M. Tech (SPDD) RISHABH JAIN 2021

ANALYSIS OF DECOLORIZATION METHODS AND ITS APPLICATION

A DISSERTATION REPORT

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE
OF

MASTER OF TECHNOLOGY

IN

SIGNAL PROCESSING & DIGITAL DESIGN

Submitted by:

RISHABH JAIN

2K19/SPD/14

Under the supervision of

Dr. N. JAYANTHI

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

JULY,2021

ANALYSIS OF DECOLORIZATION METHODS AND ITS APPLICATION

A MAJOR-2 PROJECT REPORT

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE
OF

MASTER OF TECHNOLOGY
IN
SIGNAL PROCESSING & DIGITAL DESIGN

Submitted by:

RISHABH JAIN

2K19/SPD/14

Under the supervision of

Dr. N. JAYANTHI

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

JULY,2021

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

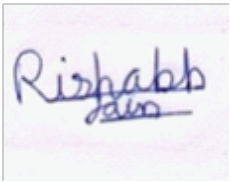
DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I **RISHABH JAIN** student of M.Tech (SPDD), hereby declare that the Dissertation titled "**ANALYSIS OF DECOLORIZATION METHODS AND ITS APPLICATION**" which is submitted by me to the Department of Electronics and Communication Engineering, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, has been done by me. This work has not have previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition



Place: Delhi **RISHABH JAIN**

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

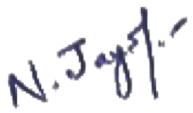
DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Dissertation titled "**ANALYSIS OF DECOLORIZATION METHODS AND ITS APPLICATION**" which is submitted by **RISHABH JAIN**, **2K19/SPD/14** of Electronics and Communication Department, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is a record of the work carried out by the student under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.



Place: Delhi

Date: 22 JULY,2021

Dr N. JAYANTHI

SUPERVISOR

ASSISTANT PROFESSOR

Department of Electronics and Communication

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

ABSTRACT

Contrast Preserving Decolorization is rgb-to-gray conversion of a color image, which tries to preserve the contrast in a Gray-scale image, as it was in the original color image. In this dissertation, various contrast preserving decolorization techniques such as SPDecolor and CPrgb2gray have been analysed and they have been applied in the field of binarization of historical documents. Pre-processing is an essential step in the binarization of historical documents in order to preserve its information content, especially in the case of low contrast images. This work addresses the problem of contrast preservation of the historical documents using contrast preserving decolorization based binarization process. For certain types of degradation such as faded/faint text , the proposed SPDecolor based pre-processing technique may perform better for local binarization techniques over the global binarization techniques. This dissertation also compares various local binarization algorithms using both conventional and contrast preserving pre-processed techniques. Improved experimental results are obtained on the DIBCO(2009-11) dataset for proposed SPDecolor based framework against CPrgb2gray and conventional color-to-gray binarization methods.

ACKNOWLEDGEMENT

I want to thank Dr. N. Jayanthi, my Project guide, who helped me throughout the project and supported, motivated and encouraged throughout the project.

I want to thank Prof. N.S. Raghava, Head of Department (ECE), DTU, who gave me the permission to use all required equipment and necessary format to complete the project.

I want to thank all the people who have helped me throughout the project with their time , resources and various other things to complete the project.

I want to thank my parents who encouraged me and provided inspiration throughout the project.

CONTENTS

Candidate's Declaration i

Certificate ii

Abstract iii

Acknowledgement iv

Contents v

List of Tables vii

List of Figures viii

List of Symbols, Abbreviations x

CHAPTER 1 INTRODUCTION 1

CHAPTER 2 LITERATURE REVIEW 3

2.1 Decolorization 3

2.2 Binarization 4

CHAPTER 3 DECOLORIZATION METHODS 5

3.1 Contrast Preserving Decolorization(CPrgb2gray) 5

3.1.1 Parametric Color-to-Gray Model. 5

3.1.2 Bimodal Contrast-Preserving Objective Function 5

3.1.3 Weak color order 6

3.2 Semi-Parametric Decolorization(SPDecolor) 9

3.2.1 Semi-parametric Decolorization model. 9

3.2.2 Analyzing the Three Subspaces 9

3.2.3 Modified Semi-optimized Method Solver 12

3.3 Decolorization Results & Discussion 13

3.3.1 Contrast measuring metrics 13

3.3.2 Comparison on Cadik's Dataset 14

3.3.3 Comparison on CSDD Dataset 16

CHAPTER 4 Application- Binarization of Historical Documents 20

4.1 PROPOSED METHODOLOGY 20

4.2 Contrast preserving decolorization(Pre-processing) 20

4.3 Binarization 21

4.4 Binarization Results & Discussion 22

5.2.1 Dataset Tested 22

5.2.2 Visual and statistical binarization Results 22

CHAPTER 5 CONCLUSION 37

REFERENCES 38

LIST OF TABLES

TABLE 3.1 CCPR of Cadik Dataset

TABLE 3.2 CCPR of CSDD Dataset

TABLE 4.1 Comparison of Statistical metrics for various DIBCO Images

TABLE 4.2 Comparison of Statistical metrics for Local binarization techniques on DIBCO(2009-11) dataset.

LIST OF FIGURES

Fig. 3.1 The SSIM measure value between subspace Z1 and Z3 , and between subspace Z2 and Z3 of Cadik dataset imaged

Fig. 3.2 The average gradient entropy of basis images in Z1 ,Z2 and Z3 of Cadik dataset images

Fig. 3.3 The CCPR of basis images in Z1 ,Z2 and Z3 of Cadik dataset images

Fig. 3.4 Images from the Cardik Dataset with various decolorization method, (a) Input Image (b) Rgb2gray output (c) CPrgb2gray output (d) SPDecolor output

Fig. 3.5 Images from the CSDD Dataset with various decolorization method, (a) Input Image (b) Rgb2gray output (c) CPrgb2gray output (d) SPDecolor output

Fig. 4.1 Flow of the proposed method

Fig. 4.2 a) Input image (8th image from DIBCO-2011 Printed) (lightly printed) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick, h) Otsu , i) Spd-Otsu, j) Cprgb2gray - Otsu

Fig. 4.3 a) Input image (2nd image from DIBCO-2011 handwritten) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick, h) Otsu , i) Spd-Otsu, j) Cprgb2gray - Otsu

Fig. 4.4 a) Input (7th image from DIBCO-2018)) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick, h) Otsu , i) Spd-Otsu, j) Cprgb2gray - Otsu

Fig. 4.5 a) Input (5th image from DIBCO-2018) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick, h) Otsu , i) Spd-Otsu, j) Cprgb2gray - Otsu

Fig. 4.6 a) Input image (1st image from DIBCO-2009 Printed) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick, h) Otsu , i) Spd-Otsu, j) Cprgb2gray - Otsu

Fig. 4.7 i) Input image (5th image from DIBCO-2009 Printed) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Fig. 4.8 i) Input (1st jmage from DIBCO-2010HW)) (lightly printed text) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Fig. 4.9 i) Input image (3rd image from DIBCO-2010 handwritten) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Fig. 4.10 a) Input (3rd image from DIBCO-2011HW)) (thin strokes, coloured background with some dark spots) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Fig. 4.11 a) Input image (2nd image from DIBCO-2011 handwritten) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Fig. 4.12 a) Input image (3rd image from DIBCO-2011 Printed b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Fig. 4.13 a) Input image (8th image from DIBCO-2011 Printed) (lightly printed) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Fig. 4.14 a) Input image (6th image from DIBCO-2011 handwritten) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

LIST OF ABBREVIATIONS

SPDecolor	Semi-Parametric Decolorization
CP	Contrast Preserving
PSNR	Peak Signal to Noise Ratio
NRM	Negative Rate Metric
CSDD	Complex Scene Decolorization Dataset
CCPR	Contrast Preserving Ratio

CHAPTER 1

INTRODUCTION

Translation of color-to-Gray is an important need for numerous everyday applications in computer vision, image processing etc. A benefit of this translation is it allows the single channel algorithms application on the colour images, e.g for edge detection, Canny operator is used. Additional application comprise monochrome print, image recognition etc. All these things have encouraged the advancement of several color-to-gray conversion methods earlier. As the color-to-gray translation is to map a 3-Dimension vector to a 1-Dimension scalar, so it's a dimensionality-reduction procedure, and so it suffers some information loss. So as an outcome, numerous advanced algorithms have been evolved to smartly use the limited range in gray scales to display the input color image details and contrasts.

The technique Contrast Preserving rgb2gray(CPrgb2gray)[28] and Semi-Parametric Decolorisation (SPDecolor)[10] uses the 2^{nd} order multi-variance polynomial model. CPrgb2gray is Parametric based decolorization, while SPDecolor is Semi-parametic decolorization. The objective of the Semi-Parametric based decolorization is that it improves the modelling and numerical lacks of the gradient error penalty underneath the 2^{nd}-order multi-variance polynomial construction in CPrgb2gray[28]. Firstly, the gradient error data-fidelity measure extremely emphases to penalize the pixels that have larger value and overlooks the pixels with minor value [40], resulting an overfitting from the pictorial view point. Second, the gradient error penalty w.r.t the weight parameter is very non-linear, and the solution is estimated and local, even if a result is found. These two limitations are reflected on the phenomenon that it occasionally produces poor outcomes, especially for colour images containing plentiful colors and patterns.

In these decolorization methods , gradient error guided optimization with a 2^{nd}-order multi-variance polynomial model K_r, K_g, K_b, K_rK_g , K_rK_b, K_gK_b, K_rK_r , K_gK_g, K_bK_b has been adopted. The key difference between the work in CPrgb2gray[28] and SPDecolor[10] is that in former all the weights have been optimized while in latter the author have pre-fixed the weight of the first three line channels K_r, K_g, K_b and only optimize the six weights of the K_rK_b, K_rK_g, K_gK_b, K_rK_r, K_gK_g, K_bK_b . The thought in SPDecolor is that, it treats the 9 terms in the 2^{nd} order model with dissimilar importance, that aims to find improved results in enhancing the non-linear cost function. It is noteworthy that this 2-step or iterative enhancement approach is widely used in newer applications of image such as video/image reconstruction and denoising.

The application where I have used this SPDecolor method is in Binarization of historical documents. The ancient documents are a source of great knowledge on topics such as philosophy, literature, science, medicine, religion etc. These ancient knowledge may be of great use in future research and development work. So to preserve these texts digitization of these documents is an essential . Also digitization will help to remotely accesses these document in any part of world and to easily search and index the document. This digitization of document is mainly done with Optical character recognition (OCR) system. In OCR systems Image Binarization is the first step. Binarization separates text(foreground) from background. As these documents are very ancient, so due to aging and many other factors such as quality of the paper or ink , humidity etc , there is lot of challenges in digitizing these documents . The challenges such as stains , smear , wrinkles due to humidity, ink bleed through , lightly printed text are very common in these ancient documents . So this type of degradation makes digitization challenging.

In recent past numerous approach, has been put forward to binarize the document images based on global and local thresholding methods. [1]- [3]and [15] [16]. Several researchers also worked on the restoration of ancient degraded documents for better readability of the text which in turn increases the OCR performance [4] – [9]. Color-to-gray transformation is a necessary pre-processing aid for binarization. Because of the dimensionality reduction from three dimensional colour space value to one dimensional gray values, definitely it suffers from the information loss. However, since many decolorizing methods have been employed I found that Semi-parametric decolorization (SPDecolor)[10] preserves the appearance of the original colour image in comparison with various previous technique as a pre-processing step for the process of digitization.

The most important theories of optimization of color-to-grayscale conversion involved global contrast among colours [13] or local contrasts among pixels [12]. Both methods are slow processes prone to local minima because of a large number of variables and thus prove impractical. A novel approach for the binarization of document images using Semi-Parametric based color to gray transformation as a pre-processing tool to preserve and enhance the color contrast, which helps in better binarization is put forward in this work. The proposed method is tested mainly on DIBCO(2009-11) datasets [19-21]. Experiments on the dataset show the improvement in the binarized output achieved by our proposed framework based on the various binarization metric such as F-measure, PSNR and NRM.

CHAPTER 2

Literature review

2.1 Decolorizaton

Earlier, lots of advances has been done in algorithms and theory for decolorization that is perception-driven, follow-on in a great quantity of methods.
These methods in the works can be largely classified into three categories:

a) Approaches of local mapping manipulating the color pixel values of local distribution.

b) global mapping approaches, together with the methods of objective minimization exploiting the differences between mapped values of gray pixels and original values of color pixel.

c) hybrid approaches.

In 1st category, the mapping of pixel values from color-to-gray is typically varying spatially, which is local distribution dependent. For example, Eschbach *et al.* [29] showed a technique conserving chromatic edges, which was apprehended by addition of components of high-frequency luminance channel chromaticity. Nemcsics *et al.* [30] rebuilt the image of grayscale from the color image gradients, by measurement of luminance & color differences as the contrast of gradient coloured space. Myszkowski *et al.* [31] did the image disintegration into numerous components of frequency and then grouping weights adjustment through chromatic channels. Though they were able to conserve local features, color constant areas can be transformed unevenly if there is change in mapping in the regions..

In the second class of global mapping methods, mapping of grayscale output value of pixel is done, independent of its location, by color pixel value of the input. For example, Tumblin *et al.* presented color contrasting between pair of pixels [32]. Westall *et al.* did enforcing of constraints straight on different pairs of color & a function that is quadratic is constructed by which gray images were solved[33]. Lee *et al.* [34] presented a parametric nonlinear model for mapping that is global type taking inspiration from [32]. Cho *et al.* [35] also took inspiration from [33] and balanced the gradient among pixels, or between some landmarks that are predetermined & pixels. Chen *et al.* [36] did the introduction of 3 visual cues and then they were induced into a global energy function whose optimisation by variational method was done. Jia *et al.* [28] gave relaxation to the severe order constrain and then to preserve maximally the original color contrast ,a 2nd order multivariate polynomial parametric model was adopted.

The third category of methods are hybrid and they frequently comprise a multi-stage procedure or many constraints. Like , Ng *et al.* presented a brightness preservation and local variance maximization algorithm through which decolorization is achieved [39]. Total variation regularization has been employed in the decolorization process to lessen, at the nearby pixel locations, the difference among the local transformations. The newly introduced optimization method that is saliency-guided region-based [37] uses a parametric color-to-gray mapping function that is two-stage, that considered the joint local and global information.

2.2 Binarization

Document binarization techniques, which are being proposed regularly over the past years, can be broadly classified as global thresholding techniques and local thresholding techniques [24]. A single threshold value is used by the Global thresholding method to binarize an image. Some example of global thresholding method are Otsu[25] , Kitter[26] etc. these algorithms works fine when the text is well contrasted against the background and text covers most part of image. But when there is not a uniform illumination in image then these method does not perform well. Local thresholding techniques such as ; Sauvola[15] , Niblack[27] uses a threshold value that depends on local neighbourhood pixels. These local methods does not perform well when there is blurriness in the image and also where there is large bleed-through regions in text. If there is a large variation in illumination then the Niblack method doesn't adapt to it . Nick algorithm[16] shows improved results over the Niblack algorithm results but there is an addition of unwanted noise.

CHAPTER 3

DECOLORIZATION METHODS

Here two contrast Preserving Decolorization Methods have been analysed:-

1. CPrgb2gray
2. SPDecolor

3.1 Contrast Preserving Decolorization(CPrgb2gray)

Among the global color-to-gray mapping approaches, a centre of attraction is 2nd order multi-variance polynomial model about the 3 channels of color image. Historically, the classical rgb-to-gray model treats that the output grayscale is a constrained linear combination of color image channels R, G and B, that is, $y = \sum \; w_c K_c$, where Kg ,Kr , Kb represents channels Green, Red, and Blue. In standard matlab function rgb2gray(),(Wr =0.2989, Wg =0.5870, Wb= 0.1140) is fixed weights for images, with the consideration that the green channel information is more human vision sensitive. Lu *et al.* [28] developed a polynomial two-order multi-variance model about the 3 channels of the color image for contrast preserving color-to-gray conversion.

3.1.1 Parametric Color-to-Gray Model

The function for decolorization $y = f(\mathbf{c})$ produces a grayscale value y for each color c = (*r, g, b*) value. Here the mapping scheme adopted is global, which using the same mapping function throughout for conversion.

The polynomial space of degree n for color **c** = (*r, g, b*) **is defined mathematically as**

$$\prod_n = \text{span}\{K_r{}^{d_1} K_g{}^{d_2} K_b{}^{d_3} : d_i = 0, 1, 2, 3, 4 \ldots\ldots d_1 + d_2 + d_3 \leq \text{n}\} \qquad (3.1)$$

where Πn is a polynomial space belonging to monomials family.
So ,the mapping function is

$$y = \sum_{b_c \in Z} w_c b_c \qquad (3.2)$$

where b_c is the c^{th} monomial basis of $\prod_n$, c={r,g,b}

The mapping function is uniquely determined by weights w_c
As here n=2 so ,

$$\prod_2 = Z = K_r,\ K_g,\ K_b,\ K_r K_b,\ K_r K_g,\ K_g K_b,\ K_g K_g,\ K_r K_r,\ K_b K_b$$

3.1.2 Bimodal Contrast-Preserving Objective Function

In this section, the color contrast preserving objective function based on a weak color order constraint has been described. Here the objective function, which uses color contrast ,is weak color order based . $g_x - g_y$ represent the difference between gray pixels x and y, so the conventional energy function, that is L2-norm based [34, 32] can be written as:

$$\min_g \sum_{(x,y)=P} \left(g_x - g_y - \delta_{x,y}\right)^2 \qquad (3.3)$$

Where 'g' is the estimated image that can be with [34] or without [32] a parametric form, 'P' represent neighbouring pixel pair . Here x and y pixel pair is an ordered pair. Color contrast $\delta_{x,y}$ have a signed value and represents the color pair difference.
Color contrast for CIELab is given as

$$|\delta_{x,y}| = \sqrt{(L_x - L_y)^2 + (a_x - a_y)^2 + (b_x - b_y)^2}$$

The sign of *Lx – Ly* typically determine the sign of $\delta_{x,y}$. So, there could be a contrast loss problem if we enforce this type of color order. Also, for human visual system ,where there is ambiguity in color order , it may be noncompliant . So this model modifies this condition and now uses a selection mechanism that is bimodal.

Equation (3.3) can be interpreted in view of probabilistic inference. It implies that the grayscale difference of two pixels *x* and *y* follows a Gaussian distribution with mean *δx,y*. We can interpret Equation (3.3) according to probabilistic inference which says that a Gaussian distribution is being followed by difference ($g_x - g_y$) which has a mean *δx,y*.

$$\prod_{x,y} N_\sigma(\delta_{x,y}, \sigma^2) \propto \prod_{x,y} \exp\left\{-\frac{|\Delta g_{x,y} - \delta_{x,y}|^2}{2\sigma^2}\right\} \qquad (3.4)$$

Here the peak of the gausssian distribution is at *δx,y*, and is single mode. By this, it is inferred that the contrast is constrained, and the sign of gray-level is determined. It may be noted here that sign does not represent physical meaning. And thus the difference can be $-\delta x,y$ or $+\delta x,$y and that develops a proper contrast preserving and flexible constraint. So this is the motivation for this work that automatically select color orders using this bimodal distribution

3.1.3 Weak color order

There are many color pairs that are ordered (if we consider brightness as parameter), like pure white color is always brighter than any other color and there is no issue about its order to anyone. So, for such type of color pairs where there is no ambiguity, we can use Equation 3.4 that have a single-peak distribution.

$$\mathbf{c}_x \leq \mathbf{c}_y \Leftrightarrow r_x \leq r_y \& g_x \leq g_y \& b_x \leq b_y \qquad (3.5)$$

The sign of *δx,y is* directly applied to *gx – gy ,* if Equation (3.5) is satisfied. Lest the sign is not specified prior, and a selection process has been put to search the appropriate color order optimally. The likelihood term is thus defined for one pixel pair as:

$$\frac{1}{2}\{N_\sigma(\delta_{x,y}, \sigma^2) + N_\sigma(-\delta_{x,y}, \sigma^2)\} \qquad (3.6)$$

As two types of color order definition has been stated above, a map has been built to distinguish them and is stated:

$$\alpha_{x,y} = \begin{cases} 1.0 & \text{if } r_x \leq r_y, g_x \leq g_y, b_x \leq b_y \\ 0.5 & \text{otherwise} \end{cases} \qquad (3.7)$$

In color order , that is unambiguous,the prior definition of N_σ ($\delta_{x,y}$, σ^2) is enforced if $\alpha_{x,y} = 1$. Otherwise bimodal distribution is followed by the color difference, that optimally selects the negative or positive sign .
Then the objective function is:

$$\prod_{(x,y)\in p} \{\alpha_{x,y} N_\sigma(\delta_{x,y}, \sigma^2) + (1-\alpha_{x,y}) N_\sigma(-\delta_{x,y}, \sigma^2)\} \quad (3.8)$$

$(x, y) \in p$, where p is the four-neighbour set.
Equation (3.8) can be maximized when its negative logarithm term is minimized and the minimized version is stated as

$$\min_g En(g) = [-\sum_{(x,y)\in P} \ln\{\alpha_{x,y} N_\sigma(\delta_{x,y}, \sigma^2) + (1-\alpha_{x,y}) N_\sigma(-\delta_{x,y}, \sigma^2)\}] \quad (3.9)$$

Here $N_\sigma()$ stands for Gaussian distribution and $\alpha_{x,y}$ for Weak color order . Substituting in Equation (3.9) the parametric gray model described in Section 3.1.1, a function is formed consisting of unknown coefficients *{ωi}*. As the global non-linear mapping is used, only nine parameters need to be estimated. The difference of two gray pixels can then be expressed w.r.t the parameters *{ωi}*:

Parametric model stated previously is substituted in Equation 3.9 , and as a result a function having unknown coefficients {ωi} is formed . Here only nine parameters is estimated as the mapping used here is global non-linear. Since

$$\Delta g_{x,y} = g_x - g_y = \sum_{l_\epsilon \in Z} w_c l_{c(x,y)}, \; l_{c(x,y)} = b_{cx} - b_{cy} \quad (3.10)$$

which reduces in finding parametric model solution as follows:

$$\min_{w_c} En(w) = [-\sum_{(x,y)\in P} \ln\{\alpha_{x,y} N_\sigma(\sum_{l_\epsilon = Z} w_c l_{c(x,y)} - \delta_{x,y}) + (1-\alpha_{x,y}) N_\sigma(\sum_{l_\epsilon = Z} w_c l_{c(x,y)} + \delta_{x,y})\}] \quad (3.11)$$

So, the Energy function is minimized as:-

$$\min_{w_c \in (z\})} En(w) = -\sum_{(x,y)\in P} \ln \left\{ \alpha_{x,y} \exp\left(-\frac{\left|\sum_{l_c \in \{z\}} w_c l_{c(x,y)} - (\delta_{x,y})\right|^2}{2\sigma^2} \right) + (1-\alpha_{x,y}) \exp\left(-\frac{\left|\sum_{l_c \in \{z\}} w_c l_{c(x,y)} + (\delta_{x,y})\right|^2}{2\sigma^2} \right) \right\} \quad (3.12)$$

Numerical Solution

Inside energy function in equation 3.12 , partial derivative is taken w.r.t {Wc} and putting them as zero obtains a linear equation system. In simplifying the presentation, a term $J_{x,y}$ has been defined as follow:

$$J_{x,y} = \frac{\alpha_{xy} N_\sigma(\delta_{x,y}, \sigma^2)}{\alpha_{x,y} N_\sigma(\delta_{x,y}, \sigma^2) + (1-\alpha_{x,y}) N_\sigma(-\delta_{x,y}, \sigma^2)} \quad (3.13)$$

With algebraic operations that are limited, deriving partially on Wj , i.e. $\frac{\partial En(w)}{\partial w_j}$ is stated as:

$$\sum_{(x,y)\in P} \sum_{c} w_c \, l_{c(x,y)} l_{j(x,y)} - (1 - 2J_{x,y}) l_{j(x,y)} \delta_{x,y} = 0 \qquad (3.14)$$

By putting $\frac{\partial E \ (w)}{\partial w_j} = 0$ so in total 9 equations are obtained . the problem in solving it due to $J_{x,y}$ terms, which comprise functions that are nonlinear about w. Fix-point iteration approach has been applied on *w* that linearize the agreeing equations. i.e., in order to solve for w_c^{k+1} in the (k+1)th iteration, the nonlinear term $J_{x,y}^{k}$ is generated by the formerly estimated w_c^{k} , which produces the following equation:

$$\sum_{(x,y)\in P} \sum_{c} w_c^{k+1} l_{c(x,y)} l_{j(x,y)} = (2J_{x,y}^{k} - 1) l_{j(x,y)} \delta_{x,y} \qquad (3.15)$$

Through the above approximation, In Eq. (3.15) by varying parameter j ,9 equations are obtained. Therefore, in every iteration, the structure shows linearity w.r.t w_c^{k+1} and so it can be solved successfully.

3.2 Semi-Parametric Decolorization(SPDecolor)

3.2.1 Semi-parametric Decolorization model

- Here the two-order multi-variance polynomial model from the subspace modelling is being analysed.
- The key difference between the work in [28] and the method here is that weights have pre-fixed of the first three line channels $K_r,\ K_g,\ K_b$ and only optimized the six weights of the $K_rK_g,\ K_rK_b,\ K_gK_b,\ K_rK_r,\ K_gK_g,\ K_bK_b$
- The impression behind is that it treats the nine terms in the two-order model[28] with dissimilar importance, which prefers to find improved solutions in optimizing the non-linear cost function

3.2.2 Analysing the Three Subspaces

A main downside in work done by Lu *et al.* [28] is that it treats entire basis likewise. And, it uses iterative optimization method to resolve the derived model, taking more execution time. Here, the grayscale image as the addition of 3 subspaces is conveyed:

$$y = \sum_{b_{c_1}=z_1} w_{c_1} b_{c_1} + \sum_{b_{c_2}=z_2} w_{c_2} b_{c_2} + \sum_{b_{c_3}=z_3} w_{c_3} b_{c_3} \qquad (3.16)$$

$$Z_1 = \{K_r, K_g, K_b\};\ Z_2 = \{K_rK_g,\ K_rK_b,\ K_gK_b\};\ Z_3 = \{K_rK_r,\ K_gK_g,\ K_bK_b\}$$

where , Z1, Z2, and Z3 belongs to the subspaces traversed by the monomials. For simplicity, the elements in the 3 subspaces are named as basis images. As they have been previously explained and defined by linearly combining the channel images, the quest of the mapped function inside Equation 3.16 turns to determining the weight parameters (Wc). Now the variation of the three subspaces is analysed.

In fact, the differences among the three subspaces are vast. Primarily, as the elements of image inside that of subspace Z1 is the 1st order, and also the image basis inside Z2 and Z3 subspaces are that of 2nd order and truly it is the product (dot) of that inside Z1 subspace , so elements in Z1 subspace is showing less correlation, and the subspace Z1 holds more contents. Moreover, the similarities among the three subspaces is quantitatively investigated via the Structural Similarity (SSIM) measure , one of the most popular indexes that measures structural similarity in between 2 images [42].
The average Structural Similarity (SSIM) measure values amongst subspace Z1 and Z3 , and amongst subspace Z2 and Z3 are 0.5813, and 0.7470, correspondingly. By studying the above two phenomena, it is recommended that the 1st-order subspace Z1 and the later second-order subspaces Z2 and Z3 should not be treated in the equivalent manner.

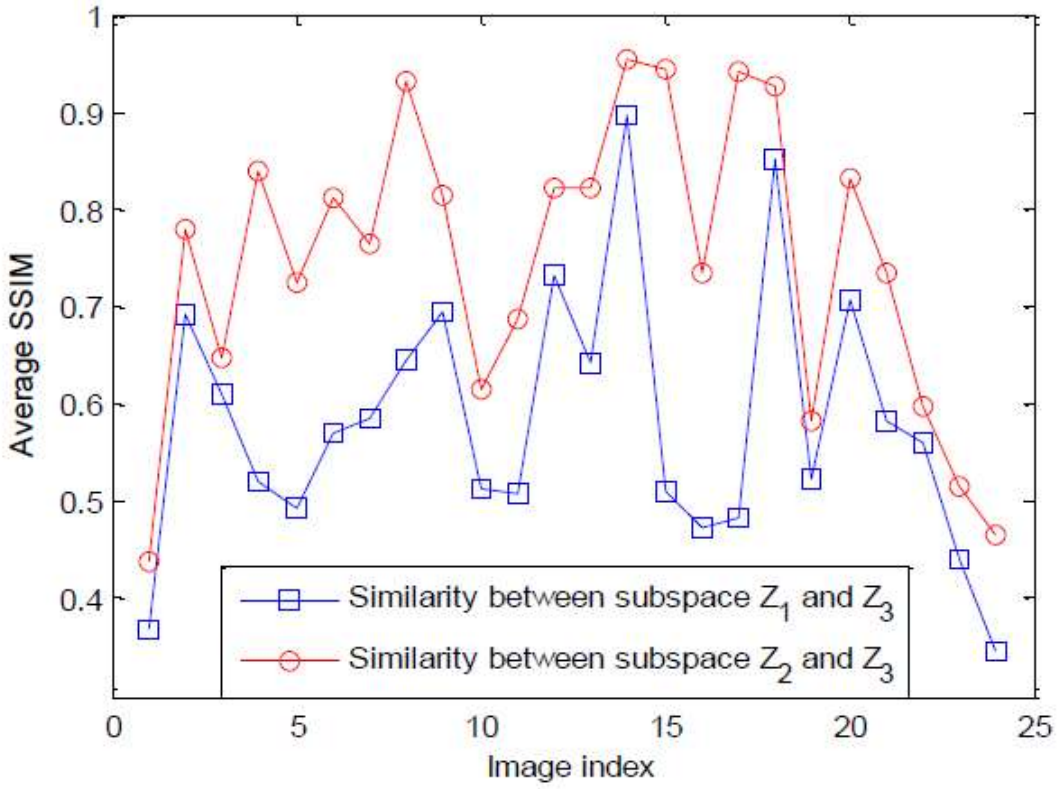


Fig. 3.1 The SSIM measure value between subspace Z1 and Z3 , and between subspace Z2 and Z3 of Cadik dataset imaged[1].

Secondly, as the gradient preservation is the utmost significant issue in decolorization. So the gradient data of the three subspaces is studied. Now color images from dataset of Cadik is taken as illustration. Quantity wise, the basis images average gradient entropy of the subspaces Z1 , Z2 and Z3 are 1.0005, 0.9627, and 0.9147, respectively. It can be detected that the subspace Z1 traces at the chief importance.

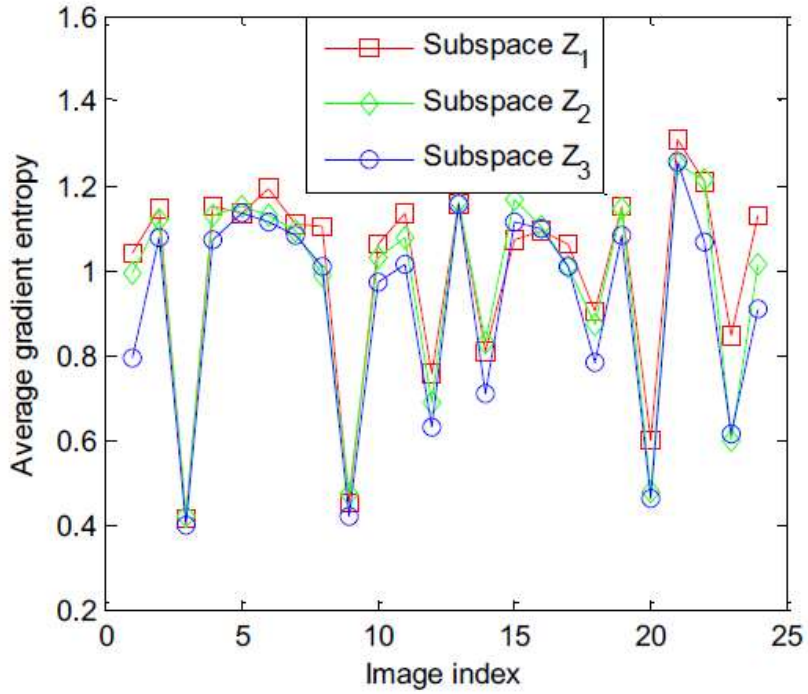


Fig 3.2 The average gradient entropy of basis images in Z1 ,Z2 and Z3 of Cadik dataset images [15][1]

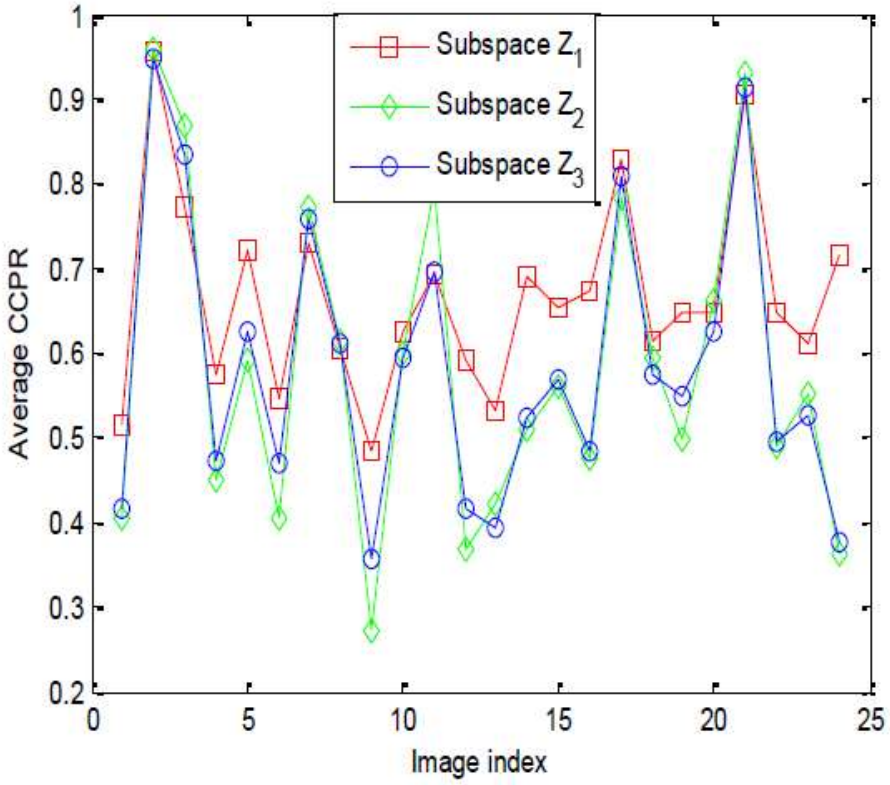


Fig 3.3 The CCPR of basis images in Z1 ,Z2 and Z3 of Cadik dataset images [15][1]

3.2.3 Modified Semi-optimized Method Solver

Under this analysis, Z1 the 1st subspace inside the 2nd order multi-variance parametric model is of utmost importance and additional two subspaces to be secondary are likewise related, the semi-optimized method to decolorization becomes

$$y = y_1 + y_2 \qquad (3.17)$$

$$y_1 = \sum_{b_{c_1}=Z_1} w_{c_1} b_{c_1} \qquad (3.18)$$

$$y_2 = \sum_{b_{c_2}=\{z_2,z_3\}} w_{c_2} b_{c_2} \qquad (3.19)$$

Initially, predefine the weight factors for the bc1 belonging to Z1 .
In this the weights set are (Wr =0.2989, Wg =0.5870, Wb= 0.1140) which is the weight of standard rgb2gray() function.

Secondly, after the immediate grayed image y1 is achieved, techniques for optimization as mentioned in [28] is used, that finds from the immediate grayed image y1 and the (2nd and 3rd) subspaces, the best optimal resultant image. In this work, the bimodal function Equation 3.11 as a reference is used, the semi-optimized model is thus offered as follows:

$$\min_{w_c\in(z_2,z_3\}} En(w) = -\sum_{(x,y)\in P} \ln\left\{\alpha_{x,y}\exp\left(-\frac{\left|\sum_{l_c\in\{z_2,z_3\}} w_c l_{c(x,y)} - (\delta_{x,y} - y_1)\right|^2}{2\sigma^2}\right)\right.$$
$$\left. +(1-\alpha_{x,y})\exp\left(-\frac{\left|\sum_{l_c\in\{z_2,z_3\}} w_c l_{c(x,y)} + (\delta_{x,y} + y_1)\right|^2}{2\sigma^2}\right)\right\} \qquad (3.20)$$

Where y1 is got from Equation 3.18. Naturally, this method can be recognized as a modification method.

Inside energy function in Equation 3.20, partial derivative is taken w.r.t {Wc} and putting them equal to zero obtain a linear equation system. In simplifying the presentation, a term has been defined as follow:

$$J_{x,y} = \frac{\alpha_{xy}(\delta_{x,y} - y_1)N_\sigma(\delta_{x,y},\sigma^2) - (1-\alpha_{x,y})(\delta_{x,y} + y_1)N_\sigma(-\delta_{x,y},\sigma^2)}{\alpha_{x,y}N_\sigma(\delta_{x,y},\sigma^2) + (1-\alpha_{x,y})N_\sigma(-\delta_{x,y},\sigma^2)} \qquad (3.21)$$

With algebraic operations that are limited, deriving partially on Wj , i.e. $\frac{\partial En(w)}{\partial w_j}$ is stated as:

$$\sum_{(x,y)\in P}\sum_{c} w_c\, l_{c(x,y)} l_{j(x,y)} - J_{x,y} l_{j(x,y)} = 0 \qquad (3.22)$$

By putting $\frac{\partial En(w)}{\partial w_j} = 0$ so in total 6 equations are obtained . As in [28], the problem in solving it due to $J_{x,y}$ terms, which comprise functions that are nonlinear about w. Fix-point iteration approach has been applied on *w* that linearize the agreeing equations. i.e., in order to solve for w_c^{k+1} in the (k+1)th iteration, the nonlinear term $J_{x,y}^k$ is generated by the formerly estimated w_c^k , which produces the following equations:

$$\sum_{(x,y)\in P} \sum_c w_c^{k+1} l_{c(x,y)} l_{j(x,y)} = \sum_{(x,y)\in P} J_{x,y}^k l_{j(x,y)} \quad (3.23)$$

Through the above approximation, In Equation (3.23) by varying parameter j ,6 equations is obtained. Therefore, in every iteration, the structure show a linearity w.r.t w_c^{k+1} and so it can be solved successfully.

3.3 Decolorization Results

In this section, comparison of results of SPDecolor, CPrgb2gray and rgb2gray is done on Cadik and CSDD dataset. The parameters value of sigma is 0.01 as used in matlab program.

3.3.1 Contrast measuring metrics

Contrast Preserving Ratio(CCPR)

CCPR is a metric which quantitatively evaluates the contrast preserving capability of the decolorisation algorithms. It works on the principle that if the color difference δ is less than that of threshold τ, than that difference could not be perceived by human naked eye. Contrast-preserving decolorisation maintains human perceivable color change. So CCPR is presented as:-

$$CCPR = \#\{(x,y)|(x,y) \in \Omega | g_x - g_y | \geq \tau \} / \| \Omega \| \quad (3.24)$$

where Ω is the set containing all neighboring pixel pairs with their original color difference $\delta x,y \geq \tau$.

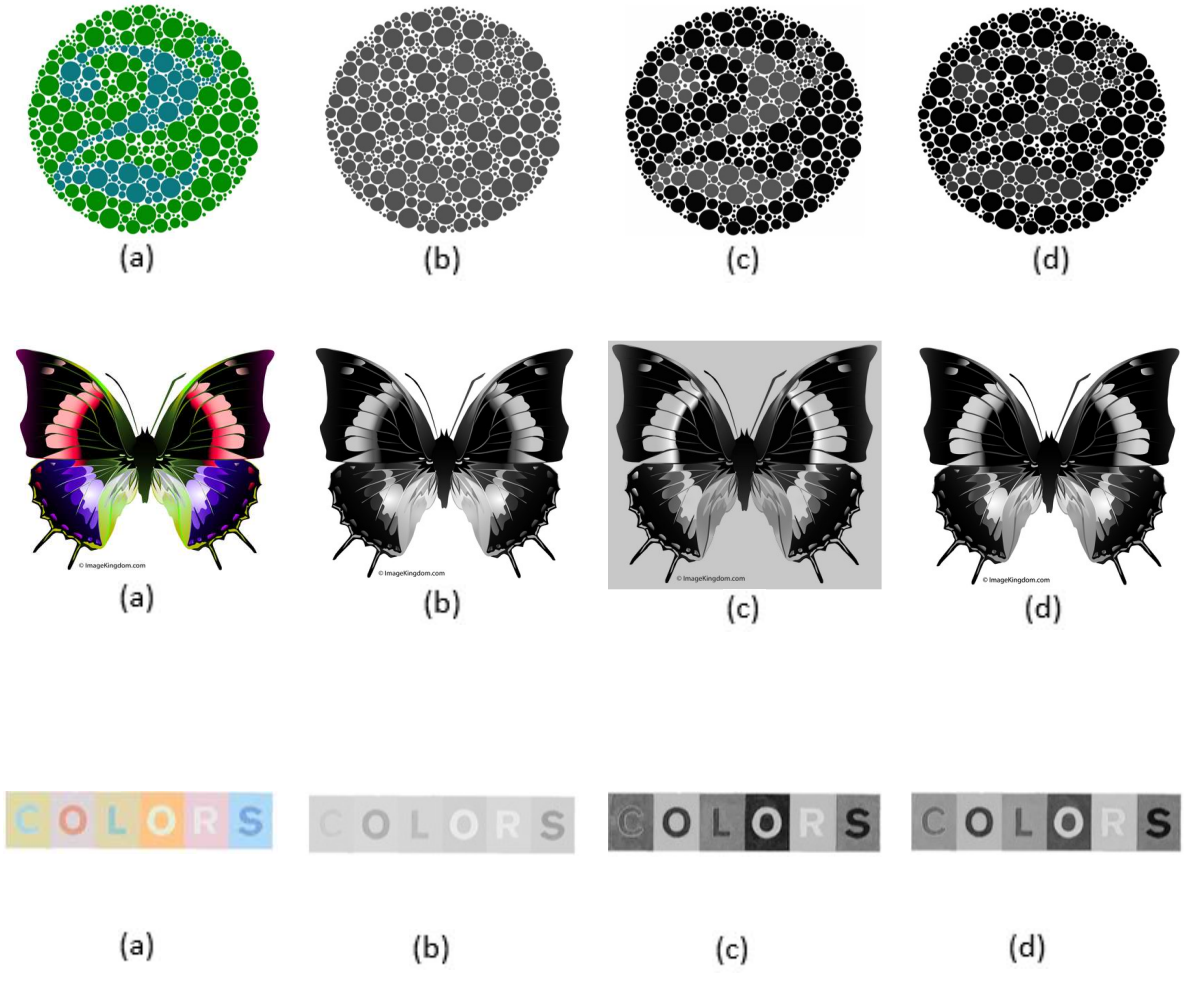
$\| \Omega \|$ is the number of pixel pairs in Ω. $\#\{(x, y)|\ (x, y) \in \Omega, |gx - gy| \geq \tau\}$ is the number of pixel pairs in Ω that are still distinctive after decolorization.

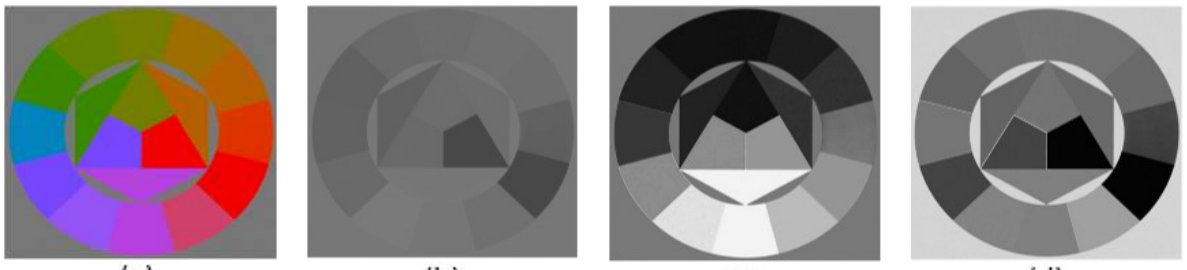
Using CCPR different methods are evaluated quantitatively on all the images of Cardik and CSDD dataset. The value of threshold τ *is varied* from 1 to 15 to calculate the average CCPR.

3.3.2 Comparison on Cadik's Dataset

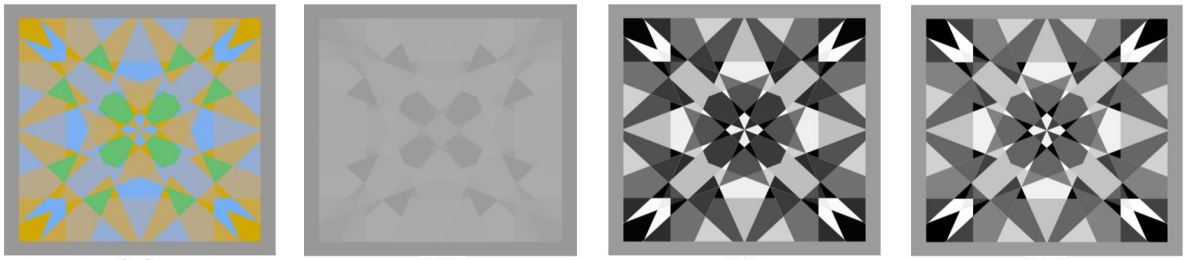
The result on Cadik dataset is shown in fig. 3.4. Its 2st column contains rgb2gray results, 3rd column contains cprgb2gray results and 4th column contains SPDecolor results. From results it can be seen that, for some images rgb2gray does not produces good results and just generates flat results. Both cprgb2gray and SPDecolor finds solution using second order multI-variance polynomial function. It can be seen From results that SPDecolor and CPRgb2gray performance is better. In SPDecolor it can be seen that it's feature preservation of color image is impressive and also its order preservation capability is good. While CPrgb2gray sometimes results in overfitting as can be seen in 6th and 8th image that as seen visually the result has more contrast .

The performance metric for quantitatively evaluating the decolorization algorithm in contrast preserving terms is CCPR. CCPR for whole cadik dataset is calculated and also the value of τ is varied from 1 to 15. It can be seen that SPDecolor outperforms other methods and has 14 threshold values that gives best result while Cprgb2gray has 13 such best thresholds. When comparing these results with rgb2rray it can be seen that these methods has better structural similarity between grayscale and color imaged. Table 3.1 shows the CCPR result on Cadik dataset.

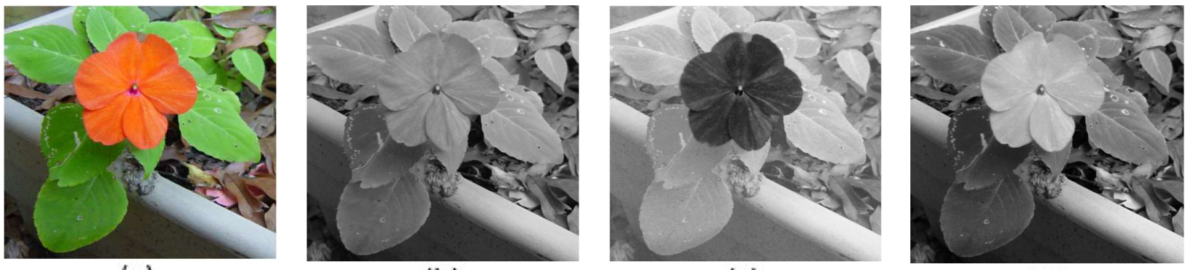




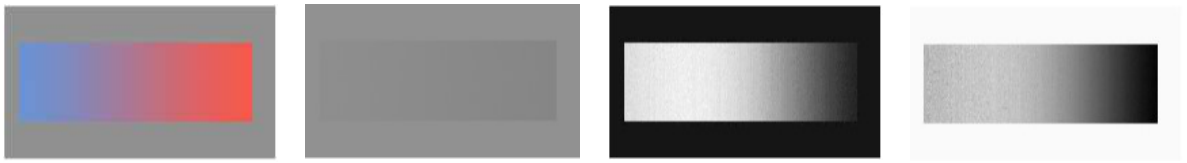
(a) (b) (c) (d)



(a) (b) (c) (d)



(a) (b) (c) (d)



(a) (b) (c) (d)

(a) (b) (c) (d)

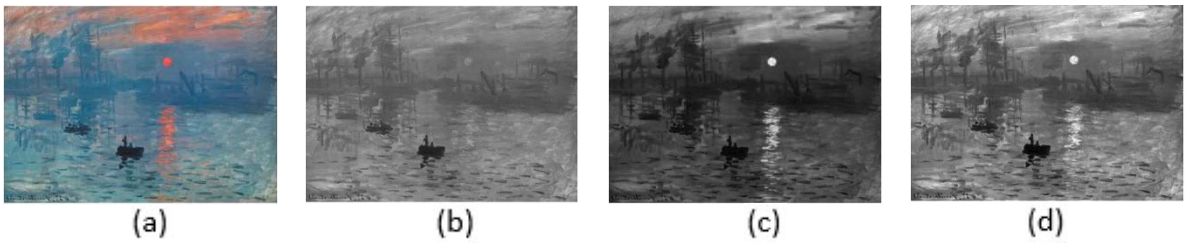


Fig. 3.4 Images from the Cadik Dataset with various decolorization method, (a) Input Image (b) Rgb2gray output (c) CPrgb2gray output (d) SPDecolor output

Table 3.1 CCPR of Cadik Dataset

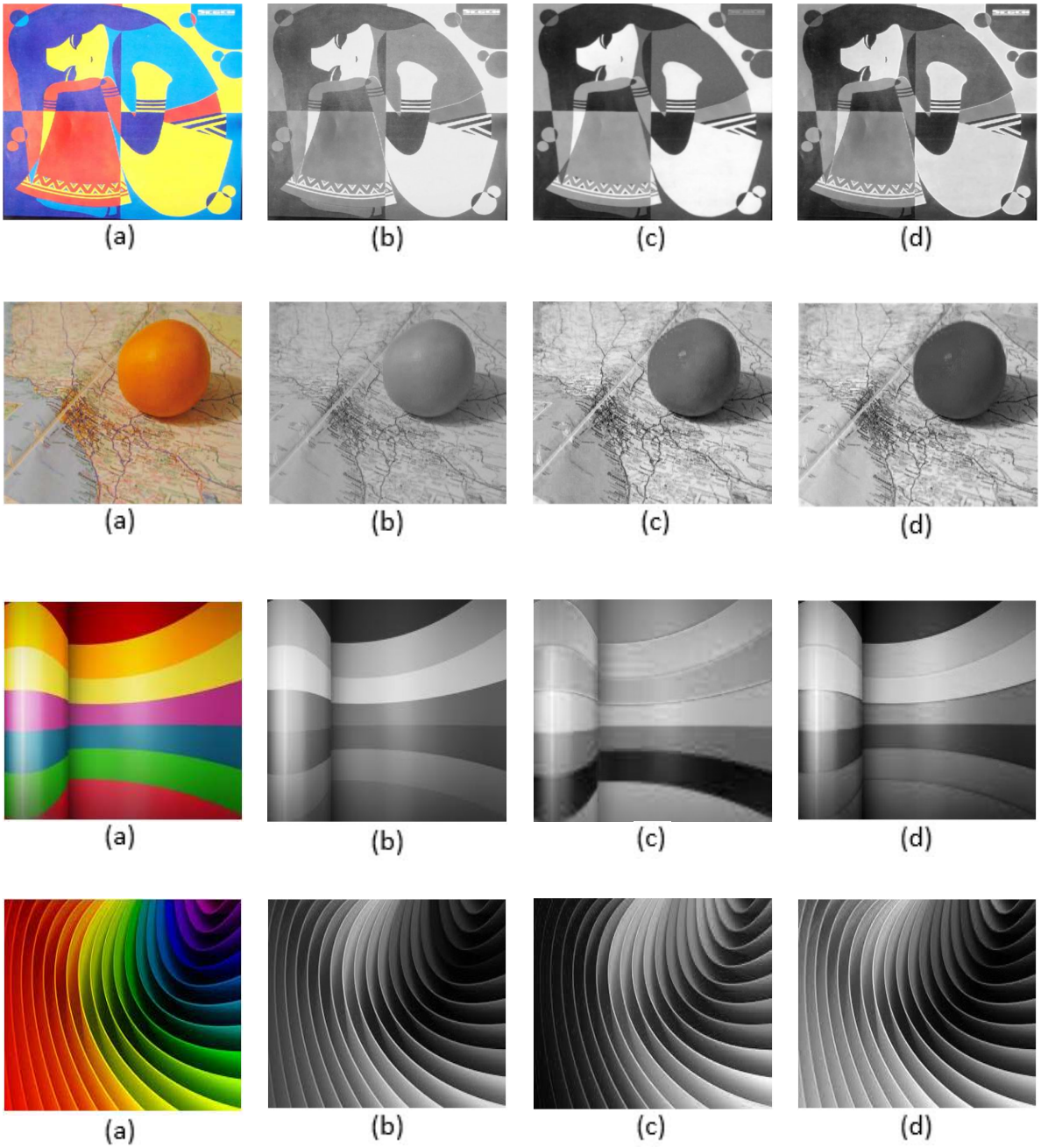
	Rgb2gray	CPrgb2gray	SPDecolor
1	0.68	**0.76**	**0.76**
2	0.64	0.73	**0.74**
3	0.62	0.72	**0.73**
4	0.60	**0.72**	**0.72**
5	0.58	**0.72**	0.71
6	0.56	**0.70**	**0.70**
7	0.55	**0.69**	**0.69**
8	0.54	**0.68**	**0.68**
9	0.53	**0.67**	**0.67**
10	0.53	**0.66**	**0.66**
11	0.52	**0.65**	**0.65**
12	0.51	**0.64**	**0.64**
13	0.50	**0.63**	**0.63**
14	0.50	**0.62**	**0.62**
15	0.49	**0.61**	**0.61**

3.3.3 Comparison on CSDD Dataset

THE CSDD dataset contains 22 images that are different and which has abundant colors & patterns . The result on CSDD dataset is shown in fig. 2st column contains rgb2gray results. 3rd column contains cprgb2gray results and 4th column contains SPDecolor results. In complex scenes SPDecolor is better than all other methods as can be seen from results. CPrgb2gray produces grayscale artefacts in some portion on images, that caused high local contrast than the original images. It can also be seen that rgb2gray does not provide necessary local and global contrast. SPDecolor results are good both in terms of accuracy as well as color orders.

Table 3.2 shows the CSDD dataset quantitative comparison. It is seen here that CPrgb2gray does not perform like it did in Cadik dataset. This shows that 2nd order multivariance polynomial model , without any extra constraints, does not produce good results. While SPDeclor uses 2 steps optimization strategy and hence the results are better. The reason for this superior result can be inferred from the fact that it reduces the solution space , by fixing the weights of the 1st subspace and the optimizing the weights of 2nd and 3rd subspace. Rgb2gray results lack contrast while it provides stable results.

Average CCPR for whole CSDD dataset is calculated and also the value of τ is varied from 1 to 15. It can be seen that SPDecolor outperforms other methods and has 12 threshold values that gives best result. When comparing these results with rgb2rray it can be seen that these methods has better structural similarity between grayscale and color imaged.



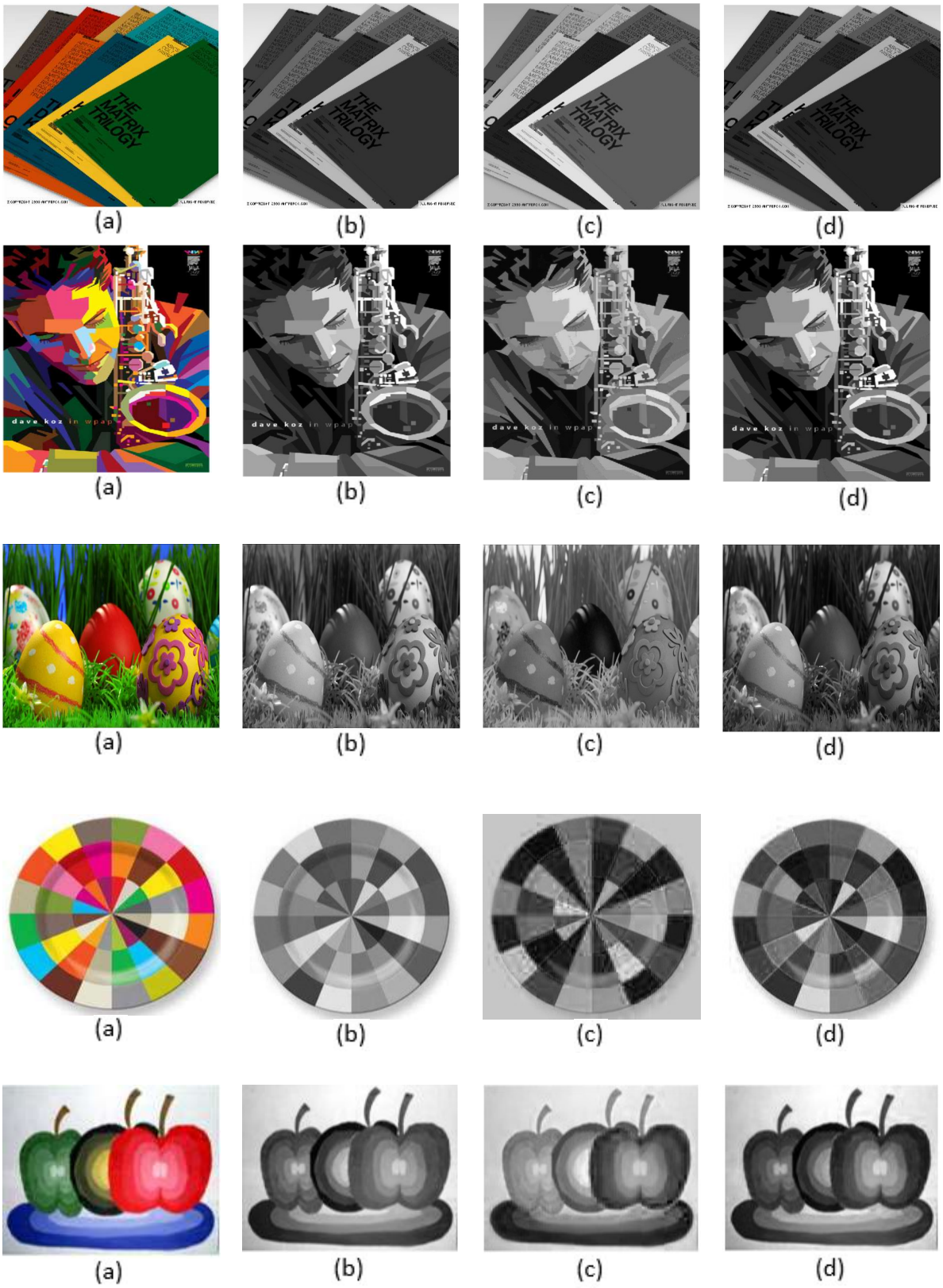


Fig. 3.5 Images from the CSDD Dataset with various decolorization method, (a) Input Image (b) Rgb2gray output (c) CPrgb2gray output (d) SPDecolor output

Table 3.2 CCPR of CSDD Dataset

	Rgb2gray	**CPrgb2gray**	**SPDecolor**
1	0.71	0.72	**0.74**
2	0.64	0.65	**0.68**
3	0.60	0.62	**0.64**
4	0.58	0.59	**0.62**
5	0.55	0.57	**0.60**
6	0.54	0.55	**0.58**
7	0.52	0.53	**0.56**
8	0.50	0.51	**0.54**
9	0.49	0.49	**0.53**
10	0.47	0.48	**0.51**
11	0.46	0.46	**0.50**
12	0.45	0.45	**0.48**
13	0.44	0.44	**0.47**
14	0.43	0.42	**0.46**
15	0.41	0.41	**0.45**

CHAPTER 4

Application- Binarization of Historical Documents

The contrast preserving decolorization techniques discussed previously has been used as an application in the field of binarization of documents. Here, these techniques has been applied as a pre-processing step before the standard global and local binarization thresholding algorithm .

4.1 METHODOLOGY

This section presents our proposed method.

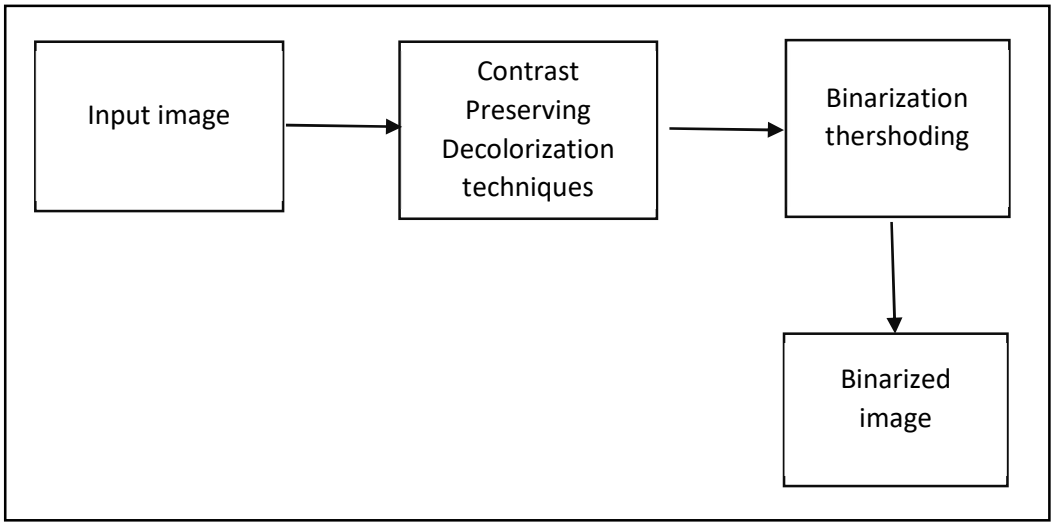


Fig. 4.1 Flow of the proposed method

Figure 4.1 shows the steps in the proposed method. Firstly, we apply our contrast preserving pre-processing step to input image, which converts the input image to grayscale, preserving its contrast, and then we apply the conventional binarization technique which results in a better binarized image

4.2 Contrast preserving decolorization(Pre-processing)

Contrast preserving decolorization is a pre-processing step used in our proposed method.The purpose of using a contrast preserving color-to-gray conversion, instead of conventional colour to gray conversion, is that it helps in restoring degradations which were not possible using conventional method.

In this proposed methodology , 2 types of contrast preserving docolorization techniques has been tested.

1. CPrgb2gray
2. Semi-parametric decolorisation(SPDecolor)

Detailed explaination of both the above techniques have been done in previous chapter.

4.3 Binarization

It is a process of separating text (foreground) from the non-text (background) regions. We experimented on Otsu[25], Sauvola[15] and Nick[16] thresholding algorithms to produce the binarized images from the Obtained SPDecolor based gray scale image. The above stated algorithms for calculating the threshold value to decide the pixels belonging to foreground and background is given below in equations (4.2), (4.3) and (4.4). Global threshold T in Otsu's binarization is given below in the equation (4.1).

T = threshold in MIN (Within-class variance) (4.1)

The within-class variance is calculated using equation (4.2).

$$\sigma_w^{\mathrm{M}} = w_b\sigma_b + w_f\sigma_f^2 \qquad (4.2)$$

Where σ_b and σ_f represents the variance of background and foreground pixels. And *wb* and *wf* are the background and foreground pixels weights respectively. Value in the equation (3.29) is the 'sum of weighted variances' that calculated *t* i.e. the threshold value. This scheme is iterated for 1 to L gray levels, so that all threshold values can be found. Global threshold value T is selected from among all value that has the least sum of weighted variances. A Pixel value which lower than the threshold becomes background, while those above threshold becomes foreground.

$$\mathrm{T}_{\mathrm{Sauvola}} = \mu * \left[1 + k\left(\frac{\sigma}{c} - 1\right)\right] \qquad (4.3)$$

The constant (k) and window size (w) are the manually selected . For Sauvola[15] binarization [23].

window size (w)= 15 and k=0.5 is endorsed .

$$\mathrm{T}_{\mathrm{Nick}} = \mu + k\sqrt{\left[\frac{1}{\mathrm{N}}(\sum p_i^2 - \mu)\right]} = \mu + k\sigma \qquad (4.4)$$

Window size (w)= 19 and k= -0.2 is recommended for Nick]16] binarization [23].

Here σ and μ are the standard deviation and mean respectively, K and C represents thresholding
constants for Sauvola [15] and Nick[16] method respectively, and are set according to the image contrast.

4.4 Binarization Results

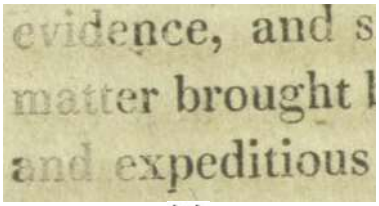
4.4.1 Dataset Tested

The Standard Local binarization approaches and its pre-processed version are tested mainly on the standard DIBCO(2009-2011) Datasets[19-21]. The images in DIBCO datasets contain various kinds of degradation . The reason for testing on these DIBCO Datasets is that it satisfies our aim for testing the proposed method on various types of degradation in historical document. [19] comprises five printed and five handwritten images and their ground truth images.[20] comprises ten handwritten images and their ground truth images. [21] comprises eight printed and eight handwritten images and its ground truth images. These datasets contain various types of degraded images that are challenging to binarize using normal binarization process.

4.4.2 Visual and Statistical binarization results

We have shown the visual results of contemporary Local binarization algorithms such as Nick[16], Sauvola[15], and contemporary Global binarization algorithms such as Otsu[25] based on conventional color-to-gray conversion and SPDecolor based binarization algorithms. We mainly concentrated on the pre-processing step that causes information loss

Figures 4.2, 4.3, 4.4, 4.5 shows comparison of various binarization methods on the document images, having various types of degradations. Figure 4.2a shows the input image that is degraded due to faded/faint text. Figure 4.3a shows the input image that has faint text due to thin strokes type degradation. Figure 4.4a shows the image that has variable colour background , faint text due to thin strokes type degradation. Figure 4.5a shows the input image that contains noise due to page borders and it also contains non-uniform & faint text type degradation. In all these types of degradation , it is seen that SPDecolor pre-processed local binarization techniques shows a better output compared to conventional local binarization technique.
It is also seen that SPDecolor Pre-processed binarization technique does not show any significant improvement in output for Global binarization Techniques .
Also, for the above degraded images, SPDecolor pre-processed local binarization technique performs better that the Global binarization technique



(a)

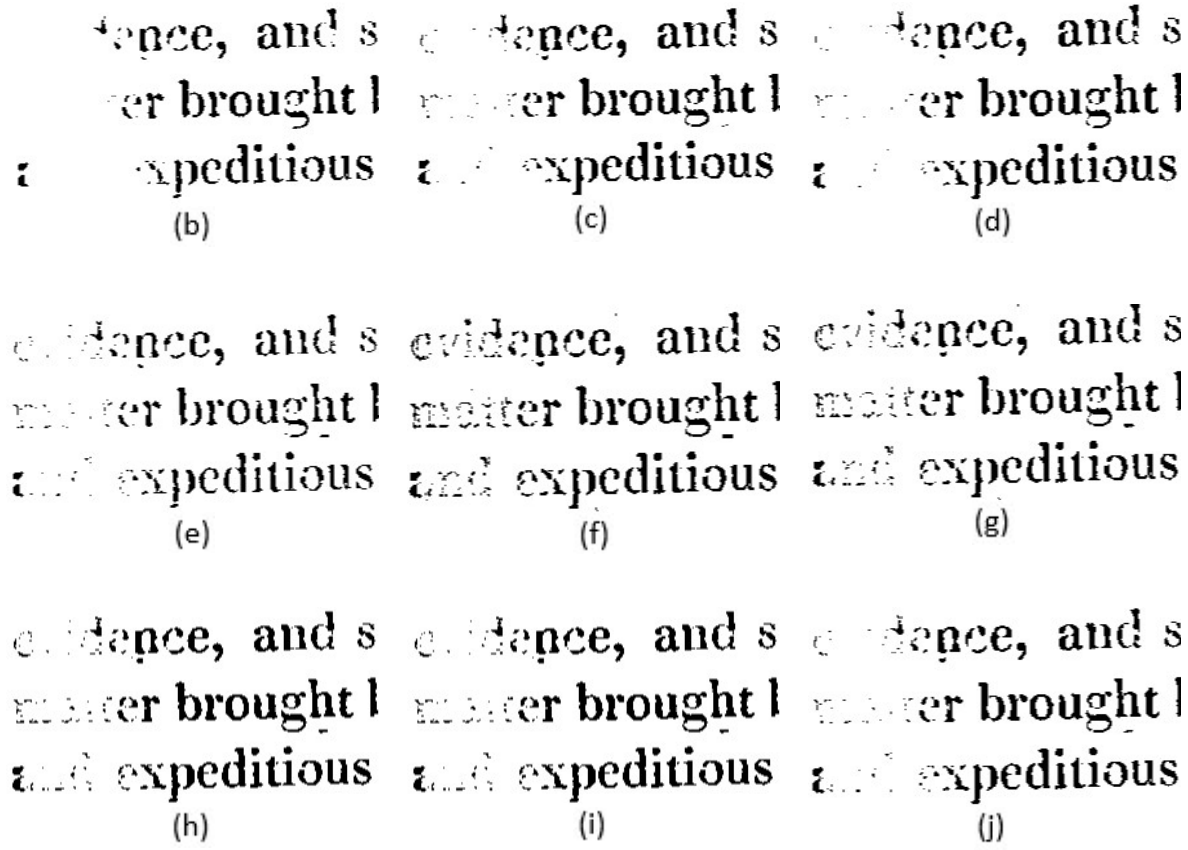
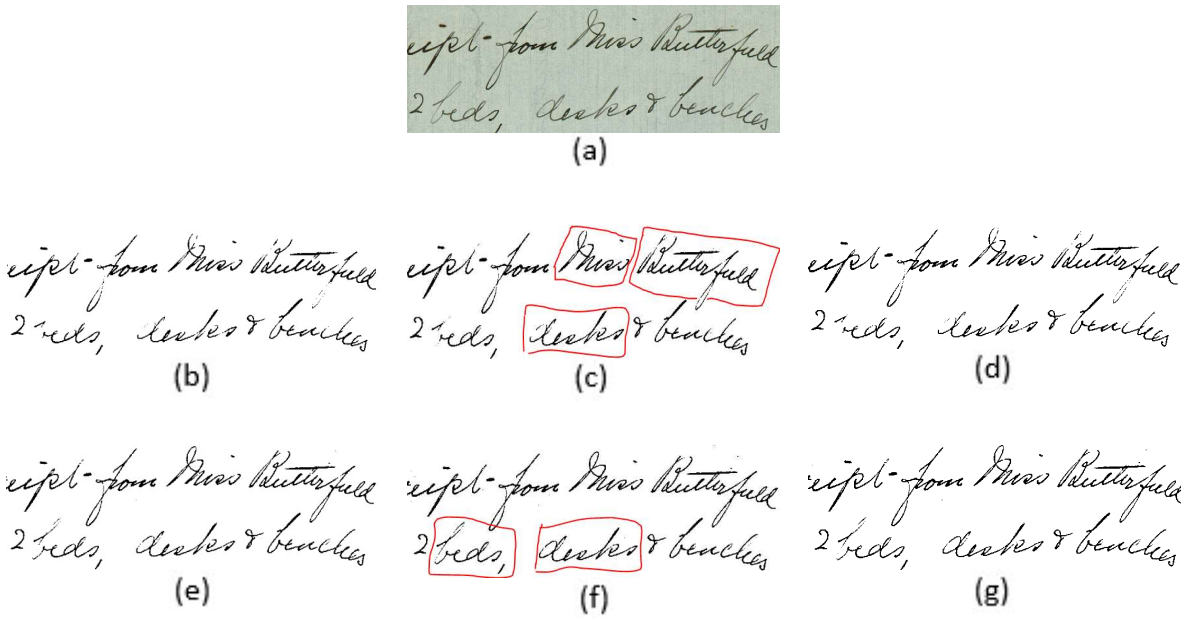


Fig. 4.2 a) Input image (8th image from DIBCO-2011 Printed) (lightly printed) b) Sauvola , c) Spd - Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick, h) Otsu , i) Spd-Otsu, j) Cprgb2gray -Otsu



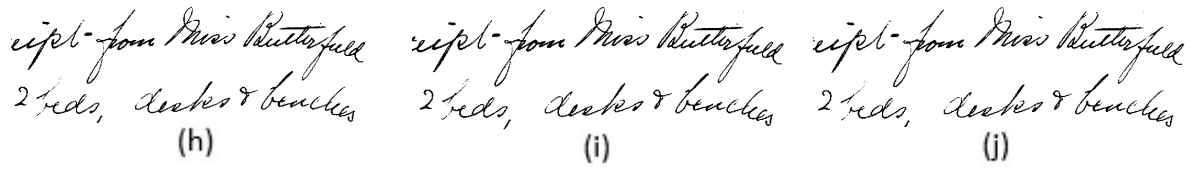
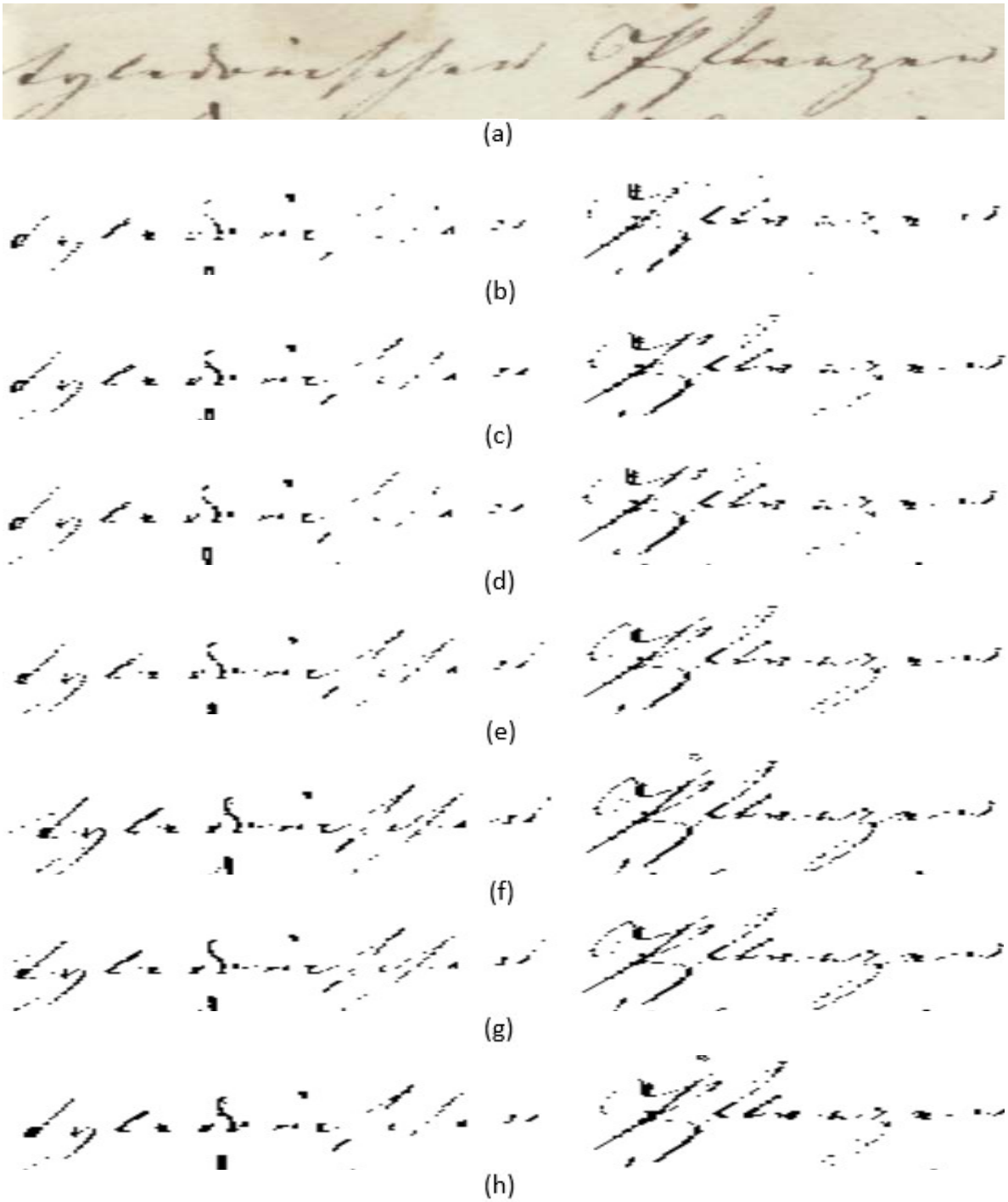


Fig. 4.3 a) Input image (2nd image from DIBCO-2011 handwritten) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick, h) Otsu , i) Spd-Otsu, j) Cprgb2gray -Otsu



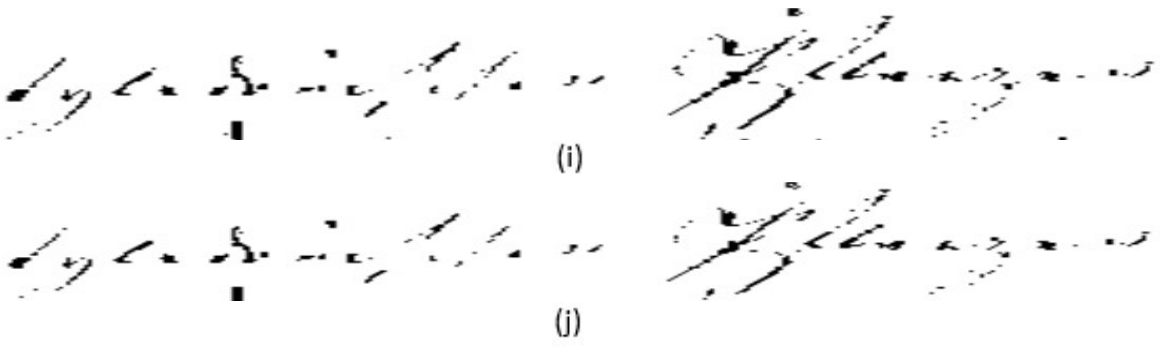
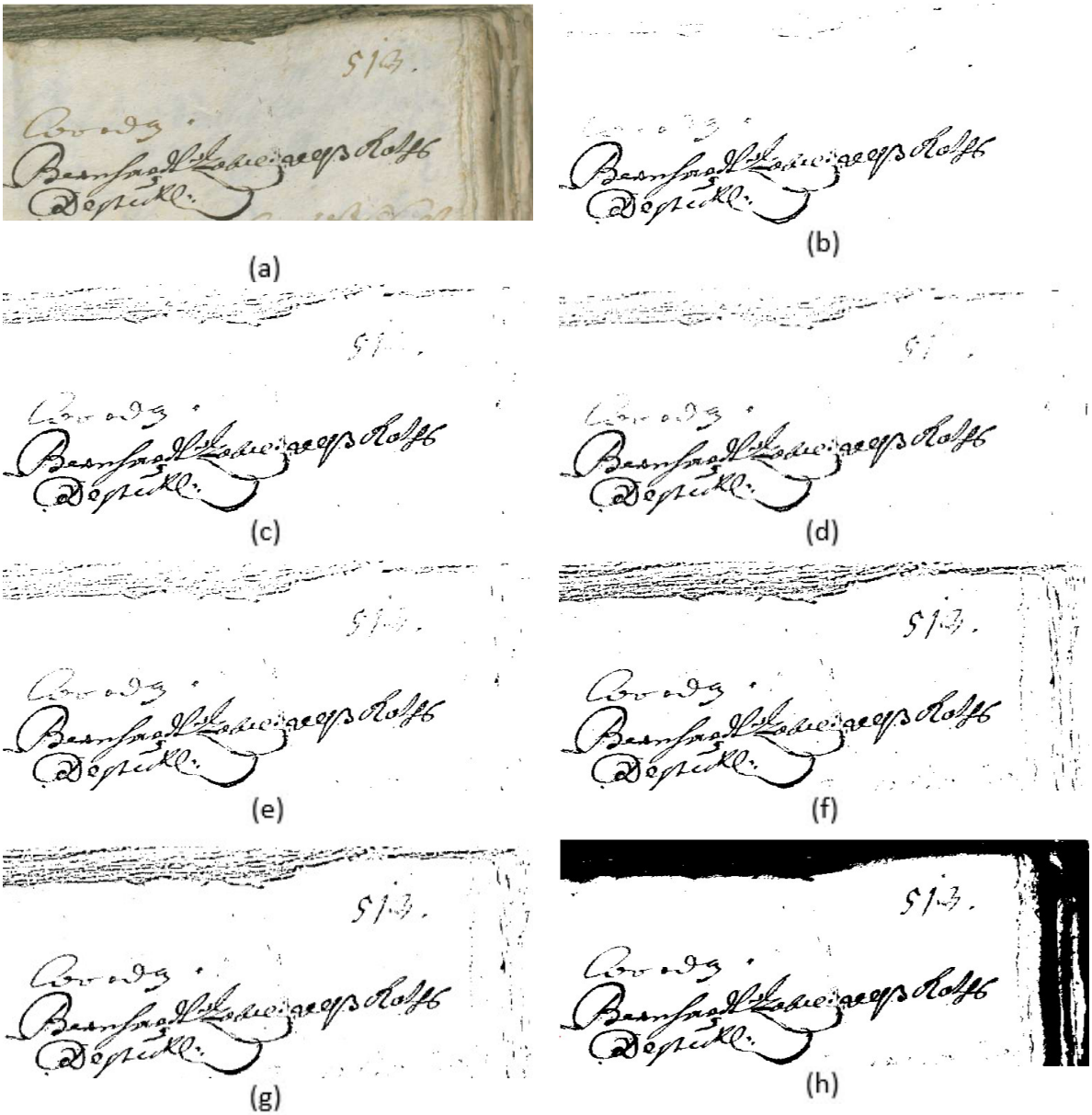


Fig. 4.4 a) Input (7th image from DIBCO-2018)) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick, h) Otsu , i) Spd-Otsu, j) Cprgb2gray -Otsu



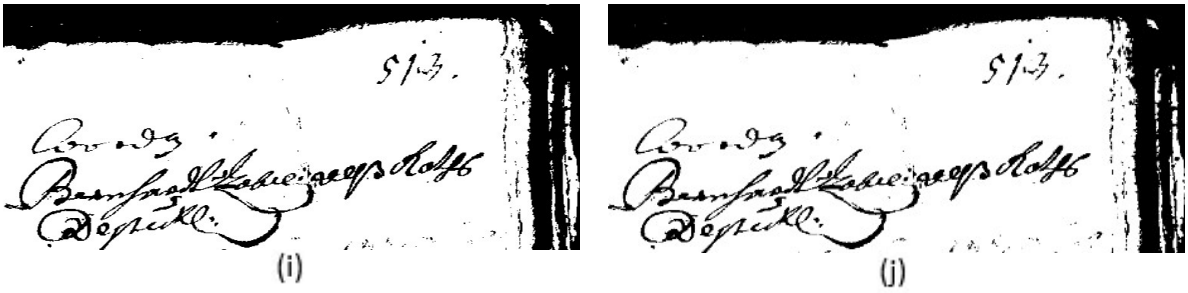


Fig. 4.5 a) Input (5th image from DIBCO-2018)) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick, h) Otsu , i) Spd-Otsu, j) Cprgb2gray -Otsu

Table 4.1 shows the binarization evaluation metrics such as F-measure, PSNR and NRM for the above stated images. From these metrics, it can be inferred that the proposed SPDecolor pre-processed binarization technique perform better for the local binarization algorithms.

Table 4.1: Comparison of Statistical metrics for various DIBCO Images

	Method	F-measure	PSNR	NRM
2nd Image DIBCO-2011 (Handwritten)	Sauvola	82.33	20.16	14.87
	Spd-Sauvola	86.69	21.21	11.40
	CP-Sauvola	80.52	15.31	15.72
	Nick	86.12	21.07	12.07
	Spd-Nick	**89.54**	**22.15**	**9.16**
	CP-Nick	79.86	15.21	16.32
	Otsu	88.93	21.91	9.50
	Spd-Otsu	87.96	21.26	9.88
	CP-Otsu	81.31	15.46	15.2
8th Image DIBCO-2011 (Printed)	Sauvola	76.86	14.74	18.57
	Spd-Sauvola	81.42	15.46	14.99
	CP-Sauvola	82.01	15.42	14.68
	Nick	72.63	14.17	21.39
	Spd-Nick	**83.06**	**16.01**	**14.28**
	CP-Nick	82.08	15.19	14.39
	Otsu	76.86	14.74	18.57
	Spd-Otsu	81.42	15.46	14.99
	CP-Otsu	80.12	15.48	14.91
7th Image DIBCO-2018	Sauvola	40.43	17.64	37.03
	Spd-Sauvola	56.43	18.49	29.66
	CP-Sauvola	50.97	18.14	32.32
	Nick	59.24	18.70	28.26
	Spd-Nick	**71.9**	**19.52**	**19.98**
	CP-Nick	65.95	19.06	23.35
	Otsu	56.76	16.80	20.14
	Spd-Otsu	57.29	16.91	20.07
	CP-Otsu	55.36	16.71	21.22
5th Image DIBCO-2018	Sauvola	48.23	14.84	33.95
	Spd-Sauvola	**52.58**	**14.89**	30.7
	CP-Sauvola	50.86	14.83	31.83
	Nick	47.56	14.54	33

Spd-Nick	48.61	13.66	28.98
CP-Nick	48.31	13.07	30.06
Otsu	26.81	8.64	29.06
Spd-Otsu	27.25	8.58	28.21
CP-Otsu	27.14	8.54	**28.14**

From the above results it is inferred that this preprocessing technique is efficient for local binarization methods only. So this contrast preserving pre-processing techniques is applied to local binarization techniques and are tested on DIBCO(2009-2011) datasets both statistically and visually.

Figure 4.6 shows the result binarization when the input is 1st image from DIBCO 2009(printed) dataset. SPDecolor pre-processing results in a better-binarized output in the case of Nick and Sauvola than the conventional Technique while there is no change in result for Otsu binarization with or without SPDecolor pre-processing. While CPrgb2gray pre-processed Sauvola, Nick and Otsu binarizations does not produce much significant change in output when compared with conventional binarizations.

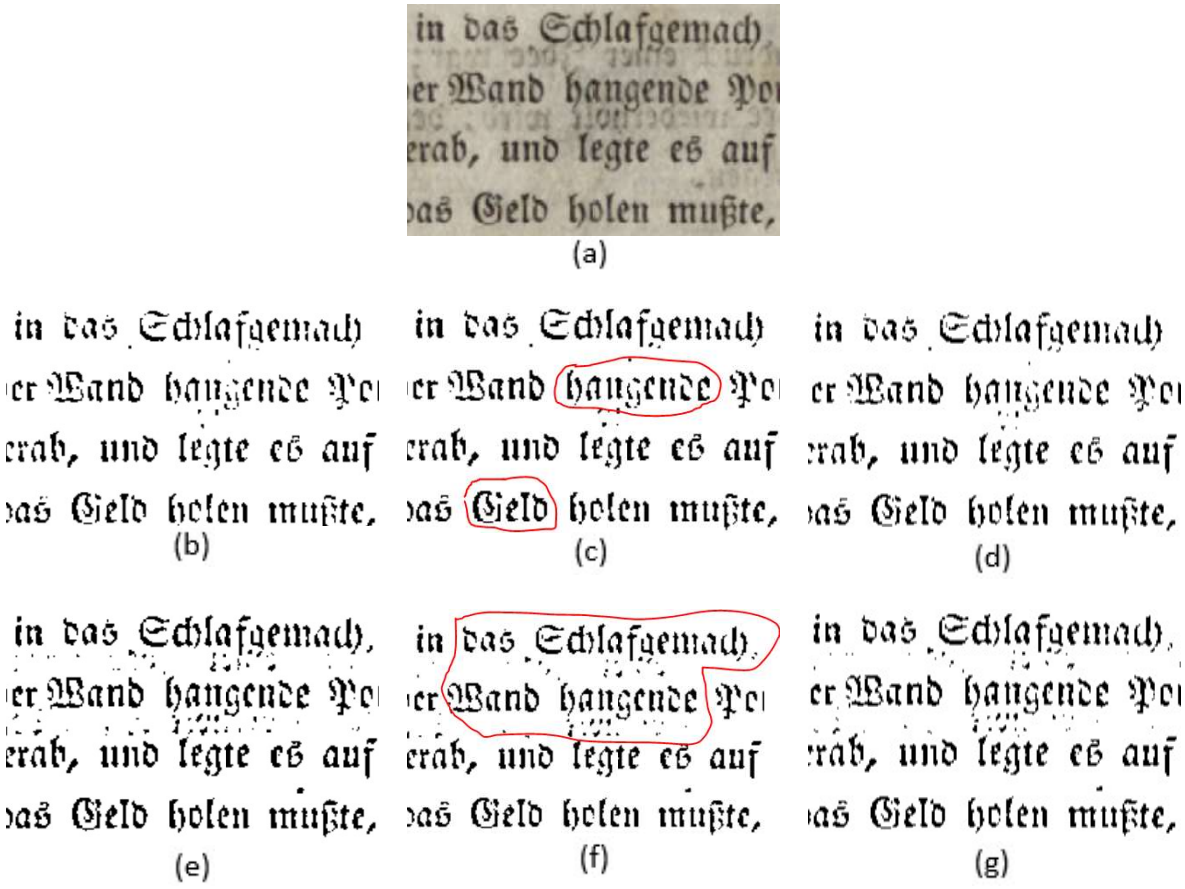


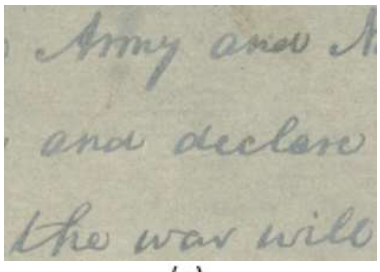
Fig. 4.6 a) Input image (1st image from DIBCO-2009 Printed) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Figure 4.7 shows the result binarization when the input is 5th image from DIBCO 2009(printed) dataset. SPDecolor pre-processing results in a better-binarized output in the case of Nick and Sauvola than the conventional Technique,while the CPrgb2gray pre-processed Savoula and Nick binarization output is degraded , when compared with its

conventional binarization methods which shows that SPDecolor pre-processing is better than cprgb2gray pre-processing for this type of degraded image.

Fig. 4.7 a) Input image (5th image from DIBCO-2009 Printed) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray - Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Figure 4.8 shows 1st image from the 2010 handwritten DIBCO dataset. The image contains dark greyish background with the text written with light blue ink. Sauvola and Nick produce a lot of degradation in the text, while when we pre-processed it with SPDecolor the results were better and now the text is visible and readable. While CPrgb2gray pre-processed Nick result is good, and it can be seen that the noise content which was there in SPDecolor pre-processing is not there in Cprgb2gray pre-processed nick . But the SPDecolor pre-processed Savoula output is better than the CPrgb2gray pre-processed Sauvola output



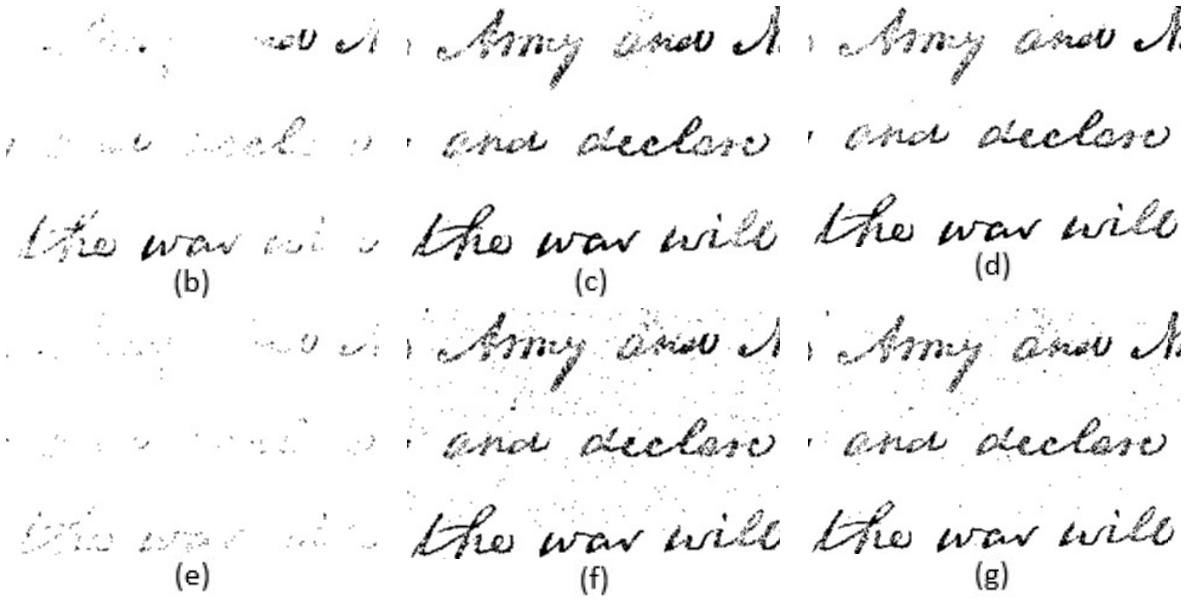


Fig. 4.8 a) Input (1st jmage from DIBCO-2010HW)) (lightly printed text) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Figure 4.9 shows 3rd image from the 2010 handwritten DIBCO dataset. The image contains greyish background with small blue dots all over in the background and some part of the text written with light hand . SPDecolor pre-processing results in a better-binarized output in case of Nick and Sauvola compared to the conventional Techniques. While CPrgb2gray pre-processed Sauvola, Nick binarizations output is degraded when compared with its conventional binarization methods.

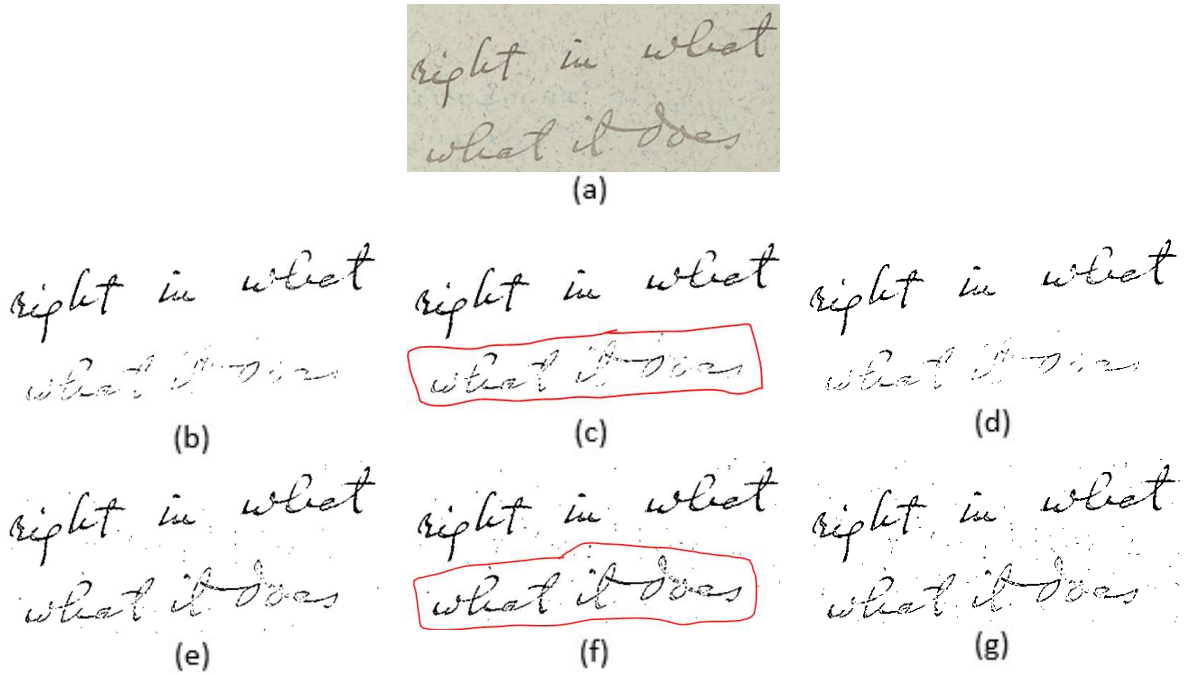


Fig. 4.9 a) Input image (3rd image from DIBCO-2010 handwritten) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Figure 4.10 shows 3rd image from the 2011 handwritten DIBCO dataset. The image contains colored background with some dark spots and the text is written with thin strokes. The conventional Sauvola and Nick method produces some degradation in the text while the SPDecolor pre-processed method produces better results. While CPrgb2gray preprocessed Savoula and Nick binarizations output is degraded when compared with its conventional binarization methods

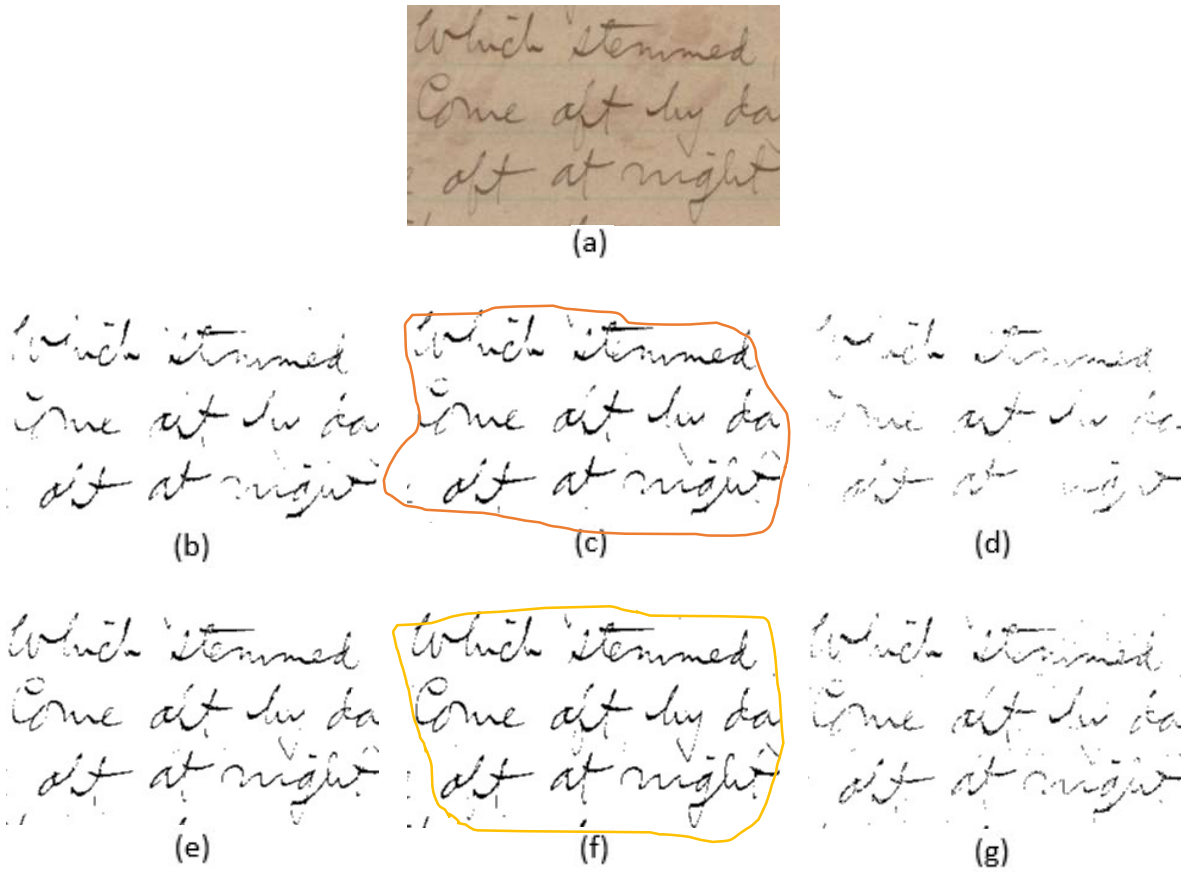


Fig. 4.10 a) Input (3rd image from DIBCO-2011HW)) (thin strokes, coloured background with some dark spots) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Figure 4.11 shows 2nd image from the 2011 handwritten DIBCO dataset. The image contains colored background with some part of text which is written with thin stroke and light hand. SPDecolor pre-processing results in a better-binarized output in the case of Nick and Sauvola than the conventional Technique. While CPrgb2gray pre-processed Savoula, Nick and binarizations output is better than tha conventional methods , SPDecolor pre-processing result is best.

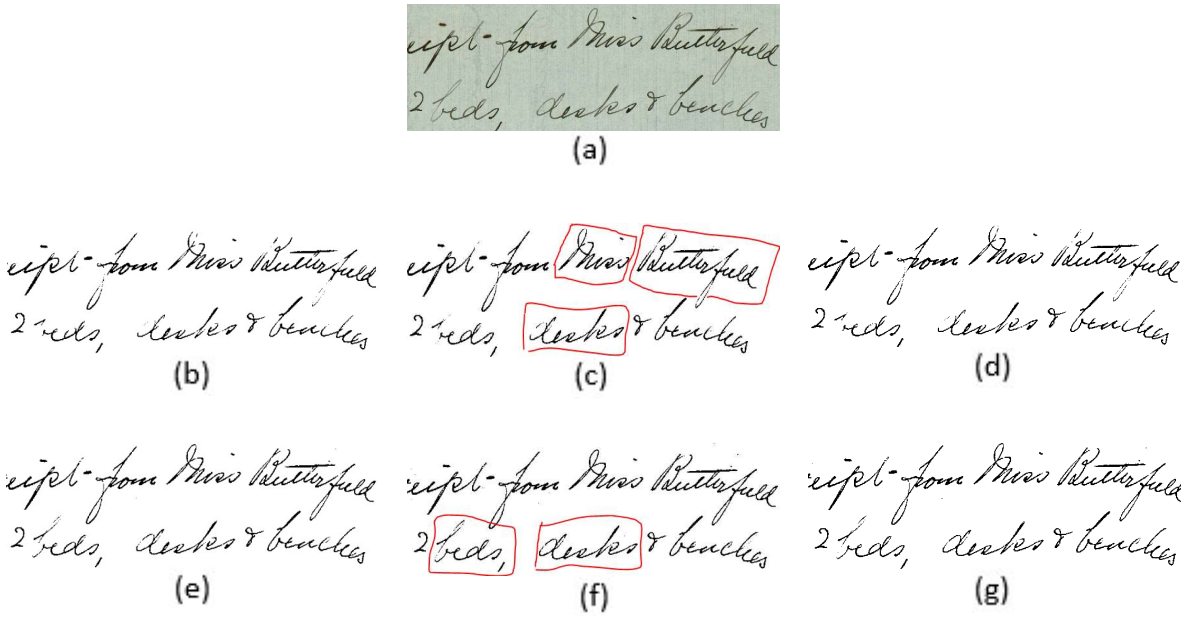
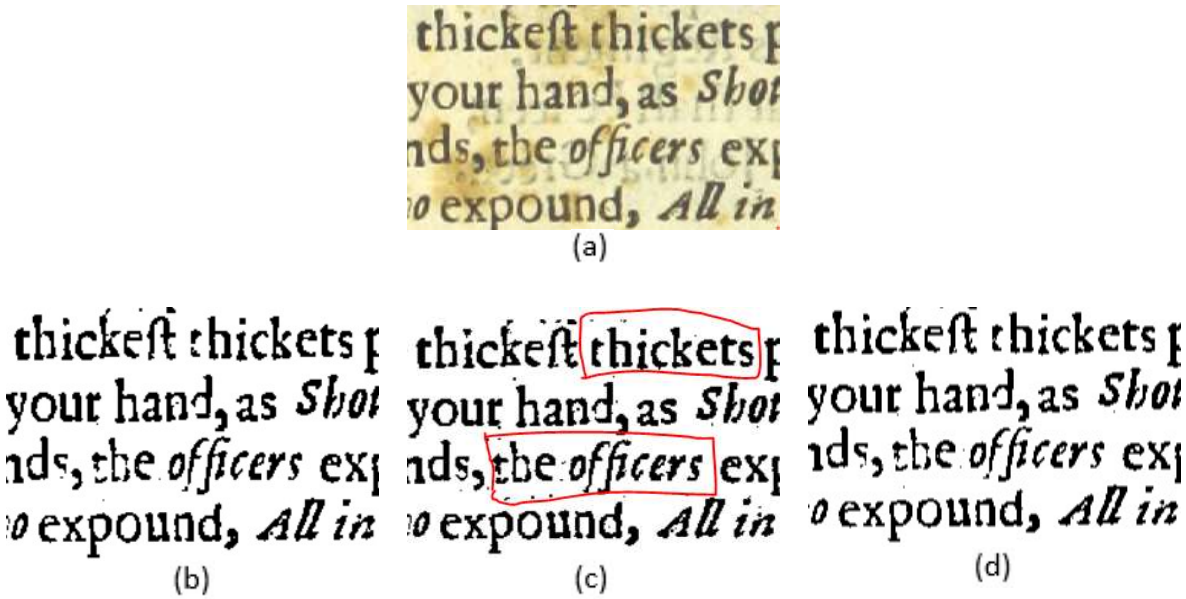


Fig. 4.11 a) Input image (2nd image from DIBCO-2011 handwritten) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Figure 4.12 shows 3rd image from the 2011 printed DIBCO dataset. The image contains colored background with ink-bleed. From the results, it can be seen that SPDecolor pre-processed Sauvola produces better results than the conventional Sauvola.CPrgb2gray pre-processing does not produce any significant change in output when compared with conventional method.



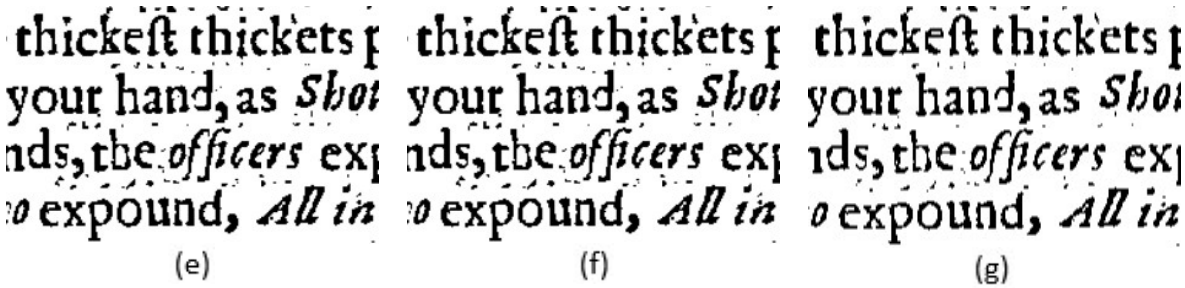
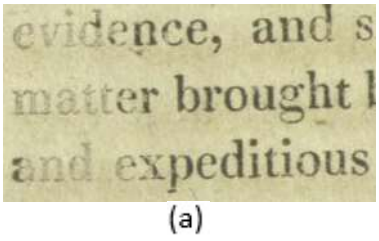


Fig. 4.12 a) Input image (3rd image from DIBCO-2011 Printed) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Figure 4.13 shows 8th image from the 2011 printed DIBCO dataset. The image contains some part of the text that is faded. From the results, it can be seen that SPDecolor pre-processed Nick and Sauvola produces better results than the conventional method .While CPrgb2gray pre-processed Savoula, Nick and binarizations output is better than than conventional methods , SPDecolor pre-processing result is best.



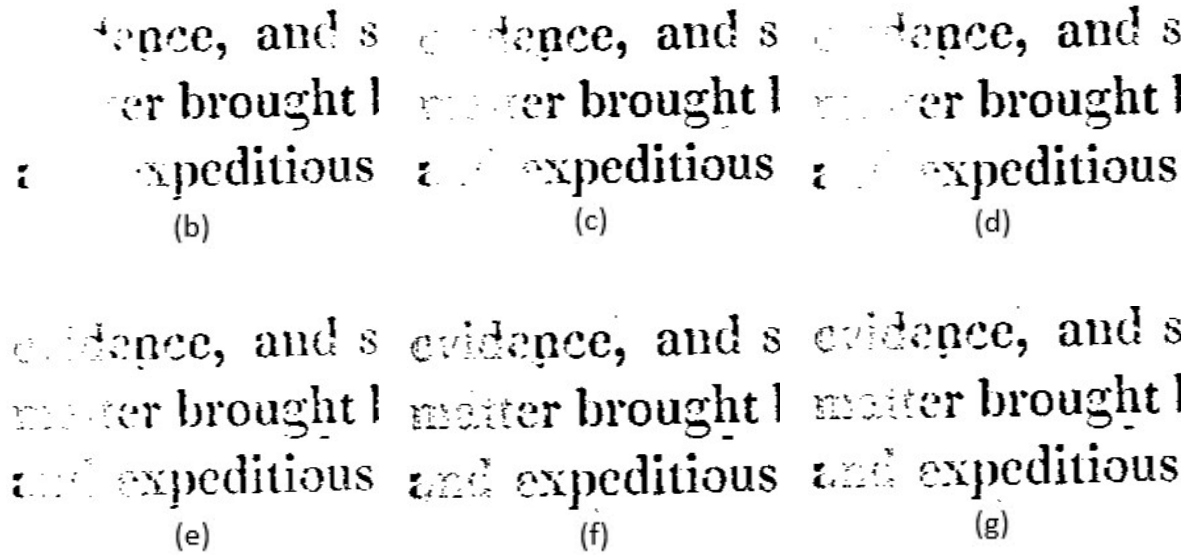


Fig. 4.13 a) Input image (8th image from DIBCO-2011 Printed) (lightly printed) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

Figure 4.14 shows 3rd image from the 2011 handwritten DIBCO dataset. The image contains color background and the faint text is written with thin strokes . The conventional Sauvola

and Nick method produces some degradation in the text while the SPDecolor pre-processed method produces better results. While CPrgb2gray preprocessed Savoula and Nick output is degraded when compared with its conventional binarization methods

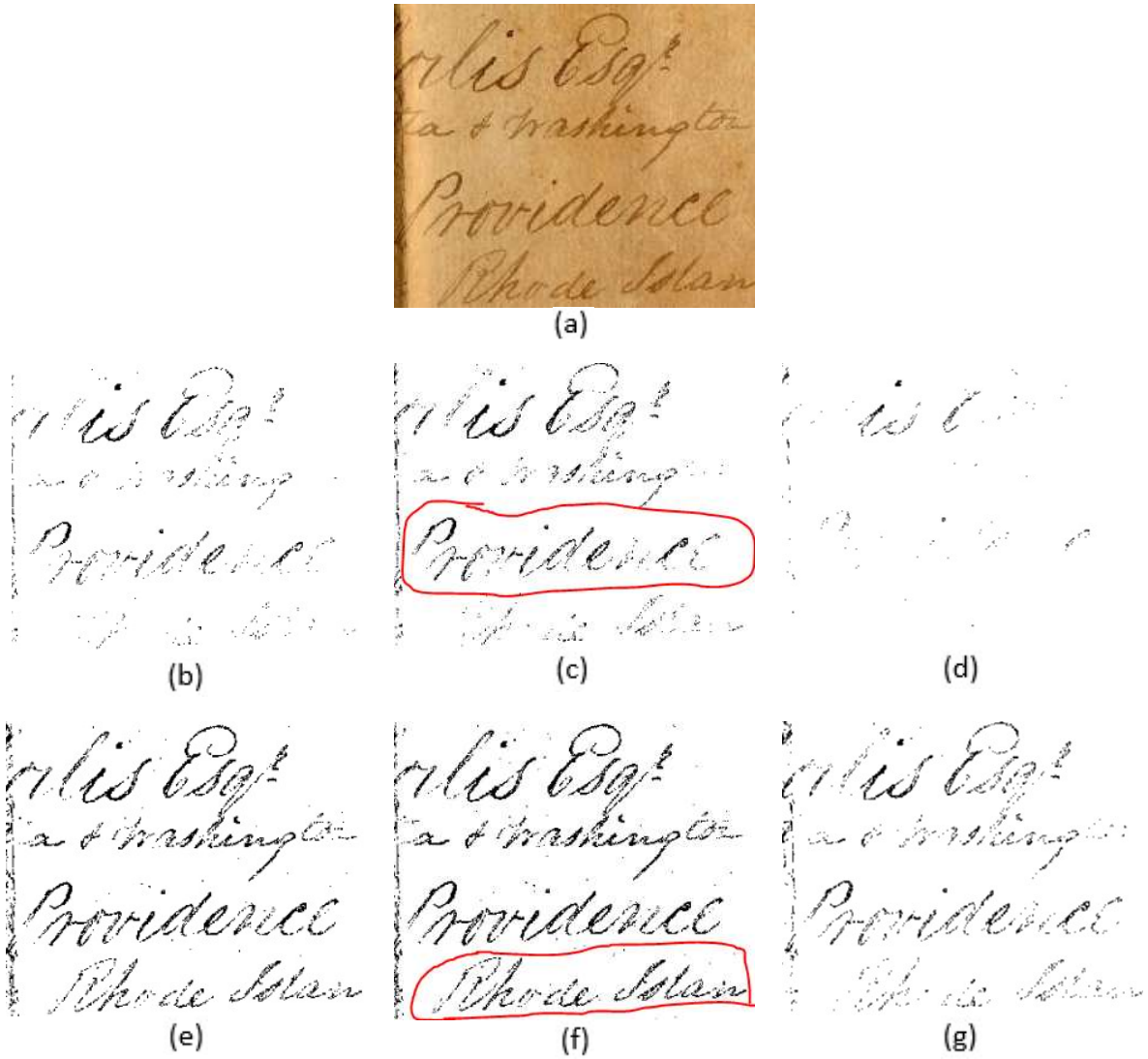


Fig. 4.14 a) Input image (6th image from DIBCO-2011 handwritten) b) Sauvola , c) Spd -Sauvola, d) Cprgb2gray -Sauvola, e) Nick, f) Spd-Nick , g) Cprgb2gray -Nick

By using benchmark ICDAR measures such as NRM , PSNR and F-measure we have quantitatively evaluated different binarization results for the above stated datasets.

F-Measure (FM)

FM can be defined with the help of Precision and Recall as:-

$$\text{Recall} = \frac{TP}{FN + TP} * 100 \qquad (4.5)$$

$$\text{Precision} = \frac{TP}{FP + TP} * 100 \qquad (4.6)$$

True Positive(TP) of an image is when both ground truth and pixels of image is treated as foreground. False Positive (FP) of an image is when ground truth of image is treated as background and image pixel is treated as foreground. False Negative(FN) of an image is when the ground truth of image gets treated as foreground and image pixel gets treated as background.

So, formula for calculating F-measure is

$$FM = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (4.7)$$

The greater is the F-measure metric the better is that technique.

Peak SNR (PSNR)

PSNR stands for Peak signal to noise ratio and it compares the binarized image with original image on the basis of quality.
Mean Square Eror (MSE) is used for the calculation of PSNR.

$$MSE = \sum_{x=1}^{M} \sum_{y=1}^{N} \frac{(I_1(x,y) - I_2(x,y))^2}{N * M} \qquad (4.8)$$

$$PSNR = 10^* \log_{10}\left(\frac{D^2}{MSE}\right) \qquad (4.9)$$

The value of D =1 as it the difference between the pixel value of background and foreground in a binary image. The greater is PSNR metric of a method ,the better is that method.

Negative Rate Metric (NRM)

NRM is mean of the negative rate of false positive (NRFP) and negative rate of false negative (NRFN) .
The Negative rate of false negative (NRFN) is stated using False Negative (FN) and True Positive(TP) as:

$$NR_{FN} = \frac{FN}{FN + TP} \qquad (4.10)$$

Negative rate of false positive (NRFP) is stated using False Positive (FP) and True Negative(TN) as:

$$NR_{FP} = \frac{FP}{FP + TN} \qquad (4.11)$$

And so NRM is:-

$$NRM = \frac{NR_{FN} + NR_{FP}}{2} \qquad (4.12)$$

Table 4.2 shows the comparison of various binarization evaluation metrics, such as F-measure , PSNR and NRM on DIBCO 2009-11 dataset, for local binarization techniques . The metrics value for the conventional Nick and Sauvola for these dataset, has been taken from [23].From these metrics it evident that the SPDecolor pre-processed binarization technique perform better than the convention binarization technique.

Table 4.2 Comparison of Statistical metrics for Local binarization techniques on DIBCO(2009-11) dataset.

Dataset	Method	F-measure	PSNR	NRM
DIBCO-2009 Printed	Sauvola	68.69	11.89	23.15
	Spd-Sauvola	**84.9**	17.01	12
	CP-Sauvola	72.79	12.86	19.28
	Nick	76.35	12.924	17.67
	Spd-Nick	84.82	**16.75**	**11**
	CP-Nick	80.05	14.01	15.22
DIBCO-2009 Handwritten	Sauvola	51.13	15.48	29.6
	Spd-Sauvola	67.4	18.6	21.9
	CP-Sauvola	65.31	16.9	22.8
	Nick	73.3	16.43	18.1
	Spd-Nick	**76.32**	**18.83**	**15.7**
	CP-Nick	74.9	17.4	17.9
DIBCO-2010 Handwritten	Sauvola	40.36	30.13	14.07
	Spd-Sauvola	63.68	51.29	17.09
	CP-Sauvola	58.37	16.23	26.86
	Nick	60.22	47.2	14.64
	Spd-Nick	**72.01**	**60.04**	**17.34**
	CP-Nick	66.06	16.12	22.83
DIBCO-2011 Printed	Sauvola	67.76	12.48	23.91
	Spd-Sauvola	**86.94**	**18.24**	**9.6**
	CP-Sauvola	78.97	15.12	12.9
	Nick	76.04	13.92	17.13
	Spd-Nick	79.47	16.2	11
	CP-Nick	77.8	14.92	13.6
DIBCO-2011 Handwritten	Sauvola	62.46	14.7	25
	Spd-Sauvola	81.02	**18.92**	13.6
	CP-Sauvola	76	16.24	18.3
	Nick	74.4	15.4	18.1
	Spd-Nick	**81.49**	18.37	**11**
	CP-Nick	79.61	15.97	14.3

Chapter 6

CONCLUSION

In this dissertation , Contrast preserving decolorization techniques such as SPDecolor and CPrgb2gray have been analysed and they have been applied in the field of binarization of historical documents. These contrast preserving techniques have been applied as pre-processing step in the normal binarization process. The performance of these technique has been assessed mainly on DIBCO (2009–2011) dataset. Experimental results on DIBCO dataset images show a preferable performance of the SPDecolor based approach, relative to the conventional color- to-gray conversion for contemporary binarization methods, in terms of obtained PSNR, NRM and F-measure metrics. From the experiments done so far, we can conclude that this SPDecolor based pre-processing improves the result of local binarization techniques compared to conventional local binarization techniques. Also, for certain types of degradation such as faded/faint text, this SPDecolor pre-processing works better for local binarization techniques than the Global binarization techniques. So, the this approach is more suitable for the local binarization techniques.

References

[1] Sridhar Cherala Pritirege,"Palm Leaf Manuscript/Color Document Image Enhancement by using improved Adaptive binarization method",Sixth Indian Conference on Computer Vision, Graphics & Image Processing,978-0-7695-3476-3/08 $25.00 © 2008 IEEE DOI 10.1109/ICVGIP.2008.64

[2] A.Venkata Srinivasa Rao et.al , "Adaptive Binarization of Ancient Documents." 2009 Second International Conference on Machine Vision, IEEE Computer society.

[3] P. K. Sahoo et.al, A survey of thresholding techniques, Computer Vision, Graphics Image Processing, 41 (2), 233&260,1988

[4] Anna Tonazzini et.al," Independent component analysis for document restoration",International Journal of Document Analysis and Recognition. Digital Object Identifi er (DOI) 10.1007/s10032-004-0121-8 IJDAR (2004) 7: 17–27.

[5] Nafchi, Hossein Ziaei et.al, "Historical document binarization based on phase information of images." Asian Conference on Computer Vision. Springer Berlin Heidelberg, 2012.

[6] Avekash Gupta et.al," Enhancement of old Manuscript Images." Ninth International Conference on Document Analysis and Recognition, 2007. ICDAR 2007.

[7] B. Gangamma et.al," A Combined Approach for Degraded Historical Documents Denoising using Curvelet and mathematical morphology". Computational Intelligence and Computing Research (ICCIC), 2010,

[8] Zhixin Shi et.al," Digital Enhancement of Palm Leaf Manuscript Images using Normalization Techniques",Center of Excellence for Document Analysis and Recognition (CEDAR), 5th International Conference On Knowledge Based ComputerSystems, 2004 December 19-22, 2004 Hyderabad, India; 2004.

[9] Jayanthi, N., et al. "Novel method for manuscript and inscription text extraction." Signal Processing and Integrated Networks (SPIN), 2016 3rd International Conference on. IEEE, 2016.

[10] Qiegen Liu, Peter X. Liu,, "Semi-Parametric Decolorization with Laplacian-based Perceptual Quality Metric , IEEE Transactions on Circuits and Systems for Video Technology

[11] Grundland, Mark, and N. Dodgson. "Decolorize: fast, contrast enhancing, color to grayscale conversion." Color Research and Application 16.6 (2009): 385-393.

[12] Gooch, Amy A., et al. "Color2gray: salience-preserving color removal." ACM Transactions on Graphics (TOG) 24.3 (2005): 634-639.

[13] K. Rasche, R. Geist, J. Westall, Re-coloring images for gamuts of lower dimension, Computer Graphics Forum 24 (2005) 423–432 (Proceedings of EUROGRAPHICS).

[14] Y. Kim, C. Jang, J. Demouth, and S. Lee, "Robust color-to-gray via nonlinear global mapping," ACM Trans. Graph., vol. 28, no. 5, pp. 1–4, Dec. 2009.

[15] Sauvola, Jaakko, and Matti Pietikäinen. "Adaptive document image binarization." Pattern recognition 33.2 (2000): 225-236.

[16] Khurshid, Khurram, et al. "Comparison of Niblack inspired Binarization methods for ancient documents." IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, 2009.

[17] Lu, Cewu, Li Xu, and Jiaya Jia. “Real-time contrast preserving decolorization.” SIGGRAPH Asia 2012 Technical Briefs. ACM, 2012.

[18] http://mscollab.hypotheses.org, Eleventh- to Twelfth-Century Irish Manuscript.

[19] DIBCO-2010 Dataset, http://users.iit.demokritos.gr/~bgat/H-DIBCO2010/

[20] DIBCO-2009 Dataset. http://users.iit.demokritos.gr/ bgat/DIBCO2009/benchmark/

[21] DIBCO-2011 Dataset, http://utopia.dutch.gr/ ipratika/DIBCO2011/benchmark/

[22] Natarajan J, Sreedevi I (2017) Enhancement of ancient manuscripts images by log based binarization technique. Int J Electron Commun 75:15–22

[23] Rani, U., Kaur, A. & Josan, G. A new binarization method for degraded document images. Int. j. inf. tecnol. (2019). https://doi.org/10.1007/s41870-019-00361-3

[24] Ntirogiannis K. Document image binarization. http://cgi.di.uoa.gr/ ~phdsbook/files/Ntirogiannis%20Kostas.pdf

[25] Otsu N. A threshold selection method from gray level histograms. IEEE Trans Syst Man Cybern

1979;9:62–6.

[26] Kittler J, Illingworth J. Minimum error thresholding. Pattern Recognit 1986;19:41–7

[27] Niblack W. An introduction to digital image processing. Englewood Cliffs: Prentice Hall; 1986.

[28] C. Lu, L. Xu, J. Jia, “Contrast preserving decolorization,” In: IEEE International Conference on Computational Photography (ICCP), 2012, pp. 1-7.

[29] R. Eschbach and R. Bala “Spatial color-to-grayscale transform preserving chrominance edge information,” *In Color Imaging Conference*, 2004, pp. 82-86.
[30] A. Nemcsics, L. Neumann, M. Cad´ık, , “An efficient perception-based adaptive color to gray transformation,” *In Computational Aesthetics*, 2007, pp. 73-80.
[31] K. Myszkowski , K. Smith, P.E. Landes, J. Thollot, “Apparent greyscale: a simple and fast conversion to perceptually accurate images and video,” *Comput. Graph. Forum*, vol. 27, no. 2, pp. 193-200, 2008.
[32] J. Tumblin, A. A. Gooch, S. C. Olsen, and B. Gooch, “Color2gray: salience-preserving color removal,” *ACM Trans. Graph.*, vol. 24, no. 3, pp. 634-639, 2005.
[33] J. Westall, K. Rasche, R. Geist, “Re-coloring images for gamuts of lower dimension,” *Comput. Graph. Forum*, vol. 24, no. 3, pp. 423-432, 2005.
[34] S. Lee ,Y. Kim, C. Jang, J. Demouth, and, “Robust color-to gray via nonlinear global mapping,” *ACM Trans. Graph.*, vol. 28, no. 5, pp. 1-4, 2009.
[35] N. I. Cho ,J. G. Kuk, J. H. Ahn, “A color to grayscale conversion considering local and global contrast,” in *Proc. ACCV*, 4(2010), pp. 513-524.
[36] C. W. Chen ,M. Song, D. Tao, C. Chen, X. Li, “Color to gray: Visual cue preservation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1537-1552, 2010.
[37] B. Sheng, H. Du, S. He, et al., “Saliency-guided color-to-gray conversion using region-based optimization,” *IEEE Trans. Image Process.*, vol. 24, no. 1, pp. 434-443, 2015.

[38] J. Jia, C. Lu, L. Xu, “Real-time contrast preserving decolorization,” *ACM Siggraph Asia Technical Berief,* 2012.
[39] M. K. Ng ,Z. Jin, F. Li, “A variational approach for image decolorization by variance maximization,” SIAM J. Imaging Sci., vol. 7, no. 2, pp. 944-968, 2014.
[40] Y. Wang, Q. Liu, P.X. Liu, W. Xie, D. Liang, “GcsDecolor: Gradient correlation similarity for efficient contrast preserving decolorization,” IEEE Trans. Image Process., vol. 24, no. 9, pp. 2889-2904, 2015.
[41] M. Cad´ık, “Perceptual evaluation of color-to-grayscale image conversions,” *Comput. Graph. Forum*, vol. 27, no. 7, pp. 1745-1754, 2008.
[42] E. P. Simoncelli ,Z. Wang, A. C. Bovik, H. R. Sheikh, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600-612, 2004