An Image Segmentation for the embedded Low Level CMOS Sensor Based on Image Enhancement

Chen-Chung Liu,
Department of Electronic Engineering,
National Chin-Yi University of technology,
ccl@ncut.edu.tw,

Chih-Chin Chang
Department of Electronic Engineering,
National Chin-Yi University of Technology
Linked_15@hotmail.com

Abstract— **In this paper, we propose an embedded system for extracting assigned objects from an RGB image taken by the low level CMOS lens, which is installed in the embedded system. The system uses ARM9 PXA270 as the development platform and combines image preprocessing, image enhancement technology and the fuzzy calculating method to extract the assign objects. The peripheral hardware of ARM9 PXA270 are key pad, four colors LCD, low level CMOS lens, and touch screen TFT-LCD. Although the system only use a low level CMOS sensor (cost cheaper), but it can take high quality images like a high level CMOS sensor (more expensive) done. One can just use a touch pen to draw a close contour on the touch screen of the embedded system to enclose the object that one wants to extract. Then the assigned objects are going to be extracted. In order to test our system's practical function, we have experimented with various sequences of color images. The experimental results proved that our system can quickly and efficiently extract the desired object from the RGB color image.**

Keywords— **fuzzy, ARM, PXA270, embedded system, image enhancement**

1 INTRODUCTION

In this article, we will talk about the utilization of image segmentation (or objects extraction) in 4 areas including: video object extraction [1], image compression [2], tracking systems [3], and pattern recognition [4]. In video object extraction, to extract the moving object from a video clip, image segmentation must be used to divide the video clip into several homogeneous regions (according to various parameters including colour, brightness, and saturation). In image compression, image segmentation is used to divide the input image into individual homogeneous regions to achieve better compression. In tracking systems, image segmentation is used to divide the homogenous regions of the traced object according to their color, luminance or texture. A merging technique based on motion estimation is then used to obtain the complete object in the next frame. In pattern recognition, to simplify the matching process, image segmentation system divides the object into sub-regions and characterizes each sub-region separately.

In recent years, closed circuit television systems (CCTV) are improved with each new day. The image recognition functions of nowadays closed circuit television systems are also be improved. The prices of the closed circuit television systems are also increasing sharply corresponding to the performance upgrading, such that the majority of community buildings or the schools use the closed circuit television systems with medium performance. Only some rich ones can use the infrared camera that can captures photos in the dark environment. Although the infrared camera can capture photos in the dark environment, but these images' color contrast is not obvious enough. In general, human is not easy to distinguish the foreground and the background from a complex background image with human eyes [5].The traditional closed circuit television systems always use personal computers to do the procession for these images captured by the surveillance systems. Personal computers should not work well when the input images are simultaneously transmitted from several closed circuit television systems. So, we propose a real time image enhancement technology in an embedded system; Let each surveillance image is processed by an Advanced RISC Machine (ARM) and then be continuously displayed with a TFT-LCD.

The embedded system has fast been developed and is already popularized to the human daily life. For example: The mobile, personal digital assistant (PDA), the electronic products for vehicle usage , the router, the digital printer, the fax machine…and so on are all the products of embedded system. Recent statistics shows that less than one percentage processors are used for the microprocessors of personal computers, and more than 99 percentage processors are used for embedded systems. This indicates the importance of the embedded system in nowadays human living. In the multitudinous applications of the processors, there is only a few processors are used in surveillance systems. We propose that an embedded surveillance system can improve the stability of the surveillance system and can reduce the equipment cost of the surveillance system.

The quality of the images taken with a low level CMOS sensor is frequently bad. The improvement of these low level images is almost handled by the image processing algorithm such as the brightness enhancement, contrast enhancement, image smoothing…, and so on. In the image contrast enhancement, the most popular method is the histogram equalization [6]. The histogram equalization has the drawback of time complexity and over enhancement, such that the histogram equalization method is not suitable to be used in an embedded system. We modify the image enhancement method that was proposed by Hwa-Hyun Cho [6]. The modified method has overcome the time complexity problem and the over enhancement problem. While the low level images are upgraded, then we can use a touch pen to select the objects displayed on the touch screen. Our system uses fuzzy algorithm [7] to divide the selected objects and the background. On the other hand, we also use MINIGUI [8] to create a graph user interface (GUI) and to display images. MINIGUI is an embedded system graph interface developed by Feynman Software company, MINIGUI also offers the full functions database of WIN32, such that programmer can easy to program his embedded system. In this paper, the experiment board is Creator-Xscale-PXA270, offered by Microtime company, and is shown in figure 1. The embedded operation system is Linux，the kernel is edition 2.6, and the images are displayed with a TFT-LCD.

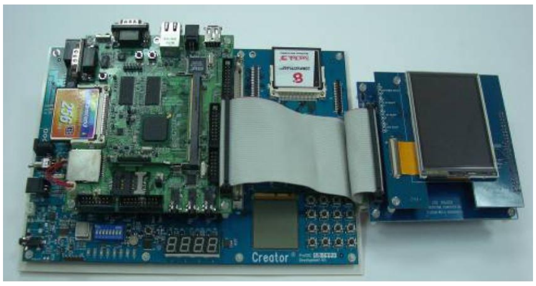


Fig.1. The Creator-Xscale-PXA270 experiment board offered by Microtime company.

The remainder of this paper is organized as follows: Section 2 presents the hardware peripherals architecture. The proposed color objection algorithm and its detail description are presented in Section 3. Empirical tests are presented in Section 4. Finally, we conclude this paper in Section 5.

2 Hardware Peripherals Architecture

Creator-Xscale-PXA270 has also provided very many hardware peripherals; they are shown in figure 2. They are the combination of key pad, four colors LCD and TFT-LCD, and CMOS Sensor. Where the CMOS Sensor is used for continuous images extracting, TFT-LCD is used for image displaying and image processing, key pad is used for the extraction of single frame image because extracting single frame image with key pad is more convenient than that with touch screen, and the four colors LCD is used to display the relative message of operation.

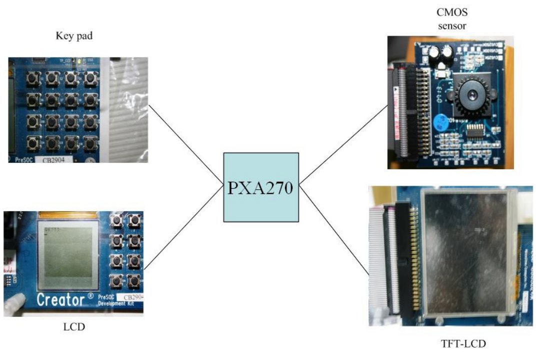


Fig. 2. The Creator-Xscale-PXA270 system and its hardware peripherals

3 The Image Extraction Algorithm

The overall segmentation procession of our proposed scheme is shown in Fig. 3. The input RGB color image is transformed into I color model and decomposed into H plane, S plane and I plane firstly. Owing to the H, S, I planes are measured with different units individually, they all need different methods for distinguishing the objects and the background. In order to create a precision color objects extraction scheme, the system also includes the brightness increasing of image, comparison enhancement of image, image smoothing, object range selecting, object extraction, and noise deleting. The details of our segmentation algorithm are described in the following.

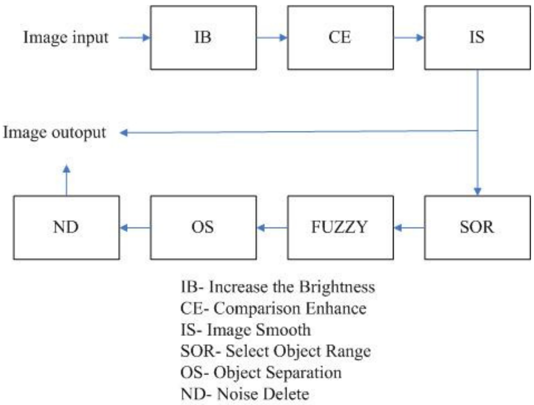


Fig.3. The flow chart of the object extraction system

3.1 Brightness Increasing of Image

The basic understanding of color perception of the human visual system shows the fact, which human eyes pick out the color of an object illuminated by white light is the reflection of selected wavelengths of light by that object [9]. The appearance color of an object can be considered as the object absorbing all colors except the colors that are reflected. Although there are an unlimited number of colors, colors of the range of human visual sensations can be produced by the mixtures of finite visual lights of various wavelengths. Human eyes have three cones shaped light detectors [10]. Each cone is sensitive to a special range of colors. One is sensitive to primarily red, a second to primarily green, and a third to primarily blue. The three sensitivity ranges overlap in fact. Each cone transmits a nerve pulse to the brain at a rate proportional to the light's intensity of that cone detects. By combining the signal rates transmitted from three cones, the brain attempts to conclude what color and brightness of the light must be. Thus, to combine various quantities of three colors (red, green and blue) of light can turn out all colors that can be perceived for human.

In televisions, computer monitors, and colored image projection systems, only the three colors are enough to adequately represent any of the unlimited visible colors [11]. For measuring or reproducing color, a number of three dimensional color models have been defined, among which most popularly used is RGB (Red, Green, and Blue) model. The RGB model is a physical system and the image in RGB model is the most suitable for the color representation of a color image. Most imaging devices (CCD cameras, CRT monitors, etc) mimic the retina and utilize three separate channels calibrated for the detection of red, green and blue phosphors to render colors at every pixel on the screen [12]. But it is not suitable for image processing applications, because its *R*, *G*, and *B* components are highly correlated. Besides, the distance in RGB color space does not stand for the perceptual difference in a uniform scale. In image processing and analysis, these *R*, *G*, and *B* components are often transformed into other color models. The I (Hue, Saturation, and Intensity) color model is the most representative of the perceptual systems, which is widely used in image processing. The advantages of I model are: hue and saturation are good correlation with the human perception of colors, and its separability of chromatic values from achromatic values. The I is the reducing redundancy models of the RGB model, its components H(hue), I(intensity), and S(saturation) are given by some color transform from RGB color space [13- 15]. The transformation from the RGB model into I model is represented by the following formulas,

$$H = \begin{cases} \theta & if \quad B \le G \\ 360 - \theta & if \quad B > G \end{cases} \qquad (1)$$

$$\theta = \cos^{-1}\left\{ \frac{1/2[(R-G)+(R-B)]}{[(R-G)^2+(R-B)(G-B)]^{1/2}} \right\} \qquad (2)$$

$$S = 1 - \frac{3[\min(R,G,B)]}{(R+G+B)} \qquad (3)$$

$$I = (R+G+B)/3 \qquad (4)$$

The I color model uses polar coordinates, as shown in Fig.4. In the double-cone shape figure, the saturation is proportional to radial distance (the center of the I disk corresponds to zero saturated colors) and it corresponds to the relative purity of a color. The hue is defined as an angle measured from the reference line drawn from the center of the I disk to the red vertex; it stands for the dominant wavelength in a mixed light and indicates a dominant color as perceived by human. Finally, the intensity or perceived lightness is the distance along the axis perpendicular to the I disk plane, and it indicates the brightness of a color.

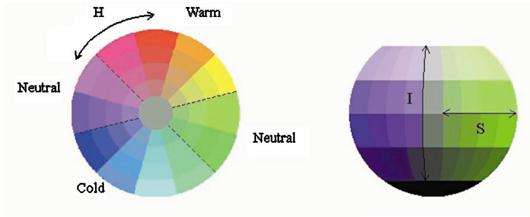


Fig.4.The I color model used polar coordinates.

Due to the CMIOS sensor in our system is a low level sensor, the image displayed in the screen is always almost in dark even we enlarge the exposure time as possible in a very bright environment. We can easy increase the brightness of an image just by adding a positive number K to the I component of the image, K is 30 in our experiment. Then, the modified I image is transformed into RGB model and we obtain a brighter image. Figure 5 shows the effect of the brightness increasing.

(a) (b)

Fig.5. The effect of the brightness increasing, (a) original image, (b) brightness increasing image.

3.2 The Enhancement of The Image Contrast

In some images transmitted from a closed circuit television system, the features of interest only occupy a relatively thin range of the gray scale. Contrast enhancement is a method to spread out the contrast of features of interest so that they take up a larger portion of the displayed gray level range without deforming to other features and the overall image quality. The goal of contrast enhancement techniques is to find a best transformation function for the original gray level to the displayed intensity, such that the contrast between adjacent structures in an image is displayed portrayed maximally. The histogram of an image is the relative frequency of gray levels occurrence in an image. Histogram procession modifies an image so that its histogram has a desired shape. This is useful in spreading out the low-contrast levels of an image in a narrow histogram, thereby achieving contrast enhancement. Image histogram equalization is a well-known automatic gray level correction which enhances efficiently image contrast by homogenizing its gray levels distribution. It is done by scaling each gray level with the cumulative histogram of the initial image. The histogram of a digital image with the gray levels in the range [0, *L*-1] is a discrete function as:

$$h(X_k) = n_k \qquad (5)$$

. Where X_k is the *k*-th gray level of the input and n_k is the number of the pixels in the image having gray level X_K. A normalized histogram is given by Eq.5, for k=0,1,…,L-1. Loosely speaking, $p(X_k)$ gives an estimate of the probability of occurrence of gray level X_k. Therefore Eq.6 called probability density function (PDF).

$$p(X_k) = \frac{h(X_k)}{n} = \frac{n_k}{n} \qquad (6)$$

Where, the sum of all components of a normalized histogram is equal to 1. Based Eq.6, the cumulative density function (CDF) as:

$$CDF(x) = \sum_{j=0}^{k} p(X_j) = \sum_{j=0}^{k} \frac{n_j}{n} \qquad (7)$$

Where x=0,1,…,L-1, $CDF(L-1) = 1$. Thus, a output (enhanced) image is obtained by mapping each pixel with level X_j in the input image into a corresponding pixel with level CDF(x) in output image. As shown in Fig.6.

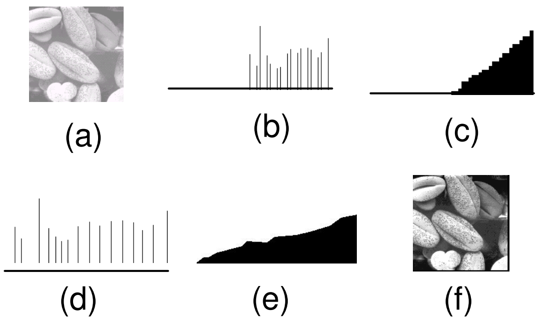


Fig.6. (a) an original image, (b) the histogram of (a), (c) the cumulative histogram of (a), (d) the histogram of figure (a) after equalization, (e) the cumulative histogram of figure (a) after equalization, (f) the contrast enhancement image from figure (a).

For embedded system, the time consumption of the histogram equalization method is too long. On the other hand, the histogram equalization method has the problem of over enhancement; it may equalize a dark region to be a over bright region, or equalize a bright region to be a over dark region. From figure 6(c) and figure 6(e), we can find that the CDF curve of the equalized image is smoother than that of original image. We use the characteristic of the CDF curve to do the image contrast enhancement, the detail steps are described in the follows.

Step 1:

Transfer the gray value [0,L-1] of each pixel of a gray image into exact one value of the five values 0, L/4, L/2, 3L/4, L-1with the following formula:

$$p'(i,j)=\begin{cases}0 & ,0\le p(i,j)<L/5\\ L/4 & ,L/5\le p(i,j)<2L/5\\ L/2 & ,2L/5\le p(i,j)<3L/5\\ 3L/4 & ,3L/5\le p(i,j)<4L/5\\ L-1 & ,4L/5\le p(i,j)<L-1\end{cases} \quad (8)$$

Then we find the PVD for each new pixel value.

Step 2:

Determine the image type to be a bright image or a dark image for each input image. If

$$|PDV(X_0 \cup X_{L/4})| > |PDV(X_{3L/4} \cup X_{L-1}| \quad (9)$$

, then the input image is a dark image. If

$$|PDV(X_0 \cup X_{L/4})| \le |PDV(X_{3L/4} \cup X_{L-1}| \quad (10)$$

, then the input image is a bright image.

Step 3:

For smoothing the CDF curve of the input image X, we minus a positive constant *y* from the CDF of X_i ,i= 0, L/4, L/2, 3L/4, L-1 for a input bright image, and plus a positive constant *y* from the CDF of X_i, i= 0, L/4, L/2, 3L/4, L-1 for a input dark image . This CDF curve smoothing method can not only reduce the time consumption but also avoid the over enhancement. The results of the CDF curve smoothing are shown in figure 7. The experiment results of contrast enhancement are shown in figure 8.

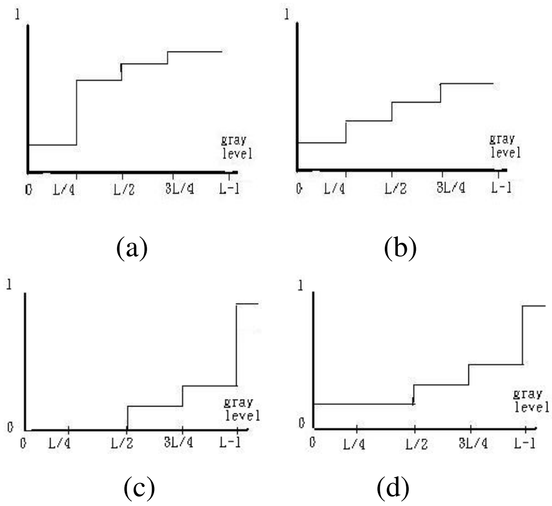


Fig. 7. (a) CDF curve of an original dark image, (b) CDF curve of the contrast enhancement dark image, (c) CDF curve of an original bright image, (e) CDF curve of the contrast enhancement bright image.

(a) (b)

Fig.8. (a) an original image, (b) the corresponding contrast enhancement image of (a).

3.3 Image Smoothing

CMOS uses the RGB color filters to extract the color information of an color image such that the input image of embedded system has many noises. These noises are going to be enlarged in the contrast enhancement step. We conduct the mean filter to filter out these noises to obtain a smooth contrast enhancement image, and the mean filter is described by the following equation,

$$p_s = \frac{p + \sum_{1}^{k} p_i}{k+1}, \qquad (11)$$

Where *P* is the original pixel value, *Ps* is the pixel value after smoothing, p_i is the pixel value of the *i-th* neighbor of *p,* and *k* is the number of the neighbors of *P*. The experiment results of image smoothing are shown in figure 9.

(a) (b)

Fig. 9. (a) an original input image, (b) the corresponding smoothing image of (a).

3.4 Objects Extracting

We display the image of a single frame on the touch screen from the input smoothed image sequence, and use the touch pen to draw a rough closed contour to enclose the object on the touch screen. We then extract 20*20 pixels p_f, $f = 1,2,...,400$, from the center of the closed contour. Finally, we use a fuzzy formula to determine whether the pixel *h* standing in the closed contour's inner part is a pixel of the object or not. Where the fuzzy formula is a negative exponential function;

$$u_c(h) = \exp(-a\,|\,h - h_f\,|) \qquad (12)$$

Where *h* is the chrominance of the testing pixel, h_f is the chrominance of p_f, and *a* is a positive empirical constant. If $u_c(h)$ is larger than 0.5, then *h* is a pixel of object, otherwise *h* is a pixel of background. The extracted object always has several noises, we can use the morphological erosion to filter out these noises.

(a) (b)

Fig.10. (a) the assigned object is enclosed by a closed line, (b) the extracted object flowerpot.

4. EXPERIMNET RESULTS

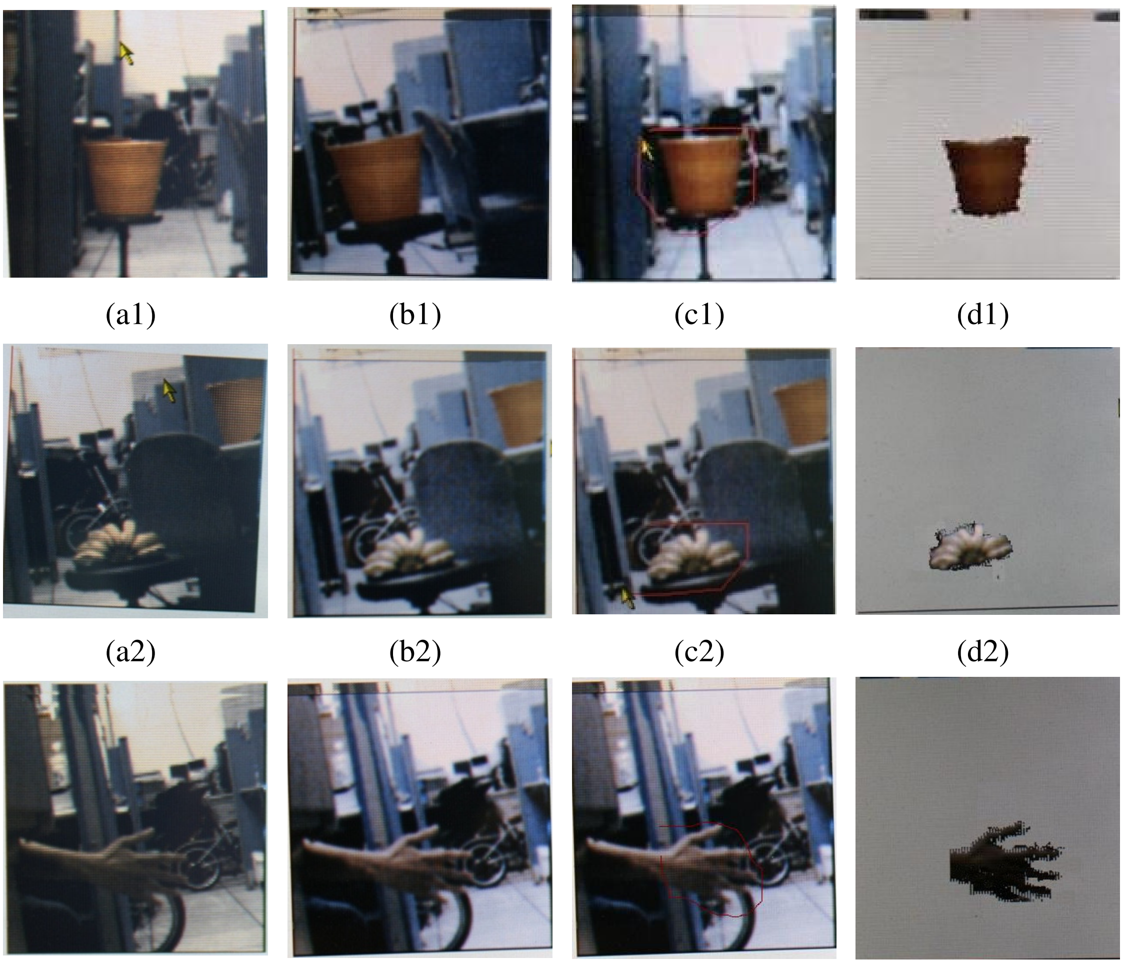
This experiment input image size is 320x240, and TFT-LCD shows size is 320x240 too, and to allow space for minigui components, so TFT-LCD actual shows size is 200x200. The time complexity of proposed method and other authors' methods is shown in Table 1, it shows that our method is the same as the Hwa-Hyun Cho method, and is simpler than the traditional method.

TABLE1

The complexity of each image enhancement method

Traditional method	**Hwa-Hyun Cho**	**Our method**
$O(n^2)$	$O(n)$	$O(n)$

Several experiment results are shown in Fig.11. (a) shows the original input image, (b) is the corresponding image that is obtained from enhancing (a), (c) indicates the assigned object which is closed by a close red, (d) shows the segmented result.



(a3) (b3) (c3) (d3)

Fig.11. experiment results: (a1) input image one, (b1) image one after enhancement, (c1) object (flowerpot) enclosed with a closed red line, (d1) extracted flowerpot. (a2) input image two, (b2) image two after enhancement, (c2) object (bananas) enclosed with a closed red line, (d2) extracted bananas. (a3) input image three, (b3) image three after enhancement, (c3) object (a hand) enclosed with a closed red line, (d3) extracted hand.

Several performance metrics for measuring the efficiency of a segmentation algorithm have been proposed. They are misclassification error (ME), relative foreground area error (RFAE) [16]. The misclassification error evaluates the inaccuracy of an algorithm; it is defined as in the following formula:

$$ME = 1 - \frac{TP + TN}{TP + FN + TN + FP} = \frac{FN + FP}{TP + FN + TN + FP} \quad (13)$$

, where TP Indicates the areas of true positive, TN represents the areas of true negative, FP indicates the areas of false positive, and FN represents the areas of false negative respectively. **Fig.12** shows the relations of TP, TN, FP, and FN areas in prediction and target area.

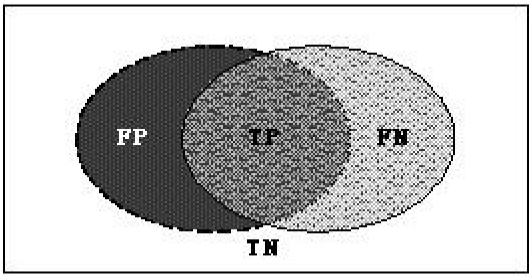


Fig.12. The relations of TP, TN, FP, and FN areas in prediction and target area.

The relative foreground area error (RFAE) evaluates the quality of an algorithm; it is defined as in the following formula:

$$RFAE = \begin{cases} \frac{TP+FN-(FP+TP)}{TP+FN} = \frac{FN-FP}{TP+FN}, \\ if\ (FP+TP) < (TP+FN) \\ \\ \frac{FP+FN-(TP+FN)}{FP+TP} = \frac{FP-FN}{FP+TP}, \\ if\ (FP+TP) \geq (TP+FN) \end{cases} \quad (14)$$

. Where FP+TP indicates the segmented object while TP+FN denotes the ground-truth object. It can be observed that RFAE is1 for the worst case of segmentation while it is zero for a perfect segmentation. In this paper, we take ME and RFAE performance metrics to measure the efficiency of we proposed segmentation algorithm. The segmentation results for several images in varying conditions are rearranged in Table 2.

TABLE2

The performance of proposed method

	Flowerpot	**Hand**	**Banana**
FP	27	31	23
FN	14	19	17
TP	2177	2397	1984
TN	2534	2686	2023
ME	0.0086	0.0097	0.0099
RFAE	0.0059	0.0049	0.003

The experiments were realized on a computer with a 624MHz Creator-Xscale-PXA270 processor and 64MB RAM, where the execution of the extraction procedure lasted between 3 and 25 seconds, depending on the size of the enclosed region and the size of extracted objects. After check Table 2, ME is less than 0.01 and RFAE is less than 0.04 in each case. This means that our proposed algorithm works with an efficient objects extraction even though the background complexity.

5. CONCLUSION

This paper purpose an efficient real time object extracting system based on contrast enhancement and Fuzzy segmentation. It is implemented in an embedded system. While it is applied in the Closed Circuit Television System (CCTV), then it not only can improve the problem of image contrast but also reduce the cost and space of CCTV. While it is applied in the mobile devices, the mobile devices can use lower grade CMOS sensor to take the same level images like as high level lens do to reduce the devices cost. . Use lower censor can take pictures like higher censor and to attain reduce cost goal. On the other hand, using ARM to develop CCTV is very suitable, because ARM not only can be embedded an operation system (OS) but also has a high degree of integration for peripheral hardware. ARM also has the watch dog function. While the system constructed with ARM cannot work abruptly or program is locked tightly, then the watch dog can reset the system. This is very important for closed circuit television system.

References

[1] H. Gao, W.-C. Siu, C.-H. Hou, *Improved techniques for automatic image segmentation*, IEEE Transactions on Circuits and Systems for Video Technology. 11(12) (2001).

[2] K. Belloulata, J. Konrad, *Fractal Image compression with region-based functionality*, IEEE Transactions on Image Processing. 11(4) (2002) 351-358.

[3] E. Ozyildiz, N. Krahnst, R. *Sharm, Adaptive texture and color segmentation for tracking moving objects*, Pattern Recognition. 35 (2002) 2013-2029.

[4] E. Navon, O. Miller, A. Averbuch, *Color image segmentation based on adaptive local thresholds*, Image and Vision Computing. 23(1) (2005) 69-85.

[5] B. A. Boghossian, S. A. Velastin, *Motion-based machine vision techniques for the management of large crowds*, Digital Object Identifier, 1999 Page(s): 961 - 964

[6] Hwa-Hyun Cho, Seong-Ho Kim, Tae-Kyung Cho, Myung-Ryul Choi, *Efficient Image Enhancement Technique by Decimation Method*, IEEE Transactions on Consumer Electronics, Vol. 51, No. 2, MAY 2005

[7] Shyang-Lih Chang, Li-Shien Chen, Yun-Chung Chung, Sei-Wan Chen, *Automatic License Plate Recognition*, IEEE Transactions On Intelligent transportation system, Vol. 5, No. 1, MARCH 2004

[8] FMsoft. (2008). Feynman Software. From http: //www. minigui. com/ index. Php ?id=home&L=0

[9] S. S. Deeb, *Genetics of variation in human color vision and the retinal cone mosaic*, Current Opinion in Genetics & Development.16 (3) (2006) 301-307.

[10] S. D.Buluswar, B. A. Draper, *Color Models for Outdoor Machine Vision*, Computer Vision and Image Understanding. 85 (2), (2002) 71-99.

[11] C. Clausen, H. Wechsler, *Color Image Compression Using PCA and Back propagation Learning*, pattern recognition. 33(2000) 555-1560.

[12] T. Bose, *Digital signal and image*

processing, J. Wiley & sons, New York, 2004.

[13] F. Ortiz, F. Torres, *Vectorial morphological reconstruction for brightness elimination in colour images*, Real-Time Imaging.10 (6) (2004) 379-387.

[14] Arthur, R.Weeks, *Fundamentals of Electronic Image Processing,* Spie Optical Engineering Press, Washington USA, 1998.

[15] F. Ortiz, F. Torres, Identification of citrus disease using color texture features and discriminant analysis, Computers and Electronics in Agriculture. 52 (1-2) (2006) 49- 59.

[16] M. Sezgin, B. Sankur, *Survey over image thresholding techniques and quantitative performance evaluation*, Journal of Electronic Imaging. 13(1) (2004) 146-165.