International Journal of Soft Computing 11 (2): 64-69, 2016
ISSN: 1816-9503

Identifying of the Space Color CIELab for the Balinese Papyrus Characters

Made Sudarma
Department of Electrical and Computer System Engineering,
Faculty of Computer System and Informatics, Udayana University,
Jimbaran Campuz, Bali, Indonesia

Abstract: The Balinese papyrus is one of the media to write the ideas from minstrels in ancient times. Currently, many of ancient literature that written in papyrus very difficult to identify because the writings were beginning to rot or fade influenced by age. This study takes the Balinese papyrus characters as its object. The improvement of the image quality in image processing of the characters here refers to the change for the quality of the image before its next development. There is the process of identifying the color space to spot the shape of the papyrus characters. The difference of the colors is so small so that it needs CIELab to process the identification out of its background and noise.

Key words: Balinese lontar, papyrus character, image improvement, noise, CIELab

INTRODUCTION

The preservation of the Balinese culture in the middle of the globalization is seen to be necessary. However, most of the Balinese young generation seems to be neglect it as they are trying things to be more updated. This is one of the recent problems that needs a quick response. In this case, the local government has been stepping up to their best to handle it. One of the steps taken seriously is preserving the traditional Balinese script, namely by papyrus digitizing. The digitizing of Balinese papyrus refers to taking and transforming of the papyrus image to the image file (Sudarma and Sutramiani, 2014). Moreover, the quality of the image of papyrus in files can be improved before they are processed further to the quality of readable and be translated into a different code.

Papyrus is a literature which has good quality, fine and beautiful that often teaches us about life philosophy, religion and many other knowledge which is very useful for us as a provision of life, so that we as young generation it is very important to understand and take care of papyrus. But the problem in understanding papyrus is that papyrus is considered sacred for Hindus in Bali. Not everyone can read and understand papyrus, considering the difficulty to read Balinese scripts on the papyrus. Besides for longer period the papyrus will be susceptible for damage.

This study is focusing on the identification of the color space of the image of papyrus, namely to distinguish the color of the characters from its background in which it is so slim that it takes CIELab to do it in order to reach the good outcome (Putra, 2010).

Former studies on papyrus that have been learned include the study to remove noise from the papyrus and to segment color images using CIELab (Prapitasari, 2012; Bansal and Aggarwal, 2011).

Literature review

Lontar: One of the preserved heritage of our ancestor is their manuscripts. The ancient manuscripts in various local languages have been found across the archipelago. Most of them are stored by common people. Some are preserved well in the local governmental agencies and traditional communities.

One of such manuscripts is lontar that are mostly found in Bali besides in Java, Lombok and Sulawesi islands. Lontar were used as the place for writing before paper was produced. In Java besides lontar used for the same purpose included nipa leaf, dluwang (the skin of kind of wood) and perkamen (skin of goat). In Sulawesi people used bamboo and rotan. In Batak, the wood skin used to write is called tribak (Fig. 1).

One of our ancestor's cultural legacies which have important value is ancient manuscripts. All over Indonesia, it is known that there are lots of ancient manuscripts written in various scripts and languages. Most of the manuscripts still stored or owned by ordinary people. The others are existed in central and regional agencies and traditional institutions.

Balinese script: According to the Pasamuhan Agung, the spelling system of Balinese characters in Latin letters should be in accordance with the Indonesian spelling system. To the regulation the following equivalents between the languages are established.

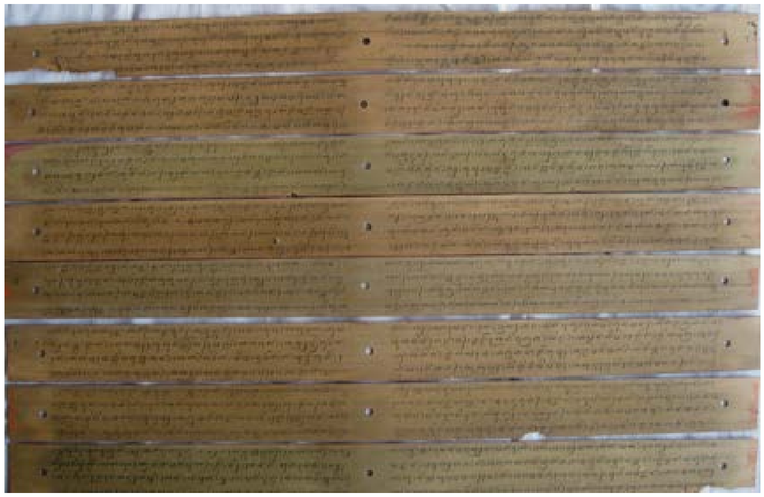

Fig. 1: The Lontar of dewa segening negara

Table 1: The balinese voice scripts

Number	Balinese script	Latin script
[illegible]	[illegible]	a
[illegible]	[illegible]	ê
[illegible]	[illegible]	i
[illegible]	[illegible]	u
[illegible]	[illegible]	e
[illegible]	[illegible]	o

Table 2: The substitution of voice scripts

Voice script	Replaced
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]

Balinese vowels: a, e, i, u, e. o (6 has been changed to pepet and taling) (Table 1). The changing transcription of the Balinese vowels into the Latin system as detailed:

- Wisarga ha (with its attachment) is to be written as the same as a kara
- Suara dirga (ulu sari, suku ilut, etc.) is to be written as the same as aksara hrasua
- The pepet and taling are equalized
- Dipthongs: ai/ia = e (taling) and au/ua = o (should be a)

According to the Balinese scholars, the words spelled in wisarga ha characters are in general of Bali origins or at least considered so. Whereas the ones in a kara are not known as from the language of either Jawa Kuna or Sanskerta. Table 2 shows the change.

Consonant: h, n, c, r, k, g, t, m, ng, b, s, w, l, p, d, j, y, ny (18 characters) (Table 3):

Table 3: Wianjana script

Number	Balinese characters	Balinese/Latin
[illegible]	[illegible]	h/a
[illegible]	[illegible]	n
[illegible]	[illegible]	c
[illegible]	[illegible]	r
[illegible]	[illegible]	k
[illegible]	[illegible]	d
[illegible]	[illegible]	t
[illegible]	[illegible]	s
[illegible]	[illegible]	w
[illegible]	[illegible]	l
[illegible]	[illegible]	m
[illegible]	[illegible]	g
[illegible]	[illegible]	b
[illegible]	[illegible]	ng
[illegible]	[illegible]	p
[illegible]	[illegible]	j
[illegible]	[illegible]	y
[illegible]	[illegible]	ny

Tabel 4: The wianjana script

Characters	Names	Terms
[illegible]	na-rambat	Murdania
[illegible]	ca-laca	talawia mahaprana
[illegible]	ga-gora	kantia mahaprana
[illegible]	ta-tawa	dantia mahaprana
[illegible]	ta-latik	Murdania
[illegible]	ba-kembang	ostia mahaprana
[illegible]	sa-sapa	Murdania
[illegible]	sa-saga	Talawia
[illegible]	pa-kapal	ostia mahaprana
[illegible]	da-madu	dantia mahaprana

- The number of Balinese characters in Latin is the same as the numbers of ha, na, ca, ra, ka, Bali (18 characters)
- The writing

[illegible] (da madu, murdania) has no particular form which is different from the Old Javanese in Latin. The names such as dantia and murdania are often regarded misguiding. It is expected that the names should represent the articulation area of the characters (Table 4).

Pengangge: Pangangge is the character that cannot be used itself. They should be put together to the vowels or consonants for their functions to change the sounds of them. The kinds are called, among them, pangangge suara, pangangge tengenan and pangangge aksara.

CIELab: CIELab is one of the color structures defined as CIE (Commicion International de 1'Eclairage/The International Commission on Illumination in 1976 (CIE 1976 L*a*b*). According to CIELab, the amount of CIE_L* to describe the brightness of the color is 0 for black and L* = 100 for white. CIE_a* dimension describes the colors of green red in which negative a* indicates green whereas CIE_a* positive indicates red. CIE_b* dimension for blue yellow is the negative b* indicates blue and CIE_b* positive indicates yellow. The transformation of RGB (Red, Green, Blue) CIELab can be done as described (Hoffmann, 2010):

$$C = CR \tag{1}$$

In which:

$$C = C^{'G} \tag{2}$$

G = 2.2 and C = R, G, B:

$$X_1 = X/X_n \tag{3}$$

$$Y_1 = Y/Y_n \tag{4}$$

$$Z_1 = Z/Z_n \tag{5}$$

$$X_1 = \begin{cases} X_1^{1/3} \text{ jika } X_1 > 0.008856 \\ 7.787X_1 + \dfrac{16}{116} \end{cases} \tag{6}$$

$$Y_1 = \begin{cases} Y_1^{1/3} \text{ jika } Y_1 > 0.008856 \\ 7.787Y_1 + \dfrac{16}{116} \end{cases} \tag{7}$$

$$Z_1 = \begin{cases} Z_1^{1/3} \text{ jika } Z_1 > 0.008856 \\ 7.787Z_1 + \dfrac{16}{116} \end{cases} \tag{8}$$

then L*a*b* to be:

$$L^* = 116Y_1 - 16 \tag{9}$$

$$a^* = 500\,(X_1 - Y_1) \tag{10}$$

$$b^* = 200\,(Y_1 - Z_1) \tag{11}$$

MATERIALS AND METHODS

Data acquisition: Data acquisition is the process of acquiring data from analog to digital, namely from the physical papyrus to become papyrus image by using scanner. Thus, first, the Balinese Lontar Bali is scanned to obtain the image. When it is done, the image is stored as a computer data before being further analysed in the processing stage.

Pre-processing

Cropping and papyrus image merging: The right, left and middle image of the lontar obtained from the scanning is cropped for the further stage. The Balinese papyrus usually has two parts. The writing on the first line is continued in the second one as can be seen in Fig. 2.

The yellow lines are the border of the image that are to be cropped. To crop, the left line is first, middle and then the right. Being cropped to be two parts, the image will be merged to be one in Fig. 3.

The color space of CIELab: On the image obtained after the cropping, the color space identification is done through CIELab. CIELab is a color model in which the L* component is a perception of the light intensity a* and b* is the component showing the number of the colors displayed (Bansal and Aggarwal, 2011). The CIELab is

Fig. 2: The image of papyrus

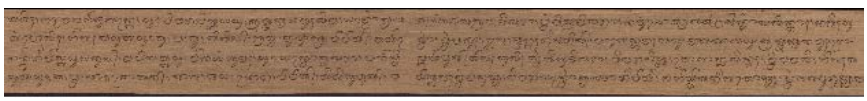

Fig. 3: The papyrus image after cropping

used because it is not all of the colors on papyrus are different from the background. Sometimes the colors are different very litlle because during the writing the papyrus writers do not use the natural types (burned candlenut) for the final processing. The CIELab is the very effective method to do the separation (Putra, 2010; Mamatha and Srikantamurthy, 2012).

The color of the image obtained is of RGB from which it is necessary to alter it into the CIELab color space using Eq. 1-11. RGB is a three dimensional color structure used for the true color type on digital image in which R is for Red, G for Green and B for Blue. Whereas CIELab (L*a*b) is the image color structure that represents L for lighting and a for red or green and b for yellow or blue (Mamatha and Srikantamurthy, 2012). The transformation produces different color structure of the image from the original one.

RESULTS AND DISCUSSION

The experiment in this study is about how to identify or spot the color space of the papyrus using CIELab. Figure 4 is the image of the printscreen system that used to the identification.

On this early display there are several menus. The one used in this experiment is the pre-processing CIELab. For inputting the image, the picture (image) search can be used. After clicking on the image search, the computer will show such as Fig. 5. Here, the selection of the papyrus is done.

In Fig. 6, the papyrus images that has been entered will be directly carried out to the process of cutting image that starting from the left side, middle, into the upper side, so that the image processed will be immediately appear under the original papyrus image.

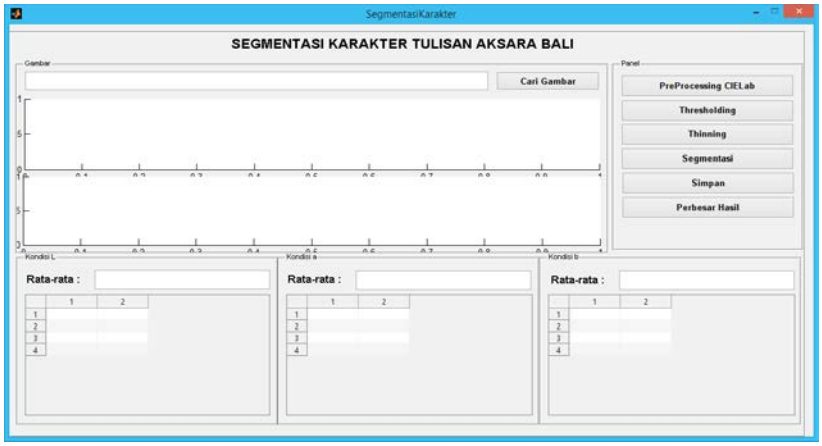

Fig. 4: The early system display

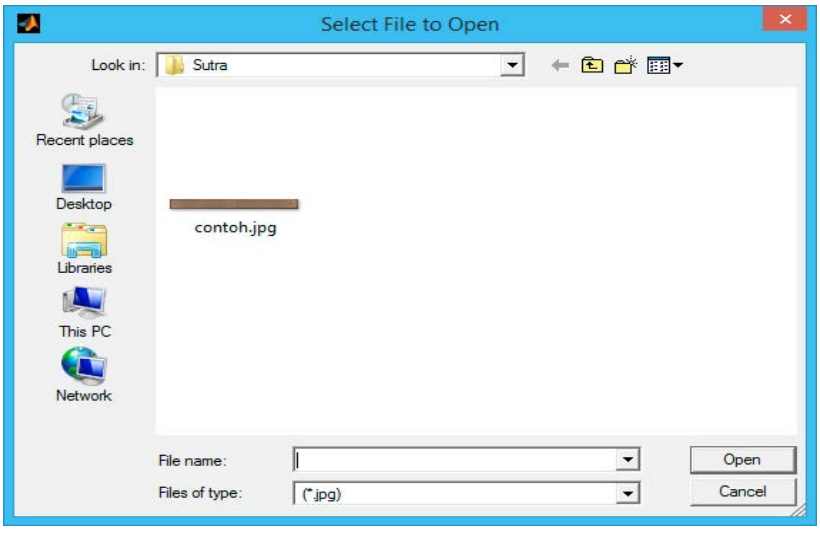

Fig. 5: Open image fiture

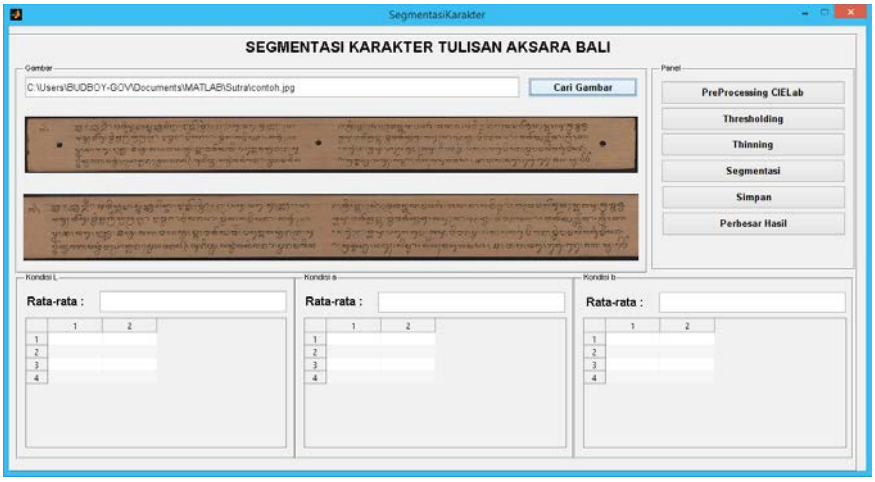

Fig. 6: The loaded and image cropping

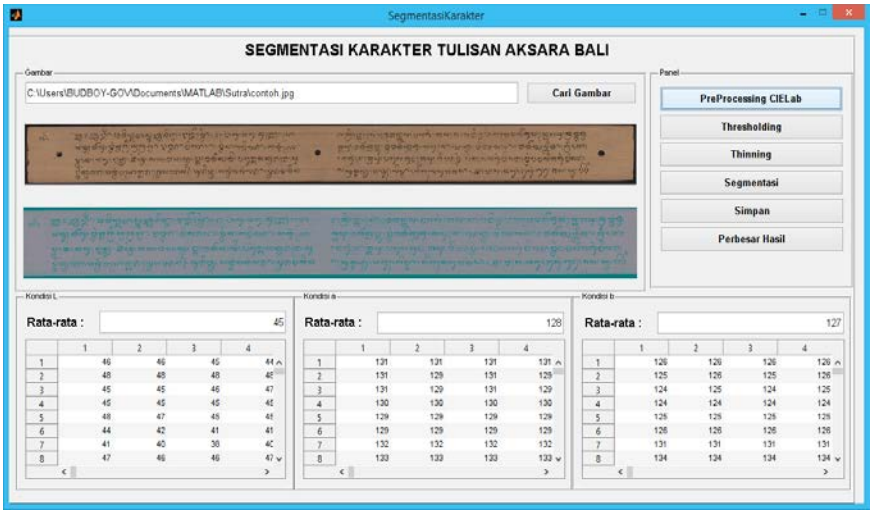

Fig. 7: Process of CIELab

The process of identifying the color space Proses using the CIELab can be done by clicking the PreProcessing CIELab. The system processes the looping as many as the number of the bit found on the image. The outcome can be seen on the table displayed by the system. From there there will be the process of obtaining the means of the L*, a* and b* values (Fig. 7).

CONCLUSION

Based on the outcome obtained from the experiment it can be concluded that the color space of CIELab can be used to identify and separate the color of the Balinese characters from the ones of the background. This is clearly informed through the values of L* = 45, a* = 128 and b* = 127 in which L* of the CIELab represents the brightness of the color as a little differentiator.

The next study can be done to analyzed from the pre-processing till the extracting to finally traslating the Balinese characters on the payrus to latin script.

ACKNOWLEDGEMENTS

The appreciation is awarded to all the colleagues who have given a lot of help both in discussion of ideas until the trial implementation stage of the device. Not forgotten

also the appreciation is delivered to the Department of Electrical and Computer System Engineering, Udayana University which has supported in this research.

REFERENCES

Bansal, S. and D. Aggarwal, 2011. Color image segmentation using CIELab color space using ant colony optimization. IJCSET, 1: 415-420.

Hoffmann, G., 2010. CIE color space. November 2, 2010. http://www.labri.fr/perso/granier/Cours/IOGS/color/ciexyz29082000.pdf.

Mamatha, H.R. and K. Srikantamurthy, 2012. Morphological operations and projection profiles based segmentation of handwritten kannada document. Int. J. Applied Inform. Syst., 4: 13-19.

Prapitasari, L.P.A., 2012. Noise Removal for Ancient Palm Leaf Manuscript of Bali. Konferensi Nasional Sistem Informasi, Sulawesi Utara.

Putra, D., 2010. Pengolahan Citra Digital. Andi Publisher, Penerbit, ISBN-13: 9789792914436.

Sudarma, M. and N.P. Sutramiani, 2014. The thinning zhang-suen application method in the image of balinese scripts on the papyrus. Int. J. Comput. Applic., 91: 9-13.