Surface-type Classification Using RGB-D

Andrew Wing Keung To, Gavin Paul and Dikai Liu

Abstract—This paper proposes an approach to improve surface-type classification of images containing inconsistently illuminated surfaces. When a mobile inspection robot is visually inspecting surface-types in a dark environment and a directional light source is used to illuminate the surfaces, the images captured may exhibit illumination variance that can be caused by the orientation and distance of the light source relative to the surfaces. In order to accurately classify the surface-types in these images, either the training image dataset needs to completely incorporate the illumination variance or a way to extract colour features that can provide high classification accuracy needs to be identified. In this paper diffused reflectance values are extracted as new colour features to classifying surface-types. In this approach, RGB-D data is collected from the environment, and a reflectance model is used to calculate a diffused reflectance value for a pixel in each Red, Green, Blue (RGB) colour channel. The diffused reflectance values can be used to train a multi-class support vector machine classifier to classify surface-types. Experiments are conducted in a mock bridge maintenance environment using a portable RGB-Depth (RGB-D) sensor package with an attached light source to collect surface-type data. The performance of a classifier trained with diffused reflectance values is compared against classifiers trained with other colour features including RGB and L*a*b* colour spaces. Results show that the classifier trained with the diffused reflectance values can achieve consistently higher classification accuracy than the classifiers trained with RGB and L*a*b* features. For test images containing a single surface plane, diffused reflectance values consistently provide greater than 90% classification accuracy; and for test images containing a complex scene with multiple surface-types and surface planes, diffused reflectance values are shown to provide an increase in overall accuracy over RGB and L*a*b* by 49.24% and 13.66%, respectively.

Note to Practitioners: This paper was motivated by the problem of inspecting inconsistently illuminated steel surfaces on a bridge structure using a robot manipulator. Existing approaches for colour-based surface classification are susceptible to illumination variance. This paper proposes the use of diffused reflectance values, which combines the use of colour and depth data to improve accuracy. In this approach, the diffused reflectance values of each image pixel are calculated by using the distance and angle between the surface represented by a pixel and the light source. The diffused reflectance values are calculated in each colour channel (Red, Green, Blue) to provide three features to classify different surface-types. This proposed approach can be applied to surface classification tasks where the light source does not uniformly illuminate the scene in the image.

Index Terms—Surface-type classification, Lambertian diffused reflectance, RGB-D

A. W. K. To, G. Paul and D. K. Liu are with the Centre for Autonomous Systems, University of Technology, Sydney, e-mail: (andrew.to@student.uts.edu.au).

I. INTRODUCTION

For a bridge maintenance robot manipulator (as shown in Fig. 1) tasked to inspect for surface rust and grit-blasting finish quality [1] [2] [3], the use of a machine vision system may require the additional setup of lighting infrastructure to operate successfully. Lighting infrastructure can either be mounted directly onto the robot manipulator or placed in carefully selected locations within the environment. The main advantage of having carefully positioned light source(s) within the environment is that the scene can be consistently lit. However, the inconvenience of adding and manually configuring lighting whenever the robot is repositioned within the bridge makes this option impractical, particularly if the environment is structurally complex. For practical use, the light source can be mounted onto the end-effector of the manipulator. But as a result, surfaces may appear inconsistently illuminated due to the robot manipulator operating in a confined work space, where the light source cannot be ideally positioned or orientated because of obstacles and robot arm constraints.
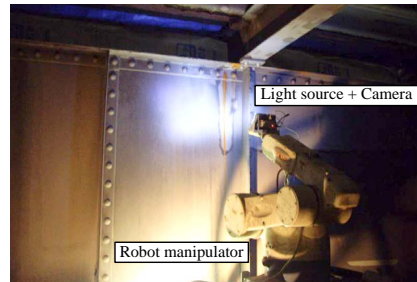

Fig. 1. Bridge maintenance robot manipulator with directional light source and camera mounted to the end-effector

Illumination inconsistency can lead to the appearance change of surface-types described by texture and/or colour features. This can cause variance in the extracted features and adversely affect the classification accuracy if the training data set does not incorporate the variance. Therefore, it is important to consider the feature variance caused by illumination inconsistency, otherwise the classifier will fail to accurately predict the classes. This paper proposes an approach to improve surface-type classification accuracy of images that contain illumination variance. The approach assumes the use of a single light source to illuminate the surfaces in the environment, such that the relative position of the light source to the camera can be estimated using a near-field light reflectance model. Based on the reflectance model, the diffused (Lambertian) reflectance values for each image pixel can be calculated and used as colour features to classify different surface-types on inconsistently illuminated surfaces.

Machine vision systems have been successfully demonstrated in factory environments to assess quality and detect defects in manufacturing processes such as marble slabs [4] and steel sheets [5] using image analysis and classification approaches. Typically, image classification is performed using uniformly illuminated images, and demonstrated applications include classifying insect wings [6], rust on sign posts [7], and aerial images [8]. But in order to accurately classify different surface-types in an environment where illumination consistency cannot be guaranteed, the implemented approaches must take into account illumination variance. For images of a single surface, invariance to illumination inconsistency has been demonstrated using local binary patterns (LBP) to extract greyscale invariant texture [9], and using L*a*b* colour space to extract colour consistent features that are tolerant to different lighting directions and intensities [10]. For images containing multiple surfaces, image adjustment techniques based on reflectance models have been proposed to directly reduce the illumination variance. For example, light attenuation adjustment is applied to images captured underwater to improve colour consistency in image stitching [11], and to provide better colour matching with images captured in a clean air environment [12]. The use of reflectance characteristics to classify different surface-types has been demonstrated in [13] and [14] based on bidirectional reflectance distribution function (brdf) parameters. The brdf parameters of a surface are estimated by illuminating the surface with multiple lights at different calibrated positions. Currently, this approach can be used for a small surface piece that is positioned within a specially designed non-portable lighting array [15].

In this paper, we propose an approach to estimate the diffused reflectance values of each pixel from an image using a reflectance model, and RGB-D data. The advantage of the proposed approach is improved portability and the capability to classify larger surface areas in an indoor environment that has low or no ambient lighting. A portable RGB-D sensor package with a light source has been designed and implemented to practically demonstrate the use of diffused reflectance values as colour features for classifying surface-types.

This paper is organised as follows. Section II provides the proposed surface-type classification approach describing the reflectance model, the process to estimate the light source position, the technique to estimate the diffused reflectance values, the process to extract alternative colour features for comparison, and the training and classifying of features using a multi-class Support Vector Machine (SVM) classifier. Section III presents the results from experiments conducted in a laboratory environment. Section IV discusses the results and Section V provides a conclusion.

II. METHODOLOGY

The proposed approach to classify surface-types using RGB-D data is shown in Fig. 2. In order to estimate the diffused reflectance values, a reflectance model is defined

and the position of the light source is calibrated with the RGB-D sensor package. Calibration RGB-D data is collected to estimate the camera-to-light source position. Upon completing calibration, the diffused reflectance values of pixels can be calculated and use to train a SVM classifier to classify surface-types. In addition, RGB and L*a*b colour space features are also extracted to train SVM classifiers to provide a comparison with the proposed diffused reflectance values. The classification results produced by each classifier trained with different colour features are compared to show the relative classification performance achieved.
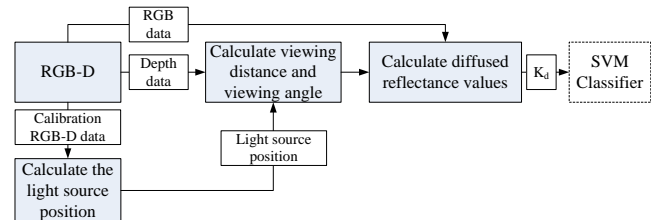

Fig. 2. RGB-D surface-type classification approach overview

A. Reflectance Model

A reflectance model defines parameters that describe the irradiance characteristics of light when reflected off a surface including attenuation, specularity and diffusion. These reflectance parameters are independent of the light source and are proposed as the features to classify surface-types. In order to estimate the diffused reflectance values of each image pixel, a reflectance model is defined and the parameters calibrated for the light source and camera setup. An overview of the implemented reflectance model is discussed in this section and the estimation of the model parameters based on [16] are presented henceforth in the subsequent two sections.

Light reflected off a surface consists of the diffused reflection and the specular reflection. The diffused reflection describes the light rays that are scattered within the surface, and the specular reflection describes the light rays directly reflected from the surface. The Torrance-Sparrow reflection model [17] describes the reflected light intensity off a surface point as a combination of both diffused and specular reflections. In this model, the Fresnel reflectance is assumed to be constant and the geometric attenuation factor for a clean air environment is 1. Thus, the reflected light intensity off a surface point is expressed as:

$$I_r = \frac{I_s}{r^2}(k_d \cos\theta_i + \frac{k_s}{\cos\theta_r}\exp(-\frac{\alpha^2}{2\sigma^2})). \qquad (1)$$

Fig. 3 illustrates the geometric model of reflection, where the light source intensity is I_s, surface-to-camera distance is l, light source-to-surface point distance is r, reflected light intensity off the surface is I_r, the light source direction vector is $\vec{L}$, the surface normal vector is $\vec{N}$, the viewing direction vector is $\vec{V}$, the angle between the light source direction vector and the surface normal vector is θ_i, the

angle between the viewing direction vector and the surface normal is θ_r, the angle between the surface normal with the bisector of the viewing direction vector and the light source direction vector is α, the surface roughness is σ, the position vectors for a surface point is $\vec{P}$, and the light source point is $\vec{S}$.
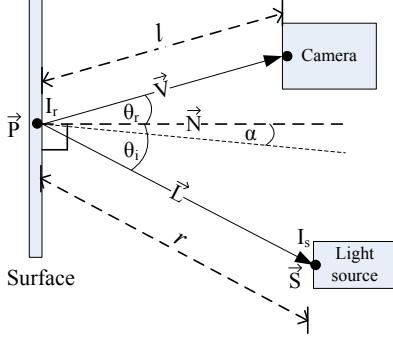

Fig. 3.   Geometric model of light reflection

The reflectance ratios (k_d, k_s) and the light source intensity (I_s) in (1) cannot be simultaneously estimated given only the observation of I_r. Therefore, a new reflectance value is defined for the combination of the reflectance ratio and light source intensity, as:

$$K_d = k_d I_s, \ K_s = k_s I_s \qquad (2)$$

and we rewrite (1) as:

$$I_r = \frac{1}{r^2}(K_d \cos\theta_i + \frac{K_s}{\cos\theta_r}\exp(-\frac{\alpha^2}{2\sigma^2})). \qquad (3)$$

It is important to highlight that the diffused and specular reflectance values are unique for each colour component, (Red, Green, Blue) where $(K_{s,R}, K_{s,G}, K_{s,B}) \in K_s$ and $(K_{d,R}, K_{d,G}, K_{d,B}) \in K_d$. However, the same process is used to estimate the reflectance value for each individual colour component. Therefore, for simplicity, the equations are presented using the set notations, (K_d, K_s). Furthermore, with the near-field use of the light source, there is a high probability that the specular reflectance regions in an image contain no useful information due to intensity readings being saturated to their maximum upper limit. Thus, we only estimate the diffused reflectance value, K_d, and the distance between the light source and the surface, r. The reflectance model equation is simplified to:

$$I_r = \frac{1}{r^2}(K_d \cos\theta_i). \qquad (4)$$

B. Light Source Position Estimation

In order to calculate r for an image pixel in the training and test images, the surface point $\vec{P}$ for the pixel and $\vec{S}$ are used (Fig. 3). The surface point $\vec{P}$, can be provided from the depth data and the light source point $\vec{S}$, which is a fixed position in the camera coordinate frame is estimated by analysing both the diffused and specular reflectance components from a calibration RGB-D data. In our estimation process we assume the calibration image: (a) contains a planar surface that consists of a single surface-type; (b) shows a surface illuminated by a near-field light source with no appreciable ambient lighting; and (c) contains a specular component and light fall-out described by a Lambertian diffuse reflectance of constant albedo.

Fig. 4a shows a greyscale example of a calibration image Ω, used to estimate the light source position. Fig. 4b shows a binary image that contains the maximum intensity region of the image identified by applying an intensity threshold, T_s, to the calibration image. The image pixel at the centroid position of the saturated region is estimated to satisfies $\alpha = 0$ and the corresponding surface point can be denoted as $\vec{P}_p$. The unit vector $\vec{L}_p$, in the direction from the surface point $\vec{P}_p$ to light source point $\vec{S}$ can be written as [16]:

$$\vec{L}_p = 2(\vec{N}_p \cdot \vec{V}_p)\vec{N}_p - \vec{V}_p \qquad (5)$$

where $\vec{N}_p$ is the unit normal vector at surface point $\vec{P}_p$, and $\vec{V}_p$ is the unit vector in the viewing direction from the surface point $\vec{P}_p$ to camera coordinate frame origin. The depth data for the calibration image (Fig. 4a) is used to produce a 3D point cloud from which a surface plane is identified using A RANSAC plane fitting algorithm [18]. The normal of the identified surface plane is taken as $\vec{N}_p$, and $\vec{V}_p$ can be calculated using the surface point $\vec{P}_p$ provide by the 3D point cloud. The fixed position of the light source point relative to the camera, $\vec{S}$, can be expressed as:

$$\vec{S} = \vec{P}_p + r_p \vec{L}_p \qquad (6)$$

where r_p is the distance between $\vec{P}_p$ and $\vec{S}$. The value of r_p is estimated by using the diffused region of the calibration image shown in Fig. 4c. The diffused region Ω_d, is extracted by identify the image pixels (i,j) in the calibration image Ω, that have intensity values I_r, below an intensity threshold T_r, as,

$$\Omega_d = \{(i,j) \in \Omega | I_r(i,j) < T_r\}. \qquad (7)$$

Least-square fitting of the diffused reflectance model of (4) and Ω_d is performed to estimate r_p.

$$r_p = \underset{t}{argmin} \sum_{(i,j)\in\Omega_d}\left(u(i,j,t) - \frac{1}{N_d}\sum_{(i,j)\in\Omega_d} u(i,j,t)\right)^2 \qquad (8)$$

where,

$$u(i,j,t) = \frac{I_r(i,j)r(i,j,t)^2}{cos\theta_i(i,j,t)}, (i,j) \in \Omega_d. \qquad (9)$$

The number of pixel points in the diffused region set is N_d, the optimisation parameter to find the value of r_p that best fits the diffused reflectance model to the diffused region in the image is t, the intensity value at a pixel point is $I_r(i,j)$, the distance between the light source and a surface point is $r(i,j,t)$, and the light source-to-surface incident angle is $\theta_i(i,j,t)$. Note that T_r and T_s are selected empirically and are dependent on the surface roughness, σ, of the calibration surface-type.

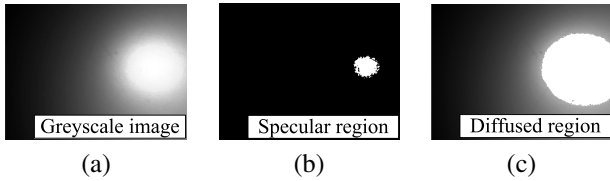| Greyscale image | Specular region | Diffused region |
| (a) | (b) | (c) |

Fig. 4. *a)* Greyscale of calibration image; *b)* Specular reflectance region in calibration image; *c)* Diffused reflectance component in calibration image

C. Diffused Reflectance Estimation

Diffused reflectance values are estimated from the RGB-D data to train and classify different surface-types. The diffused reflectance value $K_d(i,j)$ for an image pixel is expressed as:

$$K_d(i,j) = \frac{I_r(i,j)r(i,j)^2}{\cos\theta_i(i,j)} \qquad (10)$$

where the light source-to-surface point angle of incidence is $\theta_i(i,j)$, the light source-to-surface point distance is $r(i,j)$, and the reflectance intensity is $I_r(i,j)$. In order to determine $\theta_i(i,j)$, the surface normal, $\vec{N}(i,j)$, of the 3D surface point corresponding to the image pixel is estimated by analysing the eigenvectors and eigenvalues of a covariance matrix constructed from a k-neighbourhood of surrounding 3D surface points [19]. Fig. 5 shows the estimated surface normal vector, $\vec{N}(i,j)$, from which $\theta_i(i,j)$ is calculated as:

$$\theta_i(i,j) = cos^{-1} \frac{\vec{L}(i,j) \cdot \vec{N}(i,j)}{||\vec{L}(i,j)||||\vec{N}(i,j)||}. \qquad (11)$$

The light source-to-surface point distance, $r(i,j)$, is expressed as the magnitude of the light source direction vector $L(i,j)$, as:

$$r(i,j) = ||\vec{L}(i,j)||. \qquad (12)$$

The reflectance intensity for an image pixel, $I_r(i,j)$, is extracted from each colour component of the RGB image. The outcome of diffused reflectance estimation is a set of three values $(K_{s,R}, K_{s,G}, K_{s,B})$ for each pixel that can be used as the feature vector to classify surface-types.
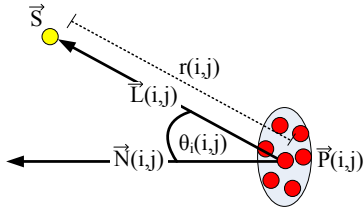

Fig. 5. Estimated surface normal vector for a surface point

D. Extraction of Colour-based Features

This section describes the extraction of colour-based features, which can be used as a performance comparison with the proposed diffuse reflectance features. The RGB intensity values have been successfully shown to classify rust from simple background surfaces [20]. However, illumination variance significantly reduces classification

accuracy when using RGB intensity values as features. Therefore, the L*a*b* colour space has been proposed in [21][10] to improve classification accuracy for images with illumination variance. The L*a*b* colour space decouples illumination into a luminance component and thus shows higher robustness towards illumination variance. The L* component corresponds to lightness ranging from black to white (0–100), the a* component measures green-red (negative-positive), and the b* measures blue-yellow (negative-positive). To calculate the L*a*b* components, the tristimulus values (X, Y, Z) are used [22].

$$L^* = 116f\left(\frac{Y}{Y0}\right) - 16 \qquad (13)$$

$$a^* = 500\left[f\left(\frac{X}{X0}\right) - \left(\frac{Y}{Y0}\right)\right] \qquad (14)$$

$$b^* = 200\left[f\left(\frac{Y}{Y0}\right) - \left(\frac{Z}{Z0}\right)\right] \qquad (15)$$

$$\text{where} f(x) = \begin{cases} x^{\frac{1}{3}} & \text{, if } x > 0.008856 \\ 7.787x + \frac{16}{116} & \text{, otherwise} \end{cases} \qquad (16)$$

For this paper, the RGB and the L*a*b* colour space are used to provide a comparison with the proposed diffused reflectance features. An RGB-based feature vector is produced directly from the intensity values of each RGB colour component to demonstrate the effect of varying illumination on classification accuracy. A second feature vector is produced using the colour component a* and b* from the L*a*b* colour space. The feature vector using a* and b* is intended to demonstrate improved tolerance to illumination variance over the RGB-based feature vector.

E. Multi-class Support Vector Machines

In order to classify features into a set of known classes, a supervised classification technique is applied. The support vector machine (SVM) is a popular classifier and has been applied in various image recognition tasks [7][23]. It has been shown to robustly perform (two-class) classification using a limited training data set. However, a multi-class SVM classifier is complicated to formulate and can be intractable. To achieve multi-class classification with SVM, the approach of combining several binary SVM classifiers is used. Several approaches are available [24], including: one-versus-all method using winner-takes-all strategy (WTA-SVM), one-versus-one method using max-wins voting (MWV-SVM), and error correcting output codes with pairwise coupling by combining posterior probabilities of individual SVMs PWC-PSVM. An empirical study comparing the different approaches concludes that WTA-SVM and MWV-SVM are competitive with no clear superiority, while PWC-PSVM consistently outperforms these two methods.

For this paper, we have chosen to use an implementation of PWC-PSVM described in [25] to perform pixel-wise classification of different surface-types in an image. Three separate multi-class SVM classifiers are trained using the feature vectors; $(K_{d,R}, K_{d,G}, K_{d,B})$, (R, G, B) and (a^*, b^*). A limited training data set has been intentionally

used to train the classifier. The dataset does not represent the whole range of lighting variance for the surface-type classes. Therefore, no data scaling is applied to the dataset before training. The classification outputs from each classifier are compared to provide a performance comparison between the feature vectors.

III. EXPERIMENTS AND RESULTS

A. Experiment Setup

The experiments presented in this paper uses an RGB-D sensor package as shown in Fig. 6; consisting of a Microsoft Kinect, a Pointgrey firefly camera with a 6 mm lens, and a 160 LED array light source. Depth images are collected from the Kinect and 640×480 pixel RGB images are collected using the firefly camera with fixed settings for: gain, brightness, aperture, shutter speed, and white balance. Extrinsic calibration is performed between the Kinect and the firefly camera to obtain a homogeneous transform matrix between the two cameras' coordinate frames.

Fig. 7 shows the experimental environment of a mock bridge channel structure, which contains surface-types including: painted surface, blasted surface, rusted surface and timber surface. From this environment, the training images used to estimate the reflectance model parameters are collected by positioning the RGB-D sensor package to capture images containing a single surface plane. Subsequent test images are from the environment using different viewpoints. It is assumed that during the process of capturing images, the light source intensity remains constant.
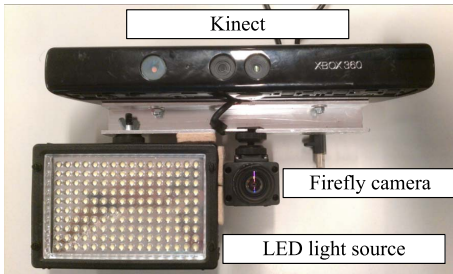

Fig. 6. Sensor package: Firefly camera, Kinect, and LED light array

B. Experiment 1

In this experiment, the classifier trained with the proposed diffuse reflectance values, K_d, is compared against the classifiers trained with the RGB and a*b* features. Test images of the four different surface-types are collected from the mock bridge structure. Fig. 8 shows how each test image contains a flat plane of a single surface-type that is captured under the illumination of a near-field LED light source. Note that manual contrast enhancement has been applied to the rusted and blasted surface test images in Fig. 8 to clearly show the different surface-types for this paper; the actual classification is performed on the original unadjusted images. The training image for each surface-type is captured at a viewing distance of 500 mm
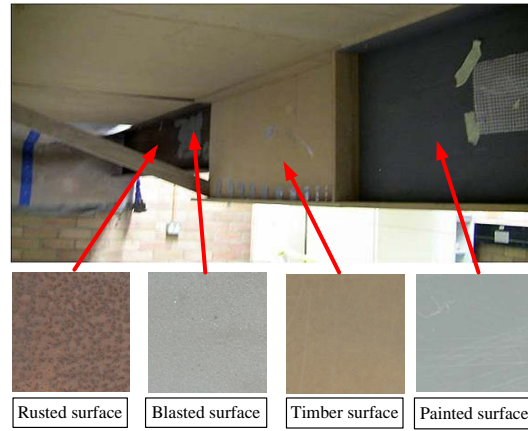

Fig. 7. Experimental environment: mock bridge channel structure

and a centre pixel viewing angle of incidence of $0°$. A subset of pixels from the training images is extracted by taking every 10th pixel along the rows and columns. This subset is then used to train the classifiers. Three additional test images are collected for each surface-type at viewing distances ranging between 500–1000 mm, and viewing angles ranging between 0–45 degrees angle of incidence between the centre pixel ray-cast to the surface normal.

The classification results produced by the three different classifiers are compared to each other. Manual cropping of the test images is necessary to omit regions with no available depth data. Fig. 9 shows the results of the K_d-based classifier exhibiting a misclassified band (yellow) on the lower portion of the images due to incomplete depth data for the RGB image. Fig. 8 shows that cropping is also necessary in test Samples 2 and 3 for the rusted and blasted surfaces. This cropping removes unrelated surface-types in the image since there is inadequate surface area with which to fill the whole image frame. Table I shows the classification accuracy results for each image. Training data resubstitution validation shows a classification accuracy of greater than 90% for K_d and a*b*. Overall across all test images, K_d demonstrated the highest accuracy with greater than 96% accuracy for all images. This was followed by the a*b*-based classifier, which also demonstrated good performance, and finally the RGB-based classifier showed the poorest accuracy due to the effect of lighting variance in the images.

In order to analyse the effect of illumination variance on classification accuracy, Fig. 9 provides a visualisation of the painted surface classification results. The colour scheme used in this figure is: blue = painted surface, teal = timber surface, yellow = rusted surface, and red = blasted surface. From the figure, it can be observed that the classification accuracy deteriorates progressively from the Training sample through to Sample 3, which corresponds to the increase in illumination variance. This deterioration is most prominent in Sample 3 for the RGB features, which shows a large region of misclassification on the left side of the image where the light fall-off is most significant. The results for K_d and a*b* also have the same mis-

classification characteristics as RGB, albeit less significant. The classification results of the other surface-types is observed to have similar misclassification characteristics due to illumination variance, particularly for the RGB feature vector. Furthermore, the effect of illumination variance on feature inconsistency can be observed by comparing the value distributions of the R colour component and the $K_{d,r}$ for the samples of painted surface-type. For the R colour component, the value distributions of the Training sample and Sample 3 have mean values of 50.89 and 27.67, and standard deviations of 19.85 and 14.94, respectively. For $K_{d,r}$, the value distribution of the Training sample and Sample 3 has mean values of 4.6 and 3.7, and standard deviations of 2.13 and 1.31, respectively. This result shows that there is less variance in the value distribution for $K_{d,r}$ between the Training sample and Sample 3. Therefore, improved classification accuracy is seen for the classifier trained with a K_d feature vector.
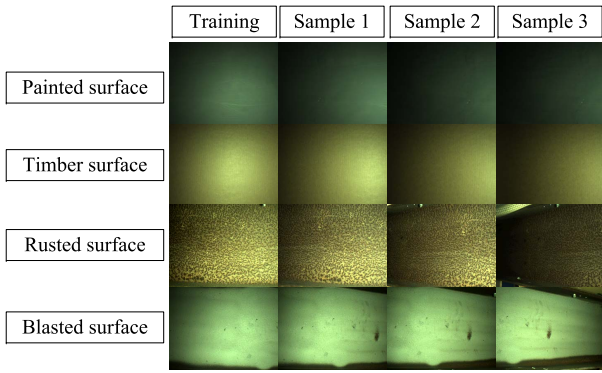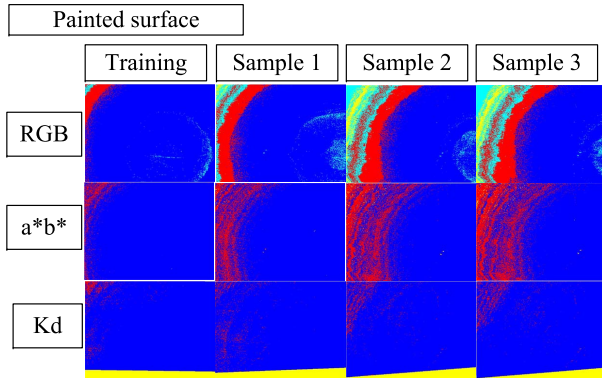

Fig. 8.   Test samples for Experiment 1

Fig. 9.   Painted surface classification results

C. Experiment 2

In this experiment, the three classifiers are demonstrated on images containing multiple surface-types and surface planes. Similar to Experiment 1, the same training images are used to extract the K_d, RGB and a*b* feature vectors to train the classifiers. Fig. 10*a* shows a test image (contrast enhanced for viewing purposes) classified in this experiment. Fig. 10*b* shows the depth image containing

TABLE I
CLASSIFICATION RESULTS OF TEST SAMPLES FOR EXPERIMENT 1

	Accuracy %		
	RGB	**a* b***	K_d
Painted surface			
Training	93.90%	96.21%	97.70%
Sample 1	77.27%	87.57%	97.19%
Sample 2	62.43%	78.13%	96.90%
Sample 3	59.34%	76.06%	96.93%
Timber surface			
Training	98.70%	96.88%	99.85%
Sample 1	93.08%	87.93%	99.96%
Sample 2	84.85%	81.27%	99.89%
Sample 3	83.15%	81.24%	99.98%
Rusted surface			
Training	54.03%	96.00%	98.98%
Sample 1	46.34%	96.355%	99.58%
Sample 2	30.12%	95.89%	98.91%
Sample 3	15.99%	95.89%	96.87%
Blasted surface			
Training	76.14%	92.27%	96.41%
Sample 1	85.82%	95.61%	98.25%
Sample 2	80.12%	95.60%	97.94%
Sample 3	56.51%	91.93%	98.36%

the two surface planes that are classified, Fig. 10*c* shows the 3D point cloud generated from the depth image, and Fig. 10*d* shows the estimated surface normals used for the calculation of the reflectance values, K_d.

Fig. 11 shows the visualisation of the classification results for the image using the same colour scheme used in Experiment 1. Table II shows the confusion matrices of the classification results for the image. The confusion matrices show that the K_d-based classifier was most successful in classifying the surface-types with the highest accuracy across all surface-type classes. The lowest classification accuracy achieved by the K_d-based classifier was observed for the timber surface (teal). From Fig. 11, it is shown that the K_d-based classifier misclassified part of the timber surface (teal) as blasted surface (red), resulting in an accuracy of 45.14% for the timber surface-type. Comparatively, RGB and a*b* classifiers both had higher misclassification of the timber surface with accuracies of 10.69% and 0.04%, respectively. Furthermore, Table III provides the average accuracy of three additional test images collected at different viewpoints within the multiple surface-types and surface plane environment. The results in the table indicates that on average, K_d has the highest accuracy followed by a*b* and finally RGB.

IV. DISCUSSION

The results presented in this paper have demonstrated that applying depth data to estimate the Lambertian diffused reflectance values of each image pixel can provide colour features for improved pixel-wise surface-type classification
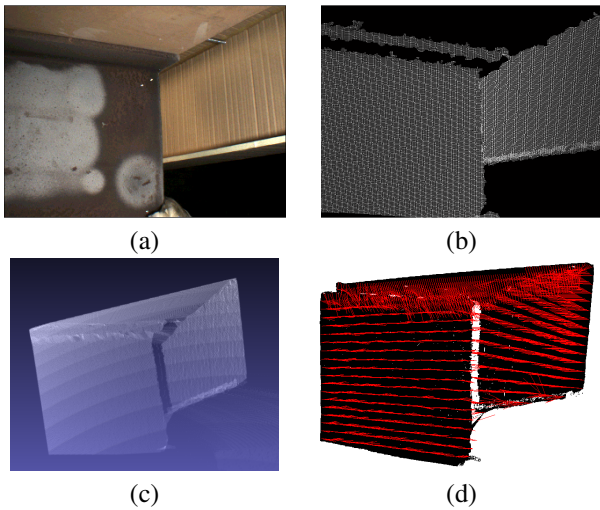
Fig. 10.   *a)* Experiment 2 test image with multiple surface-types; *b)* Depth data; *c)* 3D point cloud; *d)* Estimated surface normals for surface points
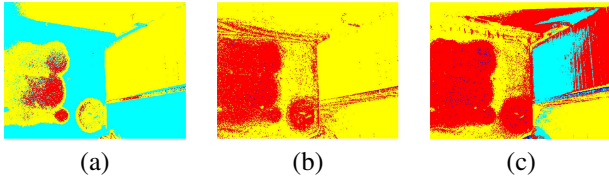

Fig. 11.   *a)* RGB classification results; *b)* a*b* classification results; *c)* K_d classification results

TABLE II
CONFUSION MATRICES FOR EXPERIMENT 2 SINGLE TEST IMAGE

Confusion matrix: RGB					
	Painted	Timber	Rusted	Blasted	Acc%
Painted	0	0	0	0	n/a
Timber	2	8823	72769	882	10.69%
Rusted	4	59313	269	2	0.45%
Blasted	0	8002	46969	14119	20.43%
Confusion matrix: a* b*					
	Painted	Timber	Rusted	Blasted	Acc%
Painted	0	0	0	0	n/a
Timber	187	34	78860	3395	0.04%
Rusted	34	3	47194	12357	79.20%
Blasted	428	0	5925	62737	90.80%
Confusion matrix: K_d					
	Painted	Timber	Rusted	Blasted	Acc%
Painted	0	0	0	0	n/a
Timber	1044	37248	7433	36751	45.16%
Rusted	34	21	47831	11702	80.26%
Blasted	1476	0	486	67128	97.16%

TABLE III
AVERAGE SURFACE-TYPE CLASSIFICATION RESULTS FOR MULTIPLE TEST IMAGES

	RGB	**a* b***	K_d
Average Timber Accuracy	12.23%	0.026%	44.05%
Average Rusted Accuracy	14.13%	80.45%	82.54%
Average Blasted Accuracy	54.43%	91.17%	96.17%
Overall Accuracy	28.79%	64.37%	78.03%

the approach for an image captured in a practical situation of a bridge maintenance environment. Both experiments showed an improvement in classification accuracy when using a classifier trained with K_d reflectance values in comparison to classifiers trained with RGB and a* b* colour space features.

The multi-class SVM was chosen to perform surface-type classification given its strength of handling non-linear and high-dimensional problems. However, in an application domain such as per pixel surface-type classification where a large sample set is processed, the overall performance benefits of the SVM classifier are reduced. Alternative classifiers suited towards a large number of sample instances that can be considered in the future for improving time efficiency includes a linear-based strategy such as liblinear, or a deterministic approach such as naive Bayes.

In comparison to a colour conversion approach that requires only the intensity values of a pixel to perform calculations, the presented approach requires intensity values and additional geometric parameters in order to perform calculations. The time required to estimate the necessary geometric parameters for each pixel including $\vec{N}(i,j)$, $\theta_i(i,j)$ and $r(i,j)$ is currently a limitation for real-time applications. To classify the 640x480 images that were used in the experiments, an additional 1–1.5 mins of computation time was required to generate the K_d values for all pixels in the image. Classification time using an SVM classifier typically required 1 min to generate results. Therefore, depending on the latency time for an application, the presented approach may be suitable for online surface-type classification.

The proposed approach is limited to the classification of planar surfaces with Lambertian reflectance that can be illuminated using a light source. Surface-types that will not work with the proposed approach include transparent surfaces, highly reflective surfaces and non-reflective surfaces. Calculating accurate surface normals for non-planar surfaces is still a challenging task for a cluttered scene and in the presence of sensor noise. A potential alternative strategy is the use of prior surface geometry knowledge of the environment to perform surface template matching. By fitting the depth data to an existing library of potential surface shapes, an accurate distance and surface normal estimation is possible for each pixel point on the non-planar surfaces.

of images with illumination variance. This paper has extended on the idea of using reflectance properties to perform surface-type classification. It has been demonstrated that it is feasible to use RGB-D data that is collected from a portable sensor package to improve surface-type classification. Experiment 1 demonstrated the approach for simple images with single surfaces and Experiment 2 demonstrated

V. Conclusion

In this paper, an approach to improve surface-type classification of images containing illumination variance caused by a single light source has been presented. The approach uses data collected from a portable RGB-D sensor package to perform reflectance model calibration and to estimate the diffused reflectance values of each pixel as classification features. The diffused reflectance values are tested against two other colour-based feature vectors (RGB and a* b*) trained with a multi-class SVM classifier. Experimental results showed that the diffused reflectance values provided the highest performance, maintaining accuracy greater than 90% for test images in Experiment 1.

The experimental results of this paper have highlighted the difficulty of accurately classifying a complex scene completely from a single observation. Therefore, it is necessary to extend this work into a probabilistic exploration framework that updates the surface-type belief of an environment through repeat observations from different viewpoints. This future work will also investigate optimising the number of observations required to confidently classify the surface-types in an environment. Further experiments will be performed in a field environment to test more surface-types.

Acknowledgment

References

[1] A. W. K. To, G. Paul, and D. K. Liu, "Image segmentation for surface material-type classification using 3D geometry information," *International Conference on Information and Automation*, pp. 1717 – 1722, Harbin, 2010.

[2] G. Paul, S. Webb, D. K. Liu, and G. Dissanayake, "A Robotic System for Steel Bridge Maintenance: Field Testing," *Australasian Conference on Robotics and Automation*, pp. 1 – 8, Brisbane, Australia, 2010.

[3] G. Paul, N. Kirchner, D. K. Liu, and G. Dissanayake, "An effective exploration approach to simultaneous mapping and surface material-type identification of complex three-dimensional environments," *Journal of Field Robotics*, vol. 26, pp. 915 – 933, 2009.

[4] J. Martinez-Alajarin, J. D. Luis-Delgado, and L. M. Tomas-Balibrea, "Automatic system for quality-based classification of marble textures," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 35, pp. 488 – 497, 2005.

[5] M. Yazdchi, M. Yazdi, and A. G. Mahyari, "Steel Surface Defect Detection Using Texture Segmentation Based on Multifractal Dimension," *International Conference on Digital Image Processing*, pp. 346 – 350, Bangkok, 2009.

[6] L. Q. Zhu and Z. Zhang, "Auto-classification of insect images based on color histogram and GLCM," *International Conference Fuzzy Systems and Knowledge Discovery*, pp. 2589 – 2593, Yantai, Shandong, 2010.

[7] E. Momma, Y. Kimura, H. Ishii, T. Ono, M. Ayoja, and T. Higuchi, "Rust classification using image analysis of steel structures," *International Conference on Sensing Technology*, pp. 409 – 413, Tainan, Taiwan, 2008.

[8] R. M. Haralick, K. Shanmugam, and H. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, pp. 610 – 621, 1973.

[9] Y. He and N. Sang, "Robust Illumination Invariant Texture Classification Using Gradient Local Binary Patterns," *International Workshop on Multi-Platform/Multi-Sensor Remote Sensing and Mapping*, pp. 1 – 6, Xiamen, 2011.

[10] P. H. Chen, H. K. Shen, C. Y. Lei, and L. M. Chang, "Support-vector-machine-based method for automated steel bridge rust assessment," *Automation in Construction*, vol. 23, pp. 9 – 19, 2012.

[11] M. Bryson, M. J. Roberson, O. Pizarro, and S. B. Williams, "Colour-Consistent Structure-from-Motion Models using Underwater Imagery," *Robotics: Science and Systems*, pp. 1 – 8, Sydney, Australia, 2012.

[12] A. Yamashita, M. Fujii, and T. Kaneko, "Color Registration of Underwater Images for Underwater Sensing with Consideration of Light Attenuation," *International Conference on Robotics and Automation*, pp. 4570 – 4575, Roma, Italy, 2007.

[13] O. Wang, P. Gunawardane, S. Scher, and J. Davis, "Material classification using BRDF slices," *International Conference on Computer Vision and Pattern Recognition*, pp. 2805 – 2811, Miami, 2009.

[14] M. Louw and F. Nicolls, "Surface Classification Via BRDF Parameters using Population Monte Carlo for MRF Parameter Estimation," *International Conference on Computer Graphics and Imaging*, pp. 145 – 154, Innsbruck, Austria, 2010.

[15] P. Debevec, A. Wenger, C. Tchou, A. Gardner, J. Waese, and T. Hawkins, "A lighting reproduction approach to live-action compositing," *ACM Transactions on Graphics*, vol. 21, no. 3, pp. 547 – 556, 2002.

[16] K. Hara, K. Nishino, and K. lkeuchi, "Light source position and reflectance estimation from a single view without the distant illumination assumption," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 493 – 505, 2005.

[17] K. E. Torrance and E. M. Sparrow, "Theory of off-specular reflection from roughened surfaces," *Journal of the Optical Society of America*, pp. 1105 – 1114, 1967.

[18] M. Y. Yang and W. Forstner, "Plane detection in point cloud data," *Technical Report, Department of Photogrammetry*, pp. 1 – 16, 2010.

[19] R. Radu, "Semantic 3d object maps for everyday manipulation in human living environments," *KI - Künstliche Intelligenz*, vol. 24, pp. 345–348, 2010.

[20] S. Lee, L. M. Chang, and M. Skibniewski, "Automated recognition of surface defects using digital color image processing," *Automation in Construction*, vol. 15, no. 4, pp. 540 – 549, 2006.

[21] P. H. Chen, Y. C. Yang, and L. M. Chang, "Automated bridge coating defect recognition using adaptive ellipse approach," *Automation in Construction*, vol. 18, no. 5, pp. 632 – 643, 2009.

[22] R. C. Gonzalez and R. E. Woods, "Digital Image Processing," *2nd Ed. Prentice Hall*, 2001.

[23] J. Dong, Y. Duan, and Z. Yang, "Three-Dimensional Surface Texture Classification Based on Support Vector Machines and Wavelet Packets," *International Symposium on Intelligent Information Technology Application*, vol. 3, pp. 124 – 127, 2008.

[24] K. B. Duan and S. S. Keerthi, "Which is the best multiclass SVM method? An empirical study," *International Workshop on Multiple Classifier Systems*, pp. 278 – 285, 2005.

[25] T. K. Huang, R. C. Weng, and C. J. Lin, "Generalized Bradley-Terry Models and Multi-Class Probability Estimates," *Journal of Machine Learning Research*, vol. 7, pp. 85 – 115, 2006.

Andrew Wing Keung To graduated with a B.Eng. in Computer Systems Engineering at the University of Technology, Sydney (UTS). Andrew is currently a PhD candidate in Engineering (Robotics) at UTS.

Gavin Paul graduated with a B.Eng. in Computer Systems Engineering at the University of Technology, Sydney (UTS). In 2010 he received his PhD in Engineering (Robotics) at UTS.

Dikai Liu is a Professor at the University of Technology, Sydney (UTS). His main research interests are field robotics and physical human-robot interaction.