

Colour Space Selection for Entropy-based Image Segmentation of Folded Substrate Images

Magdolna Apró¹, Dragoljub Novaković¹, Szabolcs Pál², Sandra Dedijer¹, Neda Milić¹

¹ University of Novi Sad, Faculty of Technical Sciences, Department of Graphic Engineering and Design, Novi Sad, Serbia, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia, E-mail address: apro@uns.ac.rs, novakd@uns.ac.rs, dedijer@uns.ac.rs, milicn@uns.ac.rs

² University of Novi Sad, Faculty of Technical Sciences, Department of Computing and Control Engineering, Novi Sad, Serbia, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia, E-mail address: sabolc.pal@rt-rk.com

Abstract: This paper is focused on analysing the effects of the chosen colour space on image segmentation accuracy for a permanent quality control of the folding process. The folding process is one of the basic operations in print finishing, but during this converting operation the printed or non-printed substrates are exposed to high tensile stresses. These stresses can cause coating cracks on the folding line, which decrease the expected aesthetic feature or even the functionality of the product. High production efficiency of the folding process could be provided by a control system for automated visual inspection. Such a quality control algorithm was proposed by the authors in previous papers. Since the proposed algorithm relies on qualitative image segmentation, it is very important to determine all the factors which influence the segmentation quality. This paper investigates the influence of colour spaces. The applied image segmentation algorithm (Maximum Entropy) works on grey-scale images, and therefore only the luminance components of the five selected colour spaces (HSI, HSL, HSV, CIE Lab and CIE xyY) were used. The segmentation quality was determined by using six different measures (quantitative and qualitative), which were combined in order to obtain a single performance measure for algorithm evaluation.

Keywords: colour space; segmentation; fold quality

1 Introduction

Folding a paper is one of the basic print finishing operations and its quality control is done by the machine operator, inspecting the folded paper mostly visually [2]. Besides such technical issues (non-precise register, double sheets folding, paper crinkling), the surface cracking along the folding line must be detected and

prevented or at least minimised during the folding process. There are different fold-ability or crack-resistance evaluation methods well known in the paper industry, for example: residual tensile strength, residual tensile stretch, residual bending stiffness, folding endurance, etc. [22]. Based on these methods under controlled conditions, a detailed investigation can be done for the fold-ability properties and fold line crack-resistances of the paper, but for real-time production quality control they cannot be applied, since they are time consuming and require special equipment. The visual control method which was used in [2], [12], [22] and [25] involves a human observer, and thus the obtained results are not repeatable, are highly dependable the observer's experience and are of a subjective nature. A simple analysis based on the white pixel analysis of the digitalised images of folding lines can be found in [3], [4], [6], [19] and [21]. The presented image analysis gave an objective quality grade for surface cracking using commercial image analysis software, but the evaluations were done separately from the production phase. Although the white pixel analysis solves the subjectivity issue, a more complex pattern analysis is needed to explore the true nature of the surface damage (quantity, distribution, size, length and width of the cracked lines, etc. [9]) and the influence of the printed colour on visual perception and therefore on the aesthetic feature.

With an objective folding quality estimation method implemented in the production process, the above described problems could be overcome. As in other fields of industry, computer vision based quality control can be applied.

A basic Objective Folding Quality Assessment (OFQA) was proposed by authors [15], which was based on a set of image analysis algorithms in combination with neural network (Figure 1).

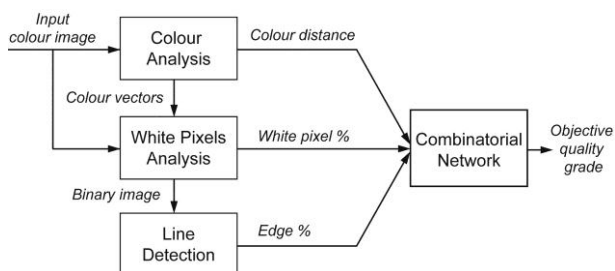


Figure 1

Block diagram of the proposed OFQA algorithm

Due the observed parameters (colour and white pixel analysis, line detection) the presented algorithm obtained objective quality measures of the SAPPI evaluation scale [2] in a good correlation with subjective grades. Although the initial results were promising, they also showed that for the industrial application, further development was necessary to improve the method's accuracy. The development and improvement process included a complete revision of some parts of the

algorithm and it revealed out that the choice of the proper image segmentation method is the most crucial step. On the observed images, in most of the cases, three different areas could be noticed: damaged area, printed area and shadow, and therefore lightness or intensity based monochrome/grey-scale auto-thresholding techniques were selected. Since the final algorithm should work autonomously with a wide range of samples, only algorithms with auto-detected thresholds were considered. In our previous paper [16], different types of grey-scale image thresholding algorithms were analysed, and the best performing algorithms were based on entropy thresholding (the so called Maximum Entropy and Renyi Entropy). As appropriate colour spaces, two colour spaces were selected: in the image processing, the widely used HSL and the perceptual uniform CIE Lab colour space. The obtained results indicated that further investigation was necessary to find the most suitable colour space for the chosen image segmentation method. The aim of this investigation was to evaluate the applicability of different colour spaces, in order to improve the image segmentation accuracy.

2 Methods

2.1 Image Segmentation

Image segmentation is an essential component of many image analysis and pattern recognition applications, conditioning the performance of subsequent analysis steps. It can be defined as the process of partitioning an image into a set of non-overlapping regions whose union is the entire image. Image segmentation techniques can vary widely according to the type of image (e.g., binary, grey, colour), the choice of mathematical framework (e.g., morphology, image statistics, graph theory), the type of features (e.g., intensity, colour, texture, motion) and the approach (e.g., top-down, bottom-up, graph-based). The simplest image segmentation technique is histogram thresholding, which assumes that the histogram of an image can be separated into as many peaks as there are different regions present in the image. The early segmentation algorithms were based on grey-level segmentation (monochrome). With the requirement changes (the emerging need of segmenting colour images) and growth of available computational power, the monochrome segmentation techniques have been extended to segment colour images. Some colour image thresholding approaches consider the 3D histograms that simultaneously contain all the colour information in the image, but since the storage and processing of multidimensional histograms is computationally expensive, most approaches consider 1D histograms computed for one or more colour components in some colour space [8, 11, 18].

In situations where the luminance (intensity) information on an image is discriminative enough, the 1D histogram approach can be used. In other words, if the intensity (grey) levels of pixels of the object of interest are substantially different from the intensity (grey) levels of the pixels belonging to the background, a 1D thresholding algorithm can be used to separate objects of interest from the background. A well-chosen colour space, whose luminance channel carries the most information about the analysed surface, could be much more effective than a segmentation based on all channels, especially if it is applied in real-time in industrial conditions [7, 11, 17].

Based on previous research of the authors [16] which was focused on the evaluation of image segmentation algorithms, Kapur et al.'s method was found to perform the best for the given class of images (images of folded substrates) and it was used further in the evaluation of colour spaces presented in this paper.

The algorithm evaluation was done with the ImageJ open source image analysing software. Since the selected method is implemented under the name "MaxEntropy" correlating to the basic concept of the thresholding method in the following, it will be referred to as Maximum Entropy [10, 23].

The Maximum Entropy thresholding method exploits the entropy distribution of the grey-levels in an image. The principle of entropy is based on uncertainty as a measure of the information contained in a source. The Maximum Entropy thresholding method considers the image foreground and background as two different signal sources, and the optimal image thresholding is achieved when the resulted image preserves as much information as possible, namely when the sum of the foreground and background entropies reaches its maximum [1, 5, 17].

Assuming that the $h(i)$ is the normalized histogram of the analysed image, the optimum threshold for the Maximum Entropy can be defined as following [23]:

$$T_{opt} = \underset{t=0..i_{max}}{ArgMax} [H_B(t) + H_W(t)] \quad (1)$$

Where $H_B(t)$, the entropy of black pixels and $H_W(t)$, the entropy of white pixels, are defined as [23]:

$$H_B(t) = - \sum_{i=0}^t \frac{h(i)}{\sum_{j=0}^t h(j)} \log \frac{h(i)}{\sum_{j=0}^t h(j)} \quad (2)$$

$$H_W(t) = - \sum_{i=t+1}^{i_{max}} \frac{h(i)}{\sum_{j=t+1}^{i_{max}} h(j)} \log \frac{h(i)}{\sum_{j=t+1}^{i_{max}} h(j)} \quad (3)$$

In order to improve the quality of the image segmentation, a smoothing filter can be integrated as a pre-filter step into the processing chain. By removing redundant

details and noise from the input image, the pre-filtering step can reduce the problem of complex textures (halftone and rosette pattern of the observed surfaces). In our previous paper [16], three different filters were tested and evaluated (the Gaussian, Mean and Kuwahara filters). Although the improvement was modest, based on the obtained results, the Mean filter was selected for further use as the smoothing pre-filter. The pre-filtering and the segmentations were done using the ImageJ software, which has a broad range of plug-ins, including the plug-in for the Maximum Entropy, named as MaxEntropy, and the Mean filter.

2.2 Colour Spaces

A colour space is a geometrical representation of colours in a space and allows for specifying colours by means of three components, whose numerical values define a specific colour. They can be distinguished according to their characteristics in the following four families [14].

Primary spaces based on the trichromatic theory, which states that any colour can be expressed as a mixture of three primaries. The primaries correspond to the three types of colour sensing elements (cones) found in the human eye. The primary spaces can be [13, 20]:

- real primary colour space with physically realizable primaries (RGB, rgb),
- and imaginary primary colour space like (XYZ and xyz), whose primaries physically do not exist.

Luminance-chrominance colour spaces represent colours in terms of luminosity (L) and two chromaticity components (Cr_1 and Cr_2). The luminance-chrominance components are derived from the RGB colour space by linear or nonlinear transformations. The luminance-chrominance colour spaces can be classified as [13, 14, 20]:

- perceptually uniform spaces (CIE $L^*u^*v^*$ and CIE $L^*a^*b^*$), which determine the correspondence between the colour distance measured in colour space and the colour difference perceived by a human observer,
- television spaces like (YIQ and YUV), where the luminosity and the chromaticity signals are separated for the signal transmission,
- antagonist (opponent) spaces (wb,rg,by and YC_1C_2) based on the opponent colour theory in order to model the human visual system,
- and other spaces (such as Irg and Yxy or CIE xyY), which cannot be directly classified in the above mentioned sub families but are applied in colour image analysis as well.

Perceptual spaces quantify the colour according to the subjective human colour perception by means of the intensity, the hue and the saturation of colour.

Perceptual spaces can be also considered as luminance-chrominance spaces, since they are consisted of a luminance and two chrominance components. The perceptual colour spaces have luminance – chrominance components expressed by polar coordinates (e.g. HSL, HSV, HSB) [13, 14, 20].

Independent axis (or statistically independent component) spaces result from other spaces by applying mathematical operations that aim at de-correlating individual components ($I_1 I_2 I_3$, $P_1 P_2 P_3$, IJK) [14].

For the needs of this investigation, five colour spaces were selected with luminosity/intensity component. Two colour spaces were selected from the luminance-chrominance group (CIE xyY, CIE Lab) and three forms the perceptual colour space group (HSL, HSV and HSI). The used image acquisition equipment for sample digitalisation obtained the RGB values (sRGB), and therefore for some of the selected colour spaces, the colour components from RGB had to be transformed into XYZ space first, using the transform matrix, and then into the target spaces, applying the adequate calculations/equations. Details about the basic characteristics and transformation equations for intensity channel of selected colour spaces are presented in Table 1 [20].






Table 1
Basic characteristics and transformation equations for selected colour spaces [20]

Colour spaces	Components	Conversion form sRGB to XYZ	Conversion the intensity component
CIE xyY	x, y – chromatic component Y – luminance	$X=m_{11}R+m_{12}G+m_{13}B$ $Y=m_{21}R+m_{22}G+m_{23}B$ $Z=m_{31}R+m_{32}G+m_{33}B$	$Y = Y$
CIE Lab	a – green-red axis b – blue-yellow axis L – luminance	where $m_{11}, m_{12} \dots m_{33}$ are transformation coefficients	$L^* = \begin{cases} 116 \times \sqrt[3]{\frac{Y}{Y^w}} - 16 & \text{if } \frac{Y}{Y^w} > 0,008856 \\ 903,3 \times \frac{Y}{Y^w} & \text{if } \frac{Y}{Y^w} \leq 0,008856 \end{cases}$ where X^w, Y^w and Z^w are tristimulus values of the reference white
HSL	H – hue S – saturation L – brightness	or	$L = \frac{1}{2}(M + m)$ where: $M = \max(R, G, B)$ $m = \min(R, G, B)$
HSV	H – hue S – saturation V – value	$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [M] \begin{bmatrix} R \\ G \\ B \end{bmatrix}$	$V = M$ where: $M = \max(R, G, B)$
HSI	H – hue S – saturation I – intensity	where M is the transform matrix	$I = \frac{1}{3}(R + G + B)$

2.3 Samples Preparation

The proposed OFQA algorithm was developed for image analysis of the captured images of printed, folded and gathered substrates. In order to bring into account as many as possible different situations from real production process, a test form was prepared based on [12] and [25]. The test form consists of five different printed colour areas and a blank area for folding, as well as areas for print quality assurance. Table 2 shows the enlarged views of the printed areas for folding behaviour (surface damage) analysis, their CMYK notation of ink coverage and their target use case.

Table 2
The printed areas for folding behaviour analysis using CMYK notation

No.	Printed area	CMYK notation	Comment\explanation
1.		C 50%	damage visibility on light halftone pattern
2.		K 50%	damage visibility on dark halftone pattern
3.		K 100%	damage visibility on dark solid tone with total ink coverage of 100%
4.		C 40% + M 40% + Y 50% + K 20%	damage visibility on simulated colour image
5.		C 80% + M 80% + Y 80% + K 80%	folding behaviour at total ink coverage of 320%

The samples were made from uncoated, glossy- and matte-coated paper with basic weights of 100 g/m², 140/150 g/m² and 170 g/m². 50 samples of each paper grade were prepared in machine and cross grain direction 48 hours after printing at standard conditions (a temperature of 22°C, a relative humidity of 55%).

The sample-preparing process included the following operations and equipments:

- a) printing of the test forms was performed on KBA Performa 74 offset machine (process colours: Sun Chemical WORLD SERIES, plates: Agfa Azura TS CtP plates, dampening solution: 3% DS Acedin DH with 8% DS IPA, anti-set-off spray powder: DS 2020 B);
- b) cutting the printed test forms into suitable format for folding was done on a Perfecta 76 high-speed cutting machine (with a clamping pressure of 20 and 25 kN);
- c) folding the test forms in machine and cross direction was performed on a Horizon AFC546AKT folding machine (only one buckle folding used, standard fold rollers: combination of soft polyurethane foam rubber and steel roller, standard roller gap adjustment according to the manufacturer recommendation [24] working speed of 50 m/min).

2.4 Test Images

After the preparation, the samples were digitalised using a commercial digital camera, a flatbed scanner and a USB digital microscope. The basic settings of used equipment are presented in Table 3.

Table 3
Technical parameters and adjustments for used equipment

Used equipment	Canon A520	CanoScan 5600F	Veho VMS-001
Type	Commercial digital camera	Flatbed scanner	USB digital microscope
Colour mode	RGB	RGB	RGB
Embedded colour profile	sRGB	sRGB	-
Resolution	resolution 180 ppi	resolution 1200 ppi	resolution 300 ppi
Bit depth	8	8	8
Format	JPEG	BMP	BMP
Other	no flash, 100% digital zoom, auto white balance focal length of 5,8 mm	-	no light source, magnification of 200X, CMOS sensor

With the obtained digitalisation process, an extensive base image set was derived from the folded samples. A subset of 12 images has been selected from the base set, covering all three digitalisation methods and four different surface textures: halftone cyan, halftone black, solid-tone black and CMYK halftone pattern. The CMYK halftone pattern printed area, with 80% of each process colour was excluded from the evaluation set since the visual appearance of 320% total coverage was very similar to solid tone of black (K 100%). The selected halftone patterns present a particular challenge for automated image segmentation due to the complex texture – the halftone printed surface with damages (see Figure 2a, b and c).



Figure 2

Examples of selected images captured with digital microscope with (a) 50% cyan, (b) 50% black halftone printing and (c) rosette pattern of C 40% + M 40% + Y 50% + K 20%

To recognise the perfect folds without any surface damage is another demanding task for an image segmentation algorithm. In order to evaluate the segmentation quality for this class, 3 images were added to the selected subset, one image for each digitalisation method. Figure 3 presents an example of such a perfect folding.

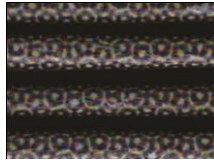


Figure 3

Example of folded paper without surface damages on the folding line

2.5 Quantitative Measures

To evaluate the segmentation performance of used colour representation, six performance measures have been used:

Misclassification error (ME) reflects the percentage of background pixels wrongly assigned to the foreground, and vice versa. For the two-class segmentation problem, ME can be simply expressed as:

$$ME = 1 - \frac{|B_o \cap B_T| + |F_o \cap F_T|}{|B_o| + |F_o|}, \quad (4)$$

where B_o and F_o denote the background and foreground of the ground truth image, B_T and F_T denote the background and foreground areas of the tested image, and $|\cdot|$ is the cardinality of the set. ME varies from 0 for a perfectly classified image to 1 for a total mismatch between reference and tested image [17].

The Hausdorff distance can be used to assess the shape similarity of the thresholded regions to the ground-truth shapes. Since the maximum distance is sensitive to outliers, Sezgin and Sankur [17] proposed a modification where the shape distortion is measured via the average of the Modified Hausdorff distances (MHD) over all objects. It can be defined as:

$$MHD(F_o, F_T) = \frac{1}{|F_o|} \sum_{f_o \in F_o} d(f_o, F_T), \quad (5)$$

where $d(f_o, F_T)$ denotes the minimal Euclidean distance of a pixel in the thresholded image from any pixel in the ground-truth image, and $|F_o|$ is the number of foreground pixels in the ground-truth image. Since an upper bound for the Hausdorff distance cannot be established, the normalization of the MHD metric can be performed by computing its reciprocal (with a small modification):

$$NMHD = 1 - \left(\frac{1}{1 + 0.2 \times (MHD - 1)} \right), \quad (6)$$

The measure derived by this formula has its optimal at 0 and its worst point at 1, as for the ME measure.

Positive false detection (PFD) is the proportion of background pixels wrongly assigned to the foreground object. Normalization (NPFDR) can be done using the overall number of pixels in the image. However, in order to maximize the covered range in [0, 1], the normalization was done using the number of background pixels.

Negative false detection (NFD) is the proportion of foreground pixels wrongly assigned to the background. Normalization (NNFD), following a similar logic as for the PFD, was performed using the number of foreground pixels.

As a derived measure the positive false-negative false detection ratio, or shorten false detection ratio (FDR) is defined, too. It serves as an auxiliary measure to make the balancedness of false detection values easy to read. It is defined as:

$$FDR = \begin{cases} \frac{PFD}{NFD}, & \text{if } PFD \geq NFD \\ \frac{NFD}{PFD}, & \text{else} \end{cases} \quad (7)$$

This measure has a minimum in 1, which is also its optimum, desired value, whereas its maximum value cannot be analytically determined. For this reason, the following method is proposed for the normalization (NFDR):

$$NFDR = 1 - \left(\frac{1}{FDR} \right) \quad (8)$$

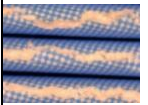
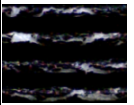

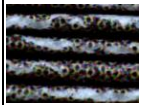
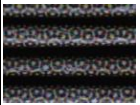
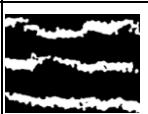
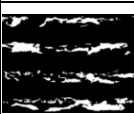
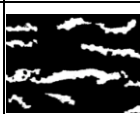
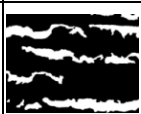

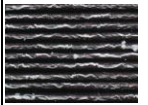




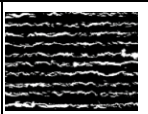
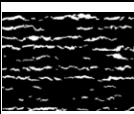
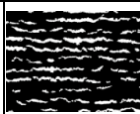
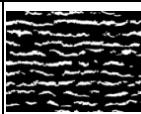



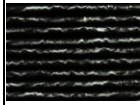


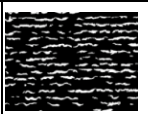
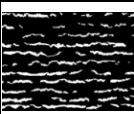
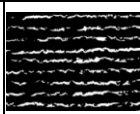
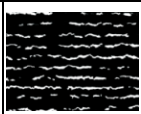

Relative foreground area error (RAE) was first proposed by [17] and is a modification of relative ultimate measurement accuracy (RUMA) measure used by Zhang (as cited in [17]). It is a comparison of object properties, more specifically the area of the detected and expected foreground. It is defined by the following formula:

$$RAE = \begin{cases} \frac{A_0 - A_T}{A_0}, & \text{if } A_T < A_0 \\ \frac{A_T - A_0}{A_T}, & \text{else} \end{cases} \quad (9)$$

where A_0 is the area of reference image and A_T is the area of the thresholded image. For a perfect match RAE is 0, while if there is zero overlap of the object areas, the RAE is 1.

All these measures require a reference or ground truth image, which was derived by hand, segmenting every sample image, marking just the cracked surfaces as foreground objects (see Table 4).

Table 4
Selected test images and their ground truth segmented pair

Type of images		Test images				
		1.	2.	3.	4.	5.
Microscope	Original					
	Ground truth					
Scan	Original					
	Ground truth					
Camera	Original					
	Ground truth					

The attempt to derive a combined measure was presented in [17] with the attempt to simplify segmentation quality evaluation. The authors proposed a simple averaging of the derived measures. However, since all of the measures carry different information about segmentation quality, simple arithmetic averaging might not give the best approximate of the overall quality measure. For this reason a further analysis of the measures were carried out during this research. As the result, the listed measures were divided into two groups: quantitative and qualitative measures. In the quantitative group there was just the ME measure. It defines the amount of misclassified pixels; hence, it is a direct measure of the overall error in detection. All the other measures belong to the second, qualitative group. In order to obtain a single, joint performance score (JPS), the following formula is proposed:

$$JPS = ME * \left(1 + \frac{NMHD + NFDR + RAE}{3} \right) \quad (10)$$

The NPDF and the NNFD were omitted from the JPS since their information is contained in the NFDR measure. However, they were used in further analysis to compare the behaviour of segmentation using different colour spaces. In order to crosscheck the derived results the evaluation was repeated by using just the ME measure. The two evaluation results were then compared.

3 Results and Discussion

A representative test image and its hand-segmented pair are shown in Figure 4a and 4b (as original and ground truth). In Figures 4c, d, e, f and g, the automatically segmented images are presented based on the luminance/intensity component from the HSI, HSL, HSV, CIE Lab and CIE xyY colour spaces, respectively.

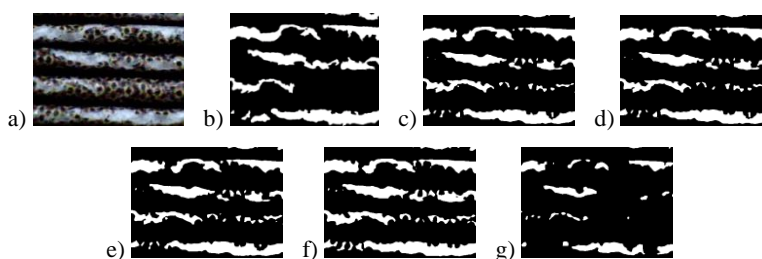


Figure 4

Example of a (a) test image, (b) its hand-segmented pair and the automatically segmented images based on (c) HIS, (d) HSL, (e) HSV, (f) CIE Lab and (g) CIE xyY colour spaces

All test images were processed by the segmentation algorithm using each of the colour spaces. The derived binary images were then pixel-wise compared to the appropriate ground truth image generating all six performance measures. Using (10), the measures were combined to form a single performance measure. These JPS values are presented in Table 5 for all test images and graphically shown in Figure 5.

Based on the average JPS values, the CIE xyY can be recognised as the best performing colour space, producing an average JPS value of 0.2293, which is significantly better than the second best performance (less than 60% of 0.3976). By further analysing each test image, it can be seen that segmentation accuracy using CIE xyY colour space results in an even performance throughout the whole test set (JPS value is around 0.2). There are three exceptions:

- *4_scan* and *4_pict* test images, which are presenting a cyan halftone substrate. Although, the CIE xyY based segmentation performed the best for these samples, significant over-detection could be observed (a detailed analysis on substrate type will be presented latter).

- 5_scan test image, which is showing a black halftone sample, was also poorly segmented with all colour spaces. As for the previous test images, significant over-detection could be observed by a slight advantage for the CIE xyY based segmentation.

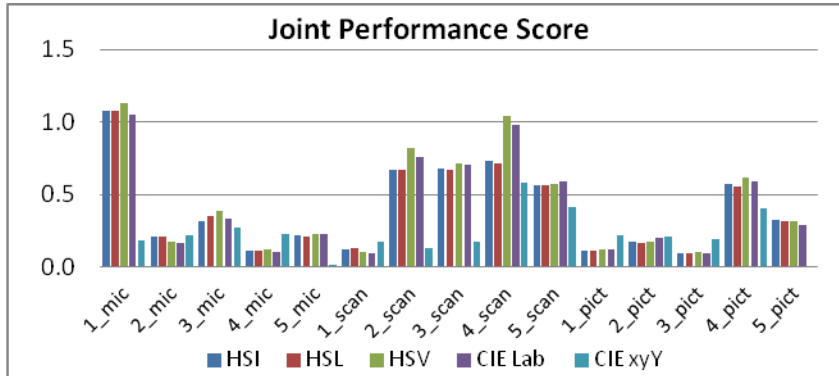


Figure 5

Joint performance score values for all test images and colour spaces

Table 5

Obtained joint performance score values for all test images

Test images	HSI	HSL	HSV	CIE Lab	CIE xyY
1_mic	1.072391	1.075942	1.127071	1.053293	0.185257
2_mic	0.208690	0.207112	0.172954	0.166763	0.218564
3_mic	0.317635	0.356105	0.386754	0.333201	0.276614
4_mic	0.112946	0.112013	0.120381	0.108577	0.228994
5_mic	0.220070	0.213945	0.229888	0.227978	0.019143
1_scan	0.124323	0.126958	0.105902	0.098954	0.178914
2_scan	0.670473	0.671548	0.821479	0.761589	0.133970
3_scan	0.674908	0.667136	0.713143	0.708855	0.174782
4_scan	0.732409	0.715503	1.044099	0.980283	0.586245
5_scan	0.566517	0.565210	0.575609	0.588636	0.415236
1_pict	0.114178	0.111927	0.123873	0.119140	0.217959
2_pict	0.171629	0.170618	0.177982	0.198934	0.211951
3_pict	0.095528	0.094056	0.103709	0.097141	0.190041
4_pict	0.570955	0.559490	0.613510	0.589825	0.402043
5_pict	0.326565	0.317093	0.317557	0.286431	0.000211
AVG	0.398614526	0.397643688	0.442260715	0.421306646	0.229328377

From the derived results (see Table 5), it can also be noticed that colour spaces HSI and HSL performed very similar for the entire set. Their average JPS measures differed by less than 0.001, whereas the highest discrete difference was 0.0385 for the 3_mic sample. Based on the JPS values, HSV can be announced as the worst performing colour space with the highest average JPS value. The CIE Lab colour space performed worse than HSI or HSL, which confirms the first results presented by the authors in [16].

In order to crosscheck the proposed joint measure, the same analysis was performed based only on the ME measure. These results can be seen in Figure 6. As can be noticed, the two charts are similar and the conclusions drawn from them are also very similar, with the differences just in magnitudes. However, it can be observed that the joint measure emphasises some differences according to favourable qualitative measures (see for example 4_mic or 2_scan samples).

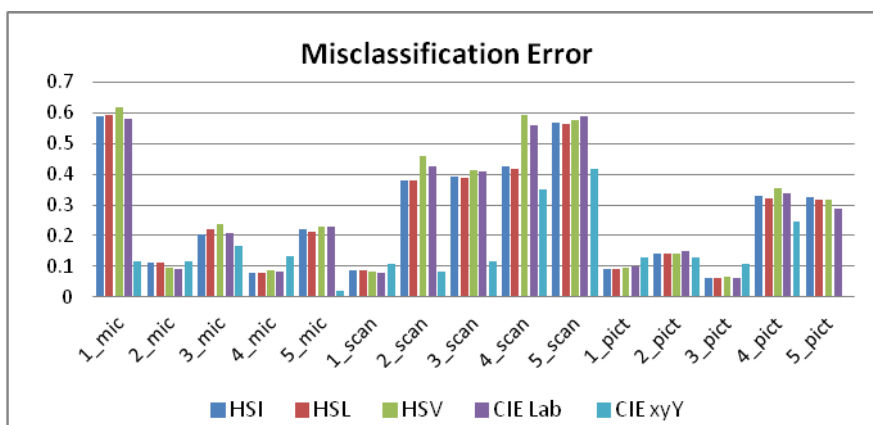


Figure 6

Obtained values for misclassification error for every test image

In order to analyse the performance of image segmentation for different digitalisation techniques and substrate types, the results of JPS were grouped accordingly. The results for image acquisition methods are shown in Figure 7. For this purpose, the results of the first 4 samples were averaged for each group. The fifth sample was omitted because it represents a perfect fold and it will be separately analysed. As can be seen, the best overall results (all colour spaces performed similarly well) were obtained for the acquisition by digital camera. This can be explained by the fact that this type of digitalisation reduces the most the effects of halftone and rosette patterns (because of the resolution and the closeness of the substrates). It is somewhat unexpected that the scanner based digitalisation resulted in the worst segmentations, since this method had ideal illumination, the best resolution and no effects of blurring or geometric distortions. Samples digitalised by scanner are over-detected, which could be explained by the method's high resolution (emphasized halftone and rosette patterns). The samples

derived by digital microscope are segmented somewhat better than those digitalised by the scanner. This is also an unexpected result, since one would expect that the halftone and rosette patterns are the most emphasized by this method. However, it seems that the high magnification helps the algorithm to correctly detect the base paper colour, hence helps avoiding misclassification of printed surfaces.

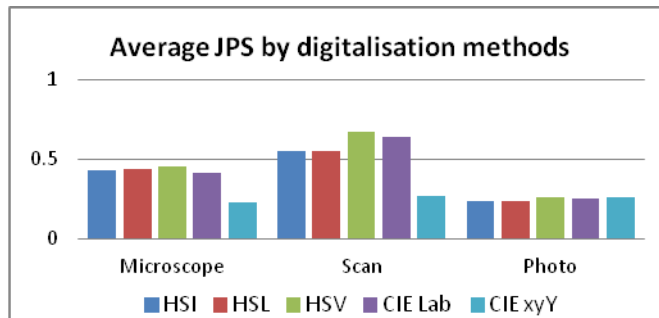


Figure 7

Average joint performance score values by digitalisation methods

Figure 8 presents performances by printed texture types. As was expected, the best results were derived for solid black printed areas (2_mic, 3_pict and 1_scan), where the difference between damages and printed surfaces is the biggest. Samples with rosette pattern are the second best segmented. At the first glance, this result is surprising because of the complex colour pattern (all four process colours are present). However, if we consider that the algorithm is working on grey-scale images, where these differences are less noticeable, then the results are less surprising. Halftone black and cyan samples were the worst segmented. This is especially true for the cyan halftone samples, where besides the halftone pattern, the relatively small difference between the grey-levels of cyan and paper colour renders the correct detection more difficult. It should also be noted that the CIE xyY based segmentation had balanced performance for all printed texture types.

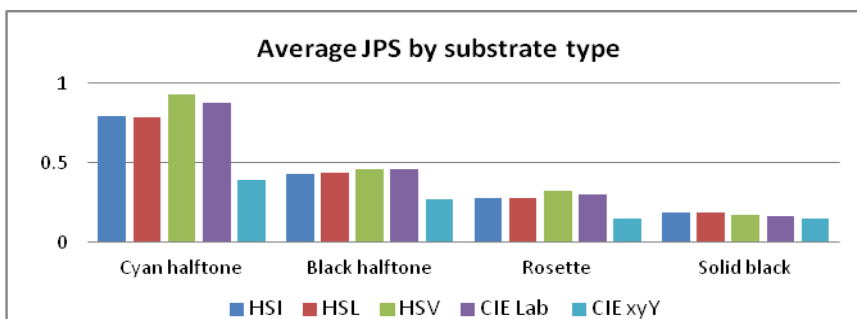


Figure 8

Average joint performance score values by printed texture type

The analysis of perfectly folded substrates (no damages visible) was set apart from the analysis of other substrates because of two reasons. On one hand, most of the measures could not be defined for these samples, making the joint performance score less relevant. On the other hand, these samples are special because there should be no foreground objects detected (there are no damages). This is a serious problem for most of the algorithms, since they are all configured to find a foreground object. Viewing this problem from the folding quality assessment point of view this would mean that there is a problem of detecting substrates, which does not require any action (the folding machine is configured well). The problem is even more serious if we consider that these substrates will be the most often presented to the OFQA. The results of this separate analysis are presented in Figure 9. As can be seen, the CIE xyY based segmentation is superior compared to the other three, having almost no over-detection for two (digitalised by microscope and digital camera) of the samples and 30% less for the scanned sample.

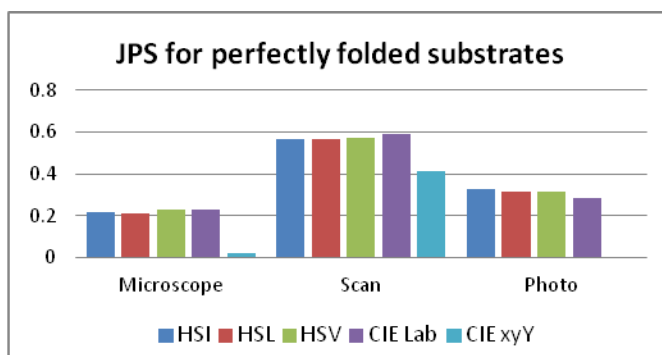


Figure 9

Joint performance score values for perfectly folded substrates

The presented results confirm the first conclusions derived by the average JPS, that the CIE xyY based segmentation performs the best from the chosen group of colour spaces. However, analysing the performances by positive and negative false detections, an oddity of the CIE xyY colour space is revealed. Namely, it tends to under-detect the damaged surfaces. This is visible in Figure 10, where the positive false detections are presented, and in Figure 11, where the negative false detections are presented.

It can be seen that the positive false detection of CIE xyY is the lowest for all samples, whereas for the negative false detection, the situation is the opposite, i.e. it produces significantly higher values. This behaviour is not always favorable, but could be successfully exploited, for example, in a two-stage segmentation process, where this step would be used to initially detect damages.

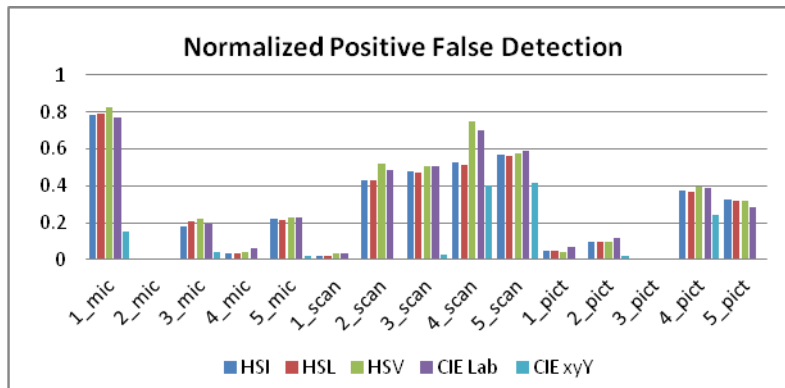


Figure 10

Normalized positive false detection values for all substrates and colour spaces

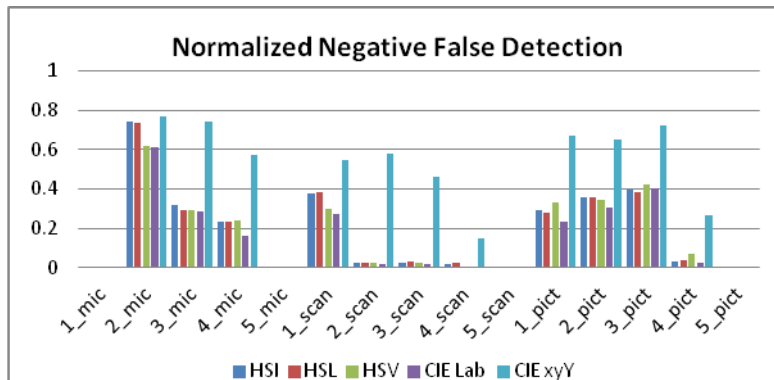


Figure 11

Normalized negative false detection values for all substrates and colour spaces

Conclusion

This paper presented a detailed evaluation of five colour spaces (HSI, HSL, HSV, CIE Lab and CIE xyY) in respect to their influence on image segmentation quality for a given set of test samples. The used segmentation algorithm was the Maximum Entropy algorithm, which is working on a 1D histogram. For this reason only the lightness (luminance) information of the colour spaces was used for evaluation. Six different measures have been derived to determine segmentation quality: misclassification error, modified Hausdorff distance, relative foreground area error, positive and negative false detection and their ratio. In an attempt to combine the measures into a unique grade, as a side effect of the research, a new combination method was proposed, which combines the metrics into a single measure. However, in order to crosscheck the results a separate verification, based only on misclassification error, was also performed. The two analyses gave similar results, showing that the segmentation shows best

performance by using CIE xyY colour space. However, it should be mentioned that segmentation based on this colour space tends to under-detect the damaged areas, while resulting in almost no false positive detections. This feature could be exploited in a two-step segmentation algorithm, where the 1D histogram analysis would be the first step. Also, the analysis of the results showed that the samples digitalised by a digital photo camera as having the lowest JPS, making this method the best choice for acquisition. As was expected, the solid black printed substrates were segmented with the least error, while the cyan halftone substrates were the hardest to segment. Based on the obtained results, further development and improvement of OFQA (Objective Folding Quality Assessment) could be achieved in refining the method's accuracy for real-time and in-line quality control on folding machines.

Acknowledgement

This work was supported by the Serbian Ministry of Science and Technological Development, Grant No.: 35027 "The development of software model for improvement of knowledge and production in graphic arts industry"

References

- [1] A. L. Barbieri, G. F. de Arruda, F. A. Rodrigues, O. M. Bruno, L. da F. Costa: An Entropy-based Approach to Automatic Image Segmentation of Satellite Images, *Physica A*, Vol. 390, 2011, pp. 512-518
- [2] Anon.: Folding and Creasing, Sappi's Technical Brochures, 2nd, revised edition, 2006 (Online) Available from: <http://www.sappi.com/NR/rdonlyres/F3F8F3B0-89B8-4528-9684-7D40C7A5817A/0/FoldingandCreasing.pdf> [Accessed 20 April 2008]
- [3] A. Yang, Y. Xie: From Theory to Practice: Improving the Foldcrack Resistance in Industrially Produced Triple Coated Paper, TappiPaperCon Conference Covington, Kentucky, USA, 2011, pp. 1845-1858 (Online) Available from: <http://www.tappi.org/Downloads/Conference-Papers/2011/2011-PaperCon-Conference/11PAP28.aspx> [Accessed 19 August 2011]
- [4] C. Barbier: On Folding of Coated Papers, Doctoral Thesis no. 56, Royal Institute of Technology, Department of Solid Mechanics, Stockholm, Sweden, 2004 (Online) Available from: http://www.t2f.nu/s2p2/S2P2_MS_9%20.pdf [Accessed 18 April 2008]
- [5] C.-I. Chang, Y. Du, J. Wang, S.-M. Guo, P. D Thouin: Survey and Comparative Analysis of Entropy and Relative Entropy Thresholding Techniques, *IEEE Proceedings of Vision, Image and Signal Processing*, Vol. 153, No. 6, 2006, pp. 837-850
- [6] C.-K. Kim, W.-S. Lim, Y. K. L: Studies on the Fold-Ability of Coated Paperboard (I): Influence of Latex on Fold-Ability during Creasing/Folding

- Coated Paperboard, Journal of Industrial and Engineering Chemistry, Vol. 16, No. 5, 2010, pp. 842-847
- [7] C. Zhang, P. Wang: A New Method of Color Image Segmentation Based on Intensity and Hue Clustering, in Proceedings of 15th International Conference on Pattern Recognition, Barcelona, Spain, Sept 3-7, Vol. 3, 2000, pp. 613-616
- [8] H. Cheng, X. Jiang, Y. Sun, J. Wang: Color Image Segmentation: Advances and Prospects, Pattern Recognition, Vol. 34, No. 12, 2001, pp. 2259-2281
- [9] International Standard: ISO 4628-4:2003, Paints and varnishes. Evaluation of degradation of coatings. Designation of quantity and size of defects, and of intensity of uniform changes in appearance. Assessment of degree of cracking
- [10] ImageJ (2011) Homepage: Image Processing and Analysis in Java, Plugins. (Online) Available from: <http://rsbweb.nih.gov/ij/plugins/index.html#segmentation> [Accessed 15 March 2011]
- [11] J. Delon, A. Desolneux, J. L. Lisani, A. B. Petro: Color Image Segmentation Using Acceptable Histogram Segmentation, Pattern Recognition and Image Analysis, Lecture Notes in Computer Science, Vol. 3523-2005, 2005, pp. 239-246
- [12] J. Eklund, B. Österberg, L. Eriksson, L. Eindenvall: Finishing of Digital Prints – a Failure Mapping, in: Proceedings of the International Congress on Digital Printing Technologies, IS&T NIP 18, San Diego, California, USA, 2002, pp. 712-715 (Online) Available from: www.t2f.nu/t2frapp_f_56.pdf [Accessed 5 May 2008]
- [13] L. Busin, N. Vandebroucke, L. Macaire, J. G. Postaire: Color Space Selection for Unsupervised Color Image Segmentation by Histogram Multi-Thresholding, in: Proceedings of International Conference on Image Processing, ICIP '04, Vol. 1, 2004, pp. 203-206
- [14] L. Busin, N. Vandebroucke, L. Macaire: Color Spaces and Image Segmentation, Advances in Imaging and Electron Physics, Vol. 151, 2008, pp. 65-168
- [15] M. Apro, D. Novaković, Sz. Pal: Objective Fold Quality Evaluation, in: Proceedings of "BlažBaromić" International Conference on printing, design and graphic communications, Senj, Croatia, 2009, pp. 21-24

-
- [16] M. Apro, D. Novaković, Sz. Pal: Evaluation of Image Segmentation Algorithms for Folded Substrate Analysis, in: *Advances in Printing and Media Technology*, Vol. XXXVIII, Edited by: N. Enlund and M. Lovreček, Iarigai, 2011, pp. 209-217, ISBN 978-3-9812704-2-6
- [17] M. Sezgin, B. Sankur: Survey over Image Thresholding Techniques and Quantitative Performance Evaluation, *Journal of Electronic Imaging*, Vol. 13, No. 1, 2004, pp. 146-165
- [18] O. Marques: *Practical Image and Video Processing Using MATLAB*, John Wiley & Sons, Inc. Hoboken, New Jersey, 2011
- [19] P. Alam, M. Toivakka, R. Carlsson, P. Salminen, S. Sandås: Balancing between Fold-Crack Resistance and Stiffness, *Journal of Composite Materials*, Vol. 43, No. 11, 2009, pp. 1265-1283
- [20] P. Colantoni, Al: Color Space Transformations, Technical report, 2004, (Online) Available from: <http://colantoni.nerim.net/download/colorspacetransform98.pdf> [Accessed 5 September 2011]
- [21] P. Rättö, J. Hornatowska: The Influence of Coating Colour Composition on the Crack Area after Creasing, *Nordic Pulp and Paper Research Journal*, Vol. 25, No. 4, 2010, pp. 488-494
- [22] R. E. Popil: Prediction of Fold-Cracking Propensity through Physical Testing, TappiPaperCon Conference, Atlanta, GA, USA, 2010 (Online) Available from: <http://www.tappi.org/Downloads/Conference-Papers/2010/PaperCon-2010-Conference/10PAP101.aspx> [Accessed 19 August 2011]
- [23] S. Jarek: Maximum Entropy Thresholding, 2004, (Online) Available from: http://ij-plugins.sourceforge.net/plugins/segmentation/Maximum_Entropy_Thresholding.pdf [Accessed 3 September 2011]
- [24] User manual for Horizon AFC546AKT cross folding machine, n.d.
- [25] V. Gidlöf, J. Granås, M. Dahlström: Functionality in Digital Packaging Printing, in: *Proceedings of the TAGA conference*, San Antonio, Texas, USA, 2004 [Online] Available from: http://www.t2f.nu/t2frapp_f_140.pdf [Accessed 5 May 2008]