https://doi.org/10.2352/ISSN.2470-1173.2018.09.IRIACV-275

Featureless-Region-Based Top Window Recognition for Automatic Industrial Monitoring Systems

Mei-Hsuan Ho, Yu-Hsuan Tsai, Chang-Tao Hsu, and Ren-Hui Chuang
Dept. of Computer Science and Information Engineering, National Chaio Tung University, Taiwan

Abstract

As a vital part for industrial automatic monitoring systems, this paper presents the algorithm for top window recognition on screen-grabbed images. For better performance in terms of accuracy, efficiency, and robustness, we introduce a novel, background-based image feature of window region, i.e., locally-maximum uniform blocks extracted from regions of uniform color, to represent the practically unique spatial structure of background regions in different windows. The recognition procedure consists of several rather simple steps, c.f., foreground-based methods, which include (i) edge extraction, (ii) uniform block extraction, (iii) uniform block matching. The template window with highest matching score is then identified the top window, with its location determined simultaneously to facilitate further analysis of window content. Experimental results indicate that the proposed approach can achieve perfect recognition results for realistic image data from the industry, and also fulfill the near real-time processing requirement of on-line monitoring tasks.

Introduction

Nowadays, factory automation is indispensable. In certain working areas in a factory, e.g., at the production line, operators may need to manually input proper commands or data to a computer, and it will be much desirable if the manual input can also be monitored automatically so that errors due to human factors can be limited. In general, we may need to monitor different contents displayed on the computer screen, such as images or text. However, some traditional software/hardware systems only provide screen images. Thus, an automatic monitoring system can only have such images obtained by a screen grabber as the system input. Furthermore, since most useful text and/or image displayed on the screen are those in the activated (working) window, i.e., the top window, the monitoring system needs to identify the top window before any needed information can be extracted. In this paper, we propose an accurate and efficient method to find the top window. An overview of the proposed automatic monitoring system is shown in Fig.1.

To recognize the top window, we need to first determine the features to be extracted in grabbed images. Over the years, several image feature extraction approaches have been proposed by utilizing different properties of an image. For example, Swain and Ballard [1] proposed a simple image recognition method based on color information, which utilizes the color histogram(s) as image feature. However, the color information cannot be used to characterizing the window images from a monochromatic computer display. Other features which are often used include edge and corner information. The edge information is one of widely used image features. A detailed study of various image edge detection techniques can be found in [2]. Corners are also popular image features for representing interesting points in image recognition. Harris corner detector is the most widely used corner feature extraction method, which has strong invariance to rotation, scaling and image noises [3]. However, edge and corner features are mostly local features which need certain clustering procedure before they can be applied in object recognition.

Some related works can be found in the field of screen content recognition. Chen and Lin [5] proposed an automatic approach to segment and correct the deformation of the screen area for images of computer displays. The screen area is first detected by contour tracing and then the deformation of the screen image is corrected through affine transformation. Kastelan et al. [4] proposed a method which aims to extract the text shown on television screen. The method first extracted edges in the test image, then detected the regions of high edge densities as text regions. Wolf and Jolion [7] proposed a text extraction method from multimedia document. They used three constraints, included gray level, morphological, and geometrical constraints, to find text location on the screen. Karatzas and Antonacopoulos [8] proposed a split and merge method to extract text information from a web image. They applied this approach on HSV color space to extract specific color text in the web images.

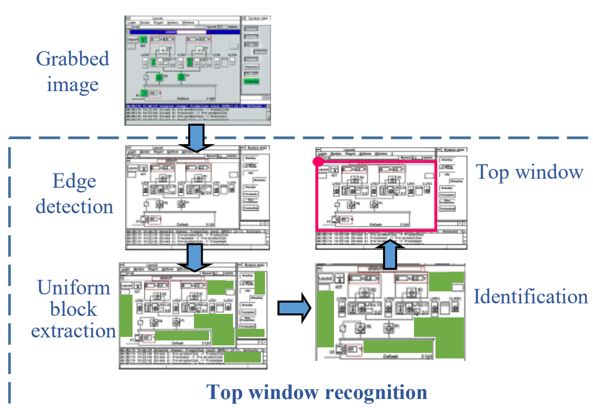


Fig. 1: System overview

While most of the existing window extraction approaches are based on the "foreground" information, which use a large number of explicit image features of window images, such as the edge or corner features mentioned above, we propose to utilize the “background” information (featureless regions) in the top window recognition process in this paper. In particular, we use featureless regions, i.e., uniform blocks (UBs), to recognize the top window displayed on the screen since such blocks can be identified robustly in very simple way. Experimental results indicate that the proposed approach, which has been deployed successfully in several industrial systems, can fully characterizes the internal structure of window images.a featureless region in an image, which is defined in this paper as a rectangular region with uniform color. Fig. 2 (a) shows an example of a test image with a single window. First, we perform edge detection to the test image and then binarize the edge information via thresholding. Then, we find a relatively large uniform region and extend the region until the region touch any edge point(s). We repeat the above procedure until these locally maximum UBs in the test image are found. The procedure of extracting the UBs from a grabbed image will be elaborated in the following.

The rest of the paper is organized as follows: In Section II, we illustrate the implementation details of extracting UB features from the grabbed images in the proposed monitoring system. In Section III, we present a technique to recognize the top window in a test image with multiple windows using these UBs. In Section IV, we present the experimental results of the proposed top window recognition method. Finally, in Section V, some concluding remarks are provided.

Extraction of Window Features

In this section, we present the technique to extract the UBs which will be used in the next section to recognize the top window in a test image with multiple windows. A UB is

Edge Detection

For edge detection, we adopt the forward-difference gradient algorithm to generate an edge image [6]. Consider a gray-scale image $f(x, y)$, the gradient magnitudes of f, denoted by $M(x, y)$, can be generated as

$$M(x,y) \approx |g_x| + |g_y| \tag{1}$$

with

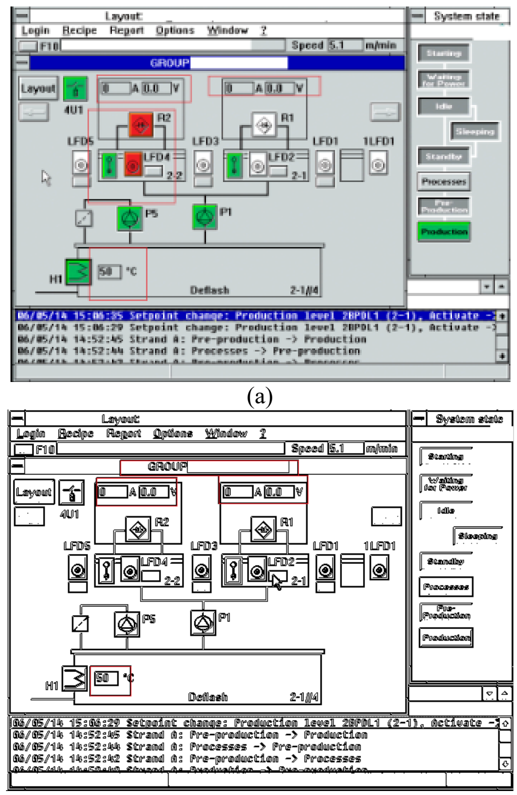
$$g_x = f(x+1, y) - f(x,y) \tag{2}$$

$$g_y = f(x, y+1) - f(x,y) \tag{3}$$

Since images in our system consist of RGB components, the representation of the gradient magnitude can be modified as:

$$M(x,y) = M_R(x,y) + M_G(x,y) + M_B(x,y) \tag{4}$$

where $M_R(x, y)$, $M_G(x, y)$, and $M_B(x, y)$ are the gradient magnitudes of R, G, and B components, respectively. An edge image can be obtained by applying simple thresholding to $M(x, y)$, as shown in Fig. 3, wherein a threshold of 60 is chosen empirically. Because of the rather simple situation involved with the grabbed images for industrial applications under consideration, the threshold can be determined quite easily.



(a)

(b)

Fig.2: An example of test image. (a)An original image. (b) The edge image of (a).

Extraction of UBs

In this paper, we only consider the recognition problem wherein there is only a single top window. From the observation of Fig. 3, an intuitive way of extracting such a top window may include: (i) detecting the title-bars of all windows, (ii) identifying those obtained in (i) which are not occluded, and (iii) identifying the window area associated with the title-bars found in (ii). However, one can see that such an intuitive approach cannot be implemented easily since the extraction of object features, e.g., the title bars, needed in (i)-(iii) cannot be performed in a simple way.

On the other hand, it is considered in this paper that the UBs, which can be identified as locally maximum rectangular background areas at different locations in Fig. 2(b), may be adopted as suitable features for the recognition of top window, with each often having fairly distinctive size, shape, and location. (In fact, in our application of factory automation, each window indeed has a unique set of UBs. For a top window with a non-unique set of UBs, additional but much simpler procedure can be adopted to determine its identity, e.g., by analyzing its title bar.Moreover, since it is assumed that there is only a single top window and all windows have fixed size and direction in our application, an extremely simple algorithm is developed in this paper to extract the UBs, as discussed in the next section.

To further simplify the extraction process, a relative large base block (BB) is selected so that only larger UBs will be considered. As illustrated in Fig.3, the procedure of the extraction of UBs, will include: (i) search for uniform region which is larger than the BB, (ii) extend the width of the block by adding columns containing no edge points, (iii) similar to (ii) but increase the height of the block with additional rows, (iv) label the area of the resultant UB as visited. The procedure then repeats to search for other UBs which do not contain edge pixels or any previously labeled areas. Thus, there is no overlap among derived UBs. Fig.4 shows the four UBs (in blue) thus extracted, with a 30×30 BB, for the edge image shown in Fig. 2(b).

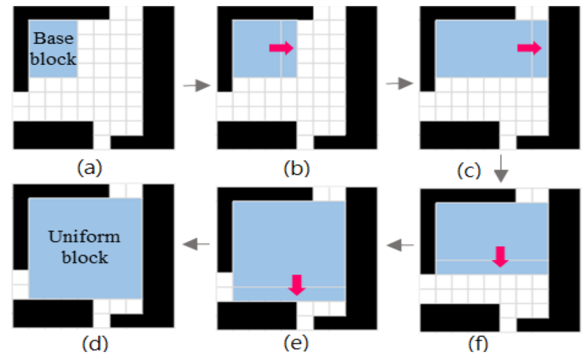


Fig.3: Extraction of UBs

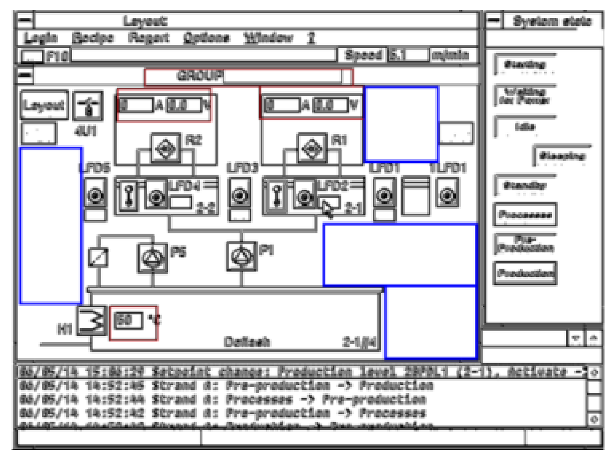


Fig.4: UBs extracted for the edge image shown in Fig.3.

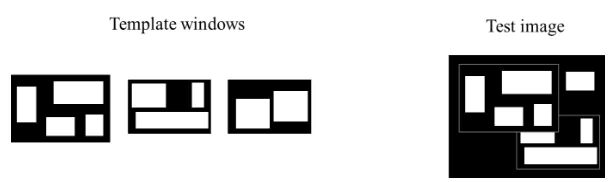


Fig 5: Block images of template windows and the test image.

Top Windows Recognition

In this section, we use a simple example to demonstrate the procedure of top window recognition and localization using the UB features described in the previous section. In particular, Fig.5 shows three template windows and a test image wherein the UBs are depicted as white blocks and gray lines are drawn to indicate the boundaries of windows in the test image. The proposed approach will first determine the corresponding template window for top window in the test image. Subsequently, the exact position of the top window is determined with respect to (anyone of) the UBs.

Identification

In this subsection, a simple block (template) matching technique is employed to identify the top window which is assumed to be free of occlusion in a test image. An overview of the central part of the matching process is shown in Fig.6, wherein the matching of each UB is performed by considering (i) its size, and (ii) its location with respect to other UBs in a template window. For (i), two UBs are simply regarded as having the same size if their widths and heights are similar, i.e., within certain threshold. The UBi and UBj are of similar sizes if

$$\left|W_j - W_i\right| < th_w \text{ and } \left|H_j - H_i\right| < th_h \tag{5}$$

where W and H represent the width and the height of blocks, respectively. Since the test image of our monitoring system is grabbed from the screen in electronic form, a very small value of 3 pixels is chosen for thw and thh. Fig.6 (upper right) shows two (light blue) candidate UBs thus identified for a (dark blue) UB in the test image. As for (ii), consider the template window shown in Fig.7. We define the relative location between two UBs as the position vector between their upper-left corners, as shown in Fig.7(a) for the (yellow) vector between UBi and UBj. Thus, the relative location of UBi with respect to other UBs in the window can be represented by the set of (four) vectors shown in Fig.7 (b). With the above definition of relative location of a UB, the identification the top window in the test image can be accomplished by first performing the following for each UB, denoted as Bi, of a template window:

Step 1. Identify candidate UBs of Bi in the test image using (i).

Step 2. For each candidate UB found in Step 1, obtain its relative location to all other UBs (as a set position vectors) in the test image, and count the vectors which are similar to those associated with Bi.

Step 3. From all candidate UBs of Bi, identify the one with the highest count found in Step 2 as the match of Bi and normalize the count as the percentage of all position vectors associated with Bi. If no candidate UB is found, a zero percentage is assigned to B_i.

For each template window, the above 3-step procedure will be performed for each of its UB, and the average of the percentage values obtained in Step 3 for all its UBs is calculated. Finally, the template window with the highest average value is identified as the top window.

In Step 3, similar to the size comparison of UBs, two position vectors are regarded as similar vectors if differences in their x- and y- components are less than a small threshold of 5 pixels. Fig.8 show an example of evaluating the similarity in relative location (represented by three red vectors) of a UB (Bi) and its candidate UBs (Ci and Cj). While 100% (3/3) similarity in location is obtained for Ci, 0% (0/3) similarity is obtained for Cj. Therefore, Ci is regarded as a match of Bi, as shown in Fig.6 (lower right). Fig.6 (lower left) shows an example of top window recognition result wherein the template window with four UBs is identified as the top window because the highest average percentage value (95%) is obtained in the above matching scheme.

Localization

After the top window is recognized, its location, i.e., its upper-left corner, in the test image need to be determined so that relevant image and text information contained in the window, which is needed in the task of automatic monitoring, can be obtained. Instead of locating the corner with certain features of the window image, as mentioned earlier, location of the UBs can be used to solve the problem easily. Ideally, the location of any UB of the top window in the test image can be used to specify the window location, using the position vector with respect to the window corner, if there is no error in the derivation of the size, shape, and the location of each of its UBs. In practice, to cope with possible errors in the image data, only the UB with the highest count in Step 3 of the above matching procedure, denoted as the *dominant* UB, will be used in the localization process.

Moreover, to ensure successful matching and localization results, a threshold of 90% of similar position vectors is required. (In the ideal case, all UBs will have the 100% match in position vectors and one of them can be selected arbitrarily for the localization.)

Fig.9 shows a simple way of locating the top window in the test image using the dominant UB. Let (x_{dom}, y_{dom}) and (x_i, y_i) denote locations of the dominant UB in the matched template window and the test image, respectively. It is easy to see that location of the top window, with the upper left corner located at (x_{main}, y_{main}), can be obtained by the following equations:

$$x_{main} = x_i - x_{dom} \quad (6)$$

$$y_{main} = y_i - y_{dom} \quad (7)$$

Experimental results

In this section, experiments are performed on two different industrial monitoring systems, denoted as System I and System II. The size of test images of two systems are 640×508 pixels and 1024×768 pixels, respectively, with both stored in BMP format, as shown in Fig.10. There are a total of 11 and 5 template windows associated with the two systems, respectively. And a total 36 and 22 test images are considered for the two systems, respectively. In order to protect the privacy of users of the proposed monitoring systems, certain text and graphical data are erased from the images depicted in this paper. Some examples of template and test images are shown in Fig. 11. The algorithm was implemented by Microsoft Visual C++ 2010 and executed on a notebook computer with i5-2419 2.3 GHz CPU and 4G RAM.

In general, BBs of different sizes and shapes can be employed to see their influence on the recognition accuracy of the top window. However, only square BBs are considered in the following for brevity. Fig.12 (first row) shows the total number of UBs obtained for various sizes of square BBs. It is readily observable that a larger BB will yield less UB features, as expected. However, (i) too many or (ii) too less UBs will actually result in degradations of the top window recognition, as shown in Fig.12 (second row). While the former corresponds to the generation of small, noise-like UBs which are less discriminative features, the latter, being less in amount, may be affected (blocked) easily by occlusion among windows.

On the other hand, it is apparent from Fig.12 that the proposed approach is very effective in recognizing the top window and can achieve 100% recognition rate, which is very important for industrial practice, for BB size ranging from 15 to 43 for System I (and 30 to 35 for System II). The smaller range obtained for System II may due to the fact that the image content of its windows are more complicated, e.g., in object shapes, as shown in Fig.10.

Conclusion and Future Work

As a vital part for industrial automatic monitoring systems, this paper presents the algorithm for top window recognition on screen-grabbed images. In this paper, we propose a method of top window recognition for industrial monitoring systems. For better performance in terms of accuracy, efficiency, and robustness, we introduce a novel, background-based image feature of window region, i.e., locally-maximum. uniform blocks extracted from regions of uniform color, to represent the practically unique spatial structure of background regions in different windows. Unlike foreground-based approach, which need to deal with more complex foreground features, the recognition procedure consists of several rather simple steps including (i) edge extraction, (ii) uniform block extraction, (iii) uniform block matching. The template window with highest matching score is then identified the top window, with its location determined simultaneously to facilitate further analysis of window content. Experimental results indicate that the proposed approach can achieve perfect recognition results for realistic image data obtained from two industrial monitoring systems, with different image resolutions and window appearance. Moreover, the proposed recognition algorithm can also achieve near real-time performance needed for on-line monitoring tasks. Using the proposed simple, yet effective, image features in similar monitoring systems which require the recognition of top window, as well as other object recognition applications, is currently under investigation.

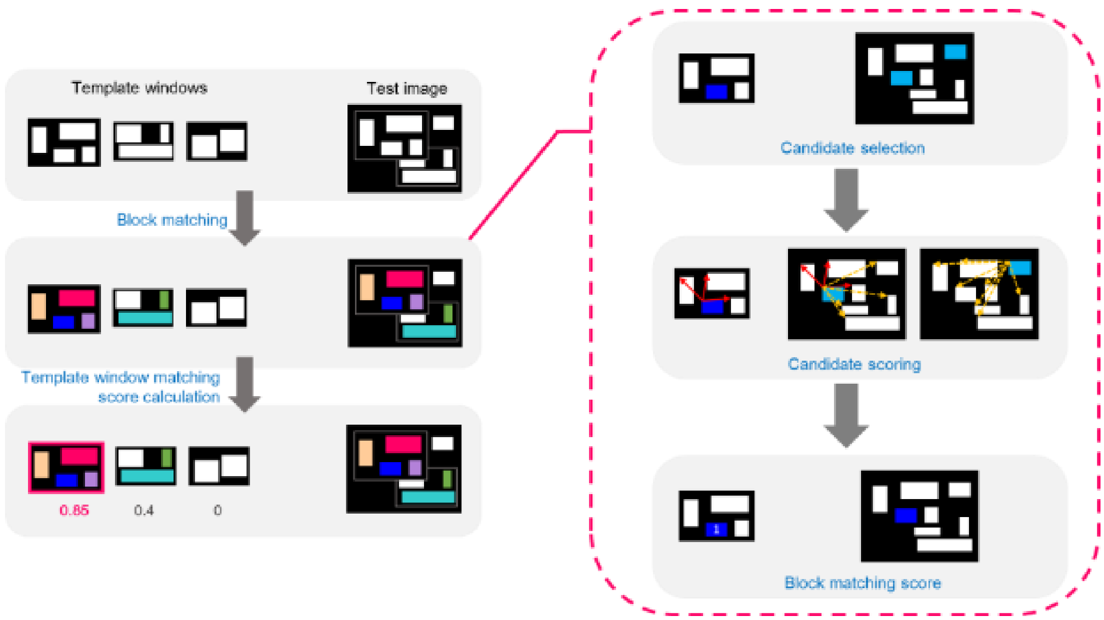


Fig. 6: Block matching for a template window, with detailed matching steps of the blue UB shown on the right (inside the dashed boundary)

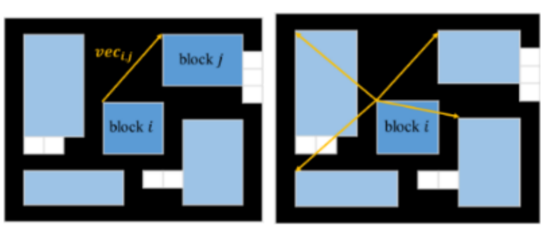


Fig. 7: A position vector (left) and the set of all position vectors of a UB of a template window (right).

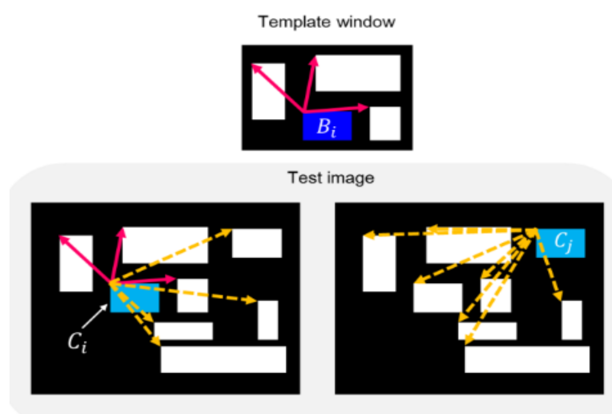


Fig. 8: An example of evaluating the similarity in relative location between a UB (B_i) and two candidate UBs (C_i and C_j).

According to our experiments, which are also omitted for brevity, if the recognition to the top window is correct, the accuracy in determining its location is always 100%. Furthermore, average execution rate of the proposed method is 12 images per second, which will certainly suit the need of various near real-time applications.

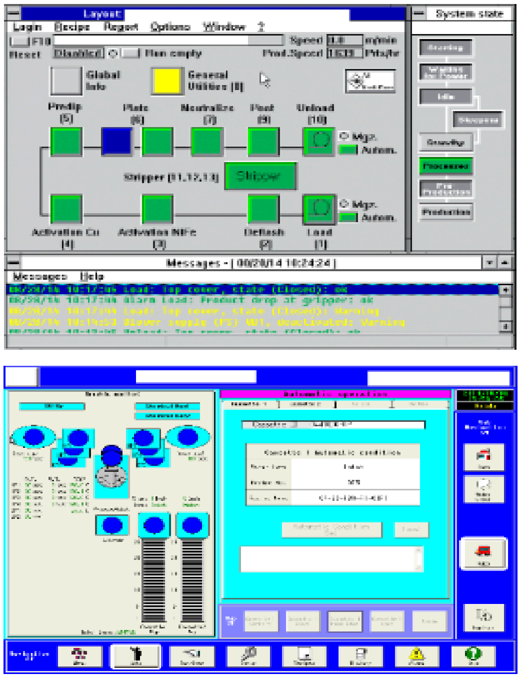


Fig. 10: Example images of system I (top) and system II (down).

References

[1] M. J. Swain and D. H. Ballard, “Color indexing,” *International Journal of Computer Vision*, Vol.7, pp. 11-32, 1991.

[2] R. Maina and H. Aggarwal, “Study and Comparison of Various Image Edge Detection Techniques,” *International Journal of Image Processing*, Vol.3, pp.1-11, 2009.

[3] C. Harris and M. Stephens, “A combined corner and edge detector,” *Alvey vision conference*, Vol.15, pp.147-151 1988.

[4] I. Kastelan, S. Kukolj, and V. Pekovic, “Extraction of text on TV screen using optical character recognition,” In *IEEE 10th Jubilee International Symposium on Intelligent Systems and Informatics*, pp.153-156, Subotica, 2012.

[5] Y. S. Chen and K. L. Lin, “Screen image segmentation and correction for a computer display,” *International Journal of Pattern Recognition and Artificial Intelligence*, Vol.28, No.4, 1454002-1- 1454002 -21, 2014.

[6] B. Jähne, *Digital image processing*, Springer-Verlag Berlin Heidelberg.

[7] C. Wolf and J.-M. Jolion, “Extraction and recognition of artificial text in multimedia documents,” *Formal Pattern Analysis & Applications*, vol. 6, issue 4, pp 309-326, Feb. 2004.

[8] D. Karatzas and A. Antonacopoulos, “Text extraction from web images based on a split-and-merge segmentation method using colour perception,” *ICPR 2004*, Cambridge, UK, August 2004, pp. 634-637.

Author Biography

Mei-Hsuan Ho received her B.S. degree in the Department of Computer Science and Information Engineering from National Taipei University, Taipei,

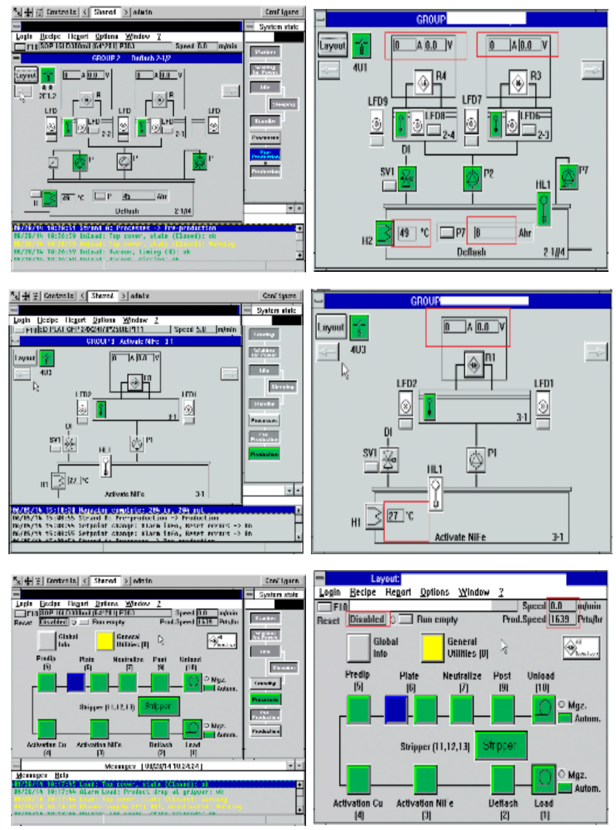


Fig. 11: The example images of test (left column) and template (right column) images.

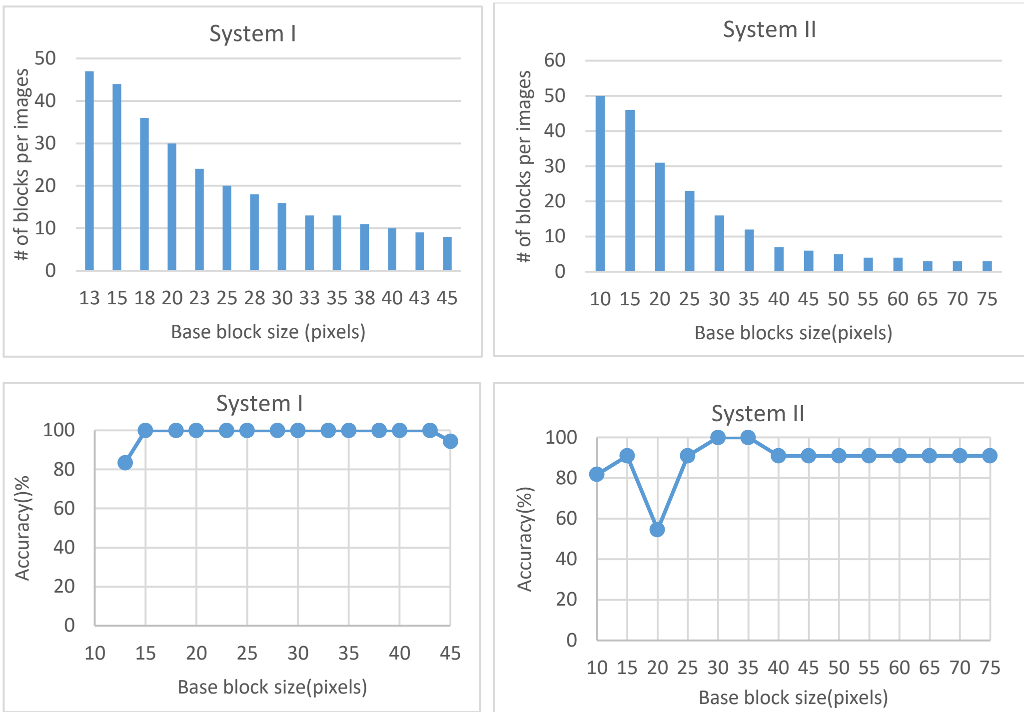


Fig.12. The numbers of uniform blocks (first row) and accuracies (second row) of two systems according to various base block si zes

Taiwan, in 2013, and M.S. degree in the Institute of Multimedia and Engineering from Chiao Tung University, HsinChu, Taiwan, in 2015. She is now working in the MediaTek Technologies, Co. as a software engineer. Her research topics focus on computer vision.

Yu-Shiuan Tsai received his BS and MS in Mathematics from the National Taiwan University, Taiwan, in 2002 and 2005, respectively. He received his Ph.D in Computer Science from the National Chiao-Tung University, Taiwan, in 2011. He is currently an Assistant Professor in the Department of Computer Science and Engineering, National Taiwan Ocean University. His current research interests include computer vision, big data analysis, and social networks.

Chang-Tao Hsu received his B.S. and M.S. degrees in the Department of Biomedical Engineering from Chung Yuan Christian University, Taiwan, in 1996 and 1998, respectively. He received his Ph.D. in the Department of Computer Science and Information Engineering at National Central University, Jhongli, Taiwan, in June 2016. His research interests include computer vision and image processing; especially in the computer vision system for advanced safety vehicle and computer vision techniques. He is now working in the National Chiao Tung University as a postdoctoral researcher.

Jen-Hui Chuang received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1980; the M.S. degree in electrical and computer engineering from University of California, Santa Barbara, CA, USA, in 1983; and the Ph.D. degree in electrical and computer engineering from University of Illinois at Urbana–Champaign, Champaign, IL, USA, in 1991.
Since 1991, he has been with the Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan, where he is currently a Professor and the Dean of the Computer Science College. His research interests include robotics, computer vision, 3D modeling, and image processing..