition, Champaign IL, June 1992, 1992, pp. 633–635 http://dx.doi.org/10.1109/cvpr.1992.223122.

[79] G. Teodoro, T. Pan, T.M. Kurc, J. Kong, L.A.D. Cooper, J.H. Saltz, Efficient irregular wavefront propagation algorithms on hybrid CPU-GPU machines, Parallel Comput. 39 (4–5) (2013) 189–211 http://dx.doi.org/10.1016/j.parco.2013.03.001.

[80] R. Unnikrishnan, C. Pantofaru, M. Hebert, A measure for objective evaluation of image segmentation algorithms, Proceedings of the 2005 IEEE Conference on Computer Vision and Pattern Recognition (CVPR '05), Workshop on Empirical Evaluation Methods in Computer Vision, 2005, pp. 34–41 http://dx.doi.org/10.1109/cvpr.2005.390.

[81] J. Miller, Y. Bellinger, Distress Identification Manual for the Long-term Pavement Performance Program. Report No. FHW-RD-03-031, 4th Edition, June 2003, (2003).

[82] F. Crete, T. Dolmiere, P. Ladret, M. Nicolas, The blur effect: perception and estimation with a new no-reference perceptual blur metric, SPIE Proc. on Human Vision and Electronic Imaging XII, 6492: 64920I. San Jose, California, USA, January 2007, 2007 http://dx.doi.org/10.1117/12.702790.

[83] L.-Q. Xu, J.L. Landabaso, M. Pardas, Shadow removal with blob-based morphological reconstruction for error correction, IEEE International Conference on Acoustics, Speech, and Signal Processing, March 18–23, 2 2005, pp. 729–732 http://dx.doi.org/10.1109/icassp.2005.1415508.

[84] S. Chambon, J.M. Moliard, Automatic road pavement assessment with image processing: review and comparison, Int. J. Geosci. 2011 (2011), http://dx.doi.org/10.1155/2011/98935 (Article ID 989354).

[85] Y. Tsai, V. Kaul, R.M. Mersereau, Critical assessment of pavement distress segmentation methods, J. Transp. Eng. 136 (1) (2010) 11–19, http://dx.doi.org/10.1061/(ASCE)TE.1943-5436.0000051.

[86] H.-Y. Ying, X.-J. Zhang, X.-Y. Wang, LS-SVM-based image segmentation using pixel color-texture descriptors, Pattern. Anal. Applic. 17 (2) (2014) 341–359, http://dx.doi.org/10.1007/s10044-012-0302-x.