

Performance Measure as Feedback Variable in Image Processing

Danijela Ristić and Axel Gräser

Institute of Automation, University of Bremen, Otto-Hahn-Allee NW1, 28359 Bremen, Germany

Received 28 February 2005; Revised 4 September 2005; Accepted 8 November 2005

This paper extends the view of image processing performance measure presenting the use of this measure as an actual value in a feedback structure. The idea behind is that the control loop, which is built in that way, drives the actual feedback value to a given set point. Since the performance measure depends explicitly on the application, the inclusion of feedback structures and choice of appropriate feedback variables are presented on example of optical character recognition in industrial application. Metrics for quantification of performance at different image processing levels are discussed. The issues that those metrics should address from both image processing and control point of view are considered. The performance measures of individual processing algorithms that form a character recognition system are determined with respect to the overall system performance.

Copyright © 2006 Hindawi Publishing Corporation. All rights reserved.

1. INTRODUCTION

Throughout the development of image processing systems, nearly all research has been dedicated to design of new algorithms or to improvement of existing ones. In the last years, a significant effort also has been devoted to quantitative performance assessment of different image processing methods [1]. In that, image processing algorithms mostly have been considered on their own and developed performance measures have been used to evaluate the effectiveness of individual algorithms or to compare the different image processing algorithms [2, 3]. However, in practice an image processing system consists of serial image processing operations combined differently depending on the overall goal of the vision system. Depending on application, it can happen that a performance measure of an algorithm if considered on its own is not a suitable performance measure if the same algorithm is encapsulated within a larger system. Therefore, it is very important to measure the effectiveness of individual algorithm within a vision system. Recently, some results on performance measures that provide a step to building vision systems that automatically adjust algorithm parameters at each level of the system to improve overall performance were published [4, 5]. In this paper, such kind of performance measure is considered but throughout the consideration of inclusion of control techniques in standard image processing system. The inclusion of closed-loop control is suggested to overcome the problems of standard open-loop image processing. The motivation is the knowledge coming from control

theory, that closed-loop systems have the ability to provide a natural robustness against disturbances and system uncertainty [6].

When control techniques are discussed in connection with image processing, they are usually done so in the context of an *active vision* or *visual servoing systems* [7, 8] which use the image processing to provide visual feedback information for closed-loop control. In contrast to active vision and visual servoing, there are only a few publications dealing with the usage of control techniques in image processing [9–11]. The intention of this paper is to give an additional contribution to the topic.

Other authors that have used classical and modern control techniques to solve image processing and machine vision problems used them for improving the reliability of applied processing techniques. For example, in [9] control ideas were used to improve subpixel analysis in pattern matching. In [12] feedback strategies were used for improving on established single-pass hypothesis generation and verification approaches in object recognition. In those publications, as well as in research on active vision, the image quality is taken for granted. The assumption is that the image at the high level of an object recognition system is of a quality good enough for successful feature extraction. In contrast, in this paper as well as in [13] the feedback control of image quality at different levels of image processing is considered. As it will be shown, the feedback structures are realized as feedbacks between the quality of the image at a particular image processing level and the parameters of image processing algorithms.

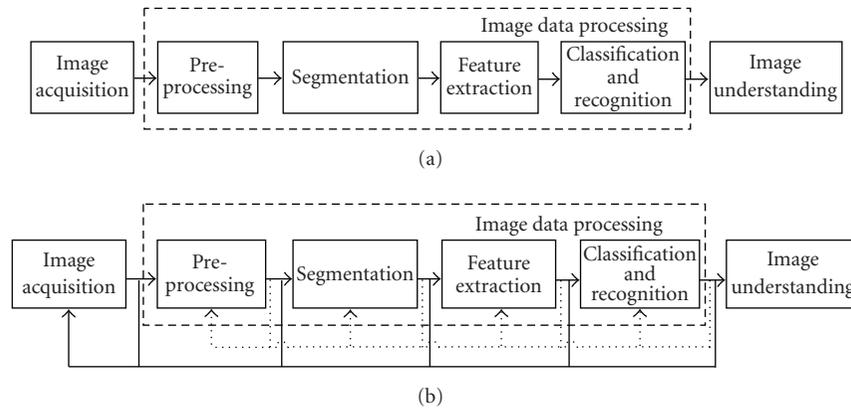


FIGURE 1: Block diagrams of standard open-loop (a) and closed-loop (b) digital image processing.

Hence, the performance measure of an individual image processing algorithm is a measure which reflects the quality of the resulted image. This measure must be appropriate not only from the image processing point of view but also from the control point of view. A basic requirement of the control is that the quality of the image has to be measured so that the control variables can be changed to optimize it. It should be possible to calculate it easily from the image and it must be understandable what the image quality actually is.

The paper is organized as follows. The key principle of the inclusion of feedback structures in image processing is given in Section 2. The main emphasis is on the choice of the control and feedback variables. In some applications, illumination condition or camera parameters may be used as control variables that influence the quality of images at different levels of image processing. In applications where this is not possible, only internal image processing variables are available for control. This and other specifics, as well as the benefit of the closed-loop control in image processing, are discussed more detailed in Section 3 throughout the demonstration of results achieved for optical character recognition in the automotive industry. In this case, our feedback mechanism treats the image acquisition as the essential image processing step, bearing in mind the influence of the quality of original (not-processed) image on the subsequent processing, which is different from other published results on feedback structures in image processing [12, 14]. Active image acquisition should provide a “good” original (not-processed) image suitable for subsequent processing. In literature there are different treatments of the control of image acquisition on its own. For example, in [15], a feedforward of camera parameters according to the appropriate quality measure is used to provide an image of good quality. We focus our attention on the illumination condition as another important factor for the success of image acquisition. In contrast to [16] where the active illumination is considered as a forward action, in this paper the error-based closed-loop control technique is discussed. Besides the image acquisition control, aiming at improvement of the quality of not-processed image and so at providing a basis for the image data processing to be supported by a more

robust input image data, the inclusion of feedback control at the segmentation level of image processing is considered. The goal is to suggest the possible method of the automatic adjustment of the internal image processing variables, based on the classical control techniques, for improvement of the overall image processing performance. The comparison of performance of proposed method of closed-loop parameter adjustment to the performances of two traditional open-loop adaptive segmentation methods is given in Section 4.

Even though the results presented in this paper are entirely associated with the acquisition and segmentation levels of image processing and application of the character recognition in industrial environment, it is believed that the applied technique will be equally suitable for both other steps and other applications of image processing.

2. CLOSED-LOOP CONTROL IN IMAGE PROCESSING

A majority of image processing applications concern the object recognition and typically consist of three subsequent parts: image acquisition, image data processing, and image understanding, as it is shown in Figure 1(a).

In a strict sense, the image acquisition does not belong to image processing [12, 17]. However, having in mind the influence of the original (not-processed) image on the subsequent processing, the image acquisition has to be considered as an essential processing level. If imperfections of the original image are introduced in the standard sequential processing steps, then the results of the subsequent steps become unreliable. Nevertheless, in many real-world applications the vision engineers will regard the images as given and use traditional preprocessing methods to improve the image quality. This is time-consuming and the results are of low accuracy due to the lost image information during the image acquisition. And, in some cases, due to the poor data, it is even impossible to improve the image quality by any standard preprocessing technique. The weak result of one processing unit directly lowers the quality of the following processing step which leads to low robustness of the overall system. Also, the processing at lower levels is performed

regardless of the requirements of the following steps. The introduction of feedback and control strategies at all levels of image processing applications as proposed in Figure 1(b) will lead to higher robustness and reliability of the image processing system [13, 18].

It is possible to include two types of closed-loops in a standard image processing system. The first one can be named *image acquisition closed-loop*. Here the information from all subsequent stages of image processing may be used as feedback to control acquisition conditions (solid lines in Figure 1(b)). The aim is to provide a “good” image for the subsequent processing steps from preprocessing, through the segmentation and feature extraction to the classification. As it will be shown in the following section, the image acquisition closed-loop may cancel the need for the traditional preprocessing techniques and so can be considered as a new image processing method.

The second type of closed-loop can be realized as the feedback between the quality of the image, representing the input of a higher processing level, and the parameters of image processing at lower level as represented with dashed lines in Figure 1(b). This closed-loop adjusts parameters of the applied processing algorithm according to requirements of the subsequent image processing step and so can be named *parameters adjustment closed-loop*. Hence, in contrast to image acquisition closed-loop, this can be treated as a *local feedback* at the related stage of image processing.

The closed-loop control in image processing differs significantly from the usual industrial control, especially concerning the choice of the actuator and the controlled variables. Generally, the actuator variables are those that directly influence the image characteristics. Hence, for the image acquisition closed-loop, depending on the application, the actuator variables can be camera’s parameters or the illumination condition. For the second type of closed-loop, the actuator variables are parameters of applied processing algorithms (e.g., coefficients and size of a smoothing filter at preprocessing level, threshold or parameters and size of filter masks for point, line, and edge detection at segmentation level, etc.). However, the choice of the controlled variable is not a trivial problem. This variable has to be appropriate from the control as well as from the image processing point of view as it was explained in Section 1. From the image processing point of view, a feedback variable must be an appropriate measure of image quality.

The problem of identifying which image data are good and which are bad has become a serious issue in the vision community [15]. To answer the question “what is the image of good quality?” is quite a difficult problem. The image quality obviously depends on the interpretation of the context. If the image of a “top model” is considered, a good image may hide some details like, for example, no perfect skin. For a surgical endoscope a good image is one showing the organ of interest in all details clearly for human interpretation. However, in the machine vision context it must be kept in mind that “what human being can easily see is not at all simple for a machine” [19]. Therefore, one of the main problems in the implementation of automated visual inspection systems

is to understand the way in which the machine “sees” and the conditions that have to be created for it to perform its task at its best. Since the correct image understanding highly depends on the result of the object recognition and classification, which represent the last of sequentially arranged image processing steps, it turns out that a “good” image is one on which the subsequent steps work well. If we consider the performance measure of image processing at a particular level as a measure of the quality of the corresponding image, then the performance of active acquisition may be a measure that reflects how good contrast of the original image is. The performance of algorithm at the segmentation level measures correctness of segmentation of the image areas corresponding to objects of interest, and so forth.

In the following, the principle of the choice of the control and controlled variable in image processing is presented for the case of character recognition in industrial environment. The authors are of the opinion that even though this choice is application dependent, once a pair of controlled and actuator variable is found the framework for the inclusion of proven error-based control techniques in different image processing applications is provided.

3. FEEDBACK CONTROL FOR IMPROVEMENT OF CHARACTER RECOGNITION IN INDUSTRIAL APPLICATIONS

3.1. The problem

Automated reading of human-readable characters, known as optical character recognition (OCR) [20] is one of the most demanding tasks for computer vision systems since it has to deal with different problems like wide range of fonts, confusing characters such as B and 8, or unevenly spaced characters. Besides the mentioned common problems concerning the nature of text information to be recognized in various industrial applications ranging from the pharmaceutical industry to the automotive industry, there are numerous specific challenging conditions that should be met. In the automotive industry, that represents one of the most frequent and important application area of the OCR, there is a great variety of identification marks on different materials to be detected. Some of them are shown in Figure 2.

Each character type has specific challenges for the character recognition system, but for all types of identification codes the main difficulty concerns the image acquisition condition. The reliable detection of the identification codes is necessary throughout the whole car manufacturing process including the painting process. Different surface types, containing the characters to be detected, ranging from the rough surface of casting to the differently colored and polished surface of the car body, lead to very different light reflection conditions. Hence, it turns out that even for the same characters the illumination condition during the image acquisition in different stages of manufacturing has to be adjusted. To investigate the possibility of fully autonomous and robust OCR system, the experiment of the imaging of differently colored metallic plates, with scratched numerical characters

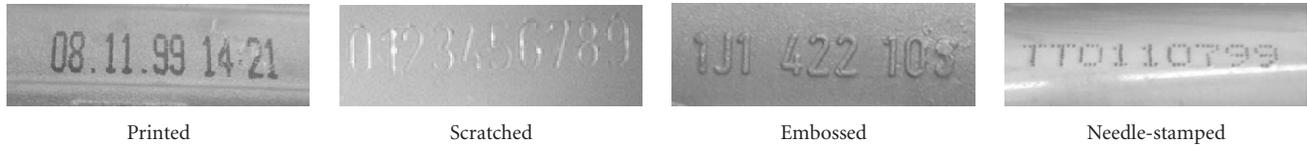


FIGURE 2: Different types of identification codes on metallic surfaces to be detected.

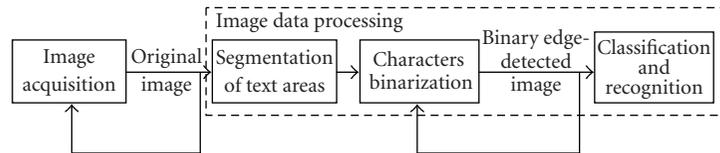


FIGURE 3: OCR system with included feedback control.

on them, in variable illumination conditions was performed [10]. The variable illumination was accomplished by varying the position of the point light source. This simple lighting arrangement fulfills the requirements for the detection of scratched or embossed characters. The scratched and embossed marks of work pieces, representing surface deformations, are required in many production processes as durable markings, resistant to subsequent processing steps. Because of their three-dimensional structure, characters created this way are often difficult to illuminate, to segment and, consequently, to detect. The using of directional front lighting is a good way to visualize surface deformations [19] since the characters appear bright in contrast due to the reflection from the characters edges. However, in case of a heavily textured surface such as it is created by certain machining methods or caused by pollution, the whole surface may appear bright since it contains a lot of microscopic deformations. The problem is also that depending on the quality of the marking process, the depth of the characters can vary demanding a different illumination even for the same characters due to different reflection conditions. Hence, to find the optimal position of the light source is of major significance for characters detection. The determination of the light source position or the appropriate combination of multiple lighting elements by the vision engineer in an iterative process [16] is time-consuming. By choosing the appropriate controlled variable and the closed-loop control strategy for image acquisition, the adjustment of the parameters of the illumination setup can be done automatically for different types of characters and surfaces. The suggested image acquisition closed-loop is illustrated in Figure 3 as the feedback between the quality of the original (not-processed) image and the illumination conditions as a crucial factor for the image acquisition.

As shown in Figure 3, alike to most other OCR systems [20] we consider the classical structure, but only as the skeleton. Hence, the system is in two major sections: image acquisition and image data processing. Processing consists of usual steps: segmentation of characters to be recognized,

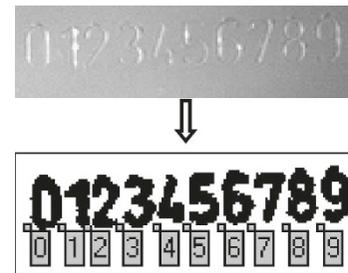


FIGURE 4: Control goal: “good” original image (top), recognized characters in corresponding binary edge-detected image (down).

their binarization, and finally classification and recognition. The novel difference in our configuration in comparison to other traditional systems lies in the inclusion of two control loops. Besides the above-mentioned image acquisition closed-loop, the local feedback at the segmentation level was introduced. This closed-loop was realized as the feedback between the quality of binary edge-detected image and the threshold value as a parameter determining the success of the characters binarization.

The control goals of included control loops are performances of the corresponding processing steps. While the control goal of the first considered closed-loop is to provide the original image of “good” quality suitable for the subsequent segmentation, consisting of edge-detection and characters binarization, the second closed-loop has as a goal to give a “good” image input for the classifier. The “good” input means that the binary edge-detected image contains the “full” clearly separated characters that resemble the characters used for the training of the classifier. That is when, by using a simple classifier, all characters can be recognized as shown in Figure 4. Hence, the overall control goal of the implemented feedback control loops is to provide a basis for the classification to be supported by a reliable data from the lower levels of image processing. Therefore,

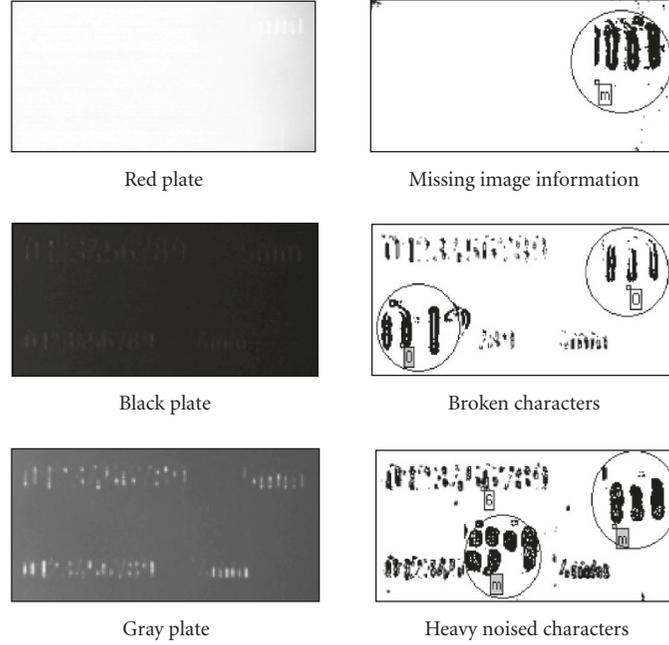


FIGURE 5: Original image of metallic plate with 26 scratched characters on it (left). Recognized characters in binary edge-detected image (right).

as it will be shown, the measures of effectiveness of individual control-loops are determined considering the number of correctly recognized characters which is the overall OCR system performance.

3.2. Image acquisition closed-loop

As it was said above, the control goal of the image acquisition closed-loop is to give the not-processed image of good quality. To find the measure of the image quality that could be used as feedback variable, some image types, representing the undesired cases for OCR relative to the illumination conditions, were first investigated (Figure 5).

The red, black, and gray plates with scratched characters on them, of 5 mm and 4 mm height and 0.5 mm width, were imaged in identical illumination conditions. Due to the different light reflection from the plates, the acquired images resulted in a too bright, too dark, and low-contrast image, respectively as shown in Figure 5. The first two images are obviously of so bad quality that even for the human being it is difficult to recognize characters on them. On the first sight, the third image is a “good” one since a human being can recognize characters on it. However, the corresponding image histogram (Figure 6(c)) is too narrow indicating the very low contrast of the edges of characters. Hence, the corresponding binary edge-detected image as well as binary images of the first two acquired images are of poor quality. The characters on them are broken, heavy noised, or simply there are no characters due to the lost image information during the image acquisition. Consequently, the result of character recognition is quite weak and unreliable. Since lost image informa-

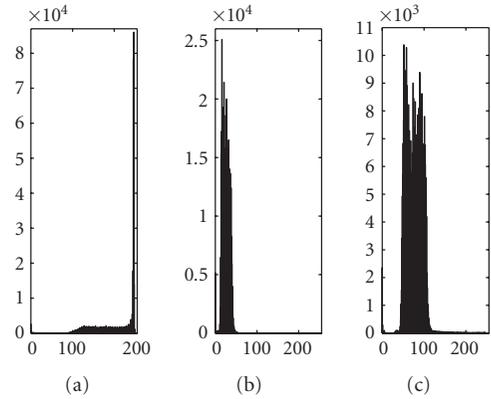


FIGURE 6: Gray-level histograms of the bright (a), dark (b), and low-contrast (c) images shown in Figure 5.

tion cannot be restored, it turns out that the image suitable for character recognition must contain the maximum information.

In the classical information theory, the measure of the average information generated by the source is the *entropy* [21]. Considering an image as a source with independent pixels, the entropy is defined as the information content of the image and is given by the following formula:

$$H = - \sum_{i=0}^{N-1} p_i \log_2 p_i \quad [\text{bits/pixel}], \quad (1)$$

TABLE 1: Division of gray-level scale into three areas.

Gray-value areas	1 (dark)	2 (middle)	3 (light)
Gray values	0 ··· 35	36 ··· 179	180 ··· 255

where

(i) p_i is the probability of occurrence of pixel value i :

$$p_i = \frac{\text{number of pixels with gray-level } i}{\text{total number of pixels in the image}}; \quad (2)$$

(ii) N is the number of pixel values (gray levels). For the usual case of an 8-bit integer image $N = 256$ when, according to (1), theoretical maximum entropy is 8 [bits/pixel].

The definition of the image entropy (1), also known as the entropy of one-dimensional (1D) histogram or 1D entropy, indicates the maximal entropy as the best measure of the image quality. Even though it is correct for some image processing applications [15], for the character recognition it is not the case. Maximal entropy corresponds to the case of all gray values equally distributed over the image pixels. That means that the image can contain too bright or dark spots which, as seen in Figure 5, cover the characters to be recognized. In order to avoid dark or light spots in an image, causing the loss of information on characters, the majority of image information should be contained in the gray levels from the middle part of the gray-level scale. However, the image using only the gray levels from the middle area, according to example shown in Figures 5 and 6(c), is of low-contrast and so not suitable for the OCR.

All previous discussions indicate that *to have an image of high contrast, suitable for further character recognition, the image histogram must be stretched over the whole gray-level scale, but the maximum of information must be carried by gray levels from the middle gray-value area*. In order to find the measure of the stretch degree of the image histogram, the following division of a gray-level scale to a dark, middle, and light area is suggested as shown in Table 1 [10].

A coefficient α is introduced, which represents the relative contribution of the entropy in the middle gray-value area to the total sum of entropies:

$$\alpha = \frac{H_2}{(H_1 + H_2 + H_3)}. \quad (3)$$

The coefficient α is used as the performance measure of the spread of the image histogram over the gray-level scale. Its reference value is $0.5 \leq \alpha$. The entropies in dark H_1 , middle H_2 , and light H_3 areas are determined according to

$$H_j = - \sum_{i=LB_j}^{UB_j} p_i \log_2 p_i, \quad j = 1, 2, 3, \quad (4)$$

where p_i is the probability of occurrence of pixel value i in the j th gray-level area and LB_j and UB_j are, respectively, lower and upper boundaries of the corresponding gray-value area.

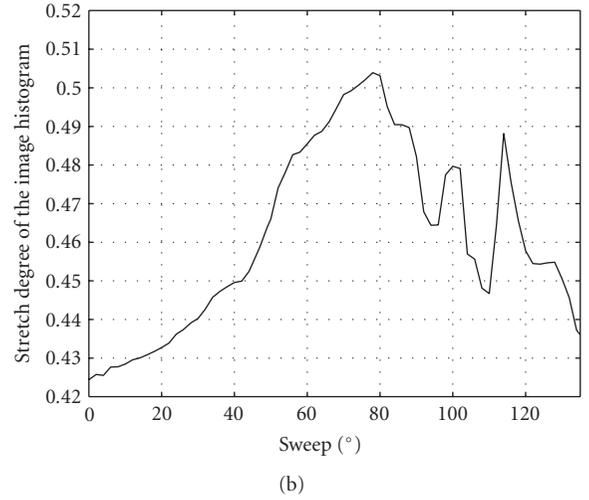
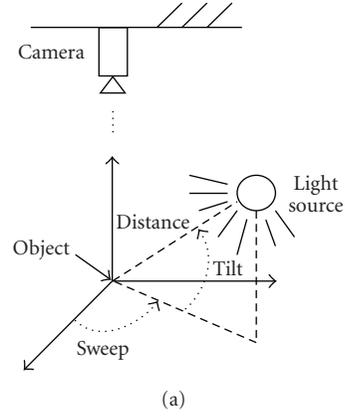


FIGURE 7: Position of the light source with respect to the imaged object (a). Stretch degree of the image histogram for different sweeps of the light source (b).

The boundaries were determined by testing the changes of the image contrast on a set of images representing the full range of lighting conditions. The idea was to overcome the drawback of one dimensional histogram of not giving any information on spatial distribution of gray levels in an image, and consequently to provide information about over- and poor-lighted image areas.

The chosen control variable was a parameter determining the light source position and consequently the light source intensity. More precisely, the sweep of the light source with respect to the imaged object was considered as variable while two other parameters that determine the position of the light source (Figure 7(a)) were kept as constant.

Figure 7(b) shows the changing of the stretch degree of the image histogram α with changing of the position of light source during the image acquisition. As it can be seen the chosen measure of the quality of image histogram, and consequently of image contrast, is *sensitive* to the control variable across the available operating range. Also, it is obvious that there is *one-to-one steady state mapping* between these two variables and that it is possible to achieve the *global*

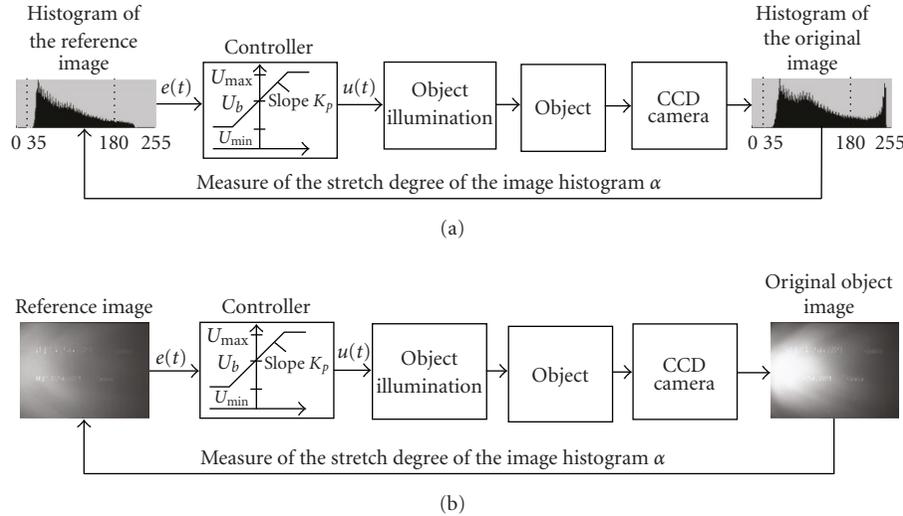


FIGURE 8: Image acquisition closed-loop.

maximum of α by changing the illumination condition. Since these basic prerequisites for successful control action to be performed are fulfilled, α was used as feedback variable in the implemented image acquisition control in our OCR system. The block diagram of the image acquisition closed-loop, which provides the image of high contrast suitable for subsequent segmentation and characters binarization, is shown in two forms. The former, shown in Figure 8(a), presents the effect of the implemented control on the image histogram, and the latter in Figure 8(b) explicitly demonstrates the result of control of image quality.

On the first sight the effect of the implemented control technique in image acquisition is the same as of the traditional image preprocessing technique known as *contrast stretching* [22]. The novel difference is that in contrast to the traditional case the implemented control technique changes, also, the contour of the histogram and so provides the avoidance of the saturated image case when the classical contrast stretching fails. The traditional contrast stretching makes the overlighted image areas larger which degrades the image quality in applications when larger bandwidth of gray levels is needed. Hence, the suggested control-based method can be regarded as a new image processing method.

Once the image of good contrast is achieved, the second feedback at the segmentation level of OCR system, which will be described in the next section, is initialized. By the on-line maintaining of the achieved good quality of not-processed image, the input image of the segmentation level can be treated as the image of constant quality. The benefit is that the process of binarization of characters to be detected may be considered as deterministic process.

3.3. Threshold adjustment closed-loop

Image segmentation is a key step in character recognition [20, 23]. If the characters to be detected are not correctly

segmented from the background, it is not possible to extract accurately the characters features needed for the classification and character recognition. Since the weak features lead to weak character recognition, it is of crucial importance to achieve the reliable segmentation of the text to be detected.

In our system, the segmentation of text area consists of two image processing operations: edge-detection and thresholding. Bearing in mind that the image acquisition closed-loop provides on-line original image of good quality, the assumption that the edges of characters are correctly identified by chosen Sobel 5×5 [22] filter mask can be taken for granted. Hence, the eventual success or failure of subsequent classification and character recognition highly depends on the thresholding step. Thresholding is an image point operation which produces a binary image from a gray-scale image (in our system from the gray-scale edge-detected image). A binary zero is produced on the output image whenever a pixel value on the input image is greater than chosen threshold. A binary one is produced otherwise. Therefore, the quality of binary image depends on the threshold. Too high threshold value yields a very small number of black pixels in the binary image and so, in the case of white background and black characters, leads to loss of information on characters to be detected. In contrast, a low threshold value yields a large number of black pixels in the binary image. In that case a lot of black pixels may be “not useful” in the sense that they do not belong to characters to be recognized. These “extra” black pixels arise due to the reflection from some deformations on the imaged plate surface which are also recognized as edges in the edge-detection step. That is why the adequate determination of the threshold value and its adaptation to environmental changes is of major importance for characters recognition. Since it is very difficult to estimate what is “too high or too low threshold value” without any feedback information on the result of image binarization, using the fixed threshold in traditional open-loop image

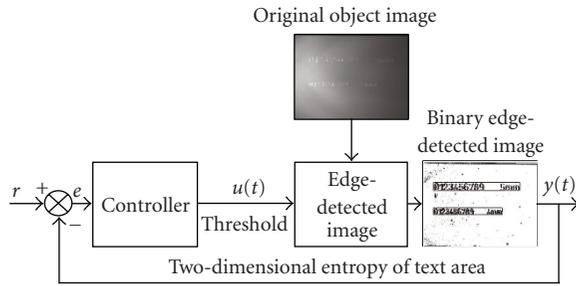


FIGURE 9: Threshold adjustment closed-loop.

processing system often gives poor character recognition results. There are publications that treat the adaptation of threshold value but mostly in open-loop and time-consuming iterative process [24, 25]. The suggestion is to apply the proven error-based control techniques in the implemented closed-loop shown in Figure 9.

Since the threshold value is the parameter which directly influences the quality of binary edge-detected image, it was considered as the control signal in the implemented closed-loop. The more compact are black pixels that form the characters to be recognized, the binary edge-detected image is of better quality. Hence, the measure of connectivity of black pixels in segmented text area was naturally imposed as controlled variable.

We introduce the two-dimensional (2D) entropy as a measure of connectivity of black pixels forming the characters to be detected. It is defined by the following formula:

$$S = - \sum_{i=0}^8 p_{(0,i)} \log_2 p_{(0,i)}, \quad (5)$$

where $p_{(0,i)}$ is the probability of occurrence of a pair $(0, i)$ representing the black pixel surrounded with i black pixels (i takes values from 0 to 8 while considering the 8-neighborhood):

$$p_{(0,i)} = \frac{\text{number of black pixels surrounded with } i \text{ black pixels}}{\text{number of black pixels in the image}}. \quad (6)$$

Figures 10(a) and 10(b) show, respectively, the images of the “good” numerical character “2” and the “broken” one together with the corresponding histograms of distribution of pairs $(0, i)$ found in the characters images.

As it is obvious, the histogram of the “full” character is very narrow in contrast to the histogram of the “broken” character. This is the expected result since the number of different pairs $(0, i)$ in the image of “good” character is smaller than in the image of “noised” character, but the probability of occurrence of found pairs $(0, i)$ is larger. It is known that random variable X with a large probability of being observed has a very small degree of information $-\log p(X)$ [21]. Hence, according to (5), the 2D entropy of a “good” character, formed of connected black pixels, is supposed to be quite smaller than the 2D entropy of a “broken” or “noised”

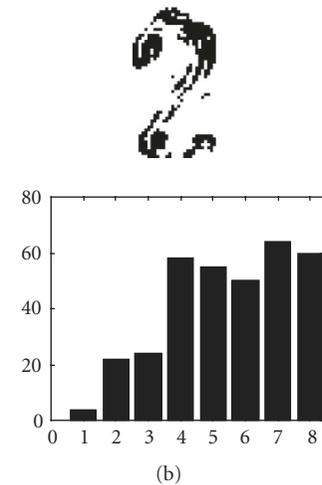
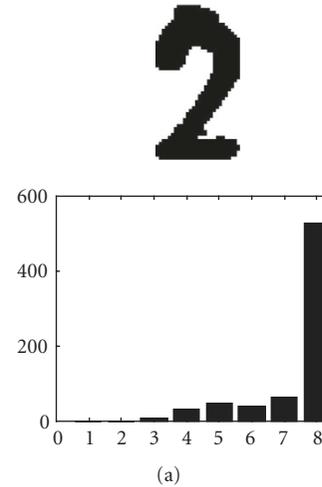


FIGURE 10: “Full” (a) and “broken” (b) numerical character “2” with the corresponding histograms of distribution of pairs $(0, i)$.

character. The results of 1.4147 and 2.698 for the 2D entropy of shown “full” and “broken” character “2”, respectively, confirm the previous statement. This provides a basis for the use of 2D entropy as a measure of the quality of a binary image containing the characters to be detected.

The case considered here assumes the black characters on white background, but the same measure can be used in the opposite case since the introduced 2D entropy is in general the measure of the connectivity of pixels representing the characters to be detected. In other words, the introduced metric (5) is a performance of the thresholding stage of image segmentation.

Figure 11 shows the changing of the 2D entropy of text area in binary edge-detected image S with changing of the threshold value.

Obviously, the 2D entropy of text area is *sensitive* to the chosen control variable across the available operating range. Also, it is evident that there is *one-to-one steady state mapping* between these two variables and that it is possible to

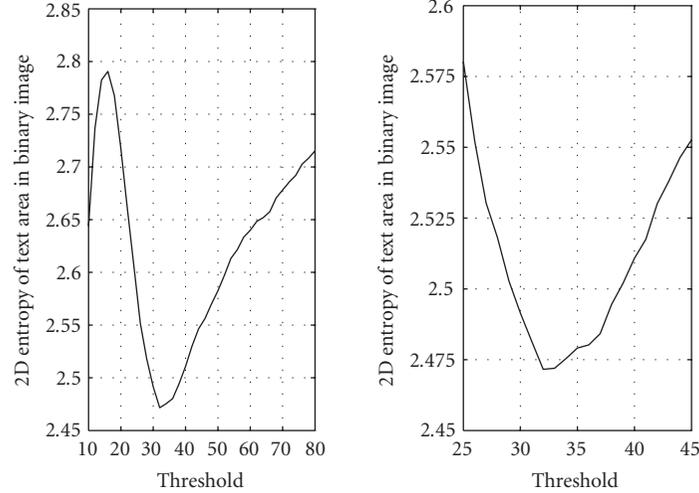


FIGURE 11: 2D entropy of text area in binary image versus threshold value.

achieve the *global minimum* of S by changing the threshold value at binarization stage of image segmentation. The satisfied basic prerequisites for successful control action to be performed prove the pair “threshold—2D entropy of text area” as a good pair “actuator variable—feedback variable” in the implemented threshold adjustment closed-loop.

The response of the threshold adjustment closed-loop and consequently of the overall character recognition system, in the experiment of imaging of a metallic plate with scratched characters on it, is presented in Figure 12.

The achieved result shows that the implemented image acquisition closed-loop rejected the disturbances before they influenced the primary control object, that is, the binary edge-detected image. In the threshold adjustment closed-loop, the number of black pixels in text area was gradually increased so that characters were gradually “filled up” as shown in Figure 12 for the case of numerical characters 0, 1, 2, and 3. The reliable character recognition was achieved after the fifth cycle of implemented threshold adaptation closed-loop.

4. COMPARISON OF THE THRESHOLDING PERFORMANCES

In this section, the performance of proposed closed-loop control-based thresholding method is compared with the performances of two traditional adaptive thresholding methods: 1D entropy-based thresholding and 2D entropy-based thresholding [24, 25]. In contrast to our method which uses feedback information on quality of binary image to adjust the threshold, those two methods present “forward actions.” The 1D entropy, that is 2D entropy, of the background and foreground of the gray-level image to be thresholded (in our system edge-detected image) is calculated. Then the threshold which corresponds to the maximum of the sum of background and foreground entropies is determined as explained in more details in the following.

4.1. 1D entropy-based thresholding

A few widely-used thresholding methods are based on the concept of 1D entropy defined in Section 3.2 [25]. According to [26], the threshold value t_0 divides the gray-level scale of the 1D histogram of the image to be segmented into two areas. One corresponds to image background and the other corresponds to image foreground, that is, to objects to be segmented. In an image the foreground area (objects) may consist of bright pixels on the dark background as in our case of edge-detected images. Then the 1D entropies of the background H_b and foreground H_f regions of an 8-bit image are, respectively, defined as

$$H_b = - \sum_{i=0}^{t_0} p_i \log_2 p_i, \quad (7)$$

$$H_f = - \sum_{i=t_0+1}^{255} p_i \log_2 p_i, \quad (8)$$

where p_i in (7) is the probability of occurrence of pixel value i in background area $i = 0, \dots, t_0$, and p_i in (8) is the probability of occurrence of pixel value i in foreground area $i = t_0 + 1, \dots, 255$.

The threshold t_0 which will provide optimal result of image binarization is the one maximizing the sum of the entropies (7) and (8). Threshold determined this way is supposed to yield a binary image with the maximum information on segmented objects.

4.2. 2D entropy-based thresholding

The main disadvantage of using of the entropy of 1D histogram is that it does not give any information on spatial characteristics of the image. In order to overcome that problem the entropy of two-dimensional (2D) image histogram has been defined [24]. 2D image histogram is the graphical presentation of the distribution of pair (i, a) representing

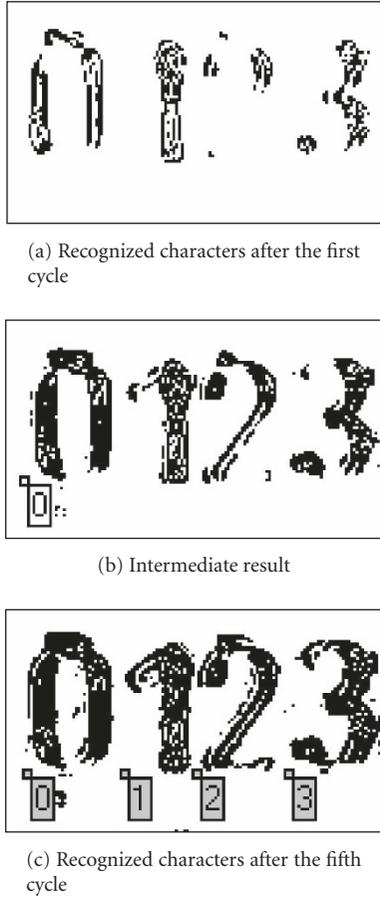


FIGURE 12: Character recognition result achieved with the OCR system with implemented closed-loops.

the pixel of gray-level i surrounded with neighborhood pixels with average gray-value a . The entropy of the 2D histogram of an 8-bit gray-level image is defined as follows:

$$H = - \sum_{i=0}^{255} \sum_{a=0}^{255} p_{ia} \log_2 p_{ia}, \quad (9)$$

where p_{ia} is the probability of occurrence of the pair (i, a)

$$p_{ia} = \frac{\text{number of pairs } (i, a)}{\text{total number of pairs in the image}}. \quad (10)$$

As in the case of 1D entropy-based thresholding here the 2D entropies of the background and foreground of a gray-level image supposed to have bright objects on the dark background are calculated as

$$H_b = - \sum_{i=0}^{t_0} \sum_{a=0}^{a_0} p_{ia} \log_2 p_{ia}, \quad (11)$$

$$H_f = - \sum_{i=t_0+1}^{255} \sum_{a=a_0+1}^{255} p_{ia} \log_2 p_{ia}, \quad (12)$$

where p_{ia} in (11) is the probability of occurrence of the pair (i, a) in background area while p_{ia} in (12) is the probability of its occurrence in foreground area.

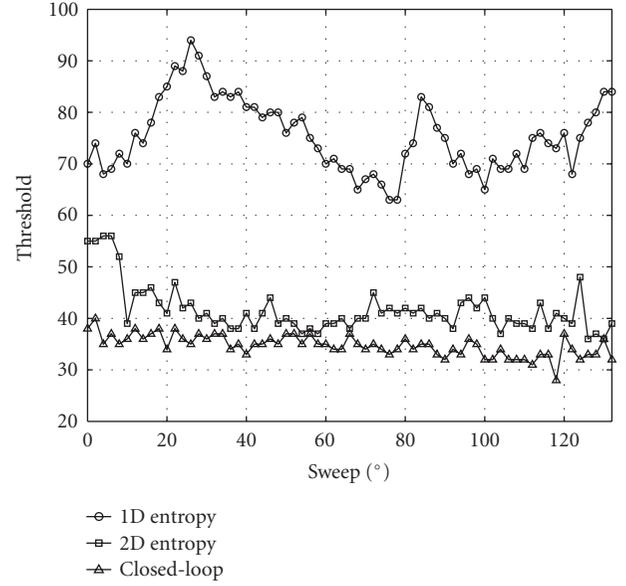


FIGURE 13: Threshold values for the images, corresponding to different sweeps of the light source, obtained using three adaptive methods.

The algorithm then searches for the values $i = t_0$ and $a = a_0$ that maximizes the sum of the background and foreground 2D entropies (11) and (12). This is where the threshold is located.

Bearing in mind that a binary image is an image of only two-pixel values, the suggested 2D entropy of the binary image (5) can be treated as the special case of the 2D entropy of a gray-level image (9).

4.3. Experimental results

The binarization of 72 images of metallic plates with scratched characters on them, captured for the different sweeps of the light source with respect to the imaged object, was performed using the two above described traditional adaptive thresholding methods and the closed-loop control-based thresholding method proposed in this paper. The optimal thresholds resulted from all three methods can be seen in Figure 13.

Figure 14 shows the 2D entropy of text area in binary images corresponding to images captured in different illumination conditions. The edge-detected image of each original image was binarized three times using the threshold values determined according to three previously described methods. The binarization results were compared using the 2D entropy of text area in binary image as the performance criteria.

Obviously the lowest values of 2D entropy of text area in binary images, representing the inputs to classifier, are obtained by closed-loop control-based thresholding. As explained in Section 3.3, the low 2D entropy of binary image leads to better recognition result as can be seen in Figure 15. Presented binary images of numerical characters 1, 2, 3,

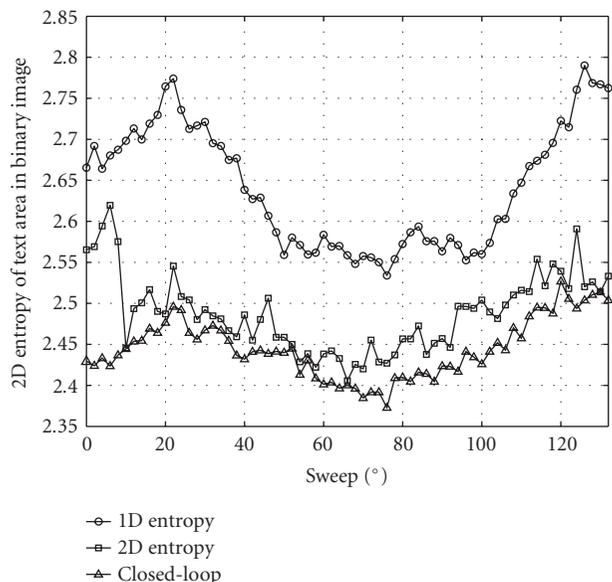


FIGURE 14: 2D entropy of binary images, corresponding to different sweeps of the light source, obtained using three thresholding methods.

and 4 correspond to the original image captured for the sweep 78° of the light source with respect to the imaged object. That image is the output of the implemented image acquisition closed-loop explained in Section 3.2. Binary image shown in Figure 15(a) is obtained using the threshold value 63 determined by 1D entropy-based thresholding method. Figure 15(b) corresponds to threshold 42 that maximizes the 2D entropy of edge-detected image as explained in Section 4.2. The method proposed in this paper resulted in the threshold value 34 which yielded the lowest 2D entropy of the text area and so the best recognition result as shown in Figure 15(c).

5. CONCLUSIONS

In this paper the performance measure of image processing has been considered through the consideration of the inclusion of closed-loop control in standard open-loop image processing system. The idea behind the incorporation of classical control techniques in a standard image processing is to improve its robustness and reliability. The specifics and benefit of inclusion of feedback control at different levels of an image processing system have been demonstrated through the results achieved for recognition of 3D characters on metallic surfaces which represents a standard industrial image processing application. The inclusion of image acquisition closed-loop and the closed-loop at image segmentation level of character recognition system has been considered. The main emphasis is on the choice of the actuator and controlled variable. It has been shown that appropriate measures of the image quality can be considered as the feedback variables in included closed-loops at particular image processing levels. The measures of image quality, represent-

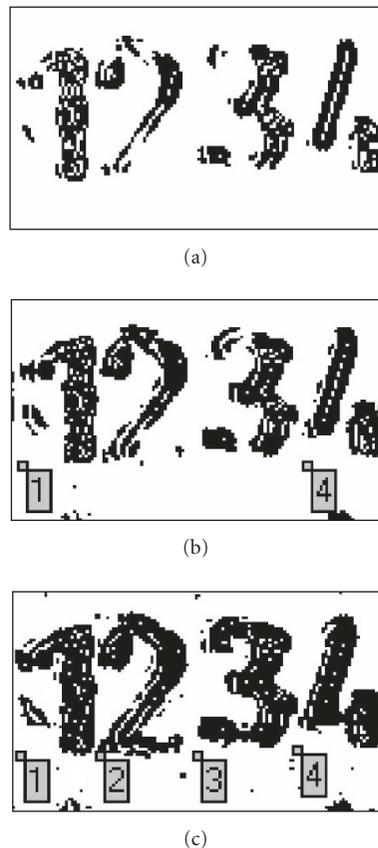


FIGURE 15: Character recognition result achieved with the OCR system with thresholding based on 1D entropy (a), 2D entropy (b), and closed-loop (c).

ing the performance measure of corresponding closed-loop image processing, have been determined with respect to the overall system performance. It has been shown that the performance measure should be appropriate from both the image processing and control point of view.

The choice of controlled and actuator variables while including the feedback control in image processing strongly depends on the image processing application and is not always straightforward due to the availability of a large number of variables that can be treated as measures of the quality of images but which are not all appropriate as feedback variables in closed-loop control. This fact differentiates the most control in image processing from the classical industrial control. However, once a pair of controlled and actuator variables is found for the specific application, the framework for the inclusion of proven error-based control methods is provided.

Experimental results on comparison of performance of the proposed thresholding method, representing the segmentation step, to the performances of traditional adaptive thresholding methods are presented. The results confirmed benefit of the using of feedback information on the quality of binary image to adjust the thresholding parameter.

REFERENCES

- [1] S. Bailey, M. Forshaw, and M. Hodgetts, "A methodology for goal oriented image processing-performance evaluation," in *Workshop on Performance Characterisation and Benchmarking of Vision Systems*, Canary Islands, Spain, January 1999.
- [2] L. Wenyin, J. Zhai, D. Dori, and T. Long, "A system for performance evaluation of arc segmentation algorithms," in *Proceedings of 3rd Workshop on Empirical Evaluation Methods in Computer Vision (CVPR '01)*, Kauai, Hawaii, USA, December 2001.
- [3] J. C. McCall and M. M. Trivedi, "Performance evaluation of a vision based lane tracker designed for driver assistance systems," Tech. Rep., Computer Vision Robotics Research Laboratory, University of California, San Diego, Calif, USA, December 2004.
- [4] P. L. Palmer, H. Dabis, and J. Kittler, "A performance measure for boundary detection algorithms," *Computer Vision and Image Understanding*, vol. 63, no. 3, pp. 476–494, 1996.
- [5] M. Mirmehdi, P. L. Palmer, and J. Kittler, "Optimising the complete image feature extraction chain," in *Proceedings of the 3rd Asian Conference on Computer Vision (ACCV '98)*, vol. 2, pp. 307–314, Hong Kong, January 1997.
- [6] K. Ogata, *Modern Control Engineering*, Prentice-Hall, London, UK, 2002.
- [7] P. Corke, *Visual Control of Robots: High Performance Visual Servoing*, Research Studies Press, Taunton, UK; Wiley, New York, NY, USA, 1996.
- [8] B. Jähne, H. Haußecker, and P. Geißler, Eds., *Handbook of Computer Vision and Applications-Vol.3. Systems and Applications*, Academic Press, London, UK, 1999.
- [9] D. Nair, L. Wenzel, A. Barp, and A. Siddiqi, "Control strategies and image processing," in *Proceedings of the 7th International Symposium on Signal Processing and Its Applications (ISSPA '03)*, pp. 557–560, Paris, France, July 2003.
- [10] A. Gräser, O. Ivlev, and D. Ristić, "First experiences with feedback structures in image processing," in *Proceedings of 24th Colloquium of Automation*, pp. 98–109, Salzhausen, Germany, 2002.
- [11] D. Kriegman, G. D. Hager, and A. S. Morse, Eds., *The Confluence of Vision and Control*, Springer, New York, NY, USA, 1998.
- [12] M. Mirmehdi, P. L. Palmer, J. Kittler, and H. Dabis, "Feedback control strategies for object recognition," *IEEE Transactions on Image Processing*, vol. 8, no. 8, pp. 1084–1101, 1999.
- [13] D. Ristić, I. Volosyak, and A. Gräser, "Feedback Control in Image Processing," *atp international-automation technology in practice*, no. 1, pp. 61–70, 2005.
- [14] J. Peng and B. Bhanu, "Closed-loop object recognition using reinforcement learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 2, pp. 139–154, 1998.
- [15] J. A. Marchant, "Testing a measure of image quality for acquisition," *Image and Vision Computing*, vol. 20, no. 7, pp. 449–458, 2002.
- [16] J. J. Racky and M. Pandit, "Active Illumination for the segmentation of surface deformations," in *International Conference on Image Processing (ICIP '99)*, vol. 1, Kobe, Japan, October 1999.
- [17] B. Jähne, *Practical Handbook on Image Processing for Scientific Applications*, CRC Press LLC, Boca Raton, Fla, USA, 1997.
- [18] I. Volosyak, O. Kouzmitcheva, D. Ristić, and A. Gräser, "Improvement of visual perceptual capabilities by feedback structures for robotic system FRIEND," *IEEE Transactions on Systems, Man, and Cybernetics—Part C*, vol. 35, no. 1, pp. 66–74, 2005.
- [19] C. Demant, B. Streicher-Abel, P. Waszkewitz, and M. Strick, *Industrial Image Processing-Visual Quality Control in Manufacturing*, Springer, Berlin, Germany, 1999.
- [20] S. Mori, H. Nishida, and H. Yamada, *Optical Character Recognition*, John Wiley & Sons, New York, 1999.
- [21] H. Maitre and J. Zinn-Justin, "Entropy, information and image," in *Progress in Picture Processing*, pp. 81–115, Springer, North Holland, Amsterdam, 1996.
- [22] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, USA, 2002.
- [23] C. A. Rahman, W. Badawy, and A. Radmanesh, "A real time vehicle's license plate recognition system," in *Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance*, pp. 163–166, Miami, Fla, USA, July 2003.
- [24] A. S. Abutaleb, "Automatic thresholding of gray-level pictures using two-dimensional entropy," *Computer Vision, Graphics and Image Processing*, vol. 47, no. 1, pp. 22–32, 1989.
- [25] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 146–165, 2004.
- [26] D. W. R. Paulus and J. Hornegger, *Pattern Recognition and Image Processing in C++*, F. Vieweg & Sohn, Braunschweig/Wiesbaden, Germany, 1995.

Danijela Ristić received the M.S. degree in automatic control and robotics from the Faculty of Mechanical Engineering, University of Niš, Serbia, in 1998. Since November 2002, she has been a Ph.D. candidate at the Institute of Automation, University of Bremen, Germany. Her current research interest is feedback control in image processing.



Axel Gräser received the Diploma in electrical engineering from the University of Karlsruhe, Germany, in 1976, and the Ph.D. degree in control theory from the TH Darmstadt, Germany, in 1982. From 1982 to 1990, he was the Head of the Control and Software Department at Lippke GmbH, Germany. From 1990 to 1994, he was a Professor of control systems, process automation, and real-time systems at the University of Applied Sciences, Koblenz. Since 1994, he has been the Director of the Institute of Automation, University of Bremen, and the Head of the Department of Robotics and Process Automation. His research interests include service robotics, brain robot interface, digital image processing, and augmented reality.

