5

Fusion of Images Recorded with Variable Illumination

Luis Nachtigall and Fernando Puente León
Karlsruhe Institute of Technology
Germany

Ana Pérez Grassi
Technische Universität München
Germany

1. Introduction

The results of an automated visual inspection (AVI) system depend strongly on the image acquisition procedure. In particular, the illumination plays a key role for the success of the following image processing steps. The choice of an appropriate illumination is especially critical when imaging 3D textures. In this case, 3D or depth information about a surface can be recovered by combining 2D images generated under varying lighting conditions. For this kind of surfaces, diffuse illumination can lead to a destructive superposition of light and shadows resulting in an irreversible loss of topographic information. For this reason, directional illumination is better suited to inspect 3D textures. However, this kind of textures exhibits a different appearance under varying illumination directions. In consequence, the surface information captured in an image can drastically change when the position of the light source varies. The effect of the illumination direction on the image information has been analyzed in several works [Barsky & Petrou (2007); Chantler et al. (2002); Ho et al. (2006)]. The changing appearance of a texture under different illumination directions makes its inspection and classification difficult. However, these appearance changes can be used to improve the knowledge about the texture or, more precisely, about its topographic characteristics. Therefore, series of images generated by varying the direction of the incident light between successive captures can be used for inspecting 3D textured surfaces. The main challenge arising with the variable illumination imaging approach is the fusion of the recorded images needed to extract the relevant information for inspection purposes.
This chapter deals with the fusion of image series recorded using variable illumination direction. Next section presents a short overview of related work, which is particularly focused on the well-known technique photometric stereo. As detailed in Section 2, photometric stereo allows to recover the surface albedo and topography from a series of images. However, this method and its extensions present some restrictions, which make them inappropriate for some problems like those discussed later. Section 3 introduces the imaging strategy on which the proposed techniques rely, while Section 4 provides some general information fusion concepts and terminology. Three novel approaches addressing the stated information fusion problem

are described in Section 5. These approaches have been selected to cover a wide spectrum of fusion strategies, which can be divided into model-based, statistical and filter-based methods. The performance of each approach are demonstrated with concrete automated visual inspection tasks. Finally, some concluding remarks are presented.

2. Overview of related work

The characterization of 3D textures typically involves the reconstruction of the surface topography or profile. A well-known technique to estimate a surface topography is photometric stereo. This method uses an image series recorded with variable illumination to reconstruct both the surface topography and the albedo [Woodham (1980)]. In its original formulation, under the restricting assumptions of Lambertian reflectance, uniform albedo and known position of distant point light sources, this method aims to determine the surface normal orientation and the albedo at each point of the surface. The minimal number of images necessary to recover the topography depends on the assumed surface reflection model. For instance, Lambertian surfaces require at least three images to be reconstructed. Photometric stereo has been extended to other situations, including non-uniform albedo, distributed light sources and non-Lambertian surfaces. Based on photometric stereo, many analysis and classification approaches for 3D textures have been presented [Drbohlav & Chantler (2005); McGunnigle (1998); McGunnigle & Chantler (2000); Penirschke et al. (2002)].
The main drawback of this technique is that the reflectance properties of the surface have to be known or assumed a priori and represented in a so-called reflectance map. Moreover, methods based on reflectance maps assume a surface with consistent reflection characteristics. This is, however, not the case for many surfaces. In fact, if location-dependent reflection properties are expected to be utilized for surface segmentation, methods based on reflectance maps fail [Lindner (2009)].
The reconstruction of an arbitrary surface profile may require demanding computational efforts. A dense sampling of the illumination space is also usually required, depending on the assumed reflectance model. In some cases, the estimation of the surface topography is not the goal, e.g., for surface segmentation or defect detection tasks. Thus, reconstructing the surface profile is often neither necessary nor efficient. In these cases, however, an analogous imaging strategy can be considered: the illumination direction is systematically varied with the aim of recording image series containing relevant surface information. The recorded images are then fused in order to extract useful features for a subsequent segmentation or classification step. The difference to photometric stereo and other similar techniques, which estimate the surface normal direction at each point, is that no surface topography reconstruction has to be explicitly performed. Instead, symbolic results, such as segmentation and classification results, are generated in a more direct way. In [Beyerer & Puente León (2005); Heizmann & Beyerer (2005); Lindner (2009); Pérez Grassi et al. (2008); Puente León (2001; 2002; 2006)] several image fusion approaches are described, which do not rely on an explicit estimation of the surface topography. It is worth mentioning that photometric stereo is a general technique, while some of the methods described in the cited works are problem-specific.

3. Variable illumination: extending the 2D image space

The choice of a suitable illumination configuration is one of the key aspects for the success of any subsequent image processing task. Directional illumination performed by a distant point light source generally yields a higher contrast than multidirectional illumination pat-

terns, more specifically, than diffuse lighting. In this sense, a variable directional illumination strategy presents an optimal framework for surface inspection purposes.
The imaging system presented in the following is characterized by a fixed camera position with its optical axis parallel to the z-axis of a global Cartesian coordinate system. The camera lens is assumed to perform an orthographic projection. The illumination space is defined as the space of all possible illumination directions, which are completely defined by two angles: the azimuth φ and the elevation angle θ; see Fig. 1.

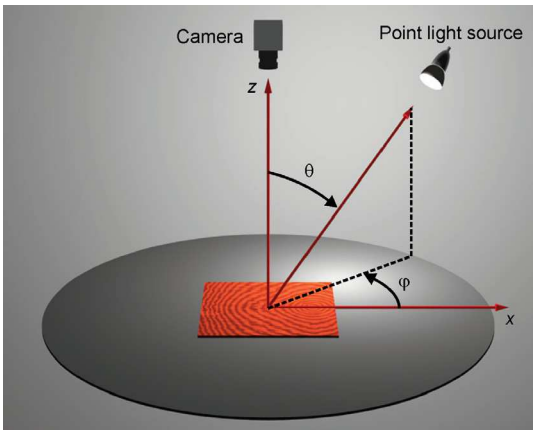

Fig. 1. Imaging system with variable illuminant direction.

An illumination series $\mathcal{S}$ is defined as a set of B images $g(\mathbf{x}, \mathbf{b}_b)$, where each image shows the same surface part, but under a different illumination direction given by the parameter vector $\mathbf{b}_b = (\varphi_b, \theta_b)^{\mathrm{T}}$:

$$\mathcal{S} = \{g(\mathbf{x}, \mathbf{b}_b), \quad b = 1, \ldots, B\}, \tag{1}$$

with $\mathbf{x} = (x, y)^{\mathrm{T}} \in \mathbb{R}^2$. The illuminant positions selected to generate a series $\{\mathbf{b}_b,\ b = 1, \ldots, B\}$ represent a discrete subset of the illumination space. In this sense, the acquisition of an image series can be viewed as the sampling of the illumination space.
Beside point light sources, illumination patterns can also be considered to generate illumination series. The term illumination pattern refers here to a superposition of point light sources. One approach described in Section 5 uses sector-shaped patterns to illuminate the surface simultaneously from all elevation angles in the interval $\theta \in [0°, 90°]$ given an arbitrary azimuth angle; see Fig. 2. In this case, we refer to a sector series $\mathcal{S}_{\mathrm{s}} = \{g(\mathbf{x}, \varphi_b),\ b = 1, \ldots, B\}$ as an image series in which only the azimuthal position of the sector-shaped illumination pattern varies.

4. Classification of fusion approaches for image series

According to [Dasarathy (1997)] fusion approaches can be categorized in various different ways by taking into account different viewpoints like: application, sensor type and information hierarchy. From an application perspective we can consider both the application area and its final objective. The most commonly referenced areas are: defense, robotics, medicine and space. According to the final objective, the approaches can be divided into detection, recognition, classification and tracking, among others. From another perspective, the fusion

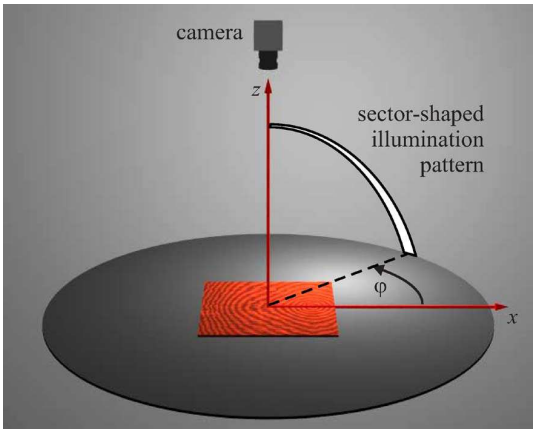

Fig. 2. Sector-shaped illumination pattern.

approaches can be classified according to the utilized sensor type into passive, active and a mix of both (passive/active). Additionally, the sensor configuration can be divided into parallel or serial. If the fusion approaches are analyzed by considering the nature of the sensors' information, they can be grouped into recurrent, complementary or cooperative. Finally, if the hierarchies of the input and output data classes (data, feature or decision) are considered, the fusion methods can be divided into different architectures: data input-data output (DAI-DAO), data input-feature output (DAI-FEO), feature input-feature output (FEI-FEO), feature input-decision output (FEI-DEO) and decision input-decision output (DEI-DEO). The described categorizations are the most frequently encountered in the literature. Table 1 shows the fusion categories according to the described viewpoints. The shaded boxes indicate those image fusion categories covered by the approaches presented in this chapter.

Application		Sensor		Data	
Application Domain	Fusion Objective	Sensor Type	Sensor Suite Configuration	Sensor Data	I/O Based Characterization
Defense	Detection	Active	Parallel	Recurrent	DAI-DAO
Robotics	Recognition	Passive	Serial	Complementary	DAI-FEO
Medical	Classification	Active/passive	-	Cooperative	FEI-FEO
Space	Tracking	-	-	-	FEI-DEO
-	-	-	-	-	DEI-DEO

Table 1. Common fusion classification scheme. The shaded boxes indicate the categories covered by the image fusion approaches treated in the chapter.

This chapter is dedicated to the fusion of images series in the field of automated visual inspection of 3D textured surfaces. Therefore, from the viewpoint of the application area, the approaches presented in the next section can be assigned to the field of robotics. The objectives of the machine vision tasks are the detection and classification of defects. Now, if we analyze the approaches considering the sensor type, we find that the specific sensor, i.e., the camera, is

a passive sensor. However, the whole measurement system presented in the previous section can be regarded as active, if we consider the targeted excitation of the object to be inspected by the directional lighting. Additionally, the acquisition system comprises only one camera, which captures the images of the series sequentially after systematically varying the illumination configuration. Therefore, we can speak here about serial virtual sensors.
More interesting conclusions can be found when analyzing the approaches from the point of view of the involved data. To reliably classify defects on 3D textures, it is necessary to consider all the information distributed along the image series simultaneously. Each image in the series contributes to the final decision with a necessary part of information. That is, we are fusing cooperative information. Now, if we consider the hierarchy of the input and output data classes, we can globally classify each of the fusion methods in this chapter as DAI-DEO approaches. Here, the input is always an image series and the output is always a symbolic result (segmentation or classification). However, a deeper analysis allows us to decompose each approach into a concatenation of DAI-FEO, FEI-FEO and FEI-DEO fusion architectures. Schemes showing these information processing flows will be discussed for each method in the corresponding sections.

5. Multi-image fusion methods

A 3D profile reconstruction of a surface can be computationally demanding. For specific cases, where the final goal is not to obtain the surface topography, application-oriented solutions can be more efficient. Additionally, as mentioned before, traditional photometric stereo techniques are not suitable to segment surfaces with location-dependent reflection properties. In this section, we discuss three approaches to segment, detect and classify defects by fusing illumination series. Each method relies on a different fusion strategy:

- Model-based method: In Section 5.1 a reflectance model-based method for surface segmentation is presented. This approach differs from related works in that reflection model parameters are applied as features [Lindner (2009)]. These features provide good results even with simple linear classifiers. The method performance is shown with an AVI example: the segmentation of a metallic surface. Moreover, the use of reflection properties and local surface normals as features is a general purpose approach, which can be applied, for instance, to defect detection tasks.
- Filter-based method: An interesting and challenging problem is the detection of topographic defects on textured surfaces like varnished wood. This problem is particularly difficult to solve due to the noisy background given by the texture. A way to tackle this issue is using filter-based methods [Xie (2008)], which rely on filter banks to extract features from the images. Different filter types are commonly used for this task, for example, wavelets [Lambert & Bock (1997)] and Gabor functions [Tsai & Wu (2000)]. The main drawback of the mentioned techniques is that appropriate filter parameters for optimal results have to be chosen manually. A way to overcome this problem is to use Independent Component Analysis (ICA) to construct or learn filters from the data [Tsai et al. (2006)]. In this case, the ICA filters are adapted to the characteristics of the inspected image and no manual selection of parameters are required. An extension of ICA for feature extraction from illumination series is presented in [Nachtigall & Puente León (2009)]. Section 5.2 describes an approach based on ICA filters and illumination series which allows a separation of texture and defects. The performance of this

method is demonstrated in Section 5.2.5 with an AVI application: the segmentation of varnish defects on a wood board.

- Statistical method: An alternative approach to detecting topographic defects on textured surfaces relies on statistical properties. Statistical texture analysis methods measure the spatial distribution of pixel values. These are well rooted in the computer vision world and have been extensively applied to various problems. A large number of statistical texture features have been proposed ranging from first order to higher order statistics. Among others, histogram statistics, co-occurrence matrices, and Local Binary Patterns (LBP) have been applied to AVI problems [Xie (2008)]. Section 5.3 presents a method to extract invariant features from illumination series. This approach goes beyond the defect detection task by also classifying the defect type. The detection and classification performance of the method is shown on varnished wood surfaces.

5.1 Model-based fusion for surface segmentation

The objective of a segmentation process is to separate or segment a surface into disjoint regions, each of which is characterized by specific features or properties. Such features can be, for instance, the local orientation, the color, or the local reflectance properties, as well as neighborhood relations in the spatial domain. Standard segmentation methods on single images assign each pixel to a certain segment according to a defined feature. In the simplest case, this feature is the gray value (or color value) of a single pixel. However, the information contained in a single pixel is limited. Therefore, more complex segmentation algorithms derive features from neighborhood relations like mean gray value or local variance.
This section presents a method to perform segmentation based on illumination series (like those described in Section 3). Such an illumination series contains information about the radiance of the surface as a function of the illumination direction [Haralick & Shapiro (1992); Lindner & Puente León (2006); Puente León (1997)]. Moreover, the image series provides an illumination-dependent signal for each location on the surface given by:

$$g_{\mathbf{x}}(\mathbf{b}) = g(\mathbf{x}, \mathbf{b}) \,, \tag{2}$$

where $g_{\mathbf{x}}(\mathbf{b})$ is the intensity signal at a fixed location $\mathbf{x}$ as a function of the illumination parameters $\mathbf{b}$. This signal allows us to derive a set of model-based features, which are extracted individually at each location on the surface and are independent of the surrounding locations. The features considered in the following method are related to the macrostructure (the local orientation) and to reflection properties associated with the microstructure of the surface.

5.1.1 Reflection model

The reflection properties of the surface are estimated using the Torrance and Sparrow model, which is suitable for a wide range of materials [Torrance & Sparrow (1967)]. Each measured intensity signal $g_{\mathbf{x}}(\mathbf{b})$ allows a pixel-wise data fit to the model. The reflected radiance L_{r} detected by the camera is assumed to be a superposition of a diffuse lobe L_{d} and a forescatter lobe L_{fs}:

$$L_{\mathrm{r}} = k_{\mathrm{d}} \cdot L_{\mathrm{d}} + k_{\mathrm{fs}} \cdot L_{\mathrm{fs}}. \tag{3}$$

The parameters k_{d} and k_{fs} denote the strength of both terms. The diffuse reflection is modeled by Lambert's cosine law and only depends on the angle of incident light on the surface:

$$L_{\mathrm{d}} = k_{\mathrm{d}} \cdot \cos(\theta - \theta_{\mathrm{n}}). \tag{4}$$

The assignment of the variables θ (angle of the incident light) and θ_n (angle of the normal vector orientation) is explained in Fig. 3.

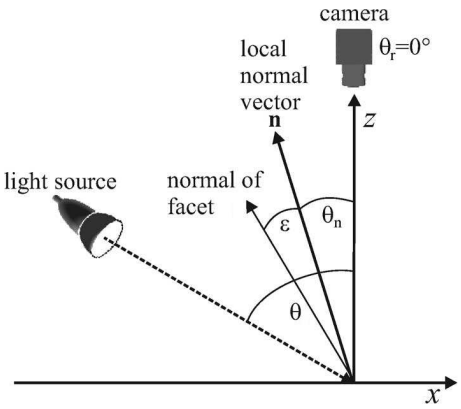

Fig. 3. Illumination direction, direction of observation, and local surface normal **n** are in-plane for the applied 1D case of the reflection model. The facet, which reflects the incident light into the camera, is tilted by ε with respect to the normal of the local surface spot.

The forescatter reflection is described by a geometric model according to [Torrance & Sparrow (1967)]. The surface is considered to be composed of many microscopic facets, whose normal vectors diverge from the local normal vector **n** by the angle ε; see Fig. 3. These facets are normally distributed and each one reflects the incident light like a perfect mirror. As the surface is assumed to be isotropic, the facets distribution function $p_\varepsilon(\varepsilon)$ results rotationally symmetric:

$$p_\varepsilon(\varepsilon) = c \cdot \exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right). \tag{5}$$

We define a surface spot as the surface area which is mapped onto a pixel of the sensor. The reflected radiance of such spots with the orientation θ_n can now be expressed as a function of the incident light angle θ:

$$L_{fs} = \frac{k_{fs}}{\cos(\theta_r - \theta_n)} \exp\left(-\frac{(\theta + \theta_r - 2\theta_n)^2}{8\sigma^2}\right). \tag{6}$$

The parameter σ denotes the standard deviation of the facets' deflection, and it is used as a feature to describe the degree of specularity of the surface. The observation direction of the camera θ_r is constant for an image series and is typically set to 0°. Further effects of the original facet model of Torrance and Sparrow, such as shadowing effects between the facets, are not considered or simplified in the constant factor k_{fs}.
The reflected radiance L_r leads to an irradiance reaching the image sensor. For constant small solid angles, it can be assumed that the radiance L_r is proportional to the intensities detected by the camera:

$$g_{\mathbf{x}}(\theta) \propto L_r(\theta). \tag{7}$$

Considering Eqs. (3)-(7), we can formulate our model for the intensity signals detected by the camera as follows:

$$g_{\mathbf{x}}(\theta) = k_{\mathrm{d}} \cdot \cos(\theta - \theta_{\mathrm{n}}) + \frac{k_{\mathrm{fs}}}{\cos(\theta_{\mathrm{r}} - \theta_{\mathrm{n}})} \exp\left(-\frac{(\theta + \theta_{\mathrm{r}} - 2\theta_{\mathrm{n}})^2}{8\sigma^2}\right). \tag{8}$$

This equation will be subsequently utilized to model the intensity of a small surface area (or spot) as a function of the illumination direction.

5.1.2 Feature extraction

The parameters related to the reflection model in Eq. (8) can be extracted as follows:

- First, we need to determine the azimuthal orientation $\phi(\mathbf{x})$ of each surface spot given by $\mathbf{x}$. With this purpose, a sector series $\mathcal{S}_{\mathrm{s}} = \{g(\mathbf{x}, \varphi_b),\ b = 1, \ldots, B\}$ as described in Section 3 is generated. The azimuthal orientation $\phi(\mathbf{x})$ for a position $\mathbf{x}$ coincides with the value of φ_b yielding the maximal intensity in $g_{\mathbf{x}}(\varphi_b)$.
- The next step consists in finding the orientation in the elevation direction $\vartheta(\mathbf{x})$ for each spot. This information can be extracted from a new illumination series, which is generated by fixing the azimuth angle φ_b of a point light source at the previously determined value $\phi(\mathbf{x})$ and then varying the elevation angle θ from 0° to 90°. This latter results in an intensity signal $g_{\mathbf{x}}(\theta)$, whose maximum describes the elevation of the surface normal direction $\vartheta(\mathbf{x})$. Finally, the reflection properties are determined for each location $\mathbf{x}$ through least squares fitting of the signal $g_{\mathbf{x}}(\theta)$ to the reflection model described in Eq. (8). Meaningful parameters that can be extracted from the model are, for example, the width $\sigma(\mathbf{x})$ of the forescatter lobe, the strengths $k_{\mathrm{fs}}(\mathbf{x})$ and $k_{\mathrm{d}}(\mathbf{x})$ of the lobes and the local surface normal given by:

$$\mathbf{n}(\mathbf{x}) = (\cos\phi(\mathbf{x}) \sin\vartheta(\mathbf{x}), \sin\phi(\mathbf{x}) \sin\vartheta(\mathbf{x}), \cos\vartheta(\mathbf{x}))^{\mathrm{T}}. \tag{9}$$

In what follows, we use these parameters as features for segmentation.

5.1.3 Segmentation

Segmentation methods are often categorized into region-oriented and edge-oriented approaches. Whereas the first ones are based on merging regions by evaluating some kind of homogeneity criterion, the latter rely on detecting the contours between homogeneous areas. In this section, we make use of region-oriented approaches. The performance is demonstrated by examining the surface of two different cutting inserts: a new part, and a worn one showing abrasion on the top of it; see Fig. 4.

5.1.3.1 Region-based segmentation

Based on the surface normal $\mathbf{n}(\mathbf{x})$ computed according to Eq. (9), the partial derivatives with respect to x and y, $p(\mathbf{x})$ and $q(\mathbf{x})$, are calculated. It is straightforward to use these image signals as features to perform the segmentation. To this end, a region-growing algorithm is applied to determine connected segments in the feature images [Gonzalez & Woods (2002)]. To suppress noise, a smoothing of the feature images is performed prior to the segmentation. Fig. 5 shows a pseudo-colored representation of the derivatives $p(\mathbf{x})$ and $q(\mathbf{x})$ for both the new and the worn cutting insert. The worn area can be clearly distinguished in the second feature image $q(\mathbf{x})$. Fig. 6 shows the segmentation results. The rightmost image shows two regions that correspond with the worn areas visible in the feature image $q(\mathbf{x})$. In this case, a subset of

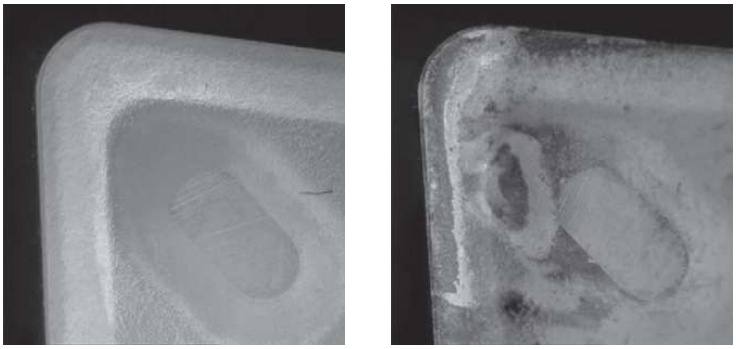

Fig. 4. Test surfaces: (left) new cutting insert; (right) worn cutting insert. The shown images were recorded with diffuse illumination (just for visualization purposes).

the parameters of the reflection model was sufficient to achieve a satisfactory segmentation. Further, other surface characteristics of interest could be detected by exploiting the remaining surface model parameters.

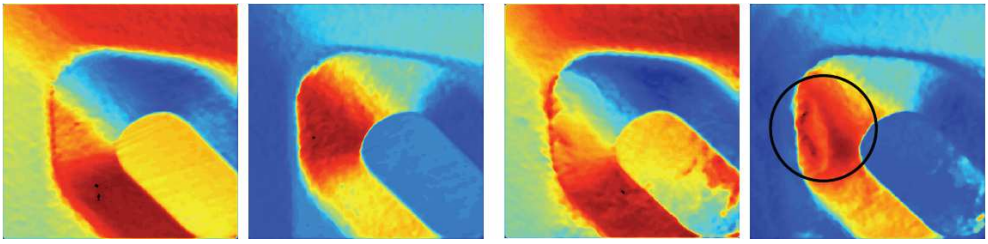

Fig. 5. Pseudo-colored representation of the derivatives $p(\mathbf{x})$ and $q(\mathbf{x})$ of the surface normal: (left) new cutting insert; (right) worn cutting insert. The worn area is clearly visible in area of the rightmost image as marked by a circle.

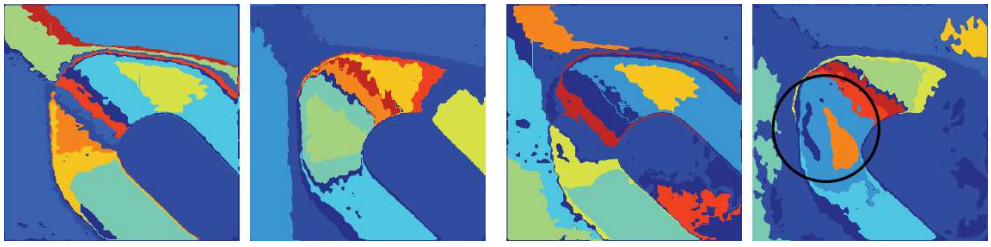

Fig. 6. Results of the region-based segmentation of the feature images $p(\mathbf{x})$ and $q(\mathbf{x})$: (left) new cutting insert; (right) worn cutting insert. In the rightmost image, the worn regions were correctly discerned from the intact background.

Fig. 7 shows a segmentation result based on the model parameters $k_{\mathrm{d}}(\mathbf{x})$, $k_{\mathrm{fs}}(\mathbf{x})$ and $\sigma(\mathbf{x})$. This result was obtained by thresholding the three parameter signals, and then combining