Matko ŠARIĆ, Hrvoje DUJMIĆ, Mladen RUSSO

University of Split

Scene Text Extraction in HSI Color Space using K-means Algorithm and Modified Cylindrical Distance

***Abstract**. Text extraction, that is segmentation of characters from background, is especially important step that greatly determines final recognition performance. Particular focus is put on this task for scene text which is characterized with wide set of degradations like complex backgrounds, uneven illumination, viewing angle etc. In this paper we introduce text extraction method based on k-means clustering with modified cylindrical distance in HSI color space. Performance of this distance is analyzed depending on different degrees of chroma reliability. For purpose of result comparison, K-means text extraction is also performed with cylindrical distance in HSI color space and Euclidean distance in RGB color space. Complementarity of tested distances is also analyzed showing possible direction for further performance improvement.*

***Streszczenie**. W artykule opisano metodę ekstrakcji tekstu z obrazu w oparciu o klasteryzację k-means oraz przestrzeń kolorów typu HSI we współrzędnych cylindrycznych. Poprawność dystansowania została poddana analizie w zależności od różnego stopnia rzetelności kolorów. Wyniki działania algorytmu ekstrakcji tekstu k-means zostały porównane dla współrzędnych cylindrycznych w przestrzeni kolorów typu HSI oraz współrzędnych Euklidesowego w przestrzeni kolorów typu RGB. (**Ekstrakcja tekstu z obrazu w przestrzeni kolorów typu HSI z wykorzystaniem klasteryzacji k-means i zmodyfikowanych współrzędnych cylindrycznych**).*

Keywords: scene text, K-means, cylindrical distance.
Słowa kluczowe: tekst w obrazie, k-means, współrzędne cylindryczne.

Introduction

Text in images and video represent interesting form of semantic image content. Information contained in text data enables different applications such as document analysis [1], automatic license plate recognition [2], sign detection and translation [3], content based image indexing [4], etc. Text efficiently describes image content and it can be easily extracted in comparison with other semantic forms.

Text is often categorized as text in documents, caption text and scene text [5]. Text in documents is relatively easy extracted and recognized because of its properties (regular structure, strong contrast, clean background etc.). For this type of text research is focused on recognition of non-latin scripts [6] and document structure analysis as line segmentation [7]. Characters artificially added to image or video frame are classified as caption text. Typical examples are subtitles or match results in sports video. Scene text is integral part of recorded object as text on bottle or city name on traffic sign. Hence this type of text doesn't have constraints in structure, shape, illumination, background etc.

Growing availability of digital cameras and camera equipped smartphones enables easy acquisition of scene text present in our environment. Despite its practicality, these devices introduce new set of problems such as sensor noise, viewing angle, blur, variable illumination etc. In combination with variety of object shapes, materials and possible backgrounds these conditions makes scene text extraction difficult task. Jung et al. [5] divide problem of text information extraction into 5 steps: detection, localization, tracking, extraction and enhancement, and recognition (OCR). In this paper we focus on extraction step in which characters are segmented from background, that is text pixels are separated from background pixels. OCR software, usually optimized for text in documents, is very sensitive to artifacts arising from poor character extraction (background portions, connected characters, missing character parts etc.). In case of scene text, that is often characterized by complex backgrounds and uneven lightning, extraction is important step that determines success of recognition stage performed by OCR.

Mancas-Thillou and Gosselin [8] divide text extraction methods on two categories: thresholding-based and grouping-based. Histogram thresholding, adaptive or local thresholding [9] and entropy based methods belong to first category. Although these methods aren't computationally demanding, they handle grayscale images or color channels independently and hence they are not suitable in case of complex backgrounds and varying colors. Second category includes region-based, learning-based and clustering-based methods. Region-based techniques most often refer to region-growing [10] and split-and-merge algorithm [11]. Their advantage is inclusion of spatial information, but efficiency heavily depends on parameter values. In learning-based methods several well–known classifiers (multi-layer perceptrons, self-organizing maps) are used for text extraction. Main problem is creation of representative training database that should cover diversity of scene text.

Clustering-based methods rely on assumption that pixels will form groups in chosen color space. K-means algorithm is most often used for text extraction because of its speed and efficiency. Mancas-Thillou and Gosselin [12] use k-means clustering in RGB color space with Euclidean distance and angle distance where better result is chosen by feedback from recognition results. Number of clusters is set to 3 representing characters, background and character edges. In [13] 4-means is exploited for text segmentation in HSV color space. In [14] results of text extraction with k-means algorithm are presented for set of different color spaces (RGB, Lab, Luv, HSV etc.). RGB color space gives best results, but it should be noted that for all color spaces Euclidean distance is used as color difference measure. Same author analyzed results of k-means obtained with Euclidean distance and 6 angle distances in RGB color space. Euclidean distance performs best overall, but angle distances are able to cope with degradations in more complex images.

In this paper we propose text extraction method in HSI color space based on k-means algorithm and modified cylindrical distance [15]. Although Euclidean distance in RGB color space proved its efficiency as clustering metric in text extraction task [14], it doesn't correspond to cylindrical nature of HSI color space. In that case cylindrical distance is more appropriate solution because it takes into account angular values. In comparison with cylindrical distance, modified cylindrical distance, presented in [15], better corresponds to human perception because it models gradual transition between achromatic and chromatic pixels through hue and saturation reliability functions. We investigate whether its usage as clustering metric can improve performance of scene text extraction. Analysis is performed in dependence on different degrees of chroma weight and results are compared with those obtained by Euclidean distance in RGB space.

Minkowski metric and angle distance

Extraction of textual information from images and video requires appropriate measure of color difference or color similarity. Regarding that color information is usually represented in vector form most straightforward solution is usage of distance measures for *m*-dimensional vectors. One of the most often used measures for vector dissimilarity is generalized weighted Minkowski metric:

$$d_p(i,j) = c\left(\sum_{k=1}^{m} \xi_k \left|x_{ik} - x_{jk}\right|^p\right)^{\frac{1}{p}} \tag{1}$$

where m represents dimension of vector $\vec{x}_i$ and x_{ik} is k-th element of $\vec{x}_i$. The nonnegative scaling parameter ***c*** determines overall discrimination power, while parameter ξ_k represents weight of component k. Different values of parameter p define special cases of Minkowski metric: the city-block distance or Manhattan distance ($p=1$), the Euclidean distance ($p=2$) and the chessboard distance ($p=\infty$). Euclidean distance is the most often used for text extraction in RGB color space. For two pixels with values (R_i, G_i, B_i) and (R_j, G_j, B_j) it is defined as

$$d_{eucl}(i,j) = \sqrt{\left(R_i - R_j\right)^2 + \left(G_i - G_j\right)^2 + \left(B_i - B_j\right)^2} \tag{2}$$

Although Euclidean distance isn't appropriate measure for perceptual color difference in RGB color space, it shows good performance in text extraction as clustering distance in k-means algorithm [14].

Besides vector magnitudes angles between them can also be exploited as color difference measure. Cosine-based similarity can be defined as normalized inner product:

$$s\left(\vec{x}_i, \vec{x}_j\right) = \frac{\vec{x}_i \vec{x}_j^{\,T}}{\left|\vec{x}_i\right|\left|\vec{x}_j\right|} = \cos\theta \tag{3}$$

where $\vec{x}_i$ and $\vec{x}_j$ represent vectors and θ represents angle between them. From this relation angle difference can be expressed as:

$$\theta = \cos^{-1}\left(\frac{\vec{x}_i \vec{x}_j^{\,T}}{\left|\vec{x}_i\right|\left|\vec{x}_j\right|}\right) \tag{4}$$

In RGB color space Euclidean distance reflects changes in intensity, while angle distance represents difference in chromaticity enabling character segmentation in presence of shadows and uneven lightning. Complementarity of these distances is exploited for scene text extraction in [12].

Cylindrical distance

Cylindrical distance [16] between two pixels in HSI color space is defined as:

$$d_{cylindrical}(i,j) = \sqrt{d_{chroma}^2(i,j) + d_{intensity}^2(i,j)} \tag{5}$$

$$d_{chroma}(i,j) = \sqrt{S_i^2 + S_j^2 - 2S_iS_j\cos\theta} \tag{6}$$

$$d_{intensity}(i,j) = \left|I_i - I_j\right| \tag{7}$$

$$\theta = \begin{cases} \Delta & if\ \Delta < 180^\circ \\ 360^\circ - \Delta & otherwise \end{cases} \tag{8}$$

$$\Delta = \left|H_i - H_j\right| \tag{9}$$

where $H \in \left[0^\circ, 360^\circ\right]$, $S \in [0, 255]$, $I \in [0, 255]$. Chromatic distance d_{chroma} (Fig. 1a) represents distance between two pixels in chromatic plane, $d_{intensity}$ is absolute value of intensity difference and $d_{cylindrical}$ (Fig. 1b) is defined as hypotenuse of right-angled triangle.

When calculating cylindrical distance instability of hue and saturation components should be taken into account [16]:

- hue is meaningless when intensity is very low,
- hue is unstable when saturation is very low,
- saturation is meaningless when intensity is very low.

These properties require distinction of chromatic pixels, that is pixels with stable saturation and hue, and achromatic pixels where intensity is only relevant attribute.

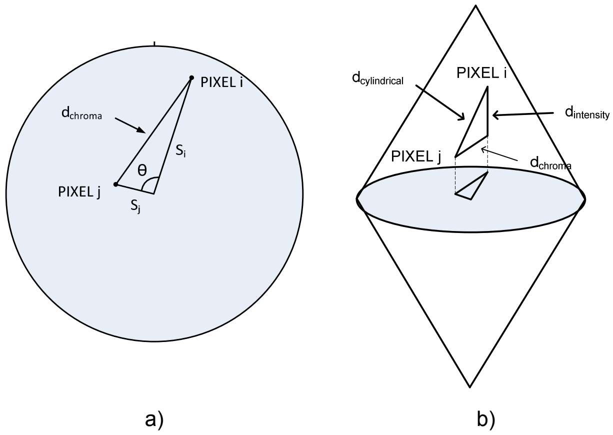

Fig. 1. a) chromatic distance b) cylindrical distance

Equations (5)-(9) give cylindrical distance for chromatic pixels. In case of achromatic pixels chromatic distance is unreliable and hence cylindrical distance is reduced to intensity difference ($d_{intensity}$)

Discrimination between chromatic and achromatic pixels can be performed by thresholding of saturation and intensity component. In [17] achromatic pixels in HSI color space are defined with following conditions:

$$\begin{aligned} &intensity > 90\%\ \max\ intensity\, value, \\ &intensity < 10\%\ \max\ intensity\, value, \\ &saturation < 10\%\ \max\ saturation\, value. \end{aligned} \tag{10}$$

Although discrete model based on threshold is used when calculating cylindrical distance [17], it should be noted that they don't correspond to the behaviour of human visual system: transition from scotopic (achromatic) to photopic (chromatic) vision goes gradually through mesopic vision. In order to circumvent this problem continuous models for differencing chromatic and achromatic pixels are proposed in literature ([18-20]). In these models transition function defines weight (reliability or importance) of hue and saturation component.

Modified cylindrical distance

Despite its disadvantages threshold-based differencing of chromatic and achromatic pixels is most often used in literature for calculation of cylindrical distance [17]. In [15] is proposed modified cylindrical distance that uses continuous model for distinction between achromatic and chromatic pixels:

(11) $d_{\_mod\_cylindrical}(i,j)=\sqrt{d_{chroma}^2(i,j)+d_{intensity}^2(i,j)}$

(12)

$$d_{chroma}(i,j)=w_{\Delta S}\left(w_\theta\sqrt{S_i^2+S_j^2-2S_iS_j\cos\theta}+(1-w_\theta)\left|S_i-S_j\right|\right)$$

(13) $$d_{intensity}(i,j)=\left|I_i-I_j\right|$$

(14) $$w_\theta=\sqrt{w_H(i)w_H(j)}$$

(15) $$w_{\Delta S}=\min\left(w_S(i),w_S(j)\right)$$

where w_θ denotes hue difference reliability, $w_H(i)$ is hue reliability for pixel *i*, $w_{\Delta S}$ is saturation difference reliability and $w_S(i)$ is saturation reliability for pixel *i*. In [15] it is showed that in comparison with original formulation of cylindrical distance modified cylindrical distance gives results closer to human perception for colors on the transition between achromatic and chromatic region. In this paper we want to investigate whether usage of this distance could improve performance of k-means algorithm in scene text extraction task. From (12) it can be seen that saturation difference reliability $w_{\Delta S}$, that is saturation reliability w_S, determines contribution of chromatic information to the value of distance. Taking into account complementarity between intensity and chroma regarding text extraction task it is interesting to observe whether changing degree of chromaticity in modified cylindrical distance could improve character segmentation. This analysis is performed by varying parameters of saturation reliability function.

K-means algorithm

Clustering algorithms are considered as most efficient methods for scene text extraction [8]. One of the main advantages is their versatility in comparison with parametric methods such as adaptive thresholding or region growing. K-means algorithm is especially popular because of its easy implementation and low computational requirements. It tries to minimize sum of distances between points and cluster centers that is represented by:

(16) $$\sum_{j=1}^{k}\sum_{i\in S_j} distance\left(x_i^{(j)},c_j\right)$$

where *distance* is chosen distance measure between point $x_i^{(j)}$ and the cluster centre c_j, S_j is set containing elements of cluster j and k is number of clusters. Algorithm consists of following steps:

1. In set of N points, corresponding to image pixels, choose k points as initial cluster centers (centroids) c_j
2. Assign each point to nearest cluster S_j based on its distance from cluster center c_j
3. For each cluster Sj compute a mean μj of each cluster and set the mean as new cluster center ($cj=\mu j$)
4. Repeat the steps 2 and 3 until the centroids no longer move

Influence of distance measure in k-means algorithm regarding text extraction is demonstrated in [14]. It is concluded that for RGB color space Euclidean distance gives best results, but it also shows complementarity with angle distances. Author also evaluates performance of k-means algorithm with Euclidean distance measure in different color spaces where RGB proves as most efficient color space. Despite that it should be noted that Euclidean distance is not appropriate measure for cylindrical color spaces (HSI, HSV, HLS etc.) and this fact motivate us to use cylindrical distance and modified cylindrical distance as clustering metrics in this paper.

Results

Evaluation was performed on test set from word recognition task in ICDAR 2011 Robust Reading Competition Challenge 2: Reading Text in Scene Images [21]. This database consists of 1189 scene text images covering broad set of problems like complex backgrounds, different layouts, poor contrast, uneven illumination, low resolution etc. Text extraction was done in MATLAB using k-means algorithm with three different clustering distances: modified cylindrical distance in HSI color space, cylindrical distance in HSI color space and Euclidean distance in RGB color space. Number of clusters is set to 2 where one is textual foreground and second represents background. Binary images, obtained as result of text extraction, were passed to Google OCR engine Tesseract 2.04 in order to finally recognize text

As it is suggested in [21], for evaluation purposes we used edit (Levenshtein) distance between ground truth string and string recognized using Tesseract, where deletions, substitutions and insertions have equal costs. This distance is normalized by dividing the number of recognized characters with number of characters in ground truth word. Another measure also used in [21] is percentage

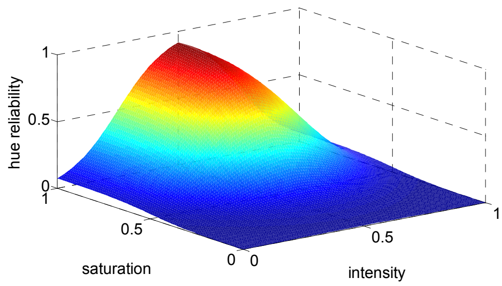

Fig. 2. Hue reliability function proposed by Aptoula [20]

of correctly recognized words, that is number of cases where normalized edit distance is equal to zero.

When calculating modified cylindrical distance hue and saturation reliability function should be chosen. We decide to use hue weight function proposed by Aptoula [20] and defined by equations:

(17) $$g(S)=\frac{1}{1+e^{-k(S-S_0)}}\qquad S\in[0,1]$$

(18) $$f(L)=\begin{cases}\dfrac{1}{1+e^{-k_L(L-L_L)}} & for\ L\le 0.5\\[2ex] \dfrac{1}{1+\mathrm{e}^{k_L(L-L_U)}} & for\ L>0.5\end{cases}\qquad L\in[0,1]$$

where k_L=10, L_L=0.25, L_U=0.75.

(19) $$w_H=f(L)\times g(S)$$

This model calculates hue reliability in dependence on saturation and lightness (intensity) component and takes into account hue unreliability for high lightness (intensity) levels (Fig. 2). Saturation reliability was modeled by sigmoid function defined with:

(20) $$w_S(I)=\frac{1}{1+\mathrm{e}^{-k(I-I_0)}}$$

Parameter *k* was set to 10 to obtain smooth transition from fully unreliable to reliable saturation component: lower values of (for example *k*=1) would result with linear dependency, while higher values (for example *k*=100) give sharp transition similar to threshold-based models. In order to analyze how changes of chroma contribution in modified cylindrical distance affects text extraction performance we vary the values of parameter I_0 what results in shifting of sigmoid function (Fig. 3). Lower values of I_0 give higher contribution of chroma component while higher values lower its influence and emphasize intensity difference.

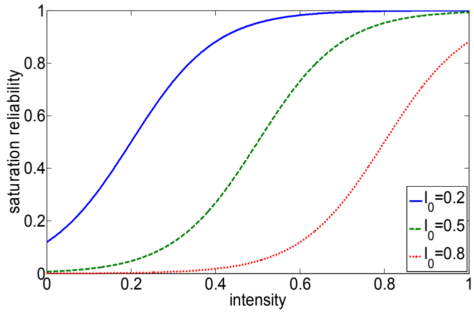

Fig. 3. Saturation reliability for k=10 and different values of I_0

Table 1. Results for modified cylindrical distance in dependence on I_0

I_0	Total edit distance	Correct recognition (%)
0.1	663.4	43.9
0.2	655.2	43.7
0.3	622.8	45.2
0.4	612.9	47.8
0.5	611.5	47.2
0.6	628.5	47.4
0.7	664.5	46.6
0.8	653.6	46.9
0.9	665.8	46.8

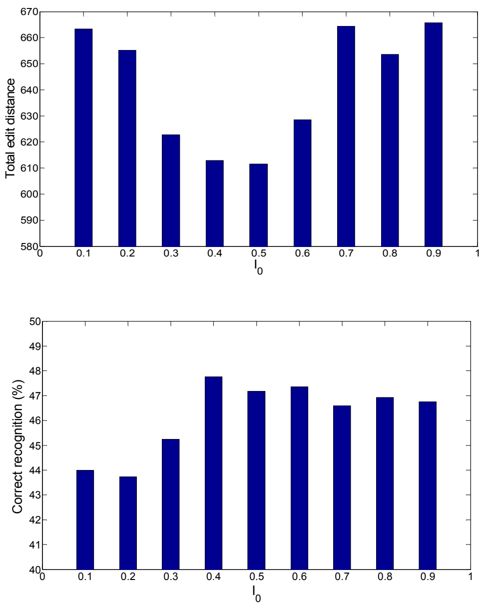

Fig. 4. Results of k-means with modified cylindrical distance in dependence on I_0. Top panel: total edit distance Bottom panel: correct recognition.

Second column in table 1 shows results for different values of I_0 reported in form of total edit distance calculated by summing normalized edit distance for each ground truth word. Third column presents percentage of correctly recognized words, that is cases where normalized edit distance is equal to zero. It can be seen that minimal values of total edit distance are obtained for I_0 around 0.5 (Fig. 4). In [15] suggested value is 0.2, but evaluation is performed with respect to human perception of color difference. This acknowledges that color distance measure close to human perception doesn't necessary lead to better text extraction as it has already been shown in case of Euclidean distance in RGB color space.

Case when I_0=0.4 yields highest correct recognition rate, but it should be noted that results are not significantly lower for values from 0.5 to 0.9 while for I_0<0.4 more obvious decrease of performance is noticed (Fig. 4). Both evaluation measures show that degree of chroma contribution determines results of modified cylindrical distance. Correct recognition rate for high I_0, when chromatic difference has only small influence to distance value, reveals importance of intensity component. Results presented in table 2 shows that modified cylindrical distance outperforms cylindrical distance in both measures proving that including of continuous models for differencing of achromatic and chromatic pixels yields improvement of character extraction Although in table 2 only best result for modified cylindrical distance is shown (I_0=0.4), same observation is valid for all tested values of I_0. In comparison with Euclidean distance modified cylindrical distance is close, but it is still worse in both measures. As it is noted in [12], Euclidean distance mostly reflects changes in intensity component which obviously has potential to correctly extract characters in many images from test set. Although cylindrical distance gives the worst results, more detailed analysis reveals its complementarity with Euclidean distance. In 277 (23%) images from test set Euclidean distance results with better extraction. On the other hand cylindrical distance gives more extracted characters in 143 (12%) images. This can be explained with presence of chromatic information that enables character segmentation in cases when intensity difference fails due to degradations. Similar relation is noticed in case of modified cylindrical distance (I_0=0.4): in 108 (9.1%) images it performs better than Euclidean distance (Fig. 5), while inverse situation appears in 145 (12.2%) images from test set (Fig. 6). For modified cylindrical distance complementarity is manifested to a lesser extent. Despite overall efficiency of Euclidean distance, in cases when it fails modified cylindrical distance is possible solution. This happens in examples where certain kind of degradations, like uneven illumination and shadows, affect intensity component. On the other hand, modified cylindrical distance incorporates not only intensity distance, but also hue and saturation difference through chromatic distance. In [22] it is showed that hue is invariant to uniform scaling and uniform shifting in RGB space. According to [23] saturation is invariant to uniform scaling in RGB color space. Taking into account that uniform RGB scaling is caused by shadowing and shading, while RGB shifting is produced by highlights, it is clear that hue and saturation are invariant to these degradations.

Table 2. Results for different clustering distances in k-means algorithm

Distance measure	Total Edit distance	Correct Recognition(%)
Modified cylindrical distance in HSI (I_0 =0.4)	612.9	47.8
Cylindrical distance in HSI	764	40
Euclidean distance in RGB	591.17	48.5

Based on described properties, chromatic distance has ability to correctly extract characters in presence of highlights and shadowing. Parameter I_0 enables fine tuning of chroma contribution in final value of modified cylindrical distance, that is balancing between dominant intensity distance and equal weights of chroma and intensity differences. This property of proposed method makes possible handling of wider set of degradations in comparison with clustering using Euclidean distance.

a) b) c)

Fig. 5 Examples in which Euclidean distance gives better result than modified cylindrical distance a) input image b) result of k-means with Euclidean distance c) result of k-means with modified cylindrical distance

a) b) c)

Fig. 6 Examples in which modified cylindrical distance gives better results than Euclidean distance a) input image b) result of k-means with Euclidean distance c) result of k-means with modified cylindrical distance

Conclusion

Scene text extraction method based on k-means with modified cylindrical distance in HSI color space is proposed in this paper. We investigated how results are influenced by the saturation reliability function that in fact determines chroma contribution to the overall distance value. It is shown that as clustering metric modified cylindrical distance clearly outperforms cylindrical distance in text extraction according to total edit distance and correct recognition rate. This observation confirms that including of hue and saturation reliability functions in modified cylindrical distance yields improvement not only in aspect of perceptual color difference, but also in character segmentation task. Comparison is also made with results obtained using k-means and Euclidean distance in RGB color space where modified cylindrical distance has slightly poorer performance.

Further analysis shows complementarity between tested distances: modified cylindrical distance can handle cases in which Euclidean distance fails to correctly extract characters from background. This indicates that chromatic infromation enables text segmentation in presence of various degradations. It is obvious that one color distance measure can't cover variety of scene text degradations and hence in future research we will investigate strategies for combining results obtained with two or more color distances

REFERENCES

[1] Tang, Y. Y., Lee, S. W., Suen, C. Y. ‚Automatic document processing: a survey, *Pattern Recognition*, vol. 29, no. 12, p. 1931–1952.

[2] Iwanowski, M., Automatic car number plate detection using morphological image processing, *Przeglad Elektrotechniczny (Electrical Review)*, vol. 81, pp. 58-61, 2005.

[3] Watanabe, Y., Sono, K., Yokomizo, K., Okada, Y., Translation camera on mobile phone, in *Proceedings of International Conference on Multimedia and Expo*, 2003, pp. 177-180.

[4] Saric,M., Dujmic, H., Papic,V., Rozic, N., Radic, J., Player Number Recognition in Soccer Video using Internal Contours and Temporal Redundancy, in *Proceedings of the 10th WSEAS International Conference on Automation & Information (ICAI'09) 2009*, pp. 175-180.

[5] Jung, K., Kim, K., Jain, A., Text Information Extraction in Images and Video: A Survey, *Pattern Recognition*, vol. 37, no. 5, pp. 977-997, 2004.

[6] Razzak, M. I., Mirza, A. A., Ghost Character Recognition Theory and Arabic Script Based Languages Character Recognition, *Przeglad Elektrotechniczny (Electrical Review)*, vol. 87, no. 11, pp. 234-238, 2011.

[7] Brodic, D., Methodology for the evaluation of the algorithms for text segmentation based on errors type, *Przeglad Elektrotechniczny (Electrical Review)*, vol. 88, no. 1b, pp. 259-263, 2012.

[8] Mancas-Thillou, C., Gosselin, B., Natural Scene Text Understanding, *Vision Systems: Segmentation and Pattern Recognition*, Vienna, Austria: I-Tech Education and Publishing , 2007, ch. 16, pp. 307-332.

[9] Gatos, B., Pratikakis, I.,Kepene, K., Perantonis, S. J., Text detection in indoor/outdoor scene images, *Proc. First Workshop of Camera-based Document Analysis and Recognition*, 2005, p. 127–132.

[10] Lienhart, R., Effelsberg, W., Automatic text segmentation and text recognition for video indexing, *Technical Report,* University of Mannheim, 1998.

[11] Karatzas, D., Antonacopoulos, A., Colour text segmentation in web images based on human perception, *Image and Vision Computing*, vol. 25, no. 5, pp. 564-577, 2007.

[12] Mancas-Thillou, C., Gosselin, B., Color text extraction with selective metric-based clustering, *Computer Vision and Image Understanding*, vol. 107, no. 1-2, pp. 97-107, 2007.

[13] Garcia. C., Apostolidis, X., Text detection and segmentation in complex color images, in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 2000, p. 2326–2330.

[14] Mancas-Thillou, C., Natural Scene Text Understanding, PhD Thesis, Faculté Polytechnique de Mons, 2006.

[15] Saric, M., Dujmic, H., Rozic, N., Including of continuous model for discriminating chromatic and achromatic pixels in cylindrical distance, *Automatika*, vol. 51, no. 3, pp. 241-254, 2010.

[16] Tseng, D. C., Chang, C. M., Color segmentation using perceptual attributes, in *Proc. of the 11th Internat. Conf. on Pattern Recognition*, 1992, pp. 228-231.

[17] Plataniotis, K. N., Venetsanopoulos, A. N., Color image processing and applications, Springer-Verlag New York, 2000.

[18] Vadivel, A., Sural, S., Majumdar, A. K., Human color perception in the HSV space and its application in histogram generation for image retrieval, in *Proc. SPIE, Color Imaging X: Processing, Hardcopy, and Applications*, 2005, p. 598–609.

[19] Romani, S., Sobrevilla, P., Montseny, E., On the reliability degree of hue and saturation values of a pixel for color image classification, in *Proceedings of IEEE Internat. Conf. on Fuzzy Systems*, 2005, pp. 306-311.

[20] Aptoula. E., Lefévre, S., On the morphological processing of hue, *Image and Vision Computing*, vol. 27, no. 9, 2009.

[21] Shahab, A., Shafait, F., Dengel, A., ICDAR 2011 Robust Reading Competition Challenge 2: Reading Text in Scene Images, in *Proc. 11th International Conference of Document Analysis and Recognition*, 2011, pp. 1491-1496.¸

[22] Perez, F., Koch, C., Toward color image segmentation in analog VLSI: algorithm and hardware, *International Journal of Computer Vision*, vol. 12, no. 1, pp. 17-42, 1994.

[23] S. Romani, P. Sobrevilla, and Montseny.E., "On the reliability degree of hue and saturation values of a pixel for color image classification," in Proceedings of IEEE Internat. Conf. on Fuzzy Systems, 2005, pp. 306-311.

***Authors**: Ph. D. Matko Saric, Ph. D. Hrvoje Dujmic, Ph. D. Mladen Russo, University of Split, Faculty of electrical engineering, mechanical engineering and naval architecture, Rudjera Boskovica 32, 21000 Split, Croatia, E-mail: msaric@fesb.hr, hdujmic@fesb.hr, mrusso@fesb.hr*