ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

IDENTIFICATION TOOL OF RUPIAH BANKNOTES FOR BLIND PEOPLE WITH AUDIO OUTPUT USING OPTICAL CHARACTER RECOGNITION (OCR) METHOD

Albert Gunadhi, Daniel Ivan, Diana Lestariningsih, Widya Andyardja and Yuliati
Department of Electrical Engineering, Faculty of Engineering, Widya Mandala Catholic University Surabaya,
Jl. Kalijudan Surabaya, Indonesia
E-Mail: albertgunadhi@gmail.com

ABSTRACT

Visually handicapped person apparently can't learn through some phisycal experiences or observations that required sight, in that case they have to get the information by direct access of another sensory perceptions such as touch or hear. Even in their financial management, visually handicapped persons are rely on somebody who has a normal vision whom they really trusted to organize their money. Therefore the writer brought up an idea to make a device whose its function is to identify a value of a money with audio output. It is expected that help visually handicapped person when have a transaction using cash effectively, quicker and also to avoid miscalculation. To identify a value of a money, the writer using numerical character as parameter. This device using image processing and optical character recognition as the method and also uses some of the important components such as windows tablet and webcam. Before it could recognize anything, the image have to have image pre-processing and some of segmentation to continue in the training character. After training character has done, next it will be continue with recognition process. When a character has been recognized, there will be a loop audio that mention a value of a money.

Keywords: identification device, optical character recognition, image processing.

1. INTRODUCTION

Rupiah banknote which widely distributed already has a code for visually handicapped/ blind people so they able to identify and recognize the value of money easily through touch sensory only if condition of the money is kind of new, but the fact is not every money that circulating around society are still smooth and new. If it happen, they will going to have a trouble to identify or need longer time. Generally in Indonesia, mostly in financial terms a blind people is rely on normal people who they believe to organize their finances. Besides, a blind people very rely on their memory to remember their money, also they often separate value of money into different pockets.

In this research, writer brought an idea to make a device whose its function is to identify a value of a banknote with audio output for blind/ visually handicapped people using optical character recognition (OCR) method. In our opinion this method is suitable to recognize numerical character that be visible on money. Those characters simply become parameter that we use to recognized the value of money. This device is expected to help visually handicapped even blind people when have a transaction using cash effectively, quicker, and also to avoid miscalculation. Problems that appear in this research are:

A. Use of OpenCV library for image processing.
B. Application of Tesseract OCR for conversion of characters of image into editable text.
C. Do some Tesseract training character.
D. How to generate a new language that contains characters which become identification parameter.
E. how to match the data that have been obtained with recorded sound file as an output.

In this research, writer has applied boundaries so it could be more specific and focused. The boundaries are:

a. The tool was made to identify the nominal value of money.
b. Banknote is assumed as real money.
c. Money that will be detected currently only for Rupiah banknotes with several conditions which are not torn out and visually still fit for use.
d. The tool only recognizes numeric characters as identification parameter.
e. Recognition accuracy target is 90%.

Purposed that want to be achieved in this research are:

a) Image processing program can run properly as expected.
b) *Optical Character Recognition* system can do training characters to generate a new language as well run recognition process.
c) Create an identification tool for blind people with audio output when the money has been successfully identified.

1.1 Computer vision: Optical character recognition

1.1.1 Computer vision [1]

Computer vision is one of branch in computer science which studies how computers can analyze and recognize objects observed in this case banknotes. Just like human vision, *computer vision*is expected to extract some information from seen object to perform certain tasks.

VOL. 11, NO. 13, JULY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Optical character recognition or OCR in this research is part of computer vision. In OCR, the computer asked to see the text contained in the image, and then describe seen object as an output text. To facilitate computer vision towards the object and describe seen object, computer vision combined image processing and pattern recognition. Image processing aims to generate a good image so the pattern objects which contained in image are easily recognizable. Only few topics of image processing theory will be discussed in this paper as guidelines such as:

1.1.2 RGB image to grayscale conversion [2]

Converting color images to grayscale are necessary in order to easier to remove noise on the image. Grayscale image is a digital image that has only one channel on each pixel value with 8 bit color depth. In the beginning the image is composed of three layers (layer) is R-layer, the G-layer and B-layer. To convert the color images that have a matrix of each value of r, g, and b intograyscale value gs, then the conversion can be done through the following equation:

$$gs = 0.299*r + 0.587*g + 0.114*b \tag{1}$$

with this gray color format the color red = green = blue, has been generated. this value is used to indicate the level of intensity so that colors that have this type of image is black, gray and white.

Figure-1. Result of RGB image to grayscale conversion.

Figure-2. Result of Otsu threshold on Rp 100.000.

1.1.3 Image smoothing [3]

Gaussian filter has been widely used in the field of image analysis, especially for the process of smoothing, blurring, eliminating detail and eliminates noise. Gaussian function of one-dimensional (1-D) can be expressed as:

$$G(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-x^2 / 2\sigma^2\right) \tag{2}$$

Meanwhile in 2-D form, Gaussian distribution is:

$$G(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-x^2 + y^2 / 2\sigma^2\right) \tag{3}$$

1.1.4 Otsu threshold [4]

Threshold is an image segmentation technique (grayscale/color) based on the value of the image and then converting it into a binary image (black and white) by changing each pixels according to whether the pixel is inside or outside a certain range. A certain range is the threshold value T. Threshold arrange all pixels that have intensity above threshold value to foreground which is white or equal to 255 (1) and all remaining pixels to background which is black (0).

Threshold values to be searched is expressed by K. K values ranging from 1 to L, with L up to 255. The probability of pixel i is expressed by

$$P_i = \frac{n_i}{N} \tag{4}$$

Which ni declare the number of pixels with gray level i, and N denotes the number of pixels in the image.

1.1.5 Optical character recognition (OCR) [5]

Optical Character Recognition (OCR) is a process of converting a scanned image into editable text. Scanned image is referred to in this research that the image / images acquired into the computer through the camera captured. This image contains characters, texts or symbols to be: processed by a computer, identified and then converted into character codes such as ASCII or Unicode other. After this conversion, the characters in the image is no longer shaped image that can't be edited, but instead has become a text that can be edited, copied, and used for any purpose. Figure-3 describes the general process performed by OCR.

In the OCR system, first thing to do is image preprocessing. Preprocessing on the image usually includes grayscaling, noise removal, and thresholding. After preprocessing completed the next step is the segmentation that aims to separate areas of observation. Next step is character normalization, which is a process to change region dimension of each characters, such as thickness of character, and then do feature extraction process to get features of each character in order to make difference between character. Last step of OCR process is recognition; in this step algorithm will compare features that want to be recognized with saved data.

1.1.6 Tesseract OCR

Tesseract is a open source system Optical Character Recognition (OCR) that can be used by various operating systems. OCR engine was originally developed at Hewlett-Packard (HP) between 1984 and 1994 (Smith, 2007).

Tesseract is designed to recognize white text on a black background or black text on a white background. This led to the design leads to a connected component

VOL. 11, NO. 13, JULY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

analysis / connected component (CC) and the operation of the outline components. The first step after the analysis of CC is to find a blob on the text region. A BLOB is a putative unit that can be classified, which could be one or more components that overlap horizontally.

After decided which outline would form the blob, text line seeker will detect text line based on nearby characters vertically in line of text.

Figure-3. OCR Process in general.

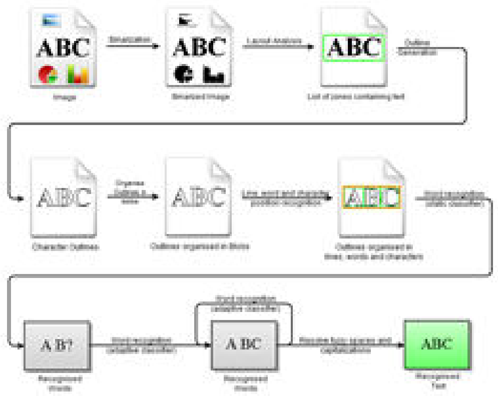


Figure-4. Block diagram Tesseract Word Recognition.

1.1.7 Unicode

Input image of analysis process is an alphabet or numeric character which will generate a code. This code will be use to search trained data in tesseract box files. Result of recognition are codes that recognized by engine. This code is related to one of the characters (letters / numbers) based on the character set used. Unicode is typically in the form of hexadecimal between 0 through 9 and A through F.

1.2 OpenCV [5]

OpenCV (Open Computer Vision) is a developer library developed by Intel Corporation. This library consists functions of computer vision and API (Application Programming Interface) for image processing and pattern recognition; both high level and low level, and as a real time optimization applications.Features of OpenCV are:

a. Manipulation of image data (memory allocation, copying, setting, conversion)

b. Image and video I / O (file and camera basedinput, image / video fileoutput)

c. Matrix and vector manipulation routines along linear alegabra (products, eigenvalue)

d. Imageprocessing fundamental (filtering, edge detection, color conversion, histogram)

e. Camera calibration (calibration patterns, estimation of the fundamental matrix, homography estimation, stereo correspondence) Analyst movement (opticalflow, segmentation, tracking).

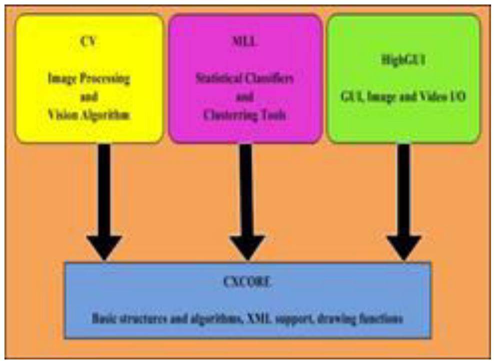


Figure-5. LibraryOpenCV structure.

Figure-6. Webcam M-Tech WB-200

Figure-5 shows the structure of OpenCV. The library consists of CV, MLL, HighGUI, CXCORE. Function library structure, among others:

a. CV: for image processing and algorithm of the vision.

b. MLL: a statistical classifiers and clustering tools

c. HighGUI: as a GUI (Graphical User Interface), Image and Video I/O

d. CXCORE: as the basic structure of the data, XML (Extensible Markup Language) support and graphics functions.

1.3 Webcam [6]

The camera is use to take images of objects, the result of the acquisition of the camera will be processed by image processing so that the result will be a binary image. Binary image necessary for data processing using optical character recognition could run optimally. These researches use an M-Tech WB-200 webcam.

This research uses a webcam for the sake of affordable price and easy implementation, so it will run with tablet properly.

Table-1. Webcam specification.

No.	Specification	Information
1	*Pixel*	5.0 MP
2	Resolution	640 x 480 *pixel*
3	Interface	USB 2.0
4	Framerate	15 fps
5	Compatible	Windows XP / 7 / 8
6	Dimesion	60 x 24.6 x 24.6 mm

1.4 Tablet Acer IconiaW4 [7]

Image processing can't use a microcontroller because data processing in the microcontroller can't run quickly so have to using the tablet as an interface for fast data processing as image processing in real time.

Figure-7. Tablet Acer Iconia W4.

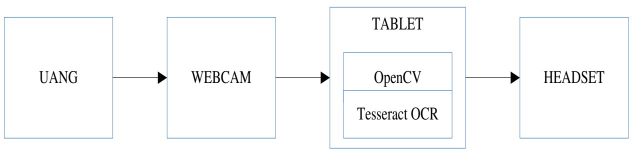


Figure-8. Device block diagram.

Figure-7 shown tablet Acer Iconia W4. Specifications of Acer Iconia W4 are shown in Table-2.

Table-2. Tablet specifications.

No.	Specification	Information
1	OS	Windows 8
2	Size	8 inches
3	System chip	Intel Atom Z3740
4	Processor	Quad-core 1.33 GHz
5	Graphic Processor	Intel Graphics Media Accelerator 3650
6	System Memory	2048 MB RAM
7	Battery Capacity	4960 mAh

2. METHOD

This chapter will discuss about design of the system both in hardware and software. In general, block diagram of system shown in Figure-8. Figure-8 represents a system of this banknotes identification tool in general. Image of banknote which is became identification object will be acquired by a webcam that connected to tablet as main processing site. Using OpenCV library, image that acquired from webcam will be processed. Steps of image processing will be shown in Figure-9. Color image will be converted to a grayscale image. If still there's a noise on the grayscale image, it will be reduced by Gaussian smoothing. And then will be continued with threshold. Purpose of this process is to segmented image based on threshold value image itself, then convert it into binary image by changing each pixels according to whether the pixel is inside or outside a certain range. A certain range is the threshold value T. Threshold arrange all pixels that have intensity above threshold value to foreground which is white or equal to 255 (1) and all remaining pixels to background which is black (0). Then to define the edges of objects usingCanny edge detection.

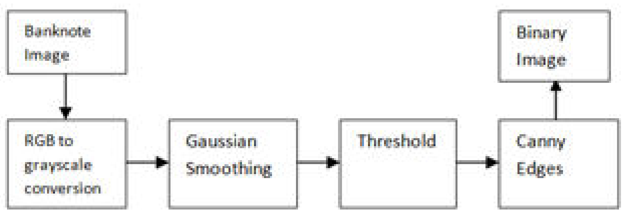


Figure-9. Block diagram of image processing.

2.1 Dimentional design

Dimension of this tool can be shown in figure 10. Dimension of this tool is 22.5 cm x 15.1 cm x 8.5 cm (length x width x height).

VOL. 11, NO. 13, JULY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

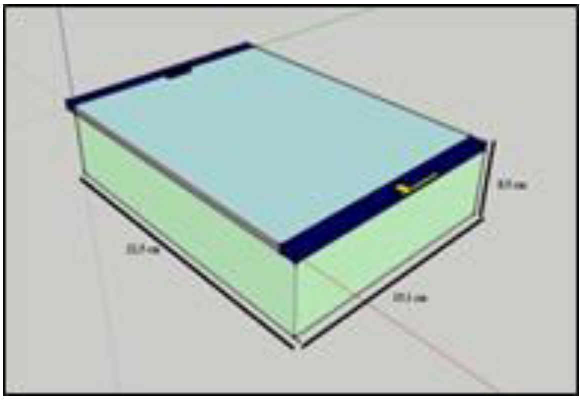


Figure-10. Tool's dimension.

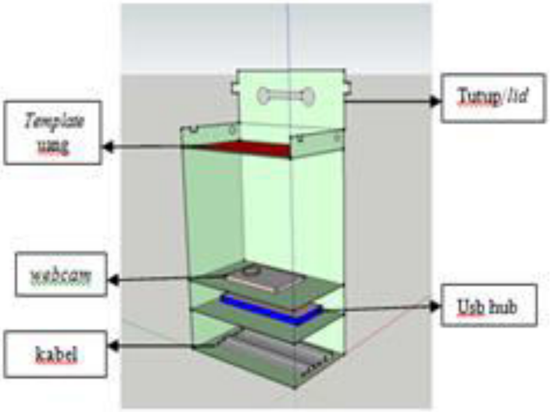


Figure-11. Rear view of the tool.

If seen from the rear, the casing consists of several panels, among others, a cable (base panel), a USB hub (second panel), a webcam (third panel), panel for template money, and cover / lid. The distance between the camera lens with a template of money made by 12 cm to optimize the sight distance between the camera and the object. Parameter identification were applied in this research is numerical characters contained in some parts on both sides. For example, ten thousand Rupiah banknotes shown in the following Figure-12.

Figure-12. Identification parameteronRp 10.000.

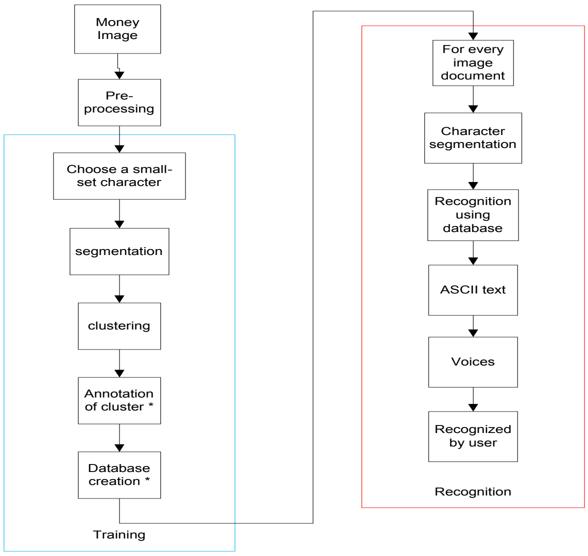


Figure-13. Block diagram of processed.

The position of the camera is placed at the edges so that the format of this parameter could be applied to all denominations of banknotes Rupiah. Data naming is:

a. *-amori represents first parameter.

2.2 Software design

Software design aims to simplify in the making of a identification device so that can be used well. Broadly speaking, the system is designed as shown in the following Figure-13.

Figure-13 is a block diagram of the processing of the object overall. In general, the program's success in converting the characters from the image (scanned image) to format text (editable text) for the further purpose divided into three sections including a character extraction, train OCR systems, and recognition (recognition).

At first, capture image is done by webcam and then pass into extraction process. Next step, image will be processed with image pre-processing. Pre-processing is the transformation of raw data input to assist computational capabilities and feature extraction as well as to reduce noise By using theory of color to grayscale conversion that shown in the equation 1, after converted into gray image then performed image smoothing. Smoothing method that used in this process is Gaussian blur/ Gaussian smoothing, and then will be continued with threshold. Once the image has become a binary image it will continued with the process of creating a set of characters based on each character. This character set can be formed from the contour and bounding box for each character.

VOL. 11, NO. 13, JULY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

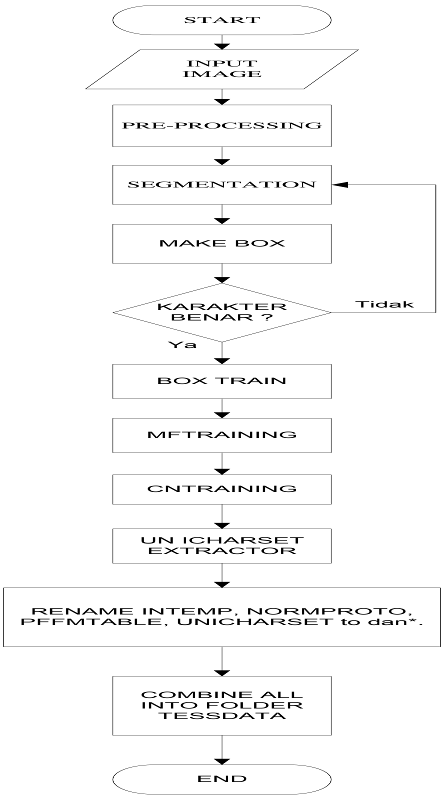


Figure-14. Flow chart of Training System.

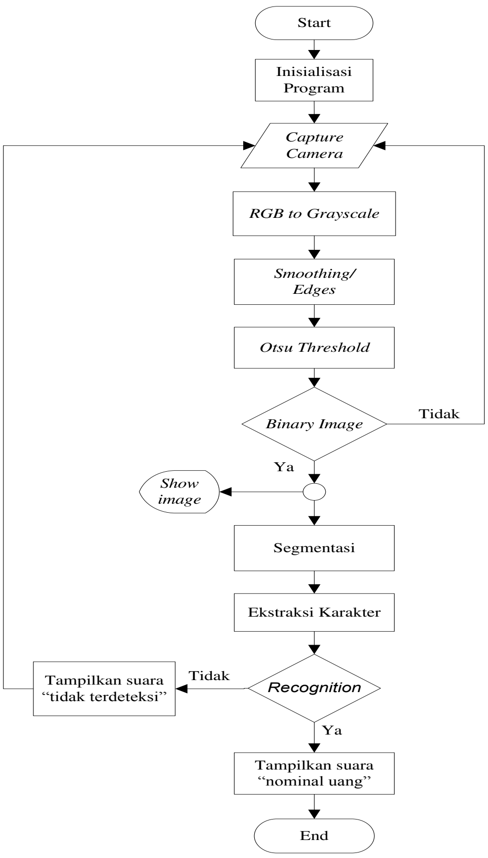


Figure-15. Flow diagram of Object Recognition.

System training procedure described in this following steps:

1. To generate training images, at least it needs 10 samples minimum for each character.
2. Make *boxfiles*. Box files is a text files contains list or character sequences of image due to train with 1 character for each line along with coordinates of bounding box. Scripts to make box in command line: tesseract 1000-amori.tif 1000-amor batch. nochop makebox

   Sometimes misread result of character will happened but it's quite normal, for example: read '1' as 'I', read '0' as 'O' or read '5' as 'S'. All we have to do is to edit wrong character until it become true, save it, and also change the extension with *.box. for example : 1000.box.

3. Next step is to run tesseract in training mode. Script to run training in command line is :
   tesseract 1000-amori.tif junk nobatch box.train

4. When all the characters from the extracted training process has been completed, the next step is to process cluster. Cluster process aims to create a prototype. Shape of the characters can be combined / clustering through mftraining and cntraining program..Scripts of both processes are :
   mftraining 1000-amori.tr 1000-bmori.tr ...

   cntraining 1000-amori.tr 1000-bmori.tr ...

   Mftraining program will generate output inttemp and pffmtable. Data inttemp contains the form of a prototype while pffmtable contains the number of characteristics / features expected in a character.

5. After the successful clustering next step is to generate the data unicharset use unicharset_extractor program.Script to run unicharset extractor is:
   Unicharset_extractor 1000-amori.box 1000-bmori.box...

6. After all the process is complete, copy the files inttemp, normproto, pffmtable, unicharset, DangAmbigs, freq-dawg, user-words, and word-dawg and rename it with additional information, for example: dan.normpoto, etc.
7. *Database*done.

Flow Diagram of Recognition System

Figure-15 shown a flow diagram using C++.C++do some initializing program such as variable declaration, GUI interface, adding OpenCV library and also Tesseract OCR library. *OpenCV library version 2.1 is used during the pre-process performed on the image, among others: RGB conversion to grayscale, smoothing, edges, and thresholding segmentation which is part of the character extraction.*

VOL. 11, NO. 13, JULY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Result from the extraction of character will continue with the process of recognition / recognition which uses a database that was created earlier. If recognition is successful, it will generate ASCII text data. ASCII data will be combined with data previously recorded sound.

3. EXPERIMENTAL RESULT AND DISCUSSIONS

Measuring and testing devices was performed to determine the performance of the tools that have been designed and made. Measurements and tests performed include:

- Testing Image Processing
- Testing thresholding on each parameter
- Testing Training Character
- Testing Recognition
- Testing Voice

3.1 Image processing experimental

Image processing experimental is necessary to examine if image processes that has been performed is able to generate a good image and also to reduce noises. In this case, noise may occur due to glare, uneven illumination, the template fingerprints or smudges on the money that led to recognition fail. Therefore, the dilation process is carried out to enhance the image and reduce noise output. Character recognition very influenced by the output image, the better the image and the greater the probability of recognition. The following table shows the comparison of the output image through the dilation process.

Table-3. Output image with dilation.

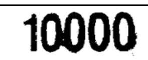
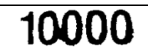
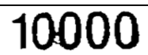
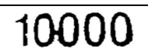
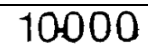
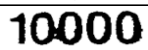
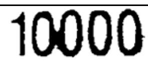
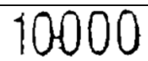
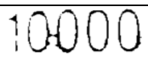
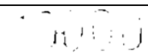
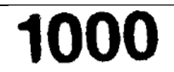
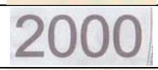
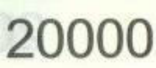
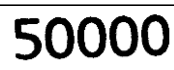
ITERASI	KERNEL	RESULT
1	2x2	
	3x3	
	4x4	
	5x5	
	6x6	
2	2x2	
	3x3	
	4x4	
	5x5	
	6x6	

Table-4. Threshold test results in each value.

Money	RGB Image	Binary Image
1.000		
2.000		
5.000		
10.000		
20.000		
50.000		
100.000		

From the table above can be explained that the process of dilation can reduce noise that appears significantly. In this research, the authors chose to use a 3x3 kernel with a number of iterations = 2 because it is more suitable when applied in all value for money that successful recognition.

3.2 Experimental result of thresholding

Thresholding testing performed to see if the threshold value has been able to convert all of the colors on the identification parameters of banknotes.

From the results in Table-4, it can be seen that the threshold value can be used on all parameters rupiah denomination banknotes. Good image conversion results will largely determine the quality of the training data and character recognition.

3.3 Training character

Testing character training performed in several stages. The following stages of testing character training.

a. Makebox testing

This process is creation of a file box containing the characters from the image of the first character of each row is set along with the coordinates of the bounding box of each character. This step is done once according to the number of parameters in each value of banknotes that will produce data text that contains characters from the object. After all makebox process applied to each fragment, then edit the characters are less precise.

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Table-5. MakeBox test result.

Uang parameter	Makebox result	Edited	Boundingbox			
1.000-amori	1	1	37	37	120	203
	0	0	139	38	268	208
	0	0	278	41	407	211
	0	0	414	43	542	210
2.000-amori	2	2	38	38	158	206
	0	0	169	40	292	208
	0	0	27	41	420	213
	0	0	431	47	551	218
5.000-amori	S	5	8	13	156	214
	O	0	159	18	301	213
	O	0	307	16	445	212
	O	0	451	17	583	207
10.000-amori	1	1	53	37	102	200
	0	0	122	41	206	206
	0	0	224	44	307	209
	0	0	325	50	410	214
	O	0	430	54	514	219
20.000-amori	2	2	24	26	128	209
	0	0	140	32	245	215
	0	0	254	37	360	218
	0	0	368	42	472	223
	O	0	480	49	584	224
50.000-amori	S	5	38	42	133	203
	O	0	142	42	244	206
	O	0	253	46	352	208
	O	0	359	48	460	211
	O	0	470	52	569	214
100.000-amori	1	1	35	38	69	185
	0	0	88	37	169	190
	0	0	189	35	269	187
	0	0	284	34	364	186
	O	0	383	34	461	181
	O	0	477	31	557	179

b. Training box test

All data characters have been corrected subsequently done the training process.

Table-6. Training box test result.

Uang/ parameter	Tess training tester	***Recognize*time** (ms)
1000-bmori	1 0 0 0	31.712
2000-bmori	2 0 0 0	51.288
5000-bmori	5 0 0 0	42.778
10000-bmori	1 0 0 0 0	47.701
20000-bmori	2 0 0 0 0	79.419
50000-bmori	5 0 0 0 0	41.432
100000-bmori	1 0 0 0 0 0	96.446

Table-6 shown training box result only takes time less than 10 seconds on parameter 1.

3.4 Pengujian recognition

Testing recognition / identification on any denominations with a data sample as much as ten data every denomination. Samples used in the testing can be seen in appendix 2. This recognition test results shown in Figure-16.

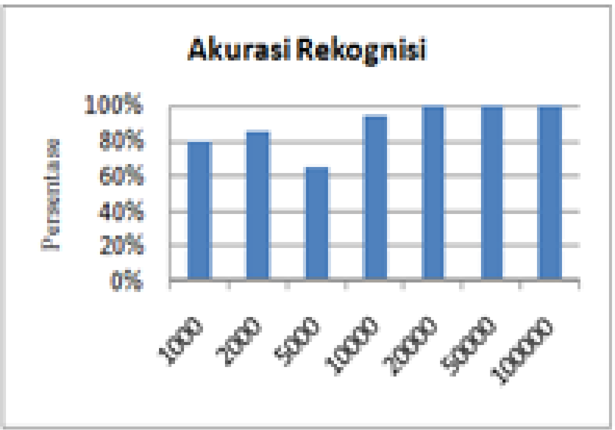


Figure-16. Recognition accuracy.

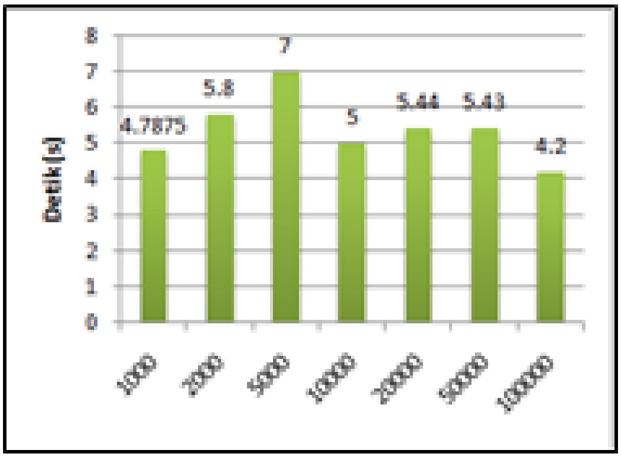


Figure-17. Time of audio output.

There are seven different value in Indonesia Rupiah Banknote. In this research writer takes 10 samples for each value, so there are total 70 samples. Samples were sorted based on visual condition in the order 1 for really shabby banknotes up to 10 for relatively new condition. Evaluation of recognition accuracy will done per character as well as the processing time required. There are 330 characters for all samples, whereas there are 31 false recognition characters. Thus we can conclude character recognition accuracy for all samples is sebesar$\frac{330-31}{330}$x 100 % = 90.60%. Recognition errors may occurred due to condition of the sample had worn so reading of numeric characters in certain parts of the parameters resulted in the introduction of failure.

Time recognition process for all samples take average $\frac{2.5577}{70}$= 0.03653 detikThus it can be seen that the average time required to identify each one character is 0.00775 seconds.

3.5 Audio test

Apart from recognition, a measure of the test success is tool able to generate a sound with a maximum time limit of sixty seconds since money was put on the template. This recognition test results shown in Figure-17 is the average time it takes to identify and make a sound.

From the graphic above most of samples tested by identification tool can mention value of money within a period of less than 10 seconds. From the analysis and observation, whensample is entered; the tool can mention the value of money, but havinga bit problems when there is glare or uneven illumination while the program is running. This can be solved by adding morphology in image processing program and set the number of kernel and iterations that is suitable for the entire samples.

4. CONCLUSIONS

From the results of the design, test and measurement, it can be concluded as follow matters;

a) Numerical characters recognition on a contrast section that becomes a parameter identification using OCR method has an accuracy rate of 90.60% for sample testing money with a very shabby condition until relatively new condition.

b) Rate of processing time required to identify each sample with 1 strategy is 0.0365 seconds. While time is needed to mention the value of the money <10 seconds when the money is inserted into the template with the perfect position.

c) Dilation program which is part of morphology in image processing is very influential in reducing noise that appears when recognition process was running.

REFERENCES

[1] Priska A. 2012. PerancanganAplikasi Pengenalan Karakter Korea pada Platform Android menggunakan Metode OCR Adaptive Classifier. http://www.thesis.binus.ac.id/Doc/bab2/2012-00525-mtif%20bab2.pdf, diakses tanggal 14 Oktober 2014.

[2] Putra Darma. 2009. Sistem Biometrika, CV. ANDI OFFSET, Yogyakarta.

[3] Anil K. J. 1991. Fundamental of Digital Image Processing, Prentice Hall.

[4] Somers H. 2003. Computers and Translation, John Benjamins Publishing Co., Amsterdam

[5] Unknown. 2014. OpenCV Archive. www.opencv.willowgarage.com, diakses 20 December 2014.

[6] Unknown. 2015. Webcam M-Tech. http://www.alnect.net/products.php?/2/71/155/1642/Accessories-Parts/Computer-Accessories/Webcam/Webcam-M-Tech-WB-200, diakses 1 Maret 2015.

[7] Acer. 2015. Acer Iconia W4 Series, www.acer.com/sg/ac/en/content/, diaksestanggal 27 Februari 2015.

[8] Smith Ray. 2007. An Overview of the Tesseract OCR Engine. Proc 9th Intl. Conf. on Document Analysis and Recognition. 2: 629-633, Google Inc., California.

[9] Smith R., Antonova D., Lee D. 2009. Adapting the Tesseract Open Source OCR Engine for Multilingual OCR. MOCR 9thProc Intl. Workshop on Multilingual OCR, Artcl. 1, Google Inc., California.