# The Effect Of Binarization Algorithms Considering Color-To-Gray Scale Conversion Methods On Historic Document Images

**Paramasivam M E, Sabeenian R S, Dinesh P M**

**Abstract:** Character recognition from historic document images has been a challenge for computer scientists. The background of these documents contain enormous degradations, which gets visualized as noise. A few key preprocessing steps before character recognition are color-to-gray scale (C2G) transformation, binarization and segmentation. Numerous algorithms have been proposed, however there has been no generalized method for binarization invariant of the type image. Many C2G methods have been projected, but each of which directly mapped the noise from color image on to its gray scale version. This paper has tested the effect of C2G transformation on local and global binarization methods using images of DIBCO 2013 dataset. We have analyzed how variation of degradations, along with color contributions affects the binary image formation. Results show that not all binarization methods provide outputs for varying gray scale images. The gray scale image obtained by gamma correction based C2G conversion, eradicated noised to a maximum extent and hence has supported any kind of binarization algorithms. The qualitative measure for all obtained binary images was computed using F-Measure with their respective ground truths. To understand the contribution of each color channel in RGB color space, contrast-per-pixel (CPP) was computed. The identical F-Measure values for images with equal CPP, invariant of C2G transformation used has also been examined in this paper. To conclude, we have tested the interweaving relation between two major areas of image processing viz., Natural Image Processing and Document Image Processing.

**Index Terms**: Image processing, Image analysis, Image color analysis, Image enhancement, Document Image Processing, Natural Image Processing, Color-to-Gray Scale Conversion, Document Image Binarization, Contrast Measure, Contrast-Per-Pixel, Binarization Metrics, F-Measure.

## 1. INTRODUCTION

HISTORIC document images have a number of challenges embedded in it viz., removal of bleed-through & stains, retrieval of faded characters, machine understanding of multiple font colors and font styles used in the same image and many more. Dimitrios and et.al [1] have shown few samples of degraded old paper backgrounds. Text Extraction from the images with almost all such state-of-affairs has been a demand for more than a decade. Sahare and Sanjay [2] have reviewed various text extraction algorithms on scene and document images.

The process of character recognition involves expertized engines, but the input to these must be isolated characters. A detailed survey on various character recognition systems have been carried out by Yamamoto [3], Tang and et.al [4], Shivaprasad and et.al [5], Pal & Chaudhuri [6] and Ramakrishnan & Petta [7]. The step before recognition involves a number of pre-processing methods such as skew correction, noise reduction, segmentation of characters and many more. The recognition engines try to map isolated character images to their corresponding Unicode and finally give an editable blue-print of the document image.


- *Paramasivam M E is in the Department of Electronics and Communication Engineering, Sona College of Technology, Salem, TN, India (Phone:91-427-4099 774; fax: 91-427-4099 888; e-mail: sivam@sonatech.ac.in, sivam.sct@gmail.com).*
- *Sabeenian R.S. is in the Department of Electronics and Communication Engineering, Sona College of Technology, Salem, TN, India (e-mail: sabeenian@sonatech.ac.in, sabeenian@gmail.com).*
- *Dinesh P M is in the Department of Electronics and Communication Engineering, Sona College of Technology, Salem, TN, India (e-mail: dinesh@sonatech.ac.in, pmdinesh4253@gmail.com*


### 1.1 Document Image Binarization

Document Binarization falls in the domain of Document Image Processing (DIP) and researchers in this domain have shown least anxiety for analyzing the effect of various color-to-grayscale techniques on color document image. A gray scale image with noise, when binarized shall contain clumps of black regions representing noise. The segmentation process mis-associates these clumps present in the image as characters and henceforth the efficiency of recognition engines decrease. Cheriet [8] and many other scientific works claim that efficient segmentation is best accomplished by proper binarization of the image. Over past few years, new algorithms for binarization of historic document images have been proposed in contests like DIBCO [9] [10], H-DIBCO [11], IMPACT and many more [12] [13] [14].

### 1.2 Color-to-Gray (C2G) scale conversion

The advent of advanced electronic displays and printing devices has increased the usage of color images in day-to-day life. However, for computational purposes, it is preferred to have a gray scale rather than a color image. Thus, the process of converting color to a gray scale image becomes a predominant process in image processing. The most familiar method used is the rgb2gray command available in MATLAB [15]. This process of color-to-gray scale conversion falls in the domain of Natural Image Processing. Section **Error! Reference source not found.** details about a few of the C2G methods used in this paper. Document Image Processing and Natural Image Processing are viewed as two separate arenas in image processing. We have tried to emphasize the clutching relation between these two separate arenas. This paper has approached to understand:

- the effect of certain C2G techniques on binarization methods.
- how certain binarization techniques have provided consistent fair results invariant of gray-scale image produced by different C2G methods
- the unique behavior of images with each pixel holding

equal color contributions in RGB model.

## 2 IMAGE ALGEBRA

An image is a distribution of integers over a two dimensional space ($\mathbb{R}^2$). The possible values at a particular point (pixel) can be expressed as a set denoted as $\mathbb{S} = [1,2,\dots L-1, L]$. The total number of elements in $\mathbb{S}$ would be $2^n$, where n indicates the number of bits required to store each pixel in the image. A case in which n is 8, then the image be called as 8-bit image. If there are 'M' and 'N' pixels along each row and column respectively, then the size of the image can be interpreted as M x N and thus the space definition can be represented as $\mathbb{R}^{M\times N}$. A gradient variation of values in the set $\mathbb{S}$ over the $\mathbb{R}^{M\times N}$ is termed as a 'Gray Scale Image'.

### 2.1 Color Image

The spectrum of light has wide variety of colors, with each color falling on a unique wavelength spectrum [16]. The contribution of each color at a particular point can be indicated with integer values which when mapped on a $\mathbb{R}^{M\times N}$ space would increase the dimension to $\mathbb{R}^{M\times N\times D}$. Such representations on a higher dimension space are termed as Hyper Spectral Images. Detailing of such representations is beyond the scope of this paper, as we have restricted the third space $\mathbb{D}$ to have a maximum size of 3, for the sake of simplicity.

#### 2.1.1 RGB Color Space

The contribution of primary colors (Red, Green and Blue) at a particular pixel in a $\mathbb{R}^{M\times N}$ space would produce a color image ($\mathcal{C} \in \mathbb{R}^{M\times N\times 3}$). This mapping of primary colors on to a co-ordinate system is termed as RGB color space representation. On an algebraic perspective, the color image can be expressed as a concatenation of three matrices, with each matrix embodying the integer values of primary colors.

$$\mathcal{C}(x,y) = ([\mathcal{R},\mathcal{G},\mathcal{B}]^T : \mathcal{R},\mathcal{G},\mathcal{B} \in \mathbb{R}^{M\times N}, \forall(x,y)) \quad (1)$$

## 3 C2G METHODS

Color Image Processing is a tedious task for simulation and real-time effectuation. A mapping function f to transform all the vital contents from $\mathbb{R}^{M\times N\times 3}$ (higher order dimension) to $\mathbb{R}^{M\times N}$ (lower order dimension) space is termed as color-to-gray scale conversion and is expressed as

$$f(\mathcal{C},\mathcal{L}) = (f: \mathcal{C} \to \mathcal{L}, \mathbb{R}^{M\times N\times 3} \to \mathbb{R}^{M\times N}, \forall(x,y)) \quad (2)$$

where $\mathcal{L}(x,y) \in \mathbb{R}^{M\times N}$ represents a Gray Scale Image.

Chirstopher Kanan [17] has detailed about a dozen of such mapping functions and amongst those listed, half a dozen of them deal on weighted color channel for gray scale conversion. The forthcoming section shall brief a few C2G techniques used in this paper.

### 3.1 GIMP

Susstrunk [18] and Spaulding [19] have described on matching the human perception of brightness present in the RGB Color space using a weighted sum of color channels.

$$\mathcal{L}_{GIMP}(x,y) = [0.3\,\mathcal{R} + 0.59\,\mathcal{G} + 0.11\,\mathcal{B}] \quad ()$$

$\forall(x,y)$, where $\mathcal{R},\mathcal{G},\mathcal{B} \in \mathbb{R}^{M\times N}$ represents the primary color channels.

#### 3.1.1 RGB Color Space

International Lighting Commission CIE (Commission Internationale de l`Eclairage) introduced a number of standard color spaces, apart from the RGB model. Andreas Koschan & Mongi Abidi [20] elaborated the thought process behind these color spaces. With a focus on obtaining a gray scale version of the color image, we have considered the following color spaces listed below:

- YIQ & YUV Color Space
- Hue, Saturation and Intensity (HSI)
- Hue, Saturation and Value (HSV)
- $Y'C_bC_r$ Color Space

### 3.2 Average

#### 3.2.1 YIQ & YUV Color model

The analog TV transmission (NTSC) required a compressed form of image representation for which the YIQ color model [21] was used. The Y denotes the Luminance, while I and Q represent the in-phase and quadrature components. Further developments to the YIQ model led to the YUV model [21] in which U and V are the relative difference between the Red-Blue and Green-Magenta color components. The YUV model is widely used in PAL television transmissions.

This Luminance (Y) component in $\mathbb{R}^{M\times N}$ space is expressed as:

$$\mathcal{L}_{Average}(x,y) = \left[\tfrac{1}{3}(\mathcal{R}+\mathcal{G}+\mathcal{B})\right], \forall(x,y) \quad (4)$$

#### 3.2.2 HSI Color Space

The Hue, Saturation and Intensity color space model [15] gives an alternative way of representing colors based on the level of each color along the radial axis. Hue and Saturation indicate the purity and white dilution of colors available at a pixel. The Intensity component is an arithmetic average of primary color channels. The Luminance component of YIQ and YUV model is analogous to Intensity in HSI model, thereby representing the same form in $\mathbb{R}^{M\times N}$ space. Since, the method perceives grays scale values by computing the average of integer values in color channels and hence we have named it as Average.

### 3.3 Max-Min

#### 3.3.1 HSV Color Model

In HSV Color space, the Hue and Saturation are the same as in HSI model. The Value component computes the average of the most and least dominant color channels in a RGB color space. The gray image obtained by implementing Equation 5 is termed as Max-Min in this paper.

$$\mathcal{L}_{Min-Max}(x,y) = \tfrac{1}{2}(\max(\mathcal{R},\mathcal{G},\mathcal{B}) + \min(\mathcal{R},\mathcal{G},\mathcal{B})) \quad (5)$$

### 3.4 Maximum

For a given color image ($\mathcal{C} \in \mathbb{R}^{M\times N\times 3}$) in RGB color space, the highest value among the three color channels is evaluated for each pixel. These values together form a gray scale image in $\mathbb{R}^{M\times N}$ space. We have termed the method as Maximum.

$$\mathcal{L}_{Max}(x,y) = [\max(\mathcal{R},\mathcal{G},\mathcal{B}), \forall(x,y)] \quad (6)$$

### 3.5 Luminance

#### 3.5.1 $Y'C_bC_r$ Color Space

With the YUV and YIQ color model supporting the analog version, $Y'C_bC_r$ is employed for digital representations [22]. When compared to other color spaces, this model ensures a gamma corrected version of each color channel ($\mathcal{R},\mathcal{G}$ and $\mathcal{B}$)

before transformation.With $C_b$ and $C_r$ representing the Chrominance components, $Y'$of this model varies from that of $Y$ in YUV and YIQ model. In order to distinguish between the analog and digital versions we have added the prime representation. Gray scale image obtained by this method in $\mathbb{R}^{M\times N}$ space has been termed as Luminance.

$$\mathcal{L}_{Lum}(x,y) = [16 + 65.739\,\mathcal{R}' + 129.057\,\mathcal{G}' + 25.064\mathcal{B}'] \quad (3)$$

### 3.6 Optimization with Linear Model

The Sections **Error! Reference source not found.** to **Error! Reference source not found.** have primarily dealt with maximizing the visual perception parameters (contrast, brightness, etc…). Qui and et.al [23] have approached on optimal transformation of statistical parameters (variance) for a weighted color transformation. A linear model of gray scale image is expressed as

$$\mathcal{L} = [\alpha.\mathcal{R} + \beta.\mathcal{G} + \gamma.\mathcal{B}], \forall(x,y) \quad (8)$$

To determine the values of $\alpha, \beta$ and $\gamma$, the optimization constraints are fixed as

$$(\alpha,\beta,\gamma) = \arg\max(\sigma_{\mathcal{L}}^2) \text{ and } \alpha+\beta+\gamma = 1 \quad (9)$$

where $\sigma_{\mathcal{L}}^2$ represents the variance of grayscale image ($\mathcal{L}$).

Micheal Ng & Zhengmeng [24] added up a regularization term to the constraints defined in Equation 8 with a motive to maximize the contrast transformation from the higher to lower space. This paper makes use of the improvised model proposed by Micheal Ng & Zhengmeng and we have termed it as Linear - Optimize.

### 3.7 Optimization with Non-Linear Model

Mark Grundland & et.al [25] and Jiaya Jia & et.al [26] have refereed the C2G process as a `De-colorization' approach. Kim and et.al [27] proposed a non-linear global mapping by converting the color image to CIE L C H color space. The functional model is based on trigonometric polynomial. The optimization problem was defined such that there would be minimal image gradient during the $\mathbb{R}^{M\times N\times 3} \rightarrow \mathbb{R}^{M\times N}$ transformation. We have named this method as Decolorize in our paper. The methods detailed above have focused on preserving contrast of the gray scale image, which is also contributed by noise present in the color image. For document images, the goal of binarization is targeted towards minimizing the contrast in text class. This diversified goal has led to limited performance of binarization methods till date. To the best of our knowledge, we have not found any method providing a combined solution of noise reduction and C2G mapping, which would be much useful particularly for Historic Document Image Processing.

## 4 IMAGE BINARIZATION

Binarization of a gray scale image is a common preprocessing task before segmentation. The goal of this process is to hook up contents required for segmentation (foreground information) to the minimum gray value and all other unrelated information (background information) to the highest gray value. The process being termed as binarization, some literatures accost the highest and lowest gray level to 1 and 0 gray levels respectively. To list a few, the contributions of Gatos & et.al [28] and Mohamed Cheriet [29] have satisfied the needs of effective binarization to a great extent. On an algebraic perspective, the weight to represent each pixel is reduced to one in a $\mathbb{R}^{M\times N}$ space. The transformation function (g) can be expressed as

$$g(\mathcal{L},\mathcal{I}) = [g{:}\mathcal{L}\rightarrow\mathcal{I}, n = 8 \rightarrow 1, \forall(x,y)] \quad (10)$$

The binarization process tries to make the darker pixels darker and lighter ones lighter. Furthermore, it also tries to maximize the contrast of the image.

### 4.1 Document Image Binarization

Given a set of integer values ($\mathbb{S}$) at a particular pixel on a $\mathbb{R}^{M\times N}$ space, binarization divides the set into two classes, $C_0 = [1,2,\dots t]$ and $C_1 = [t+1, t+2, \dots L-1, L]$. In a document image, $C_0$ represents the textual information perceived by human vision system (HVS), while $C_1$ embarks the noise and background of the image.

Cheriet and et.al [30] has algebraically represented the binary image obtained from grayscale image as

$$\mathcal{I}(x,y) = \begin{cases} 0, \mathcal{L}(x,y) < t \\ 1, \mathcal{L}(x,y) > t \end{cases} \quad (11)$$

The threshold 't' can be determined, either over the entire $\mathbb{R}^{M\times N}$ space or on a local space. While the former aims at determining the single threshold for the entire image, the latter computes 't' for each window slided over the image. A detailed survey on various binarization techniques has been carried out by researchers [31], [32], [33] and [34].

### 4.2 Global Thresholding

Global Thresholding divides histogram of the entire image into two, such that it represent two classes $C_0$ and $C_1$. The class $C_0$ shall contain a group of lower gray values, while the other class $C_0$ groups up higher gray values.

#### 4.2.1 Otsu Method

Otsu [35] computed the optimal threshold based on Discriminant Analysis of Statistics as described in Fukunaga [36]. The optimal thresholds can be estimated by maximizing any one of the following ratios

$$\lambda = \frac{\sigma_b^2}{\sigma_w^2}; \eta = \frac{\sigma_b^2}{\sigma_t^2}; \kappa = \frac{\sigma_t^2}{\sigma_w^2} \quad (12)$$

where $\sigma_b^2$ is the between-class variance, $\sigma_w^2$ represents the within-class variance and $\sigma_t^2$ denotes the total variation of the two classes $C_0$ and $C_1$. The newly figured threshold is expressed as $t^* = \arg\max(\eta)$ and the expression holds good for $\lambda$ and $\kappa$.Global thresholding miserably fails on images with varying brightness and contrast. The noise present in the image is of wide varieties thereby providing unsatisfactory results for Otsu Algorithm.

#### 4.2.2 Multiple Thresholding

Otsu assumes there are two peaks in the gray scale histogram, one representing foreground and other the background. Optimal threshold $t^*$ lies in the valley between two peaks. The histogram for images having complex backgrounds / enormous noises shall have multiple peaks, thereby making Otsu infeasible to find a single $t^*$.Gatos and Papamarkos [37] proposed an algorithm which received the number of peaks to be looked on the histograms as a input from the user. This algorithm iteratively checks for maximum number of peaks in the histogram, using Hill Clustering approach proposed by Du-Ming Tsai & Ying-Hsiung Chen [38]. To determine the global threshold amongst the different valley values, the Golden Search Algorithm was utilized. This technique performs to its best when gray values of the two

classes are very close.

### 4.3 Local Thresholding

Global methods extract objects for simple and uniform backgrounds at high speeds while, local thresholding eliminates varying background at a price of longer processing time. In case of document images, statistical values are extracted for each sliding window along the entire image and thresholds are computed based on these statistical values. For document images, the size of window is usually fixed to fit at-least size of two characters present in the document image, as proposed by Gatos, Pratikakis and Perantonis [39].

#### 4.3.1 Niblack Method

This is one of the most simplest statistical method of local thresholding proposed by Wayne Niblack [40]. Pixel-wise thresholds for rectangular windows are calculated based on the equation given as

$$t_{Niblack} = \mu_{..} + k.\sigma_w \quad (13)$$

where $t_{Niblack}$ denotes the threshold for a window having mean $\mu_{..}$ and standard deviation $\sigma_w$. $k = -0.2$, was determined by performing a number of experiments.

#### 4.3.2 Sauvola Method

J. Sauvola and M. Pietikainen [41] improvised the Niblack method by including a dynamics of standard deviation (R) in the expression.

$$t_{Niblack} = \mu_w + \sqrt{\left(1 - k\left(1 - \frac{\sigma_w}{R}\right)\right)} \quad (4)$$

The value of R is usually fixed to 128 and k to 0.5. This algorithm exhibits its best performance for images having background gray values marginally far away from text pixels.

#### 4.3.3 NICK Method

Each above listed methods has its own advantages and disadvantages. Khurshid and et.al [42] tried to wed up the advantageous characteristics of all these techniques to form an better approach and named it as NICK.

$$t_{NICK} = \mu_{..} + k \sqrt{\frac{\sum p_i^2}{\cdot\cdot} - p_w} \quad (15)$$

while k retains the Niblack constant factor of $-0.2$, $p_i$ indicates the pixel values in the window size $j \times k$. The $\mu_w$ used in the above equations 14 and 15 represent the windowed mean, as used in Niblack Method.

This method provides satisfactory results for images with drastic background variations. One example for such degradation in historic document images is the presence of mild to heavy oil stains at a few regions.

#### 4.3.4 Wolf Binarization

While Niblack algorithm over-performs by binarizing the background and noise, Sauvola's method provides disappointing results at regions where the pixel values of classes $C_0$ and $C_1$ are very close. Addressing this issue, Wolf [43] added some more factors, such as the normalized contrast and total mean of the image to the thresholding equation expressed as:

$$t_{Wolf} = (1 - k) \times \mu_w + k \times \mu + k \times \frac{\sigma_w}{R}(\mu_w - \mu) \quad (5)$$

The value of k is same as that in Sauvola, while $\mu$ uses the maximum gray value in the image, R is the maximum standard deviation obtained over all local neighborhoods in the $\mathbb{R}^{M \times N}$ space.

The method give better results when compared to those described in Sections 4.3.1 and Section 4.3.2. The performance of Wolf Binarization degrades at places of sharp background changes, and is the best for small noise patches in background.

#### 4.3.5 Bradley

Wellner [44] proposed an adaptive thresholding technique where each pixel is analyzed with the average of its surrounding pixels. The method trusts on the scanning order of the pixels which is a great disadvantage. Roth and Bradley [45] approached the problem at the cost of an additive iteration. The concept of Integral images is used here for computing the threshold value on an adaptive basis.

## 5 EVALUATION

Evaluation of binarized document images is an important pipeline in the process of analysis and recognition of characters from historic documents. Even though the human vision system is capable of evaluating the obtained output, a computation parameter for quicker and machine oriented evaluation is required. Manmatha and et.al [46] directly fed the results of binarization to OCR engines, such as ABBYY FineReader (in particular for English Characters) and based on the number of successful recognition of characters, the binarization method was rated. However, such methods seem to be a bumpy approach to evaluate binary images. Nitrogiannis and et.al [47] have used Skeletonized Ground Truth ($\mathcal{SG}$) Image along with the actual ground truth ($\mathcal{AG}$) for computing the precision and recall metrics. For a binary image ($\mathcal{I}$), obtained by applying various algorithms in Section 4, the metrics can be defined as given below.

### 5.1 Precision

Precision refers to the percentage of ground truth with the obtained binary image, expressed as

$$Precision = (\sum \mathcal{AG} \times \mathcal{I})/(\sum \mathcal{AG}) \times 100 \quad (17)$$

### 5.2 Recall

Recall is defined by the equation given below.

$$Precision = (\sum \mathcal{SG} \times \mathcal{I})/(\sum \mathcal{SG}) \times 100 \quad (6)$$

### 5.3 F-Measure

In the diagnostic testing of binary data, F-Measure (some texts also call it as $F_1$ score) is a sub-contrary mean of Precision and Recall, expressed as

$$F - Measure = (2 \times Precision \times Recall)/(Precision \times Recall) \quad (19)$$

### 5.4 Contrast-Per-Pixel (CPP)

In a RGB color space, each color channel can be separately visualized as a gray scale image. For example, the green and blue components can be nullified and red component alone can be modeled a single gray scale image. Paramasivam & Sabeenian [48] conducted this experiment of visualizing individual color channels as gray scale image for the DIBCO 2013 datasets and found that in most cases the HVS identified any one of the color channel conveying almost all the vital information of the color image. There were also certain images in which there was no any indispensable difference amongst the three channels. To identify an equivalent computational parameter of this perception by the HVS, we have used the Contrast-Per-Pixel (CPP) of a color channel as a basic parameter. For a given gray scale image $\mathcal{L}(x, y)$, the contrast of each pixel in an image may be defined as $\Phi$ where,

$$\Phi = \sum_{i=1}^{x}\sum_{j=1}^{y}\frac{\mathcal{L}\oplus s}{g} \quad (20)$$

and the template s in Moore Neighborhood is

$$\tau(x,y) = \begin{cases} 1, if\ \tau = (x,y) \\ -1, otherwise \end{cases} \quad (7)$$

The average of all these variations over entire the image gives Contrast-Per-Pixel (CPP) $\frac{1}{xy}\sum_{i=1}^{x}\sum_{i=1}^{y}\Phi$. Higher the value of the CPP indicates more contrast varying content in the image.

## 6 EXPERIMENTAL RESULTS & DISCUSSION

The DIBCO 2013 [11] dataset was considered for experimentation. The dataset has a total of sixteen images with equal number of handwritten (File Names: HW01 to HW08) and printed (File Names: PR01 to PR08) document images. The corresponding actual ground truths for each of the images are also available in the dataset. Each image was subjected to C2G methods listed in Section 3. The resultant gray scale images were in-turn administered to binarization algorithms mentioned in Section 4.2. The 49 resultant binary images were evaluated using F-measure metrics. The values thus obtained for each of the images has been tabulated in the Tables 1 to 16. The CPP values of primary color channels have been mentioned in each table captions. Highest F-Measure for a given image in the tables indicates closest resemblance of output to its corresponding ground truth.

### 6.1 Influence of C2G Transformation

#### 6.1.1 GIMP

This paper has used seven different C2G conversion methods. The commonly used rgb2gray command of MATLAB utilizes a weighted color conversion. The F-measure values under this method have not turned up with appreciable values for any of the sixteen images. Referring to Equation 3, maximum weight has been given to the green channel when compared to the other two. This weighing has negated the goal of binarization to capture required textual information, which is usually present in red channel as experimented by Paramasivam & Sabeenian [48].

#### 6.1.2 Average

An alternate method of gray scaling is to average up the color pixel values. In this method too, the color channel containing the text class has not been focused and hence the F-Measure values are closely in abidance to the GIMP method.

#### 6.1.3 Maximum

Paramasivam & Sabeenian [48] have demonstrated the process of visualizing each $\mathbb{R}^{M\times N}$ plane of a $\mathbb{R}^{M\times N\times 3}$ space as a separate gray-scale image. In case of DIBCO 2013 dataset, it is evident that the most dominant color channel in RGB color space possessed maximal foreground information. We have identified the dominant color channel by investigating the CPP values of each channel. The image PR05 and PR08 being rich in chromatic content have in turn exhibited higher values of CPP for each color channel. This has empowered the gray scale image obtained by maximum C2G conversion to map more foreground content, indicated by higher F-Measure values.

#### 6.1.4 Min-Max

The method has averaged the maximum and minimum color pixel values. Hence, the binarization results are very adjacent to Maximum C2G conversion. Thus, the method has not excelled by providing any best F-Measure values, due to the same reason as mentioned for GIMP and Average.

#### 6.1.5 Optimize & Decolorize

A visual perception of images HW01, HW02, HW03, PR01, PR02, PR03, PR06 and PR07 show a uniform background with very minimal degradation. Owing to this characteristic, weights for the color channels need to vary. Dynamic weight allocation perceptually for every image has been made possible by methods Optimize and Decolorize. The HW03 color image contains two different font colors with one color very close to background region. The linear optimization method for determining the weights to color channels has afforded the best gray scale image. To support this argument, the Maximum C2G method has provided poor binarization results, such that it has missed out the weak color font (lower gray values). Images HW01, HW02 and PR03 possess a uniform background with very minimal degradation. The non-linear optimization method of gray-scale formation (Decolorize) has targeted weights based on the individual color contribution. The F-Measure values of other images under this uniform background category have not been ranked the highest; however, these values are marginally better when compared to other C2G methods.

#### 6.1.6 Luminance

The Luminance C2G conversion initially carries out a Gamma correction and then weighs the color channels to obtain the gray scale image. For all sixteen images of DIBCO 2013 dataset, this C2G conversion has given satisfactory F-measure values coinciding with the human vision system. Due to the fact that most of these images are filled up with variety of degradations and noise, this gamma correction reduces noise. In certain cases, the F-measure value obtained by Luminance C2G is low. This primarily is due to the presence of larger degradations viz., oil stains on the image. Eg., PR04 has a large oil stain on the right side of the image, thereby deteriorating the performance of Otsu on Luminance C2G.

### 6.2 Performance of Binarization methods

#### 6.2.1 Otsu

The Otsu method computes $t^*$ using the class variances and total variance of the gray scale image. The presence of multiple and close by peaks in the gray scale histogram, prevents Otsu in identifying the exact position of $t^*$. The gray scale image of historic documents restrain tremendous amount of noise, such that its histogram contains multiple peaks. It is therefore difficult for this method to identify an ideal threshold. The tabulated F-Measure values affirm the non-availability of metrics value in the Otsu Column.

#### 6.2.2 Multi-Threshold

The images HW04, HW05, HW06, HW08, PR01, PR02, PR06 and PR07 possess identical degradation. The characters written or printed behind the document have bleed through there by causing a humiliation in the image. Looking on the histogram of these images in gray scale, the peaks of text class and bleed through characters fall very close. With Otsu method failing to provide binarization results for such scenarios, Multiple Thresholding iteratively identifies peaks to

determine the threshold value. The C2G methods Optimize and Decolorize fix up gray values in such a way, that multiple thresholding is unable to identify valley values, thereby not providing any F-Measure value.

#### 6.2.3 Niblack

The Niblack binarization method utilizes the first order statistical features (mean and standard deviation) obtained for each window to compute the localized threshold. This method over-performs by computing very ideal thresholds for every small window. Hence, the computational metrics of Niblack method for all sixteen images using the seven different C2G methods indicates low values.

#### 6.2.4 Sauvola

The images HW01 and PR03 have a uniform background with very negligible noise on the document. Sauvola binarization utilizes the localized statistical parameters for computing the threshold. These two images being ideal for Sauvola, have provided higher values of F-Measure. The PR01 image has a small amount of degradation on a uniform background. The gamma corrected C2G method washes away these insignificant haphazardness, thereby enabling Sauvola to provide better results. This example shows the mutual dependence of C2G transformation and binarization methods.

#### 6.2.5 Wolf

While Niblack over-performed on background regions, the Sauvola works only in scenarios of ideal backgrounds. To overcome the effects of these methods, Wolf proposed a binarization capable of providing better results for image with small black spots in the background. HW02 containing such a type of degradation provides good F-Measure values for Wolf binarization.

#### 6.2.6 NICK

The NICK approach was developed by combining other binarization algorithms. With other methods computing localized threshold using first order statistical features, NICK considers pixel values also for computation. This has enabled NICK to provide better binarization for all images invariable of the C2G method used. NICK was capable to provide best results mainly for images with gradient background variations. Historic document image have such variations due to oil stains on a localized region. The images PR04, PR05 and PR08 contain oil stains and hence have enabled NICK binarization to work better when compared to other methods. Image PR06 has a heavy bleed through and even in such a scenario NICK performs well by exhibiting good F-Measure value.

#### 6.2.7 Bradley

HW03 is the only image in this dataset containing two different colors for fonts. With Optimize C2G providing the best gray scale image, it is now essential to exactly classify text pixels very close to background pixel values under foreground section. Bradley approach uses an adaptive technique for fixing the threshold based on the average of surrounding pixels. With HW03 necessitating such a thresholding approach, Bradley has provided the best results in this case.

#### 6.2.8 Images with equal color contribution

Image files HW04, HW05, HW06 and HW07 have equal values of CPP for all the three color channels, indicating a balanced primary color contribution. These images demonstrate a consistent yield for a given binarization methods irrespective of the C2G transformation. For example, in case of HW04 Image, the F-Measure for Otsu(41.13), Multi-Threshold (92.91), Sauvola (85.27), Wolf (92.27), NICK(84.78) and Bradley (68.59) are unvaried (except for meager deviations) regardless of the C2G transformation used. This is also applicable for Tables Table5, Table 6 and Table 7. Owed at equal contribution of color channels, the gray-scale image produced by various C2G techniques is almost the same. Thus, despite the variations in binarization methods, the results are same for a particular method. On the C2G perspective, the Luminance method has shown evidently good results for all these four images, due to the gamma correction carried out before $\mathbb{R}^{M\times N\times 3} \rightarrow \mathbb{R}^{M\times N}$ transformation. In case of HW05, the ink seepage (noise) is more when compared to the available text region. The ink seepage characters and text parts have gray-values very close and hence to identify a threshold value in the valley of histogram, Multi-Thresholding supports a lot. The gamma correction for HW05 does not assist in removing the ink seepage and hence Luminance method of C2G does not provide satisfactory results. To support this argument, HW07 has a nominal and equal distribution of ink seepage and hence Luminance provides appreciable results. The HW07 image has very less pixels in text class and more in background class and does not possess any bleed through. Hence, except for this image, all the other images under equal CPP category have exposed good results for multi-thresholding approach. Amongst the sixteen images in DIBCO dataset, HW05 has shown a unique performance. This is due to two major reasons, (a) equal contribution by all color channels in RGB space and (b) presence of noise to a greater extent when compared to textual region.

## 7 CONCLUSION

This paper has attempted to ascertain the relation between two fields of image processing viz., natural image processing and document image binarization. This analysis indicates that, there can have been no generalized method for binarization. The contrasting goals of the two areas, mentioned above have led to such a surprising result. Among the seven C2G transformations, depending on the degradation present in the image, each method responded distinctly from one another. Elaboration on each method's performance has been discussed in the above section. Though we cannot assure one particular C2G transformation as the best for historic document image processing, the gamma correction based Luminance has provided results for any kind of binarization algorithm. Looking on the binarization view, the local binarization method Niblack has provided drastically poor results. The primary disadvantage of Niblack is that it adds up a bulky quantity of binarization noise in the background region. Global binarizaton method Otsu, is ineffective to identify valley points when the histogram has multiple peaks or when peaks are very close to each other. The NICK algorithm has consistently provided fair F-Measure values for majority of images. Four images with equal color contribution have their unique characteristics. They have exhibited identical F-Measure values for any kind of C2G transformation and binarization method (except for Luminance). Due to the non-variational color contribution, the weighing of color channels has not provided any variation in gray scale image. However,

the gamma correction C2G transformation alone has provided a variation in F-Measure.

***Table 1:*** *F-Measure Metrics for HW01 - DIBCO Dataset 2013 (CPP of Color Channels R: 0.177, G: 0.170 and B: 0.156)*

| HW01 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | – | 59.91 | 40.70 | 85.22 | 80.61 | 81.31 | 84.34 |
| Average | – | 60.36 | 40.52 | 86.00 | 81.35 | 83.37 | 84.92 |
| Maximum | – | 4.12 | 39.96 | 81.88 | 76.87 | 77.90 | 81.95 |
| Min-Max | – | 65.94 | 40.40 | 86.10 | 81.34 | 82.40 | 85.00 |
| Optimize | – | – | 40.56 | 86.56 | 81.73 | 82.95 | 85.33 |
| Decolorize | – | – | 40.49 | 87.55 | 80.74 | 85.99 | 86.67 |
| Luminance | 82.64 | 65.22 | 40.31 | 83.60 | 80.71 | 78.82 | 82.75 |

***Table 2:*** *F-Measure Metrics for HW02 - DIBCO Dataset 2013 (CPP of color channels R: 0.370, G: 0.347 and B: 0.306)*

| HW02 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | – | 85.38 | 60.98 | 89.98 | 91.32 | 90.95 | 84.41 |
| Average | – | 81.02 | 60.30 | 89.21 | 90.91 | 90.38 | 82.97 |
| Maximum | – | 87.35 | 62.20 | 90.84 | 91.23 | 89.96 | 86.03 |
| Min-Max | – | 80.64 | 59.82 | 88.73 | 90.43 | 90.54 | 82.15 |
| Optimize | – | 87.35 | 62.20 | 90.84 | 91.23 | 89.96 | 86.08 |
| Decolorize | – | – | 61.04 | 89.53 | 91.43 | 90.74 | 83.76 |
| Luminance | – | – | 59.93 | 88.72 | 90.71 | 90.66 | 82.18 |

***Table 3:*** *F-Measure Metrics for HW03 - DIBCO Dataset 2013 (CPP of color channels R: 0.366, G: 0.349 and B: 0.324)*

| HW03 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | – | 52.34 | 50.14 | 75.28 | 73.04 | 75.79 | 80.29 |
| Average | – | 53.80 | 50.16 | 75.95 | 73.32 | 75.82 | 80.94 |
| Maximum | – | 50.44 | 46.48 | 63.97 | 63.57 | 62.91 | 68.57 |
| Min-Max | – | 52.88 | 49.96 | 74.23 | 71.70 | 75.12 | 79.93 |
| Optimize | – | – | 51.34 | 81.77 | 77.99 | 82.78 | 84.24 |
| Decolorize | – | – | 50.01 | 79.16 | 73.99 | 81.01 | 83.12 |
| Luminance | 75.43 | 52.21 | 49.75 | 74.05 | 73.14 | 73.17 | 78.62 |

***Table 4:*** *F-Measure Metrics for HW04 - DIBCO Dataset 2013 (CPP of color channels R, G and B: 0.184)*

| HW04 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | 41.13 | 92.91 | 31.31 | 85.67 | 92.27 | 84.78 | 68.59 |
| Average | 41.13 | 92.91 | 31.31 | 85.67 | 92.27 | 84.07 | 68.59 |
| Maximum | 41.13 | 92.91 | 31.31 | 85.67 | 92.27 | 84.78 | 68.59 |
| Max-Max | 41.13 | 92.91 | 31.31 | 85.67 | 92.27 | 84.78 | 68.59 |
| Optimize | 41.13 | – | 31.31 | 85.67 | 92.27 | 84.78 | 68.59 |
| Decolorize | 41.13 | – | 31.31 | 85.67 | 92.27 | 84.07 | 68.59 |
| Luminance | 94.99 | 91.98 | 31.27 | 87.45 | 92.27 | 86.64 | 72.64 |

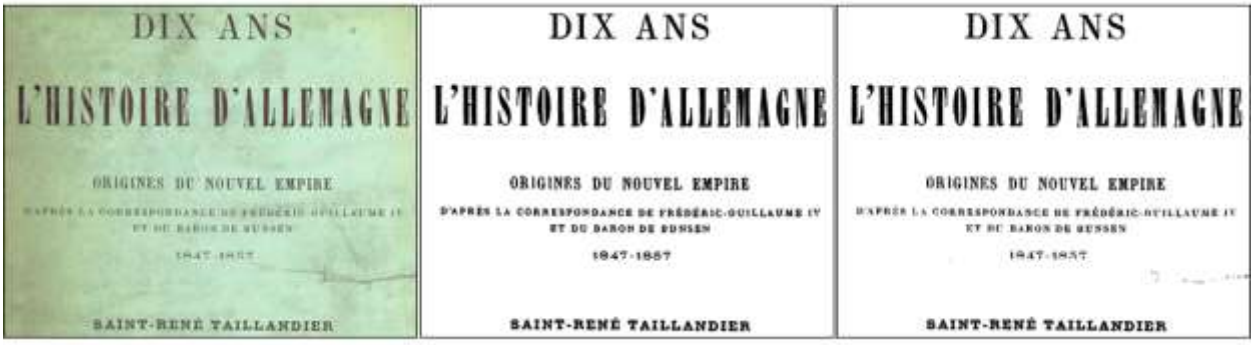


***Figure 1****. PR03 - DIBCO Dataset 2013. From Left : Original Image, Ground Truth, Decolorize C2G & Sauvola Binarization (F-Measure : 95.28 )*

***Table 5:*** *F-Measure Metrics for HW05 - DIBCO Dataset 2013 (CPP of color channels R, G and B: 0.224)*

| HW05 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | 39.42 | 90.17 | 19.19 | 42.31 | 56.76 | 44.37 | 48.35 |
| Average | 39.42 | 90.17 | 19.19 | 42.31 | 56.76 | 49.54 | 48.35 |
| Maximum | 39.42 | 90.17 | 19.19 | 42.31 | 56.76 | 44.37 | 48.35 |
| Min-Max | 39.42 | 90.17 | 19.19 | 42.31 | 56.76 | 44.37 | 48.35 |
| Optimize | 39.42 | – | 19.19 | 39.42 | 56.76 | 44.37 | 48.35 |
| Decolorize | 39.42 | – | 19.19 | 42.31 | 56.76 | 49.54 | 48.35 |
| Luminance | 57.92 | 88.20 | 19.20 | 44.99 | 56.74 | 47.62 | 51.27 |

***Table 6:*** *F-Measure Metrics for HW06 - DIBCO Dataset 2013 (CPP of color channels R, G and B: 0.217)*

| HW06 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | – | 90.20 | 49.06 | 86.36 | 89.66 | 90.33 | 91.54 |
| Average | – | 90.20 | 49.06 | 86.36 | 89.66 | 89.57 | 91.54 |
| Maximum | – | 90.20 | 49.06 | 86.36 | 89.66 | 90.33 | 91.54 |
| Min-Max | – | 90.20 | 49.06 | 86.36 | 89.66 | 90.33 | 91.54 |
| Optimize | – | – | 49.06 | 86.36 | 89.66 | 90.33 | 91.54 |
| Decolorize | 25.88 | – | 49.06 | 85.71 | 89.66 | 89.02 | 91.30 |
| Luminance | 91.40 | 91.12 | 49.09 | 87.31 | 89.65 | 90.95 | 91.88 |

***Table 7:*** *F-Measure Metrics for HW07 - DIBCO Dataset 2013 (CPP of color channels R, G and B: 0.277)*

| HW07 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | 0.46 | 21.06 | 13.11 | 52.65 | 42.26 | 53.83 | 69.58 |
| Average | 0.46 | 21.06 | 13.11 | 52.65 | 42.26 | 54.07 | 69.58 |
| Maximum | 0.46 | 21.06 | 13.11 | 52.65 | 42.26 | 53.83 | 69.58 |
| Min-Max | 0.46 | 21.06 | 13.11 | 52.65 | 42.26 | 53.83 | 69.58 |
| Optimize | 0.46 | – | 13.11 | 52.65 | 42.26 | 53.83 | 69.58 |
| Decolorize | 0.46 | – | 13.11 | 52.65 | 42.26 | 54.07 | 69.58 |
| Luminance | 41.12 | 75.83 | 13.04 | 49.98 | 42.29 | 50.57 | 67.71 |

***Table 8:*** *F-Measure Metrics for HW08 - DIBCO Dataset 2013 (CPP of color channels R: 0.212, G: 0.198 and B: 0.160)*

| HW08 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | 0.04 | 90.64 | 39.82 | 62.41 | 72.54 | 85.63 | 82.64 |
| Average | 0.03 | 91.22 | 39.62 | 61.89 | 72.06 | 83.59 | 82.25 |
| Maximum | 0.03 | 89.61 | 40.50 | 66.35 | 76.47 | 87.41 | 80.55 |
| Max-Min | 0.03 | 91.19 | 39.60 | 62.30 | 72.57 | 85.53 | 81.50 |
| Optimize | 0.05 | – | 39.72 | 62.20 | 72.38 | 85.50 | 82.32 |
| Decolorize | 0.07 | – | 39.71 | 61.96 | 72.04 | 83.62 | 82.69 |
| Luminance | 93.19 | 90.81 | 39.83 | 64.68 | 72.54 | 88.48 | 84.85 |

***Table 9:*** *F-Measure Metrics for PR01 - DIBCO Dataset 2013 (CPP of color channels R : 0.233, G : 0.208 and B : 0.178)*

| PR01 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | – | 80.53 | 35.75 | 88.67 | 88.15 | 88.83 | 79.37 |
| Average | – | 80.60 | 35.76 | 88.42 | 88.08 | 88.32 | 78.91 |
| Maximum | – | 88.25 | 36.32 | 89.01 | 88.44 | 88.82 | 81.50 |
| Max-Min | – | 82.33 | 35.87 | 88.34 | 88.02 | 88.27 | 79.07 |
| Optimize | – | – | 36.11 | 88.92 | 88.36 | 88.92 | 80.65 |
| Decolorize | 0 | – | 35.73 | 82.93 | 88.13 | 86.73 | 69.90 |
| Luminance | 87.81 | 80.54 | 35.78 | 89.11 | 88.19 | 87.67 | 80.62 |

***Table 10:*** *F-Measure Metrics for PR02 - DIBCO Dataset 2013 (CPP of color channels R : 0.329, G : 0.310 and B : 0.290)*

| PR02 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | – | 89.04 | 51.45 | 87.50 | 88.17 | 94.08 | 87.14 |
| Average | – | 89.96 | 51.17 | 86.74 | 87.49 | 94.05 | 86.87 |
| Maximum | – | 88.97 | 52.16 | 88.50 | 89.08 | 94.24 | 87.69 |
| Max-Min | – | 89.98 | 51.03 | 86.36 | 87.21 | 93.88 | 86.76 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Optimize | – | – | 51.22 | 86.83 | 87.56 | 93.98 | 86.93 |
| Decolorize | 0 | – | 51.39 | 84.74 | 88.27 | 91.91 | 51.39 |
| Luminance | 94.97 | 89.25 | 51.44 | 88.29 | 88.20 | 93.97 | 87.90 |

***Table 11:*** *F-Measure Metrics for PR03 - DIBCO Dataset 2013 (CPP of color channels R : 0.280, G : 0.306 and B : 0.259)*

| PR03 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | – | 87.12 | 51.79 | 93.99 | 93.08 | 91.36 | 87.96 |
| Average | – | 90.57 | 51.10 | 93.98 | 93.09 | 92.03 | 87.69 |
| Maximum | – | 87.11 | 51.45 | 94.25 | 93.23 | 91.47 | 86.85 |
| Max-Min | – | 89.55 | 51.17 | 94.20 | 93.27 | 91.61 | 87.56 |
| Optimize | – | – | 51.03 | 93.92 | 92.97 | 91.32 | 87.99 |
| Decolorize | 0 | – | 51.91 | 95.28 | 92.98 | 94.00 | 90.21 |
| Luminance | 90.50 | 88.15 | 51.73 | 93.34 | 93.07 | 90.44 | 86.79 |

***Table 12:*** *F-Measure Metrics for PR04 - DIBCO Dataset 2013 (CPP of color channels R : 0.251, G : 0.231 and B : 0.210)*

| PR04 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | – | 49.61 | 66.78 | 91.06 | 91.92 | 93.22 | 85.70 |
| Average | – | 48.97 | 66.78 | 91.06 | 91.92 | 93.22 | 85.70 |
| Maximum | – | 53.46 | 67.80 | 91.32 | 91.64 | 93.70 | 86.61 |
| Max-Min | – | 49.66 | 65.26 | 89.80 | 90.81 | 93.48 | 85.65 |
| Optimize | – | – | 65.85 | 90.24 | 91.34 | 93.26 | 85.62 |
| Decolorize | 0 | – | 66.75 | 89.25 | 91.91 | 93.51 | 89.07 |
| Luminance | 50.58 | 48.54 | 66.87 | 91.01 | 91.99 | 91.59 | 83.54 |

***Table 13:*** *F-Measure Metrics for PR05 - DIBCO Dataset 2013 (CPP of color channels R : 0.473, G : 0.447 and B : 0.425)*

| PR05 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | – | 84.06 | 83.01 | 89.27 | 89.63 | 94.44 | 83.01 |
| Average | – | 83.67 | 82.38 | 89.17 | 89.60 | 89.95 | 89.75 |
| Maximum | – | 86.76 | 83.11 | 89.07 | 89.29 | 94.87 | 90.00 |
| Max-Min | – | 85.35 | 81.78 | 88.93 | 89.47 | 94.18 | 89.61 |
| Optimize | – | – | 82.45 | 89.19 | 89.63 | 94.33 | 89.78 |
| Decolorize | 0 | – | 83.10 | 87.14 | 89.52 | 88.04 | 91.14 |
| Luminance | 87.97 | 86.77 | 83.06 | 89.97 | 89.60 | 94.16 | 89.00 |

***Table 14:*** *F-Measure Metrics for PR06 - DIBCO Dataset 2013 (CPP of color channels R : 0.297, G : 0.275 and B : 0.236)*

| PR06 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | – | 81.04 | 65.11 | 72.85 | 73.51 | 82.23 | 75.76 |
| Average | – | 80.34 | 64.67 | 72.51 | 73.36 | 81.93 | 75.62 |
| Maximum | – | 81.64 | 63.98 | 71.91 | 72.48 | 81.68 | 75.78 |
| Max-Min | – | 76.39 | 63.90 | 71.72 | 72.41 | 81.65 | 75.46 |
| Optimize | – | – | 64.76 | 72.64 | 73.51 | 82.10 | 75.73 |
| Decolorize | 0 | – | 65.23 | 70.56 | 73.44 | 78.05 | 74.58 |
| Luminance | 75.80 | 75.55 | 65.19 | 73.73 | 73.53 | 83.11 | 76.01 |

***Table 15:*** *F-Measure Metrics for PR07 - DIBCO Dataset 2013 (CPP of color channels R: 0.639, G: 0.613 and B: 0.553)*

| PR07 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | – | 90.27 | 88.30 | 93.73 | 93.82 | 91.53 | 89.46 |
| Average | – | 84.84 | 87.91 | 93.48 | 93.60 | 92.14 | 89.24 |
| Maximum | – | 94.04 | 88.67 | 93.80 | 93.84 | 90.97 | 89.36 |
| Max-Min | – | 84.63 | 87.69 | 93.39 | 93.50 | 91.25 | 89.10 |
| Optimize | – | – | 87.98 | 93.54 | 93.65 | 91.37 | 89.28 |
| Decolorize | 0 | – | 88.30 | 93.05 | 93.81 | 93.56 | 89.67 |
| Luminance | 93.52 | 84.45 | 88.26 | 93.82 | 93.80 | 90.55 | 89.12 |

***Table 16:*** *F-Measure Metrics for PR08 - DIBCO Dataset 2013 (CPP of color channels R : 0.474, G : 0.415 and B : 0.293)*

| PR08 | Otsu | Multi-Threshold | Niblack | Sauvola | Wolf | NICK | Bradley |
|---|---|---|---|---|---|---|---|
| GIMP | 1.44 | 76.03 | 53.02 | 64.12 | 68.11 | 74.38 | 69.94 |
| Average | 1.52 | 75.81 | 51.77 | 62.37 | 66.49 | 70.13 | 68.43 |
| Maximum | 0.30 | 78.59 | 55.77 | 66.96 | 70.78 | 79.41 | 74.19 |
| Max-Min | 0.30 | 76.91 | 51.24 | 61.28 | 65.39 | 72.24 | 68.04 |
| Optimize | 2.92 | – | 52.28 | 63.05 | 67.08 | 73.50 | 69.18 |
| Decolorize | 6.84 | – | 52.98 | 64.24 | 68.25 | 71.54 | 69.70 |
| Luminance | 74.29 | 75.58 | 53.01 | 65.17 | 68.14 | 75.77 | 72.19 |

## REFERENCES


[1] D. V. Nikolaos Ntogas, Digital Restoration by Denoising and Binarization of Historical Manuscripts Images, InTech, 2012.

[2] P. S. &. S. B. Dhok, "Review of Text Extraction Algorithms for Scenetext and Document Images," IETE Technical Review, pp. 1-21, 2016.

[3] C. Y. S. a. K. Y. S. Mori, "Historical review of OCR research and development," in IEEE, 1992.

[4] S.-W. L. C. Y. S. Yuan Y. Tang, "Automatic document processing: A survey," Pattern Recognition, vol. 29, no. 12, pp. 1931-1952, 1996.

[5] T. D. a. A. S. D. Ghosh, "Script Recognition—A Review," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 12, pp. 2142-2161, 2010.

[6] B. C. U. Pal, "Indian script character recognition: a survey," Pattern Recognition, vol. 37, no. 9, pp. 1887-1899, 2004.

[7] P. B. P. &. A. G. Ramakrishnan, "OCR in Indian Scripts: A Survey," IETE Technical Review, vol. 22, no. 3, pp. 217-227, 2005.

[8] R. Hedjam, R. F. Moghaddam and M. Cheriet, "Text extraction from degraded document images," in 2nd European Workshop on Visual Information Processing (EUVIP) 2010, 2010.

[9] B. Gatos, K. Ntirogiannis and I. Pratikakis, "ICDAR 2009 Document Image Binarization Contest (DIBCO 2009)," in Document Analysis and Recognition, 2009. ICDAR '09. 10th International Conference on, 2009.

[10] I. Pratikakis, B. Gatos and K. Ntirogiannis, "ICDAR 2011 Document Image Binarization Contest (DIBCO 2011)," in International Conference on Document Analysis and Recognition (ICDAR) 2011, 2011.

[11] A. Antonacopoulos, C. Clausner, C. Papadopoulos and S. Pletschacher, "ICDAR 2013 Competition on Historical Book Recognition (HBR 2013)," in 12th International Conference on Document Analysis and Recognition (ICDAR) 2013 , 2013.

[12] I. Pratikakis, B. Gatos and K. Ntirogiannis, "H-DIBCO 2010 - Handwritten Document Image Binarization Competition," in Proceedings of the 2010 12th International Conference on Frontiers in Handwriting Recognition, Washington, 2010.

[13] I. Pratikakis, B. Gatos and K. Ntirogiannis, "ICFHR 2012 Competition on Handwritten Document Image Binarization (H-DIBCO 2012)," in International Conference on Frontiers in Handwriting Recognition (ICFHR) 2012, 2012.

[14] F. Ghani, E. Khan and M. A. Khan, "Restoration of Old Manuscripts using Image Processing Techniques," IETE Journal of Research, vol. 46, no. 5, pp. 325-329, 2000.

[15] S. Jayaraman, S. Esakkirajan and T. Veerakumar, Digital Image Processing, NewDelhi, Delh: Tata Mc-Graw Hill Education Private Limited, 2011.

[16] R. C. Gonzalez and R. E. Woods, Digital Image Processing (3rd Edition), Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006.

[17] C. Kanan and G. W. Cottrell, "Color-to-Grayscale: Does the Method Matter in Image Recognition?," PLoS ONE, vol. 7, p. e29740, Jan 2012.

[18] S. Susstrunk, R. Buckley and S. Swen, "Standard RGB Color Spaces," in In The Seventh Color Imaging Conference: Color Science, Systems, and Applications, 1999.

[19] K. E. Spaulding, G. J. Woolfe and E. J. Giorgianni, "Image States and Standard Color Encodings (RIMM/ROMM RGB)," Color and Imaging Conference, pp. 288-294, 2000.

[20] A. Koschan and M. A. Abidi, Digital Color Image Processing, New York, NY, USA: Wiley-Interscience, 2008.

[21] W. K. Pratt, Digital Image Processing: PIKS Inside, 3rd ed., New York, NY, USA: John Wiley & Sons, Inc., 2001.

[22] M. Nixon and A. S. Aguado, Feature Extraction & Image Processing, Second Edition, 2nd ed., Academic Press, 2008.

[23] M. Qiu, G. D. Finlayson and G. Qiu, "Contrast Maximizing and Brightness Preserving Color to Grayscale Image Conversion," Conference on Colour in Graphics, Imaging, and Vision, pp. 347-351, 2008.

[24] J. Zhengmeng. and K. Michael Ng, "A contrast maximization method for color-to-grayscale conversion," Multidimensional Systems and Signal Processing, vol. 26, pp. 869-877, 2015.

[25] M. G. a. N. A. Dodgson, "Decolorize: Fast, Contrast Enhancing, Color to Grayscale Conversion," Pattern Recognition, vol. 40, no. 11, pp. 2891-2896, #oct# 2005.

[26] C. Lu, L. Xu and J. Jia, "Contrast preserving decolorization," in IEEE International Conference on Computational Photography (ICCP) 2012 , 2012.

[27] Y. Kim, C. Jang, J. Demouth and S. Lee, "Robust Color-to-gray via Nonlinear Global Mapping," ACM Trans. Graph., vol. 28, pp. 1611-1614, Dec 2009.

[28] B. Gatos, I. Pratikakis and S. J. Perantonis, "Efficient Binarization of Historical and Degraded Document Images," in Document Analysis Systems, 2008. DAS '08. The Eighth IAPR International Workshop on, 2008.

[29] R. F. Moghaddam and M. Cheriet, "RSLDI: Restoration of single-sided low-quality document images," Pattern Recognition , vol. 42, pp. 3355-3364, 2009.

[30] M. Cheriet, N. Kharma, C.-l. Liu and C. Suen, Character Recognition Systems: A Guide for Students and Practitioners, Wiley-Interscience, 2007.

[31] C. C. Fung and R. Chamchong, "A Review of Evaluation of Optimal Binarization Technique for Character Segmentation in Historical Manuscripts," in Knowledge Discovery and Data Mining, 2010. WKDD '10. Third International Conference on, 2010.

[32] O. D. Trier and T. Taxt, "Evaluation of Binarization Methods for Document Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, pp. 312-315, 1995.

[33] O. D. Trier and A. K. Jain, "Goal-directed evaluation of binarization methods," Pattern Analysis and Machine

Intelligence, IEEE Transactions on, vol. 17, pp. 1191-1201, #dec# 1995.
[34] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," Journal of Electronic Imaging, vol. 13, pp. 146-168, 2004.
[35] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," Systems, Man and Cybernetics, IEEE Transactions on, vol. 9, pp. 62-66, Jan 1979.
[36] K. Fukunaga, Introduction to Statistical Pattern Recognition (2nd Ed.), San Diego, CA, USA: Academic Press Professional, Inc., 1990.
[37] N. Papamarkos and B. Gatos, "A New Approach for Multilevel Threshold Selection," CVGIP: Graph. Models Image Process., vol. 56, pp. 357-370, Sep 1994.
[38] D.-M. Tsai and Y.-H. Chen, "A fast histogram-clustering approach for multi-level thresholding," Pattern Recognition Letters , vol. 13, pp. 245-252, 1992.
[39] B. Gatos, I. Pratikakis and S. J. Perantonis, "Adaptive degraded document image binarization," Pattern Recognition , vol. 39, pp. 317-327, 2006.
[40] W. Niblack, An Introduction to Digital Image Processing, Birkeroed, Denmark: Strandberg Publishing Company, 1985.
[41] J. Sauvola and M. Pietikäinen, "Adaptive document image binarization," Pattern Recogniton, vol. 33, pp. 225-236, 2000.
[42] K. Khurshid, I. Siddiqi, C. Faure and N. Vincent, "Comparison of Niblack inspired binarization methods for ancient documents," in Proc. SPIE 7247, Document Recognition and Retrieval XVI, 72470U, 2009.
[43] C. Wolf and J.-M. Jolion, "Extraction and recognition of artificial text in multimedia documents," Formal Pattern Analysis and Applications, vol. 6, pp. 309-326, 2004.
[44] P. D. Wellner, "Adaptive thresholding for digital desk," Xerox Research Centre, Europe, 1993.
[45] D. Bradley and G. Roth, "Adaptive Thresholding using the Integral Image," Journal of Graphics, GPU, and Game Tools, vol. 12, no. 2, pp. 13-21, 2007.
[46] V. Wu, R. Manmatha and E. M. Riseman, "Textfinder: an automatic system to detect and recognize text in images," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 21, pp. 1224-1229, Nov 1999.
[47] K. Ntirogiannis, B. Gatos and I. Pratikakis, "An Objective Evaluation Methodology for Document Image Binarization Techniques," in Document Analysis Systems, 2008. DAS '08. The Eighth IAPR International Workshop on, 2008.
[48] Paramasivam and Sabeenian, "Contrast Based Color Plane Selection for Binarization of Historical Document Images," in 2nd International Conference on on Emerging Trends in Electrical, Communication and Information Technologies (ICECIT 2015), 2015.