IMAGE DECLIPPING – SATURATION CORRECTION IN IMAGES

DİCLEHAN KARAKAYA

Master's Thesis

Graduate School

Izmir University of Economics

İzmir

2020

IMAGE DECLIPPING – SATURATION CORRECTION IMAGES

DİCLEHAN KARAKAYA

A Thesis Submitted to

The Graduate School of Izmir University of Economics

Master's Program in Electrical and Electronics Engineering.

İzmir

2020

ABSTRACT

IMAGE DECLIPPING – SATURATION CORRECTION IN IMAGES

Karakaya, Diclehan

M.Sc. in Electrical and Electronics Engineering

Advisor: Asst. Prof. Dr. Mehmet Türkan

August, 2020

High dynamic range (HDR) images present fine details in a scene and are visually more appealing than low dynamic range (LDR) images, since they contain a greater dynamic range of colors. HDR compatible displays are currently unaffordable, therefore tone-mapping algorithms have widely been used to obtain high quality images for LDR screens with a lower cost. However, tone-mapped images may contain clipped pixel regions, which should be corrected to retrieve the lost information, to acquire visually pleasing LDR images. In a single image, the recovery of color and texture information in clipped regions is challenging, yet attractive research field in image processing. Although there are several algorithms present in literature, developing a general framework for different types of image content is hard to achieve. This study proposes a single image declipping method based on linear embeddings (LE), difference of pixels and block-search. Experiments carried out on a tone-mapped HDR image dataset and raw LDR images demonstrate that the proposed algorithm is able to successfully recover saturated pixels in various types of images. Detailed statistical and visual comparisons show that this approach produces superior results on

average for both tone-mapped and (raw) LDR images when compared to existing techniques.

Keywords: Linear embedding, Declipping, Saturation correction, Color restoration, High dynamic range, Tone-mapping.

ÖZET

GÖRÜNTÜDE KIRPMA VE DOYGUNLUK DÜZELTME

Karakaya, Diclehan

Elektrik ve Elektronik Mühendisliği Yüksek Lisans Programı

Tez Danışmanı: Dr. Öğr. Üyesi Mehmet Türkan

Ağustos, 2020

Yüksek dinamik aralığa (YDA) sahip görüntüler bir sahnedeki ince detayları sunabilir ve düşük dinamik aralıktaki (DDA) resimlerden görsel olarak daha çekicidirler, çünkü renklerin dinamik aralığı YDA'da daha fazladır. Günümüzde YDA ile uyumlu cihazlar pahalı olduğundan ton haritalama algoritmaları düşük maliyetle yüksek kaliteye sahip görüntüler elde etmek için sıklıkla kullanılmaktadır. Ancak ton haritalanmış görüntüler yanmış piksel bölgeleri barındırabilir; bilgi kaybının önüne geçmek ve görsel olarak hoş görüntüler elde etmek için bu piksellerdeki kırpma düzeltilmelidir. Ton haritalanmış tek bir resimde, kırpılmış alanlardaki renk ve doku bilgisinin düzeltilmesi görüntü işleme alanında zorlu fakat çekici bir çalışma dalıdır. Literatürde birçok algoritma bulunmasına rağmen, farklı türdeki görüntüleri düzeltebilen genel bir sistem yaratmayı başarmak zordur. Bu tezde doğrusal gömme, piksel değerlerinin farkı ve blok aramaya dayanan bir tek resim kırpma düzeltme yolu önerilmiştir. Ton haritalanmış bir YDA veri seti ve ham DDA görüntüler ile yapılan deneyler, önerilen algoritmanın farklı tür resimlerdeki kırpılmış pikselleri başarılı bir şekilde düzelttiğini göstermiştir. Yapılan detaylı görsel ve istatistiksel incelemeler, önerilen yöntemin

ortalamada hem ton haritalanmış hem de (ham) DDA görüntülerde var olan tekniklerden daha iyi sonuçlar verdiğini göstermiştir.

Anahtar Kelimeler: Doğrusal gömme, Doygunluk düzeltme, Renk iyileştirmesi, Yüksek dinamik aralık, Ton haritalama.

To my family…

ACKNOWLEDGEMENTS

I would like to express my greatest appreciation to my supervisor, Asst. Prof. Dr. MEHMET TÜRKAN for his continuous support, supreme supervision, motivation and constructive comments throughout my graduate and undergraduate studies. I feel very fortunate to become his student; thanks to his guidance I have improved myself considerably for the aims I am pursuing.

Furthermore, I wish to express my deepest love to OĞUZHAN ULUCAN for his dedicated support, enthusiasm, and invaluable friendship in these challenging years. His love and motivation helped me to stay patient and goal-oriented throughout my studies.

I also wish to present my gratitude to my mother KADRIYE YÜKSEL for her encouragement in whatever I pursue, understanding and moral support. I will be grateful forever for her love.

Finally, I would like to thank the jury members Prof. Dr. Cüneyt GÜZELİŞ and Prof. Dr. Türker İNCE for their insightful comments.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The photographs we capture to immortalize the moments we love, should have an excellent quality to look as pleasant as the feelings we had in that moment. Unfortunately, undesired effects can be observed while capturing images, such as under- and over-exposed regions and color distortion. Over-exposed regions are excessively bright areas, which occur due to a too long exposure-time of the camera (Masood, Zhu and Tappen, 2009). These areas suffer from the saturation-related information loss and usually occur in scenes with high contrast (Honig and Werman, 2018). Contrast is defined as the difference between the darkest and the brightest intensity of pixels (Banterle et al., 2017). The dynamic range of a scene is increased by the contrast, corresponding to the increasing ratio between the brightest and the darkest areas. Furthermore, the dynamic range is defined as the saturation-to-noise-ratio in cameras and is indeed dependent on the sensor capacity of the camera. Hence images captured with different cameras have distinct visual quality. For instance, HDR cameras are able to produce high quality images containing 32 bits per channel and preserve details in captured scenes (Guan and Qiu, 2007). Nevertheless, HDR displays are required to project these images to avoid information loss during projection. HDR displays can process 96 bits per pixel, while LDR screens only contain 24 bits per pixel. As a result, the ability of producing and projecting high quality images makes HDR imaging devices very attractive.

With the rapid improvements in the technology, HDR-compatible equipment will soon become accessible for most of the consumers. However, the current HDR imaging technology is still unaffordable. Alternatively, HDR-like content can be obtained from an LDR exposure stack rather than directly capturing and projecting HDR images. Multi-exposure image fusion (MEF) takes a stack of input images of the same scene with different exposures and blends the stack into a single informative image, which is commonly employed to produce an HDR-like image for LDR displays (Mertens, Kautz and Van Reeth, 2009; Hayat and Imran, 2019). However, multiple input LDR images are required in MEF and their acquisition is a troublesome process (Cai, Gu and Zhang, 2018). If the camera cannot be maintained stable or hands shake during

capturing the exposures, the images may suffer from the jitter effect, which presents an output as if there is motion in the scene where in fact the scene is static (Angelov, 2012). Furthermore, the blended image may contain ghosting effects in the presence of moving objects in the scene (Mahmud et al., 2019).

In order to lower the economic burden and avoid a specialized hardware, or a human operator, an approach in obtaining high quality images is capturing a scene with HDR cameras and projecting it to LDR through tone-mapping algorithms (Pouli et al., 2013). Generally, tone-mapping algorithms are divided into two categories; global tone-mapping operators (G-TMO) and local tone-mapping operators (L-TMO) (Eilertsen, 2018). G-TMO applies the same mapping function to all the pixels in an image without taking local spatial information into account. On the other hand, L-TMO changes the function according to the pixel's local neighbourhood, hence L-TMO covers the details and local contrast better than G-TMO. However, L-TMO has two drawbacks; it is computationally expensive, and it usually tends to produce more artefacts.

Furthermore, while the employed tone-mapping operator may differ, in a typical tone-mapping algorithm, there are four different parts; *(i)* pre-processing, *(ii)* edge-preserving filtering, *(iii)* tone-curve adaption, and *(iv)* post-processing (Eilertsen, 2018). In the pre-processing stage, the HDR image is converted into an appropriate format for tone compression. Then, an edge-preserving filter such as bilateral or pyramidal is employed to decompose the image into base and detail layers. Afterwards, a tone-curve such as exponential or sigmoid is applied to transform the HDR luminance to LDR luminance. Lastly, suitable operations such as color restoration or gamma correction are adopted in the post-processing step.

In recent decades, several tone-mapping algorithms have been proposed (Reinhard et al., 2002; Reinhard and Devlin, 2005; Šikudová et al., 2015; Khan et al., 2017; Ok and Lee, 2017; Liang et al., 2018), which take advantage of several methods such as histogram of luminance, and difference compression with adaptive reference values.

Even though tone-mapping is able to produce satisfying results, some pixels displayed on LDR screens may appear brighter or darker than they supposed to be (Mantiuk et al., 2009; Chae, Lee and Sohng, 2013). Hence, detail loss and color distortion are usually observed. If these shortcomings can be eliminated, tone-mapping is a very effective approach to obtain high quality content. Therefore, the reasons behind the information loss in tone-mapping should be thoroughly investigated.

When an image is displayed on a screen with lower dynamic range than itself (e.g. through tone-mapping), the phenomenon called "clipping" occurs in pixels (Xu, Doutre and Nasiopoulos, 2010). This leads to the emergence of over- and/or under-exposed regions in the displayed image. In simple terms, a single channel or multiple (two or more) channels of pixels are represented with faulty values, which cause an information loss, e.g., detail loss, in the image (Honig and Werman, 2018). Important note here that, the only reason of this phenomenon is not tone-mapping; but capturing or projecting a high contrast scene with limited dynamic range equipment will also result in clipping. In order to obtain visually plausible images containing significant features, it is critical and strenuous to successfully recover these clipped pixels. Especially, restoring intensity values in regions with all three clipped channels is more challenging when only the single input image is available at hand.

In recent decades, with the improvements and requirements in the imaging technologies, the correction of clipped (especially over-saturated) pixels in single images has become an attractive research field in image processing and computer vision. Several studies as will be explained in Chapter 2, have been conducted to correct clipped images and to produce visually appealing output photographs.

This thesis mainly focuses on correcting saturated (over-exposed) regions of a single tone-mapped image without any side information, which indeed requires more attention since there is a shortage of studies in literature; and extends its scope via working on (raw) LDR images containing over-exposed pixels. The proposed method is a simple yet effective pixel correction algorithm, whose parts are designed

individually for pixels clipped at one, two or three channels. This individual approach is adopted to carry out a more sensitive, case-oriented procedure and obtain more robust results. The pixel regions which are clipped in only one channel are corrected via LE, and to the best of available knowledge, this is the first time LE is adopted in this research domain. Then, the pixels jointly saturated in two channels are restored by taking the advantage of channel differences in the neighbouring pixels. The pixels clipped in all three channels are later corrected through a block-search approach based on a priority function and gradient information, and lastly a post-processing is applied via bilateral filtering. The developed single image saturation correction algorithm is compared with several studies and superior results on average are obtained both statistically and visually.

The remainder of this thesis is organized as follows. In Chapter 2, significant studies present in the literature are explained. In Chapter 3, the proposed novel method is demonstrated. Experimental results and detailed visual and statistical comparisons are presented in Chapter 4. Finally, a brief conclusion with possible future directions is given in Chapter 5.

CHAPTER 2: RELATED WORK

Especially, since the beginning of the current millennium, several studies in the field of image declipping are carried out. In the literature review of this thesis mainly approaches based on a single input image are considered since they are in the same context with the proposed method.

2.1 Automatic Correction of Saturated Regions in Photographs using Cross-Channel Correlation

The algorithm developed in (Masood, Zhu and Tappen, 2009), is an important study and commonly used in other works for comparison. In the proposed method, in RGB color space all pixels having a value of more than 235 are labelled as saturated and to recover them, advantage from non-saturated neighbour pixels is taken. By computing the correlation between channels of neighbour pixels and weighting them according to their distance to the non-saturated channel of the saturated pixel, an estimated value for the saturated pixel is provided.

Assuming that only one channel is saturated in the pixels of a certain area, the ratios in the non-saturated neighbour areas can be propagated into the saturated region through a quadratic cost function as given in Eqn. 1,

$$\sum W_a^C \, (V_a - V_a^O)^2 + \left(V_a - \sum_b W_{ab} \, V_b \right)^2 \qquad (1)$$

where, V is the ratio vector between two channels at pixel a and must be minimized, b is the neighbour pixel of a, the summation is carried out over all pixels b present in the image, W_a^C is a large value for non-saturated pixels and 0 for saturated ones, and W_{ab} is a weighting function.

W_{ab} is adopted in order to ensure that color ratios are not propagated across edges. This function provides an estimation for the required similarity between the color

ratios of a and b (Eqn. 2).

$$W_{ab} = \frac{1}{m}\sum_{i} exp\left(\frac{-\left(Z_i(a) - Z_i(b)\right)^2}{2\sigma_{a_i}^2}\right) \tag{2}$$

where, the summation is carried out over all non-saturated color channels, $Z_i(a)$ and $Z_i(b)$ are the non-saturated values at pixel a's i^{th} channels derived from a 3x3 neighbourhood of a, σ_{a_i} is the variance in the 3x3 neighbourhood of a without taking into account the saturated pixels and m is the total number of non-saturated channels in a.

After acquiring the ratios, the true values of the clipped pixels are estimated via the second cost function given below (while extensions to $\boldsymbol{G}$ and $\boldsymbol{B}$ are straightforward let assume that the $\boldsymbol{R}$ channel is clipped),

$$\begin{aligned} R^* = arg\min_{R} \sum_{a} ((W_a^G\ (R_a - V_a G_a)^2 +\ W_a^B\ (R_a - V_a^2 B_a)^2 \\ +\ W_a^n\ (R_a - (R_a)^O)^2\ +\ W_a^s \sum_{b} (R_a - R_b)^2) \end{aligned} \tag{3}$$

where, $\boldsymbol{R}^*$denotes the estimated value of the first clipped channel in a, $\boldsymbol{G}_a$ and $\boldsymbol{B}_a$ are the second and third channels in a respectively, if $\boldsymbol{R}_a$ is saturated and $\boldsymbol{G}_a$and/or $\boldsymbol{B}_a$ are non-saturated the weight functions $W_a^G$and/or $W_a^B$are set to 1. Furthermore, if $\boldsymbol{R}_a$ is not saturated, W_a^n is set to 1 and $\boldsymbol{R}^*$obtains a close value to its original intensity $(\boldsymbol{R}_a)^O$. The last term in the cost function is for pixel where all three channels are saturated. In this case, W_a^s will be set to 1 to locally smooth these pixels.

For the evaluation of this approach a comparison between the ground truth images and the output images of the algorithm is carried out via computing the error. Both statistically and qualitatively good results are obtained. However, there are shortcomings in this study. When all channels contain oversaturated pixels there is no information for the color ratio calculation. Therefore, the method can only be used in

the presence of at least one non-saturated channel. Moreover, entirely clipped pixels at edges cannot be corrected. Additionally, the study focuses on face images and a manual post-processing through an image manipulation software is required.

2.2 A Multiresolution Approach to Recovering Colors and Details of Clipped Image Regions

In (Lu, Chen and Lin, 2015), a multi-resolution approach is performed to restore the color and detail information in clipped regions based on (Masood, Zhu and Tappen, 2009). As a first step, an edge-preserving smoothing algorithm via edge-preserving filter proposed in (Farbman et al., 2008) is used to extract the low frequency components, while a subtraction process is used to obtain high frequency elements from the image to restore the original details. Afterwards, a desaturation method which is based on restoring the clipped pixel's colors from its neighbours with similar gradient is carried out on the low frequency image. Differently from (Masood, Zhu and Tappen, 2009), instead of intensity similarity, gradient similarity is used in the minimization of the cost function. The incorrect color estimation is prevented through preserving the cross-channel color ratios. According to the outcomes of the study, visually plausible images are obtained thanks to the reduction of the detail loss and color bias.

2.3 Recovering Color and Details of Clipped Image Regions

In (Elboher and Werman, 2010), an automatic variation and color information restoration algorithm for clipped pixels is introduced for raw, processed color images, and video sequences. The clipped regions are recovered via taking advantage from an approach, named as *color lines*, which is introduced in (Omer and Werman, 2004). In the color lines approach, two assumptions are combined; *(i)* in real world images the RGB histogram is mostly sparse and the non-empty bins usually have non-empty neighbour histograms, *(ii)* as illumination increases the norm of the RGB color coordinates in a real world color increases. As a result of these two assumptions the color distribution of an image is presented as a set of clusters called *color lines.* Since it is possible to assign a saturated pixel into the appropriate cluster and RGB channel values increase together with illumination, an information gathered from an non-

saturated part of a cluster enables us to correct a saturated pixel.

As a first step in the declipping procedure the saturated blobs containing more than 50 pixels are detected. Then, for the restoration of these clipped pixel values which have an intensity value higher than 230, each blob is assigned to its corresponding color line. Afterwards, each color cluster is declipped via making use of its unclipped part. Only pixels which lie close to the clipped area in the RGB histogram are considered to compute the ***R***:***G***:***B*** ratio and to approximate the true values of the clipped channel. Subsequently, the saturation point which distinguishes between the clipped and unclipped parts in a cluster is computed. It is acquired via thresholding a channel and extracting the pixels having a higher value than ~250. The corresponding values of these pixels in the remaining two channels are then sorted and 1% of them in each channel are taken as the saturation point's coordinates. The slope (θ) of the non-saturated regions around the saturation point is used alongside the saturation point to declip the clipped pixels as follows,

$$r_{declipped} = s_r + \rho(r) \cdot \theta \cdot \left\| g - s_g, b - s_b \right\| \quad (4)$$

where, $r_{declipped}$ is the declipped *R* channel of a pixel, g and b present the (*G*, *B*) coordinates of this pixel, s_r, s_g, s_b present the saturation points of the *R*, *G*, *B* channels respectively, ρ is 1 or 0 when r is greater than a determined upper value or smaller than a determined lower value of the slope, respectively.

When in a cluster, all pixels are clipped in a particular channel the original values of the pixels cannot be obtained, yet the clipped region can be enhanced by subtracting the corrected image from the original input and applying the acquired map to the regions where all pixels are clipped.

The proposed algorithm fails to declip pixels when color clusters are segmented imprecisely and clipped pixels are missed during their detection.

2.4 Correcting Over-Exposure in Photographs

In (Guo et al., 2010) lightness and color are considered individually. The input image is converted into CIELab color space to process the color and lightness separately. The color information is recovered through neighbourhood propagation and confidence of colors, while lightness is corrected from the likelihood of over exposure. This method is able to correct over-exposed images successfully even when all channels are clipped. However, when an image is contains largely clipped regions, boundaries present in the image might become imprecise. Additionally, color might be propagated across different objects during the declipping process, which causes pixels to be filled incorrectly.

2.5 Correction of Clipped Pixels in Color Images

In (Xu, Doutre and Nasiopoulos, 2010), operations for clipped regions correction is performed mainly in YCbCr color space instead of RGB color domain, since chroma has less spatial variation compared to RGB. In the proposed model, firstly areas containing pixels with the value of 252.5 are identified after the noise in the image is eliminated through a bilateral filter. Secondly, the clipped pixels are divided into regions based on their chroma values. Either ***Cb*** or ***Cr*** values are employed, i.e. the one having larger variance in the area of interest is chosen. Then, regions with similar chroma near the clipped area are identified. This surrounding region may contain clipped pixels or pixels with distinct chroma values compared to the clipped region. Due to these missing pixels, normalized convolution, which can operate with irregular shaped regions is adopted for interpolation as follows,

$$\widehat{CH}(r,c) = \frac{[CH(r,c) \cdot CM(r,c)] * h(r,c)}{CM(r,c) * h(r,c)} \tag{5}$$

where, $\widehat{CH}$ is the interpolated chroma channel, CH is the chroma channel, (r,c) is the location of the pixel of interest, h is a gaussian filter, and CM is the certainty map denoting the confidence level at a pixel. CM is set to 1 if the pixel in the surrounding region is non-saturated or declipped, and 0 if it is clipped.

To reduce the error, in the beginning the pixels clipped in only one channel are corrected and the declipping procedure is continued with pixels saturated in 2 and then 3 channels. In addition, firstly pixels closer to the surrounding region are corrected since they have strong correlation with their unclipped neighbours.

After the unclipped values of the clipped chroma channels are acquired, they are converted to RGB color space. Subsequently, the following 2 equations are used to compute the new ***R***, ***G***, ***B*** values of the pixels,

$$Cb = [-0.1482 - 0.2910\ 0.4392] \times [R\ G\ B]^T + 0.5020 \quad (6)$$

$$Cr = [0.4392 - 0.3678 - 0.0714] \times [R\ G\ B]^T + 0.5020 \quad (7)$$

When 2 channels are clipped, the non-saturated third channel is used and two equations with two unknowns are solved. Similarly, when only a single channel is clipped, two values for this channel are acquired from the equations above, and they are averaged to obtain the declipped value. For pixels containing 3 clipped channels, the Y value is estimated by fitting the clipped area and its surrounding region with a 2-dimensional gaussian function and then converting the pixel into RGB color domain. As a final step, smoothing is applied to avoid sharp changeovers between pixels in the declipped regions.

The proposed method produces satisfying results, however when the surrounding region consists of only saturated pixels the approach cannot carry out the declipping process.

2.6 Correction of the Overexposed Region in Digital Color Image

In (Lee et al., 2014), the over-exposed region correction is also performed in YCbCr domain. The detection of the over-exposed regions in the image is carried out in the ***Y*** channel in two steps; a simple thresholding approach with a value of 230 is chosen and isolated over-exposed regions containing less than 40 pixels are eliminated to obtain over-exposed blobs. Then, the lightness and chrominance of these blobs are corrected

to restore information in over-exposed regions in the image. The lightness of the blobs is corrected by minimizing the energy function (Eqn. 8) while preserving the gradients of the original lightness and retaining the lightness modelled by a two-dimensional Gaussian function (Eqn.9). Gaussian modelling is preferred, since it is observed that the largest lightness values are present at the centroid of the blob and they decrease as they become distant from the centroid.

$$\hat{Y}^* = \underset{\hat{Y}}{\operatorname{argmin}} \sum_{j \in I} \left[\left(\hat{Y}_j - Y_j^{GM} \right)^2 + \xi . \left\| \nabla \hat{Y}_j - \nabla Y_j \right\|_2^2 \right] \tag{8}$$

$$Y_j^{GM} = \begin{cases} Y_{cp} \cdot \exp\left(-\frac{d_{j,cp}^2}{2\sigma_j^2} \right), j \in T \\ Y_j \qquad\qquad\qquad\quad , j \in S \end{cases} \tag{9}$$

where, $\hat{Y}^*$ is the corrected lightness, Y_j is the lightness at pixel j, GM is denotes the Gaussian model, I is the image, $\nabla \hat{Y}_j$ and ∇Y_j represent the gradients of $\hat{Y}$ and Y at pixel j respectively, ξ is a weighting factor, S is the source (non over-exposed) region, T is the target (over-exposed) region, cp is the centroid of a blob, $d_{j,cp}$ denotes the Euclidean distance between j an cp, and σ_j^2 is the variance of j.

If $j \in T$ is on the straight line connecting a blob's boundary pixels and its centroid pixel, σ_j^2 is computed as follows,

$$\sigma_j^2 = -\frac{d_{j,cp}^2}{2 \ln\left(\frac{Y_i}{Y_{cp}} \right)} . \tag{10}$$

For the chrominance restoration, clipped regions are corrected from its non-clipped neighbours. In order to avoid color fade caused by over exposure in the neighbouring pixels, the image is stretched while the original lightness is preserved. Subsequently, the clipped chrominance values are estimated through the stretched image. The stretched values are computed as follows,

$$\widehat{SV}_j = \left[R_j\ G_j\ B_j\right]^T + \zeta_j \cdot \left[\left[R_j\ G_j\ B_j\right]^T - Y_j \cdot [1\ 1\ 1]^T\right] \tag{11}$$

$$\zeta_j = \left[0.5 - \frac{1}{2 + \exp\left(-7 \cdot \left(\frac{\tilde{\sigma}_j}{100} - 0.5\right)\right)}\right] \tag{12}$$

where, $\widehat{SV}_j$ is the stretched value at pixel j, ζ_j is the weighting factor, $\tilde{\sigma}_j$ is the standard deviation of the RGB channel intensities at pixel j.

After the stretched image is acquired, the saturated pixels are declipped by propagating the stretched values via a bilateral filter as follows,

$$\widehat{CH}_\alpha^* = \frac{\left(\sum_{\varphi \in \phi} W_{\alpha,\varphi} \cdot \widehat{CH}_\varphi\right)}{\sum_{\varphi \in \phi} W_{\alpha,\varphi}} \tag{13}$$

$$W_{\alpha,\varphi} = GM\left(\left\|CH_\alpha - CH_\varphi\right\|_2\right) \cdot (\hat{Y}_\alpha - \hat{Y}_\varphi) \tag{14}$$

where, $\widehat{CH}_\alpha^*$is the color corrected pixel value, $CH_\alpha = [Cb_\alpha\ Cr_\alpha]$, $\boldsymbol{Cb}_\alpha$ and $\boldsymbol{Cr}_\alpha$are the clipped pixel α's channels in YCbCr color space, $W_{\alpha,\varphi}$is the filter weight and φ is the neighbouring pixel, $\widehat{CH}_\varphi$=[$\widehat{Cb}_\varphi\ \widehat{Cr}_\varphi$], $\phi \in \{\Psi \cup S\}$, ϕ represents a set of pixels in a 3 x 3 mask, and Ψ are the already color corrected pixels.

The proposed method is compared with the state of art methods (Masood, Zhu and Tappen, 2009; Guo et al., 2010). According to the results, visually plausible images can be obtained with lower computational cost.

2.7 Correction of Saturated Regions in RGB Color Space

In (Ju and Park, 2016), before the detection process of the clipped regions of each channel, a noise elimination process via bilateral filtering is performed. The reason of

the usage of bilateral filtering is to eliminate an isolated clipped region or a hole caused by noise in the saturated region, while preserving the details on the edges. After noise elimination, saturated regions of each channel are detected via thresholding at the value of 235. Then, the color of the distorted pixels in the saturated regions are corrected using the weighted sum of each color value at non-saturated neighbouring pixels and the process is performed at each channel individually where the weights are calculated from the color (Col), saturation (Sat) and hue values (Eqn. 15, 16). During the declipping procedure, the clipped pixels closer to the unclipped pixels are processed primarily.

$$\hat{\alpha}^{R}{}_{(r,c)} = \frac{1}{\sum_{u,v\in\vartheta} W_{(r+u,c+v)}} \sum_{u,v\in\vartheta} W_{(r+u,c+v)} \alpha^{R}_{(r+u,c+v)} \tag{15}$$

$$W_{(r+u,c+v)} = W^{Col}_{(r+u,c+v)} + W^{Sat}_{(r+u,c+v)} + W^{Hue}_{(r+u,c+v)} \tag{16}$$

where, $\hat{\alpha}^{R}$ and α^{R} are the corrected and clipped values of pixel α in the ***R*** channel respectively (extensions to ***G*** and ***B*** channels are straightforward), $W_{(r+u,c+v)}$ is the weight derived from the neighbouring pixel in the window ϑ with a size of $(2o+1) \times (2o+1)$ centred at (r, c), and $W^{Col}_{(r+u,c+v)}$, $W^{Sat}_{(r+u,c+v)}$ and $W^{Hue}_{(r+u,c+v)}$ are the weights associated with certainty map and color, saturation, and hue respectively. $W^{Col}$is computed by making use of the Gaussian function, W^{Sat} is obtained from the ***R***, ***G***, ***B*** values and W^{Hue} is calculated through the inverse cosine function. The study produced better statistical results than (Masood, Zhu and Tappen, 2009), yet the saturated regions on corrected face images are still visible.

While aforementioned studies are based on traditional methods, neural networks have recently been adopted for saturation correction (Honing and Werman,2018; Yang et al., 2018; An et al., 2018) since they have proven their potential in many image processing tasks (Karakaya, Ulucan and Türkan, 2019; Ahishali et al., 2019; Degerli et al.; 2018).

2.8 Image Declipping with Deep Networks

In (Honig and Werman, 2018) both over (threshold at 175) and under (threshold at 80) exposed regions are corrected via DeclipNet, which is inspired by generative adversarial networks. DeclipNet is an image transformation neural network, which takes the image with clipped pixels and outputs the corrected image. In the discriminator model there are 14 convolutional and transposed convolutional layers, whereas the generator model contains 10 convolutional layers. Furthermore, MSE, perceptual (PER) and adversarial (ADV) losses are introduced to the network and balanced like the following,

$$0.81MSE + 0.095PER + 0.095ADV. \quad (17)$$

During the experiments two different datasets are adopted; MSCOCO (Lin et al., 2014) for training and testing, and imageNet for validation. Artificially clipped images are used for training, while the original images are employed as ground truth. Moreover, the image size used in the network is determined as 224 x 224.

In order to evaluate the outcomes of DeclipNet, subjective tests with 18 participants are carried out. 20 over-exposed images reconstructed by DeclipNet and the algorithm introduced in (Elboher and Werman, 2010), together with the ground truth images are shown to the participants. 82.5% of the time the DeclipNet output is preferred by the participants. It is observed that DeclipNet successfully restored color and detail information even if all color channels are saturated.

2.9 Image Correction via Deep Reciprocating HDR Transformation

In (Yang et al., 2018) a model based on a convolutional encoder-decoder system is designed to recover HDR details in LDR and/or tone-mapped images. The proposed model takes advantage from two different networks to correct images, namely HDR estimation network and LDR correction network. The HDR estimation network first reconstructs the HDR content and restores all HDR details in the image, and then the LDR correction network transfers these restored HDR details into the LDR domain. The designed end-to-end model shows promising results when evaluated with the two

datasets employed in the study. On average the network reaches a structural similarity index (SSIM) score of 0.844. On the other hand, the model presents less favourable outcomes when large over-exposed regions are present in the input image.

2.10 Deep Tone-mapped HDRNET for High Dynamic Range Image Restoration

In (An et al., 2018) a super-resolution network (Yamanaka, Kuwashima and Kurita) based on convolutional neural networks is modified to restore the lost details in highlights and shadows via learning the relationship between LDR images and tone-mapped HDR images. The model is trained by a patch-based scheme using normal-exposed images in the sequence and tone-mapped HDR images via Reinhard tone-mapper (Reinhard et al., 2002). The proposed network successfully restores lost details in saturated regions, but it requires two images for the declipping process.

Aside from mentioned declipping approaches above, there are a limited number of studies which focus on the saturation correction problem of tone-mapped images (Pouli et al., 2013; Smith et al., 2006).

2.11 Color Correction for Tone Reproduction

In (Pouli et al., 2013), it is aimed to preserve the luminance of the tone-mapped image while matching its hue and saturation values with the original HDR input. In order to correct hue, images are converted into the CIE L*C*h color space where hue has a fine uniformity. The hue values of the tone-mapped image are directly set to the hue values of the original image. For the saturation correction stage, a saturation ratio between the original and tone-mapped images is obtained by means of chroma and lightness information. Experiments illustrate that the proposed method produces promising tone reproduction results.

2.12 Beyond Tone Mapping: Enhanced Depiction of Tone Mapped HDR Images

In (Smith et al., 2006), an objective perceptual metric, which measures the contrast distortion between the tone-mapped and HDR images, is proposed to avoid the

physical contrast-loss occurring during tone-mapping. Guided by the perceptual metric, chroma scaling is used to enhance contrast, thus makes details more visible, and counter-shading is employed to emphasize the foreground objects. The proposed method presents improved contrast information in tone-mapped images with satisfying results.

CHAPTER 3: PROPOSED IMAGE DECLIPPING APPROACH

This thesis proposes a declipping algorithm based on saturation correction in single images. This is indeed a challenging task because the correction of saturated pixels is performed by means of the self-information extracted from the input image. As mentioned previously, pixels with different number of clipped channels are corrected via distinct pixel- and block-based approaches. Specifically, pixels saturated in one or two channels are recovered through pixel-based methods and pixels clipped in all three channels are restored via a block-based approach. The developed algorithm iterates over three sequential passes in order to restore saturated pixels, namely one-channel correction as the first iteration, two-channel correction as the second iteration, and three-channel correction as the third iteration.

The remaining parts of this chapter explains the distinct correction methods in detail including the theory behind LE, morphological operations, scaling and sliding window method required for the one-channel correction process, the difference of pixels method used for two jointly clipped channels, the adopted gradient-based priority function for the block-search approach in three-channel declipping and the employed parameters for bilateral filtering in the post-processing step.

3.1 One Channel Correction via Linear Embeddings

In images individual pixels or patches can be reconstructed as a linear combination of their neighbours (Türkan, Thoreau and Guillotel, 2013, 2014). This reconstruction of pixels or small regions from its neighbours is generally defined as *neighbour embedding* and it is inspired from manifold learning algorithms (Roweis and Saul, 2000; Tenenbaum, De Silva and Langford, 2000; Donoho and Grimes, 2003). In this thesis, information related to a pixel clipped only in one channel is extracted using its non-clipped neighbours. The proposed approach is taking advantage of LE, which are based on the locally linear embeddings (LLE) algorithm proposed in (Roweis and Saul, 2000). LLE preserves the structure of a manifold and it is mainly used to map $\mathcal{Y}$-dimensional data into y-dimensional space, where $y < \mathcal{Y}$. High dimensional data consisting of N real-valued column vectors v_f is taken and the l nearest neighbours

$\mathcal{v}_{f_\tau}$ are identified through the Euclidean distance, where $\tau = 1, \ldots, l$, and $f \in \mathcal{Y}$. Then, in order to approximate each data point from its neighbours $\mathcal{v}_{f_\tau}$ while minimizing the reconstruction error given in Eqn. 18, the optimal weights can be computed via solving the least-squares problem given in Eqn. 19.

$$E(W) = \sum_{f} \left\| \mathcal{v}_f - \sum_{\tau} W_{f\tau} \mathcal{v}_{f_\tau} \right\|_2^2 \tag{18}$$

$$E_f(W_{f\tau}) = \left\| \sum_{\tau} W_{f\tau} (\mathcal{v}_f - \mathcal{v}_{f_\tau}) \right\|_2^2 \quad subject\ to \sum_{\tau} W_{f\tau} = 1, \forall \tau. \tag{19}$$

It is important to emphasize that, the sum-to-one constraint is required to ensure invariance hence preserving the intrinsic geometric properties in each neighbourhood, while acquiring the optimal weights. The optimum weighting coefficients are obtained as follows,

$$w_f^T = \frac{\left[\hat{V}_f^T \hat{V}_f\right]^{-1} \mathbf{1}}{\mathbf{1}^{\mathrm{T}} \left[\hat{V}_f^T \hat{V}_f\right]^{-1} \mathbf{1}} \tag{20}$$

where, $\hat{V}_f = \left[\mathcal{v}_{f1} - \mathcal{v}_f | \ldots | \mathcal{v}_l - \mathcal{v}_f\right]$ and $w_f = [W_{f1} | \ldots | W_{fl}]$.

In the pixels with only one channel clipped the fundamental LLE approach explained above is adopted for saturation correction. Given the input image ***I***, a specific binary mask[1] for each RGB channel is extracted to indicate pixels containing a value greater than 235 as suggested in (Masood, Zhu and Tappen, 2009). Then, each one-channel clipped pixel is processed through LE using its non-saturated neighbouring pixels, to estimate its original value. Let us denote three channels present in ***I*** as ***R***, ***G*** and ***B***. To calculate the saturated pixel value, the LE weight calculation depends on two neighbour assumptions: *(i)* every pixel in any channel has channel-wise neighbours at the co-located pixel locations in the others, *(ii)* in each channel, every saturated pixel

[1] The binary mask is overlapped to the RGB image and is illustrated in Fig. 1.b.

α has two selected non-saturated pixel-wise neighbours, $\mathcal{N}_1$ and $\mathcal{N}_2$. These spatial neighbours are identified by the local gradient direction, which is obtained through the gradient map of $\boldsymbol{I}$. The positions of these non-clipped neighbours according to the gradient angle are presented in Table 1 and all angles are rounded to the nearest integer while processing.

While extensions to $\boldsymbol{G}$ and $\boldsymbol{B}$ are straightforward, assume now that the saturated pixel α is in $\boldsymbol{R}$. To obtain the corrected value of $\boldsymbol{R}_\alpha$, linear reconstruction weights w_α^G and w_α^B of non-saturated pixel values $\boldsymbol{G}_\alpha$and $\boldsymbol{B}_\alpha$, respectively, are calculated by solving a constrained optimization given in Eqn. 21 as follows

$$\{w_\alpha^G, w_\alpha^B\} = \underset{\{w_1,w_2\}}{\operatorname{argmin}} \left\| \begin{bmatrix} R_{\mathcal{N}_1} \\ R_{\mathcal{N}_2} \end{bmatrix} - \begin{bmatrix} G_{\mathcal{N}_1} & B_{\mathcal{N}_1} \\ G_{\mathcal{N}_2} & B_{\mathcal{N}_2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \right\|_2^2 \quad s.t. w_1 + w_2 = 1 \tag{21}$$

where the optimum reconstruction weights w_1 and w_2 for $\boldsymbol{R}$ are calculated from the neighbouring pixels $\mathcal{N}_1$ and $\mathcal{N}_2$ from $\boldsymbol{G}$ and $\boldsymbol{B}$ in the local gradient direction. It is important to emphasize that there exists a sum-to-one constraint in order to guarantee translation and scale invariance. This constrained optimization can be directly solved by means of the inner product matrix (Roweis and Saul, 2000) (Eqn. 20). The corrected value of $\boldsymbol{R}_\alpha$ can then be calculated in Eqn. 22 as follows,

$$R_\alpha = w_\alpha^G \cdot G_\alpha + w_\alpha^B \cdot B_\alpha \tag{22}$$

where $G_\alpha$and B_α are non-saturated pixel values in the spatial location of α.

Table 1. Gradient Angles in Degrees and the Corresponding Positions of Neighbours in I for α at (r,c) are demonstrated.

Angle (degrees)	$\mathcal{N}_1$	$\mathcal{N}_2$
[0, 45] or [181, 225]	$(r-1, c+1)$	$(r+1, c-1)$
[46, 90] or [226, 270]	$(r-1, c)$	$(r+1, c)$
[91, 135] or [271, 315]	$(r-1, c-1)$	$(r+1, c+1)$
[136, 180] or [316, 360]	$(r, c-1)$	$(r, c+1)$

Due to the sensor capacity or the display range, the saturated pixel intensities are lower than their original values. The LE approach enables to estimate the true value of a clipped pixel through linear embeddings of the remaining two channels in its non-saturated neighbours. After all one-channel clipped pixel values are obtained, a new image $\boldsymbol{I'}$ is formed ranging between the minimum value of $\boldsymbol{I}$ and the highest true value of a clipped pixel. This image has to be appropriately scaled into the range between 0 and 255 in order to successfully maintain the 8-bit representation. A simple linear scaling is envisaged that maps all available pixels (x) in $\boldsymbol{I'}$ into a new scaled image $\boldsymbol{I}_s$ (Fig.1.c) via Eqn. 23 as follows,

$$x' = \frac{x}{\max(I') - \min(I')} \tag{23}$$

where $x'$is the scaled value of pixel x. Fig.1 demonstrates an example input image $\boldsymbol{I}$ with one-channel clipped pixels and the scaled image $\boldsymbol{I}_s$.

Because of the above scaling operator, both saturated and non-saturated pixels change their intensities, which has to be corrected subsequently. First of all, a simple binary mask was considered to transmit the declipped values to $\boldsymbol{I}$, however the scale has changed, and color distortions cannot be avoided in sharp changeovers. Hence, the transmission is carried out by extracting a gray-scale intensity map G of $\boldsymbol{I}$. The map is acquired by applying morphological operations to the gray-scale version of $\boldsymbol{I}$. To eliminate redundant regions, an erosion[1] operation is performed by adopting a disk-shaped structural element of size 25. Then, morphological reconstruction is conducted to recover the details and followed by a dilation[2] operation with the same structural element to expand the objects and fill their boundaries. Lastly, the complement of the reconstructed image and complement of the dilated image are used for a second morphological reconstruction. The complement of the output of this process is considered as G of $\boldsymbol{I}$, where regions containing declipped pixels have values close to 255 as can be seen in Fig.2.a.

[1] Erosion can be defined as, $I \ominus \mathcal{S} = \bigcap_{s \in \mathcal{S}} I_{-\mathcal{S}}$, where $\mathcal{S}$ is the structuring element representing the subset of pixels within the input window that we would like to consider, $\mathcal{S} = \{ s \mid s \in \mathcal{S} \}$.
[2] Dilation can be defined as, $I \oplus \mathcal{S} = \bigcup_{s \in \mathcal{S}} I_{\mathcal{S}}$.

a) $\boldsymbol{I}$ b) Clipped pixels c) $\boldsymbol{I}_s$

Figure 1. One channel correction: (a) Input $\boldsymbol{I}$ (Source: Kodak Image Set, 1991); (b) The red region demonstrates the pixels only clipped in the red channel, similarly green and blue areas present the pixels only clipped in green and blue channels, respectively; (c) Corrected and scaled image $\boldsymbol{I}_s$.

a) G b) $\boldsymbol{I}_1$

Figure 2. (a) Gray-Scale Intensity Map G of $\boldsymbol{I}$ and (b) One Channels Corrected Output Image $\boldsymbol{I}_1$.

The final step is to blend $\boldsymbol{I}$ and $\boldsymbol{I}_s$ into one image through G. In order to avoid noise, sharp changeovers and artefacts, $\boldsymbol{I}$ and $\boldsymbol{I}_s$ are first combined with a simple *sliding window* based averaging operation. In contrary to a pixel-based averaging, sliding window averaging operates on 5×5 pixels of neighbourhoods over both images and overlaps between adjacent neighbourhoods are also uniformly averaged, hence forced to be coherent. Since $\boldsymbol{I}$ and $\boldsymbol{I}_s$ contain very similar textural information at each pixel, no significant color change can be observed after this procedure. Subsequently, the intensity map G is employed to linearly blend, $\boldsymbol{I}$ and $\boldsymbol{I}_{SW}$ into the one-channel corrected output image $\boldsymbol{I}_1$ as given in Eqn. 24.

$$I_1 = G \cdot I_{SW} + (1 - G) \cdot I \tag{24}$$

$\boldsymbol{I}_1$ may contain slight color changes in the non-saturated regions due to the color transmission process (Fig. 2.b). These slight changes indeed lead to well-settled colors in the image. (The algorithm for one-channel correction is summarized in Appendix B: Algorithm 1.)

3.2 Two Channel Correction via Difference of Pixels

After having individual $\boldsymbol{R}$, $\boldsymbol{G}$ and $\boldsymbol{B}$ channels corrected through linear embeddings, pixels containing two clipped channels are considered for saturation correction. First of all, saturation masks for three different combinations of channels $\{\boldsymbol{R}, \boldsymbol{G}\}$, $\{\boldsymbol{R}, \boldsymbol{B}\}$ and $\{\boldsymbol{G}, \boldsymbol{B}\}$ are extracted (Fig. 4.a) and then connected components in these masks are processed individually to carry out a more sensitive pixel correction process.

In order to restore two-channel saturated pixels efficiently, a pixel with the highest priority is determined at a time. Hence, the pixel which has the highest confidence and is located at strong edges or structures is processed primarily in a connected component. The region containing saturated pixels is defined as the target area $\boldsymbol{\mathcal{T}}$ while the remaining part of the image is considered as the source region $\boldsymbol{\mathcal{S}}$. The border between these regions is called the contour β as depicted in Fig.3. The priority of a pixel $x \in \beta$ is calculated by the product of its *confidence* $C(x)$ and *data* $D(x)$ terms (Criminisi, Pérez and Toyama, 2004) given in Eqn. 25 as follows,

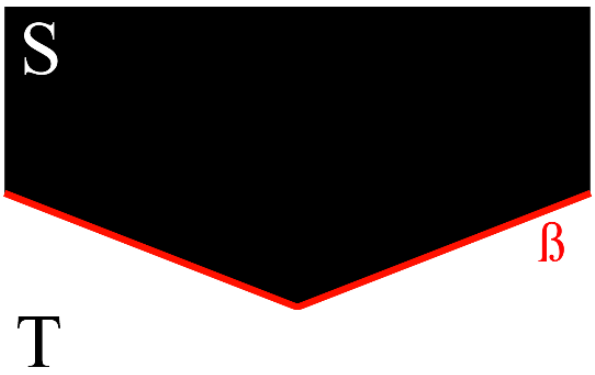


Figure 3. Priority Computation: $\boldsymbol{\mathcal{T}}$ is the Target Area, $\boldsymbol{\mathcal{S}}$ is the Source Region and β is the Contour.

$$C(x) = \frac{\sum_{\gamma \epsilon P_x \cap S} C_\gamma}{|P_x|} \qquad D(x) = \frac{|\nabla I_x^{\perp} \cdot \omega_x|}{255} \tag{25}$$

where $\boldsymbol{P}_x$ is a patch of size 5×5 pixels centred around *x*, $|P_x|$ denotes the number pixels in $\boldsymbol{P}_x$ and $\boldsymbol{\omega}_x$is the unit vector orthogonal to β at the point *x*.

The priority is computed for every non-saturated pixel on the contour per connected component, and the pixel with highest priority is used to determine the spatially closest neighbouring pixel which is clipped. The determined target pixel is then corrected and included into the source region to compute a new contour. This process is repeated until $\boldsymbol{\mathcal{T}}$ entirely becomes a part of $\boldsymbol{S}$.

The two-channel saturation correction procedure is based on intensity value differences of neighbouring pixels. It is assumed that non-saturated neighbours of the clipped pixel α possess similar intensity values with the pixel of interest. To avoid incorrect estimations at sharp color changeovers, the gradient map of $\boldsymbol{I}_1$ is extracted and only two neighbours $\mathcal{N}_1$ and $\mathcal{N}_2$ of α in the gradient direction are considered for declipping, similar to the one-channel correction process (Table 1).

While extensions {$\boldsymbol{R}$, $\boldsymbol{G}$} and {$\boldsymbol{G}$, $\boldsymbol{B}$} are straightforward, assume now that the saturated pixel α is in {$\boldsymbol{R}$, $\boldsymbol{B}$}. Hence, the **G** component of α ($\boldsymbol{G}_\alpha$) is not saturated and can be used for correction of $\boldsymbol{R}_\alpha$ and $\boldsymbol{B}_\alpha$. In simple terms, two respective pixel difference values at $\mathcal{N}_1$ and $\mathcal{N}_2$ are averaged to obtain a cross-channel offset value, which is added to $\boldsymbol{G}_\alpha$ as given in Eqn. 26.

$$\begin{aligned} R_\alpha &= G_\alpha + \frac{\left(R_{\mathcal{N}_1} - G_{\mathcal{N}_1}\right) + \left(R_{\mathcal{N}_2} - G_{\mathcal{N}_2}\right)}{2} \\ B_\alpha &= G_\alpha + \frac{\left(B_{\mathcal{N}_1} - G_{\mathcal{N}_1}\right) + \left(B_{\mathcal{N}_2} - G_{\mathcal{N}_2}\right)}{2} \end{aligned} \tag{26}$$

Note here that neighbouring pixels are assumed to be both non-saturated. In a special case when one of these neighbours is also clipped, the difference is considered only at

the non-saturated neighbour. In an extreme condition when both of these neighbours are clipped, then α remains as it is, and it will be dealt with the post-processing stage as explained in Chapter 3.4. In addition, there is no any further consideration when the difference values at $\mathcal{N}_1$ and $\mathcal{N}_2$ are both positive or both negative. However, in the case when they have different signs, the offset is calculated with the absolute difference values and its sign is determined by the largest difference magnitude.

The saturation correction process of $\{\boldsymbol{R}, \boldsymbol{G}\}$, $\{\boldsymbol{R}, \boldsymbol{B}\}$ and $\{\boldsymbol{G}, \boldsymbol{B}\}$ is repeated until $\boldsymbol{\mathcal{T}}$ entirely becomes a part of $\boldsymbol{\mathcal{S}}$. The two-channel corrected output image $\boldsymbol{I}_2$ is then used as an input to the three-channel correction stage. Fig.4 illustrates the example input image $\boldsymbol{I}_1$ with two-channel clipped pixels and the output image $\boldsymbol{I}_2$. (The algorithm for the two-channel correction procedure is provided in Appendix B: Algorithm 2.)

3.3 Three Channels Correction via Block-Search

Three-channel saturation correction is the most challenging stage since there is no information present in α. Therefore, a block-search approach is adopted, which is partly based on exemplar-based inpainting (Criminisi, Pérez and Toyama, 2004). First of all, a mask indicating pixel locations of three-channel saturation is extracted from $\boldsymbol{I}_2$ and then connected components in this mask are identified and processed individually. Subsequently, a search-region is formed for each connected component. This region is defined as an area enclosing the respective connected component, where the corrected value of α is searched. In this study, the surrounding area with an 11-pixel distance from the borders of the connected component (in $\pm x$ and $\pm y$ directions) is used as the search-region (Fig. 5).

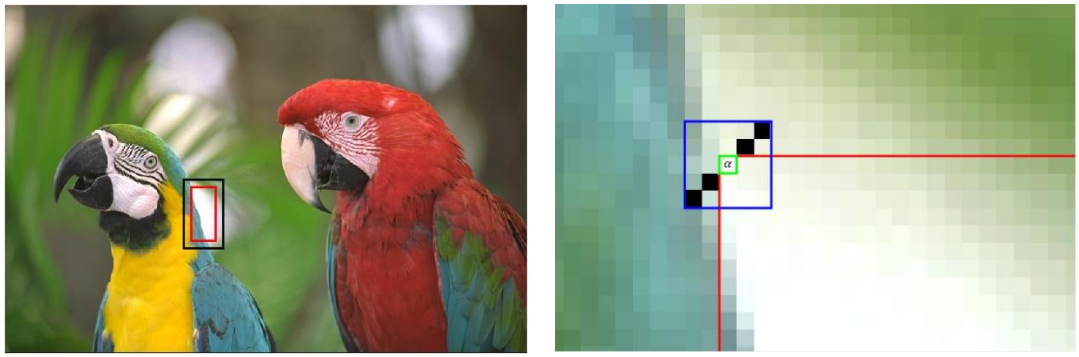
In order to obtain the declipped value of each α in the connected component, a 5×5 patch centred around α is extracted and α is with the highest priority according to Eqn. 25. The 4 neighbouring pixels are determined from the local gradient direction in the patch. These pixels will provide a simple ``directed template" for the template matching algorithm which is run over the search-region. It is important to note here that possible saturated pixels other than α are not taken into account during

computations. Furthermore, the closest 8 spatial neighbours of α are also replaced with the obtained value to avoid sharp color changeovers and probable artefacts. Fig. 6 exemplifies the input image $\boldsymbol{I}_2$ with three-channel clipped pixels and the output image $\boldsymbol{I}_3$. (The algorithm for three channel correction is provided in Appendix B: Algorithm 3.)

a) $\boldsymbol{I}_1$ with two-channel clipped pixels

b) $\boldsymbol{I}_2$

Figure 4. Two-Channel Correction: (a) {$\boldsymbol{R}$, $\boldsymbol{G}$}, {$\boldsymbol{R}$, $\boldsymbol{B}$} and {$\boldsymbol{G}$, $\boldsymbol{B}$} Jointly Clipped Pixels are represented in Cyan, Magenta and Dark Blue, Respectively; (b) Two-Channel Corrected Output Image $\boldsymbol{I}_2$.



a) Search-Region

b) Block of Interest

Figure 5. The Search Region and Block of Interest are demonstrated. (a) The Red and Black, Rectangles represent the Connected Component and Search-Region Respectively (Source: Kodak Image Set, 1991); (b) Green and Blue Rectangles illustrate α and 5×5 Block of Interest Respectively, while the Black filled Blocks demonstrate the Neighbours in the Gradient Direction (i.e. for an Angle in the Ranges of [0, 45] or [181, 225]).

3.4 Post-Processing

A post-processing step is applied after obtaining $\boldsymbol{I}_3$ to smoothen possible artefacts in pixels which could not be corrected during two-channel correction, to make local image regions more uniform after three-channel correction and to output more natural looking images. In this thesis, bilateral filters (Tomasi and Manduchi, 1998) are preferred for post-processing and several steps are performed to obtain the optimal parameters. To process perceptually similar colors and preserve perceptually significant edges, $\boldsymbol{I}_3$ is converted into CIELab color domain, named as $\boldsymbol{I}_L$. Then, all saturation masks obtained through two- and three-channel correction steps are combined into a new mask $\boldsymbol{M}$ in order to determine all processed pixel locations to be post-processed. The procedure is only applied to these pixels of interest in $\boldsymbol{I}_R$ with their CIELab values, while the rest of the image is blacked, i.e., $\boldsymbol{I}_R = \boldsymbol{M}(\boldsymbol{I}_L)$. $\boldsymbol{I}_R^B$ is obtained after $\boldsymbol{I}_R$ is processed with the bilateral filter and the post-processed pixel values are replaced in $\boldsymbol{I}_L$ to obtain $\boldsymbol{I}_O$ such that $\boldsymbol{I}_O = \boldsymbol{M}(\boldsymbol{I}_R^B) + \neg\, \boldsymbol{M}(\boldsymbol{I}_L)$. Lastly, $\boldsymbol{I}_O$ is converted back into the RGB color space.

The variance of $\boldsymbol{I}_R$ is used as the degree of smoothing for the bilateral filter. The filter has a neighborhood size of 101 and a standard deviation σ of 10 for the spatial Gaussian smoothing kernel. All these parameters are determined via experimental observations and they are fixed throughout experiments.

a) $\boldsymbol{I}_2$ with three-channel clipped pixels

b) $\boldsymbol{I}_3$

Figure 6. Three-Channel Correction: (a) Pixels clipped in all Channels are represented in Magenta and (b) Three-Channel corrected Output Image $\boldsymbol{I}_3$.

a) $\boldsymbol{I}_1$ with two- and three-channel clipped pixels

b) $\boldsymbol{I}_O$

Figure 7. Post-Processing: (a) Pixels clipped in Two- and Three-Channels are represented in Magenta and (b) the Output Image $\boldsymbol{I}_O$ in RGB.

Fig.7 demonstrates the example input image with two- and three-channel clipped pixels to be post-processed and the output image $\boldsymbol{I}_O$ in RGB. (The algorithm for post-processing is summarized in Appendix B: Algorithm 4.)

3.5 Ineffective Approaches Replaced during Algorithm Development

During the development of the declipping algorithm in this thesis several approaches are utilized to obtain a successful method. Nevertheless, before finalizing the algorithm, few approaches which were employed, failed to solve the clipping problem. For instance, in the one channel correction process, the one channel clipped pixel with the highest priority was identified and its channel intensity was replaced by the average of its non-clipped pixel-wise neighbours' corresponding channel. It was considered that; the neighbours of the clipped pixels would have a similar intensity value in the channel of interest. However, since the cross-channel correlation could not be preserved, color distortion especially at edges where sharp color changeovers occur was observed. Therefore, this approach was replaced with the method in Chapter 3.1. Furthermore, in the two channels correction procedure, $\mathcal{N}_1$ and $\mathcal{N}_2$ were identified and the saturated pixel was recovered through the color ratios in $\mathcal{N}_1$ and $\mathcal{N}_2$ according to the non-clipped channel. This method failed to successfully declip the pixels because when the ratio resulted in a large value, the declipped channel value exceeded 255 and remained clipped. Also, after one of the two saturated channels in a pixel was

declipped with the method proposed in Chapter 3.2 and this process was carried out for the entire image, the remaining clipped channel in the pixels was recovered with the method proposed in Chapter 3.1. Due to the same reason mentioned for the failed method in the one channel correction procedure, this technique was ineffective in solving the clipping problem. Moreover, in the three channels correction process, firstly a block of the size 3×3 was chosen and α was placed into one of the corners of the block so that the remaining of the block could provide the maximum number of non-clipped neighbours. The remaining of the process was carried out as explained in Chapter 3.3 and it was observed that a 3×3 block could not provide sufficient information and placing α at corners caused sharp changeovers at edges. Therefore, this approach was substituted with the method given in Chapter 3.3.

CHAPTER 4: EXPERIMENTAL RESULTS

In contrast to several related works mentioned in Chapter 2, the designed algorithm can successfully work with different types of images in terms of context and format. The dataset employed in this study consists of three different format types: *(i)* tone-mapped HDR images, *(ii)* LDR images and *(iii)* raw LDR images. The proposed method produces very successful statistical and visual outputs for all categories. In addition, it is a general framework for different types of image content such as portraits, plants and cars; whereas several existing techniques focus on a particular context, i.e., face images.

All experiments are carried out on an Intel(R) Core (TM) i7-8500U CPU @1.80 GHz 4-Core 8GB RAM machine using MATLAB R2019b. All tone-mapped HDR images are courtesy of Erik Reinhard and Tania Pouli. In addition, raw LDR images belong to the Kodak Image Dataset (Kodak Image Set, 1991), and (processed) LDR images are acquired from the webpage of Guo and the Berkeley Segmentation Dataset (Martin et al., 2001). Table 2 provides a detailed summary of the whole dataset, including the resolution of each image and the number of one-, two- and three-channel saturated pixels. In this table, the first nine images are obtained from HDR tone-mapping, the next four images are LDR and the last six are raw LDR images. It is important to note here that the dataset images vary in spatial resolution in order to observe if the resolution has an impact on the outcomes. It turns out to be that the proposed algorithm is able to successfully correct saturated images with both high and low resolutions.

Both visual and statistical analyses are conducted and reported in order to provide a comprehensive comparison and present the effectiveness of the proposed algorithm. The method is extensively compared with Xu (Xu, Doutre and Nasiopoulos, 2010), Masood (Masood, Zhu and Tappen, 2009), Guo (Guo et al., 2010), and Lu (Lu, Chen and Lin, 2015). The source code of Masood is obtained from the website of the author (Masood Algorithm, 2009) and it is adopted with its default settings. In this algorithm, a manual post-processing is applied, hence it is subjective and may lead to different outcomes. The output images of Guo are reached from the webpage of the author (Guo

Algorithm, 2010). For Xu and Lu, the statistical results are obtained from their papers. As a final note, the proposed algorithm is in its experimental form, hence the code is applied without any optimization.

Three well-known statistical quality assessment metrics: *(i)* the mean-squared-error (MSE), *(ii)* the peak-signal-to-noise ratio (PSNR) in decibels and *(iii)* the structural similarity index (SSIM) are employed alongside visual analysis results. Firstly, MSE is computed, which takes the squared differences between the reference and output images at each pixel into account (Eqn. 27) (Hore and Ziou, 2010). The lower the MSE the better the outcome.

$$MSE = \frac{1}{rs \cdot cs} \sum_{r=1}^{rs} \sum_{c=1}^{cs} (I_{O_{r,c}} - I_{r,c})^2 \tag{27}$$

where, *rs* and *cs* are the row and column sizes respectively.

Then, PSNR is obtained, which provides a quality measurement in decibel scale and it is based on MSE (Eqn. 28) (Hore and Ziou, 2010). A higher PSNR indicates a superior result.

$$PSNR = 10 \log_{10} \left(\frac{(L-1)^2}{MSE} \right) \tag{28}$$

where L is the maximum possible intensity number.

Lastly, SSIM is calculated, which measures the perceptual difference between two images and provides a score between 0 and 1, where scores closer to 1 refer to a better result (Hore and Ziou, 2010). SSIM can be computed as follows from the images' luminance (*lum*), contrast (*con*) and the structure (*str*) components,

$$SSIM = lum(I_O, I) \cdot con(I_O, I) \cdot str(I_O, I) \tag{29}$$

$$lum(I_O, I) = \frac{2\mu_{I_O}\mu_I + ct_1}{\mu_{I_O}^2 + \mu_I^2 + ct_1} \tag{30}$$

$$con(I_O, I) = \frac{2\sigma_{I_O}\sigma_I + ct_2}{\sigma_{I_O}^2 + \sigma_I^2 + ct_2} \tag{31}$$

$$str(I_O, I) = \frac{\sigma_{I_O I} + ct_3}{\sigma_{I_O}\sigma_I + ct_3} \tag{32}$$

where, μ represents the mean of the corresponding images, and σ_{I_O} and σ_I are the standard deviation of I_O and I respectively. $\sigma_{I_O I}$ is the covariance of the two images and ct_1, ct_2, ct_3 are constants added to avoid division by zero.

Table 2. Detailed Information About the Dataset.

Input Image Details		Clipped Pixel Number		
Name	Resolution	One Channel	Two Channels	Three Channels
Dino	319 x 480	5525	8120	140
Sink	384 x 576	15808	1650	66
Exploratorium	781 x 1286	120141	5287	516
Watering Can	304 x 424	3527	29	31
Butterflies	576 x 383	8410	793	11
Portraits	1080 x 1920	58796	9	0
Garden	356 x 366	3720	350	5189
Chapel	858 x 570	12976	0	0
Cars	1272 x 1917	49304	58082	846
Flowers	362 x 500	12617	1327	0
Dancer	549 x 484	64799	2764	2
House	375 x 500	2688	2238	2249
Sea Plant	450 x 600	22419	10	0
Kodim03	512 x 768	5686	1884	6
Kodim08	512 x 768	26311	962	32
Kodim17	768 x 512	618	516	47
Kodim19	768 x 512	16643	1978	2
Kodim23	512 x 768	19940	1831	2252
Kodim25	436 x 720	849	738	1283

Before carrying out a comparison with the state-of-the-art methods, 17 images from the Kodak Image Dataset are used to conduct an experiment. The images are thresholded at 200 and it is assumed that the pixel values between 200 and 255 are non-saturated. The proposed algorithm is run under this condition and it is observed that, the designed algorithm can estimate the intensity values of the pixels successfully (Table 3).

Some statistical results are provided in Table 4 and Table 5 in which the algorithm is compared with Masood and Guo, respectively. The proposed method presents mostly superior results in terms of PSNR and SSIM values, while noting that these metrics naturally may not coincide. As additional information, the statistical scores of the remaining images are provided in Table 6.

Table 3: Scores of the Experiment with Threshold of 200.

	MSE	PSNR (dB)	SSIM
Kodim01	1,063	47,864	0,999
Kodim02	2,096	44,917	0,999
Kodim03	14,602	36,487	0,993
Kodim04	3,688	42,462	0,998
Kodim05	8,687	38,742	0,997
Kodim07	4,651	41,455	0,993
Kodim09	3,166	43,126	0,995
Kodim10	15,621	36,194	0,979
Kodim11	4,215	41,883	0,990
Kodim14	8,686	38,743	0,996
Kodim16	5,105	41,051	0,997
Kodim17	4,109	41,993	0,991
Kodim18	5,022	41,122	0,998
Kodim21	7,903	39,153	0,995
Kodim22	30,730	33,255	0,955
Kodim23	22,712	34,568	0,977
Kodim25	16,452	35,969	0,988
Average	9,324	39,940	0,991

The remaining part of this Chapter is partitioned into three subsections according to the image format, where visual results and detailed experimental interpretations are provided. Whereas images on which the proposed algorithm presents less visually appealing results than desired are provided in Appendix C.

Table 4. The Comprehensive Score Comparison with Masood.

	Masood			Proposed		
	MSE	PSNR (dB)	SSIM	MSE	PSNR (dB)	SSIM
Dino	116,144	27,481	0,979	13,491	36,830	0,995
Sink	84,632	28,856	0,979	36,632	32,492	0,995
Exploratorium	143,944	26,549	0,963	67,651	29,820	0,980
Watering Can	57,929	30,502	0,985	75,261	29,365	0,974
Butterflies	335,619	22,872	0,961	40,186	32,070	0,972
Flowers	133,566	26,874	0,981	55,792	30,631	0,986
Dancer	339,887	22,818	0,843	302,617	23,332	0,861
House	203,570	25,044	0,975	4,467	41,630	0,998
Sea Plant	440,141	21,695	0,837	353,147	22,651	0,964
Kodim17	27,940	33,669	0,987	2,330	44,458	0,995
Kodim19	217,444	24,757	0,978	17,199	35,776	0,970
Average	139,652	27,695	0,977	38,568	33,930	0,983

Table 5. Statistical Score Comparison with Guo.

	Guo			Proposed		
	MSE	PSNR (dB)	SSIM	MSE	PSNR (dB)	SSIM
Dancer	125,681	27,159	0,936	302,617	23,332	0,861
House	61,080	30,272	0,978	4,467	41,630	0,998
Sea Plant	605,782	20,378	0,924	353,147	22,651	0,964
Average	264,181	25,936	0,946	220,077	29,204	0,941

Table 6.Scores of the Proposed Method.

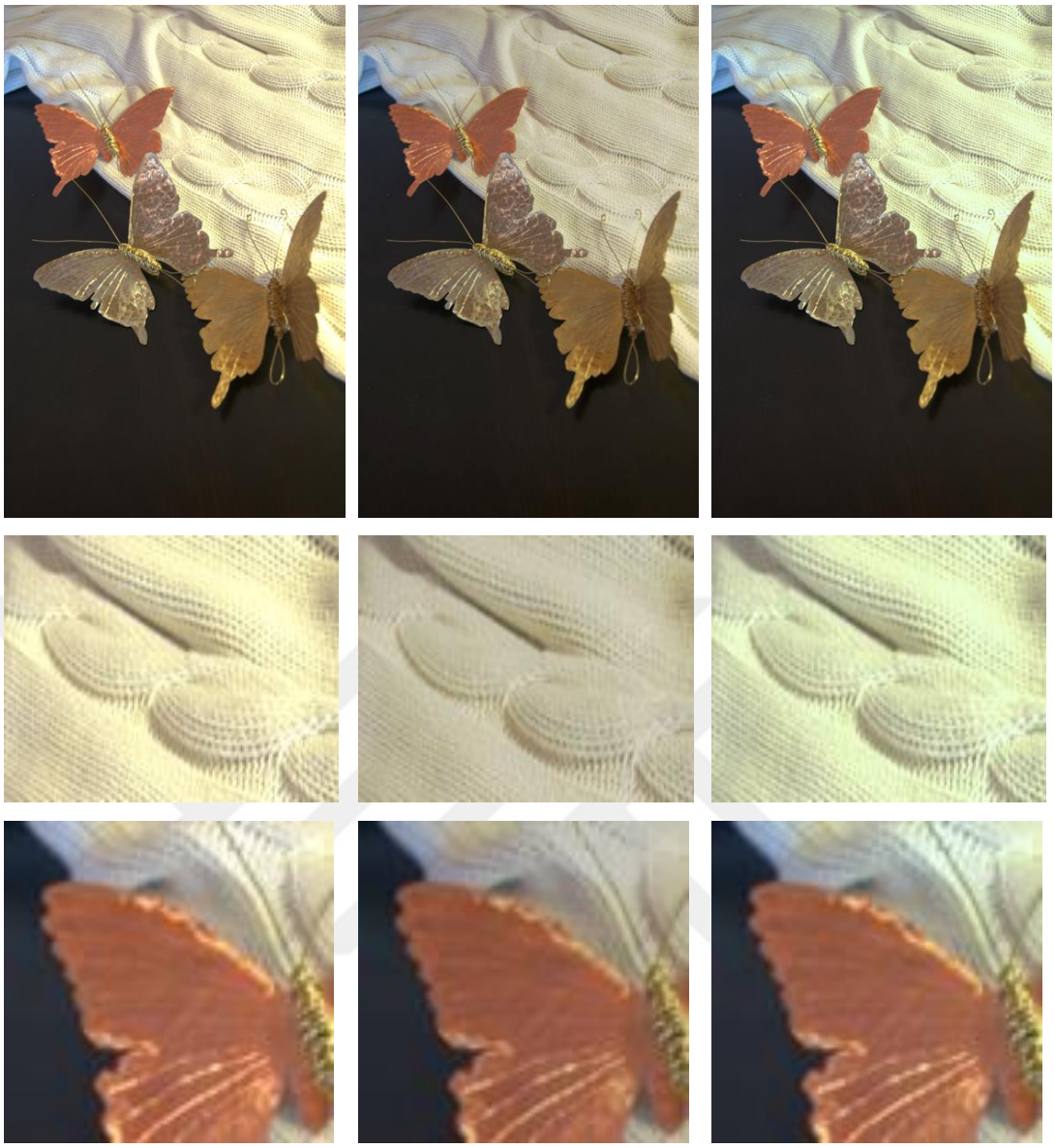
	MSE	PSNR (dB)	SSIM
Portraits	92,825	28,454	0,994
Garden	40,562	32,500	0,980
Chapel	404,358	22,063	0,966
Cars	156,616	26,182	0,962
Kodim03	5,416	40,794	0,994
Kodim08	111,627	27,653	0,964
Kodim23	11,781	37,419	0,990
Kodim25	16,591	35,932	0,988
Average	104,972	31,375	0,980

4.1 Saturation Correction in Tone-Mapped HDR Images

The first set of experiments are carried out on tone-mapped images and on average both statistically and visually superior results are obtained compared to Masood.

Figure 8 demonstrates visual results obtained for the *Butterflies* image. The proposed method has superior statistical results than Masood in Table 4. Visually, the colors of the cloth is better recovered and it is more naturally looking than Masood where the cloth appears to be darker. Furthermore, the texture of the cloth is much better preserved in the proposed method than in Masood's approach. Moreover, the light pink butterfly is successfully restored and has more vivid colors when compared to Masood.

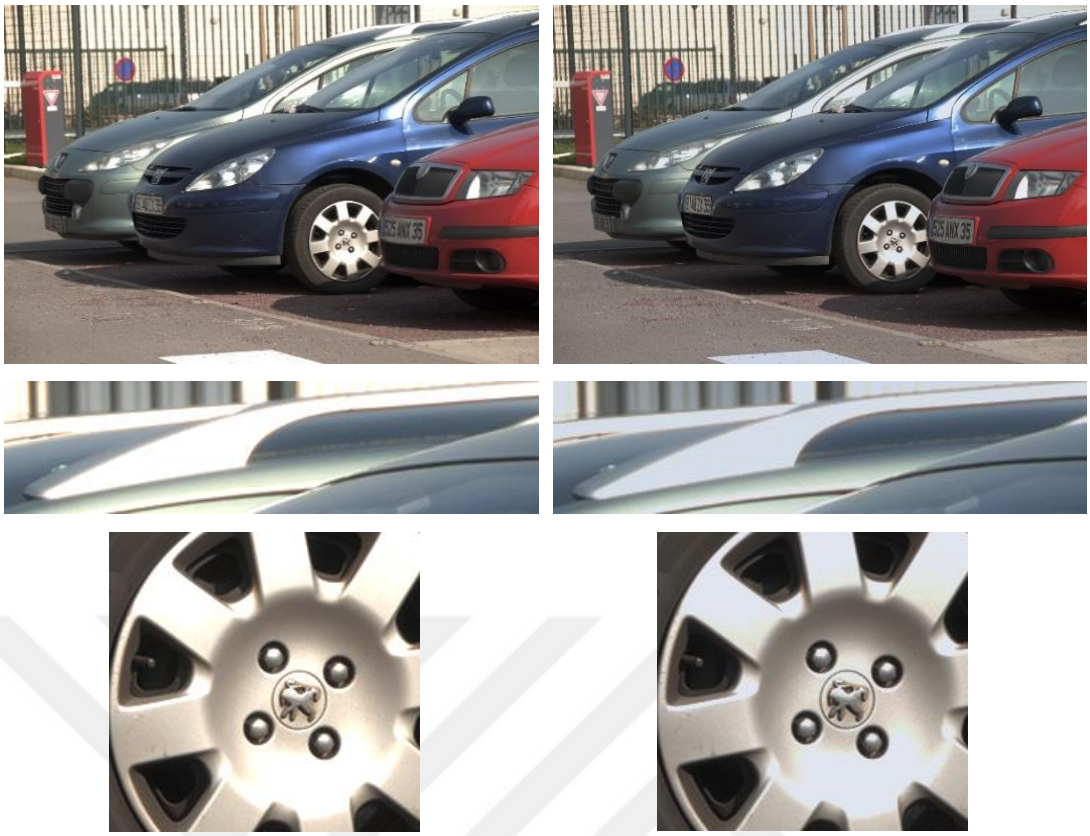
In Fig. 9, the three cars contain several clipped regions and especially pixels clipped in jointly two channels are present in this image. The top of blue and gray cars, the rim of the blue car and the headlights of all three cars are successfully recovered. Moreover, the saturated pixels on the building behind the fence are effectively corrected. Although the proposed method presents lower statistical results than the average values in Table 6, a visually appealing and plausible output image which contains well-settled colors is obtained.



a) Input b) Masood c) Proposed

Figure 8. Visual Comparison of *Butterflies* Outcomes.

Figure 10 illustrates the visual comparison for the *Watering Can* image. For this image, Masood produces statistically superior results in Table 4. However, the declipped pixels on the pipe of the watering can appear to be artificial in Masood, while the proposed method is able to correct the saturated pixels on the body of the watering can by preserving its original color. A visually more pleasing image is acquired by the proposed algorithm.



a) Input b) Proposed

Figure 9. The Result of the Proposed Method for *Cars*.

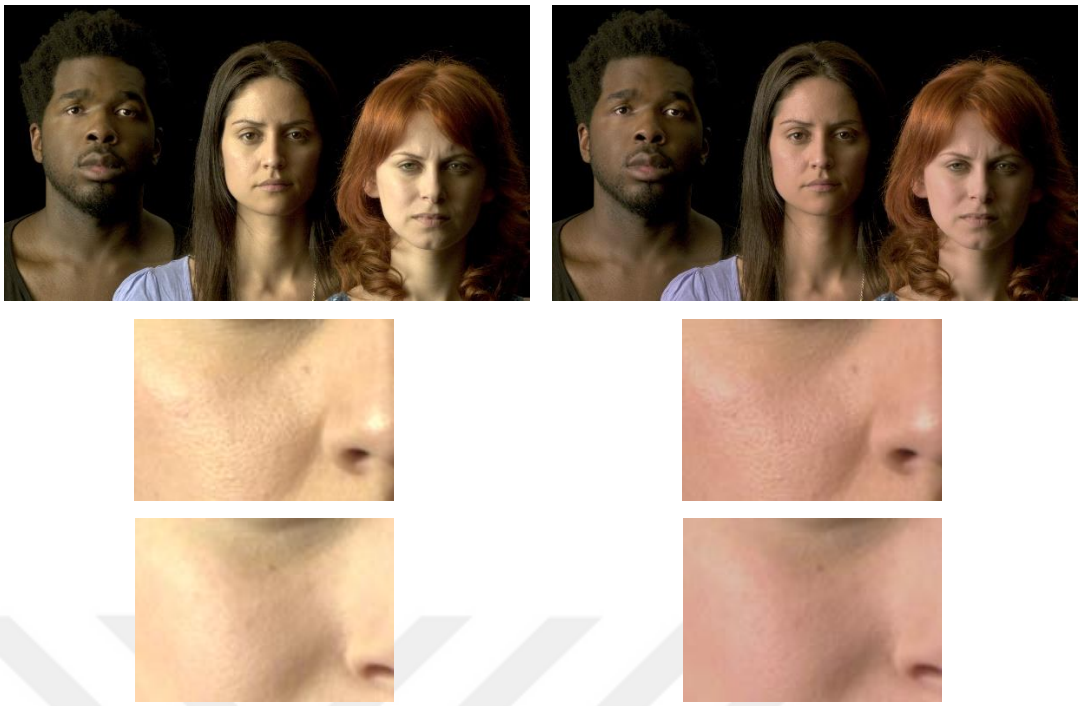
As it can further be observed in Fig. 11, the proposed method produces very efficient outcomes in human face images. According to the employed threshold, the pixels are severely saturated on the faces of women. These pixels of interest are declipped and an output image is acquired with well-settled colors and natural appearance. Furthermore, the perceptual details are successfully preserved with a high SSIM score of 0.994.

In the *Garden* image, the clipped regions on the wooden wall are recovered (Fig. 12). The color of the wood appears more natural in the output image. Also, while the saturated pixels on the bird statue are declipped, the concrete ground received a more natural-looking color. The only slight error occurred in the shadow of the roof, where mainly three jointly clipped channels are present in pixels (Table 2). A few pixels at the top of the pyramidal shadow are filled with incorrect intensity values, which did not cause an important information loss in the image.

a) Input

b) Masood c) Proposed

Figure 10. Visual Comparison of *Watering Can* Outcomes.



a) Input b) Proposed

Figure 11. The Result of the Proposed Method for Portraits.

a) Input b) Proposed

Figure 12. The Result of the Proposed Method for *Garden*.

a) Input b) Proposed

Figure 13. The Result of the Proposed Method for *Chapel*.

a) Input

b) Masood c) Proposed

Figure 14. Visual Comparison of *Exploratorium* Outcomes.

The *Chapel* image (Fig. 13) only contains one channel saturated pixels; hence it is declipped only via LE. The regions on the blue and yellow glasses of the windows, and the floor, stairs, banks and walls containing saturated pixels due to lighting are successfully declipped. The details in the image are greatly preserved. On the other hand, there is a slight color change on the roof which occurred after scaling the image back to the range between 0 and 255. Yet, this color change did not cause any significant detail loss, in fact the colors became more well-settled and natural-looking.

In the *Exploratorium* image the building, sky scene and tree trunk contain clipped regions (Fig. 14). While Masood causes color distortion at the sky scene and produces an artificial looking sky, the proposed method successfully declips the pixels and provides a visually appealing sky scene. In Table 4, it is seen that the color distortion caused an MSE of 143.944 for Masood, while for the proposed technique the MSE is only 67.651. Moreover, the proposed algorithm declips the saturated areas on the building without information loss and produces a building containing more natural-looking colors than Masood. Furthermore, while the clipped regions on the tree trunk are declipped details are preserved. From the high SSIM score of 0.980 it can also be deduced that the proposed method is more successful in preserving the details in this image than Masood (SSIM:0.963). The only slight distortion via the proposed method occurs in the small sky region seen under the dome. Since the input image contains not a sufficient amount of non-saturated color data in this area, the algorithm could not assign the correct values to most of the clipped pixels in this region.

4.2 Saturation Correction in LDR Images

In the second set of experiments, saturation correction is applied to (processed) LDR images and visually appealing outputs are obtained, in which high- and low-level details and features are greatly preserved.

Figure 15 demonstrates visual results obtained for the *Sea Plant* image, which is mainly corrected through the one channel correction via linear embeddings. While there is an obvious visual quality enhancement, the proposed technique has also a significantly higher SSIM score than Masood (Table 4) and Guo (Table 5). There exist many artefacts and distorted colors in Masood both in the foreground and background. Although no structural artefacts are present in Guo, the yellow sea plant lost its natural color. A visually unpleasant looking image is acquired. The proposed algorithm, on the other side, achieves to successfully correct over-saturated pixels and preserves specific features such as the texture and the color of the sea plant. A visually more appealing image is obtained.

a) Input b) Masood

c) Guo d) Proposed

Figure 15. Visual Comparison of *Sea Plant* Outcomes (Source: Guo Algorithm, 2010).

a) Input b) Masood

c) Lu d) Proposed

Figure 16. Visual Comparison of *Flowers* Outcomes (Source: Martin et al., 2001).

For the Flowers image in Fig. 16, the proposed method produces an output with a better PSNR of 30.631 dB which outperforms both Lu (17.993 dB) and Masood (26.874 dB). As it can be clearly observed in the figure, the declipped background color changes significantly in the result of Lu. Moreover, flowers appear artificial and their color intensity seems to be distorted. In the result of Masood, the image appears to be darker and the details in corrected regions are partially lost. Additionally, several pixels in the yellow flower on the left are restored with wrong colors. On the other hand, the proposed method successfully corrects the saturated regions by preserving details and restoring pixel values with more plausible colors, e.g., the details and colors on the petal of the yellow flower on the left.

4.3 Saturation Correction in Raw LDR Images

Several raw LDR images with distinct content are used in the final set of experiments. Visually plausible and natural-looking output images are acquired without any significant information loss.

In Fig. 17, the proposed method achieves a superior PSNR value than Masood for the *Kodim19* image (Table 4). It can be observed in this figure that the method of Masood is able to recover clipped pixels on the lighthouse (with a successful SSIM value), however visually a darker image is obtained when compared to the proposed algorithm which produces a more natural-looking image.

As it can be seen in Fig. 18, especially yellow and orange hats contain severely saturated pixels in the *Kodim03* image. The proposed method manages a successful declipping process with a PSNR value of 40.794 dB while Xu has 39.710 dB.

Figure 19 shows visual results obtained for the *Kodim17* image. The desaturated regions on the ball and crown are visually very successful in the algorithm of Masood. However, the proposed method also restores these regions successfully and it further preserves colors in the non-saturated regions much better, which leads to statistically

superior results than Masood (Table 4).

a) Input b) Masood c) Proposed

Figure 17. Visual Comparison of *Kodim19* Outcomes (Source: Kodak Image Set, 1991).

a) Input b) Proposed

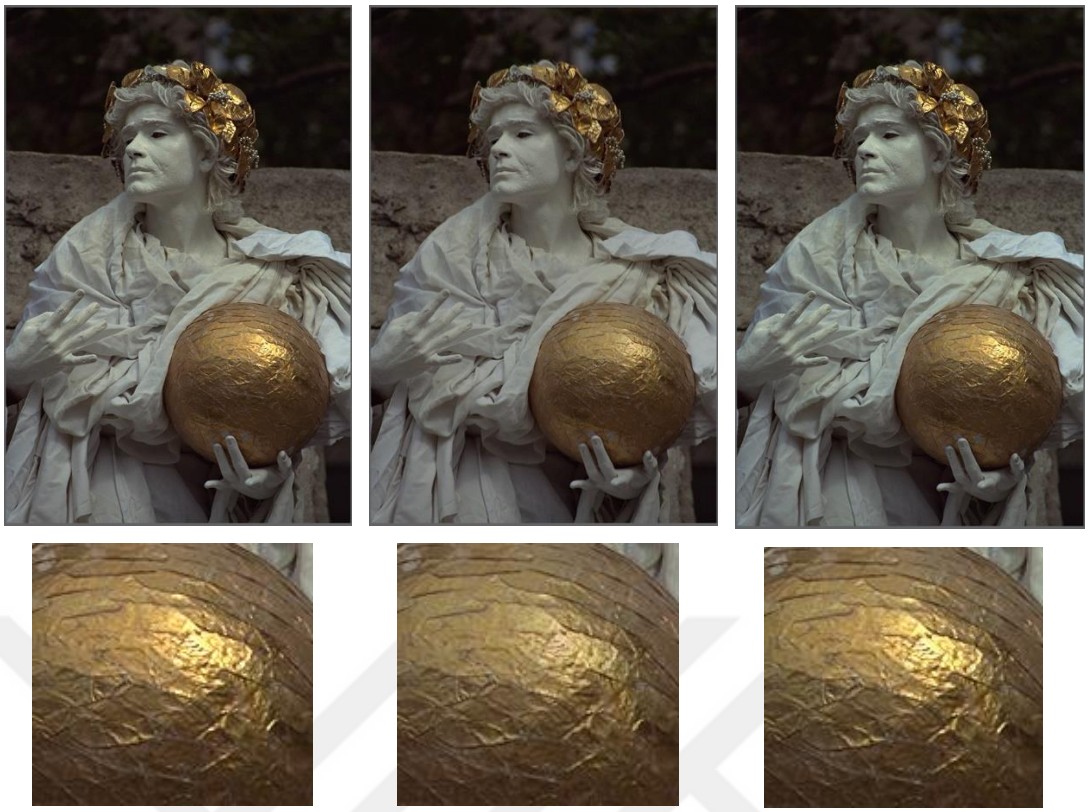
Figure 18. The Result of the Proposed Method for *Kodim03* (Source: Kodak Image

Set, 1991).

One of the most challenging images in the dataset is *Kodim23* due to its differently focused regions in the image. The background is out-of-focus compared to the foreground, which makes it more difficult to preserve details and restore faulty pixels in the image through one algorithm. Almost half of the saturated pixels are in the out-of-focus region, hence clipped areas lying at the intersection of foreground and background are more challenging since the designed method takes advantage of neighbouring pixels. As it can clearly be visualized in Fig. 20, the proposed algorithm successfully recovers the clipped regions in the foreground, i.e., the heads of parrots and the body of the left parrot, as well as the bright saturated regions in the background. Furthermore, it outperforms Xu by achieving a PSNR value of 37.419 dB, while Xu has 34.850 dB.

The clipped regions mainly caused by sunlight in *Kodim08* are successfully corrected. In this image an important observation is the declipping result in the sky. Over-exposed regions frequently occur in the sky in images and declipping these areas is challenging. The reason behind that is the neighbouring pixels of a clipped pixel in the sky are usually also saturated, therefore extracting information is strenuous and declipping algorithms tend to fail. However, as seen in Fig. 21 the proposed method produces a visually appealing result in a small sky scene, where pixels are mainly clipped in one channel. On the other hand, a minor issue emerged in the output image; the color of one of the houses is slightly distorted. Yet as it can be stated from the SSIM score of 0.964, the luminance, contrast and structure elements are greatly preserved, thus the slight color change on the house is not critical.

In Fig. 22, the *Kodim25* image contains saturated pixels on the rocks, bag, jacket and faces. The proposed algorithm successfully restores these regions and provides a visually plausible output image. It is important to emphasize that this image is mostly saturated in three channels, which is the most difficult case for image declipping. The proposed method achieves an SSIM score of 0.988 and PSNR of 35.932 dB.



a) Input b) Masood c) Proposed

Figure 19. Visual Comparison of *Kodim17* Outcomes (Source: Kodak Image Set, 1991).

a) Input b) Proposed

Figure 20. The Result of the Proposed Method for *Kodim23* (Source: Kodak Image Set, 1991).

a) Input b) Proposed

Figure 21. The Result of the Proposed Method for *Kodim08* (Source: Kodak Image Set, 1991).

a) Input b) Proposed

Figure 22. The Result of the Proposed Method for *Kodim25* (Source: Kodak Image Set, 1991).

CHAPTER 5: CONCLUSION

HDR compatible equipment will become accessible to all consumers in the near future, yet currently it is an expensive technology. Therefore, economic solutions are required to obtain high quality images. A common approach is to capture a high contrast scene with an HDR compatible camera and project the image to an LDR screen via a tone-mapping algorithm. Although tone-mapping is a successful method, clipping may occur. In this thesis, a practical declipping (saturation correction) algorithm is proposed, which is based on LE, difference of pixels and block-search.

To the best of available knowledge, LE is used for the first time in a declipping algorithm and it presents very promising results for one-channel saturation correction. While taking the advantage of channel-wise correlations, pixel differences as offsets in a neighborhood result in satisfying outcomes for two-channel saturation correction. Furthermore, a simple block-search in the gradient direction similar to the template matching approach leads to very successful results for three-channel saturation correction. Lastly, the bilateral post-processing filter avoids sharp changeovers and prevents artefacts especially at the edges. Overall, the proposed method produces both statistically and visually superior output images compared to existing methods.

This thesis has several contributions to its research field. Alongside adopting LE for the first time in this research domain, constructing a case-oriented algorithm results in a robust method, in which pixels with different number of clipped channels are processed individually. Furthermore, the proposed algorithm is able to correct different types of image content. Therefore, the method distinguishes itself from several state-of-the-art declipping approaches which focus only on a particular type of context. Aside from tone-mapped HDR images, the proposed algorithm works successfully also with raw and processed LDR images, in comparison to several existing techniques which tend to work only on raw or processed images. As a last note, a drawback of the proposed algorithm is that it partly fails to correct large clipped regions, e.g., saturated pixels in a large sky scene. This issue will be considered as a future direction of this study.

REFERENCES

Ahishali, M., Kiranyaz, S., Ince, T., and Gabbouj, M. (2019) *Dual and single polarized sar image classification using compact convolutional neural networks*, Remote Sensing. Vol. 11(11), pp. 1340.

An, G. H., Lee, S., Ahn, Y. D. and Kang, S.J. (2018) *Deep tone-mapped hdrnet for high dynamic range image restoration*, Symposium Digest of Technical Papers, Vol. 49(1), pp. 291–294.

Angelov, P. (2012) *Sense and avoid in UAS: research and applications*. United Kingdom: John Wiley & Sons.

Banterle, F., Artusi, A., Debattista, K. and Chalmers, A. (2017) *Advanced high dynamic range imaging*. 2nd edition. Boca Ratton: CRC press.

Cai, J., Gu, S. and Zhang, L. (2018) *Learning a deep single image contrast enhancer from multi-exposure images*, IEEE Transactions on Image Processing, Vol. 27(4), pp. 2049–2062.

Chae, S., Lee, S. and Sohng, K. *Color Correction for Preserved Saturation and Hue in HDR Images*. *International Conference on Circuits, Control, Communication, Electricity, Electronics, Energy, System, Signal and Simulation.* Guam, Hawaii, USA. 18 – 20 July 2013.

Criminisi, A., Pérez, P. and Toyama, K. (2004) *Region filling and object removal by exemplar-based image inpainting*, IEEE Transactions on Image Processing, Vol. 13(9), pp. 1200–1212.

Degerli, A., Aslan, S., Yamac, M., Sankur, B. and Gabbouj, M. *Compressively Sensed Image Recognition*. *European Workshop on Visual Information Processing.* Tampere, Finland, 26-28 Nov. 2018.

Eilertsen, G. (2018) *The high dynamic range imaging pipeline*. Linköping: Linköping University.

Elboher, E. and Werman, M. (2010) *Recovering color and details of clipped image regions*. International Journal of Computer Information Systems and Industrial

Management Applications, Vol. 3(-), pp. 812–-819.

Farbman, Z., Fattal, R., Lischinski, D. and Szeliski, R. (2008) *Edge-preserving decompositions for multi-scale tone and detail manipulation*, ACM Transactions on Graphics, Vol. 27(3), pp. 1–10.

Guan, J. and Qiu, G. *Display HDR Image using a Gain Map*. *IEEE International Conference on Image Processing.* San Antonio, TX, USA. 16 September – 19 October 2007.

Guo, D., Cheng, Y., Zhuo, S. and Sim, T. *Correcting Over-Exposure in Photographs*. *IEEE Conference on Computer Vision and Pattern Recognition.* San Francisco, CA, USA. 13 – 18 June 2010.

Guo Algorithm. (2010). *Correcting over-exposure in photographs outcomes* [Online]. Available at https://dongguo.org/deoverexposure/. (Accessed 02.07.2020).

Hayat, N. and Imran, M. (2019) *Ghost-free multi exposure image fusion technique using dense SIFT descriptor and guided filter*, Journal of Visual Communication and Image Representation, Vol. 62(-), pp. 295–308.

Honig, S. and Werman, M. *Image Declipping with Deep Networks*. *IEEE International Conference on Image Processing.* Athens, Greece. 7 – 10 October 2018.

Hore, A. and Ziou, D. *Image Quality Metrics: PSNR vs. SSIM*. *International Conference on Pattern Recognition.* Istanbul, Turkey. 23 – 26 August 2010.

Ju, H. J. and Park, R.-H. (2016) *Correction of saturated regions in RGB color space*, International Journal of Computer Geometry Applications, Vol. 6(2), pp. 1–13.

Karakaya, D., Ulucan, O. and Turkan, M. *A Comparative Analysis on Fruit Freshness Classification*. *Innovations in Intelligent Systems and Applications Conference.* Izmir, Turkey, 31 October- 2 November 2019.

Khan, I. R., Rahardja, S., Khan, M. M., Movania, M. M. and Abed, F. (2017) *A tone-mapping technique based on histogram using a sensitivity model of the human visual system*, IEEE Transactions on Industrial Electronics, Vol. 65(4), pp. 3469–3479.

Kodak Image Set. (1991). *Kodak Image Dataset* [Online]. Available at http://www.cs.albany.edu/~xypan/research/snr/Kodak.html. . (Accessed 02.07.2020).

Lee, D. H., Yoon, Y. J., Kang, S. J. and Ko, S. J. (2014) *Correction of the overexposed region in digital color image,* IEEE Transactions on Consumer Electronics, Vol. 60(2), pp. 173–178.

Liang, Z., Xu, J., Zhang, D., Cao, Z. and Zhang, L. *A Hybrid l1-l0 Layer Decomposition Model for Tone Mapping*. *Conference on Computer Vision and Pattern Recognition.* Salt Lake City, UT, USA. 18 – 23 June 2018.

Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollar, P. and Zitnick, C. L. *Microsoft Coco: Common Objects in Context*. *European Conference On Computer Vision.* Zurich, Switzerland. 6-12 September 2014.

Lu, Y.-H., Chen, Y.-H. and Lin, H.-Y. S. *A Multiresolution Approach to Recovering Colors and Details of Clipped Image Regions*. *International Conference on Intelligent Information Hiding and Multimedia Signal Processing.* Adelaide, SA, Australia. 23 – 25 September 2015.

Mahmud, A., Haque, R. U., Adnan, M. A. and Al Mahmud, H. M. S. *A Novel Method for Ghost Removal in High-Dynamic Range Images*. *International Joint Conference on Computational Intelligence.* Dhaka, Bangladesh, 25 - 26 October 2019.

Mantiuk, R., Mantiuk, R., Tomaszewska, A. and Heidrich, W. (2009) *Color correction for tone mapping,* Computer Graphics forum, Vol. 28(2), pp. 193–202.

Martin, D., Fowlkes, C., Tal, D. and Malik, J. *A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics*. *International Conference on Computer Vision.* Vancouver, BC, Canada. 7 – 14 July 2001.

Masood, S. Z., Zhu, J. and Tappen, M. F. (2009) *Automatic correction of saturated regions in photographs using cross-channel correlation,* Computer Graphics Forum, Vol. 28(7), pp. 1861–1869.

Masood Algorithm. (2009) . *Automatic correction of saturated regions in photographs using cross-channel correlation algorithm* [Online]. Available at http://www.syedzainmasood.com/research.html. (Accessed 02.07.2020).

Mertens, T., Kautz, J. and Van Reeth, F. (2009) *Exposure fusion: a simple and practical alternative to high dynamic range photography*, Computer Graphics forum,

Vol. 28(1), pp. 161–171.

Ok, J. and Lee, C. (2017) *HDR tone mapping algorithm based on difference compression with adaptive reference values*, Journal of Visual Communication and Image Representation, Vol. 43(-), pp. 61–76.

Omer, I. and Werman, M. *Color Lines: Image Specific Color Representation*. *IEEE Conference on Computer Visison and Pattern Recognition.* Washington, DC, USA. 27 June – 2 July 2004.

Pouli, T., Artusi, A., Banterle, F., Akyüz, A. O. and Seidel, H. P. and Reinhard, E. *Color Correction for Tone Reproduction*. *Color and Imaging Conference.* Albuquerque, New Mexico, USA. 4-8 November 2013.

Reinhard, E., Stark, M., Shirley, P. and Ferwerda, J. *Photographic Tone Reproduction for Digital Images*. *Conference on Computer Graphics and Interactive Techniques.* San Antonio, Texas, USA. 21-26 July.

Reinhard, E. and Devlin, K. (2005) *Dynamic range reduction inspired by photoreceptor physiology*, IEEE Transactions on Visualization and Computer Graphics, Vol. 11 (1), pp. 13–24.

Roweis, S. T. and Saul, L. K. (2000). *Nonlinear dimensionality reduction by locally linear embedding*, Science, Vol. 290(5500), pp. 2323–2326.

Sikudova, E., Pouli, T., Artusi, A., Akyüz, A. O., Banterle, F., Mazlumoglu, Z. M. and Reinhard, E. (2015) *A gamut-mapping framework for color-accurate reproduction of hdr images*, IEEE Computer Graphics and Applications, Vol. 36(4) pp. 78–90.

Smith, K., Krawczyk, G., Myszkowski, K. and Seidel, H. P. (2006) Beyond tone mapping: enhanced depiction of tone mapped hdr images, Computer Graphics Forum, Vol. 25(3), pp. 427–438.

Tenenbaum, J. B., De Silva, V. and Langford, J. C. (2000) *A global geometric framework for nonlinear dimensionality reduction*, Science, 290(5500), pp. 2319–2323.

Tomasi, C. and Manduchi, R. *Bilateral Filtering for Gray and Color Images*. *IEEE International Conference on Image Processing.* Bombay, India. 7 January 1998.

Türkan, M., Thoreau, D. and Guillotel, P. *Optimized Neighbor Embeddings for Single-Image Super-Resolution*. *IEEE International Conference on Image Processing.* Melbourne, VIC, Australia. 15-18 September 2013.

Türkan, M., Thoreau, D. and Guillotel, P. *Iterated Neighbor-Embeddings for Image Super-Resolution*. *IEEE International Conference on Image Processing.* Paris, France. 27-30 October 2014.

Vincent, L. (1993) *Morphological grayscale reconstruction in image analysis: applications and efficient algorithms*, IEEE Transactions on Image Processing, Vol. 2(2), pp. 176–201.

Xu, D., Doutre, C. and Nasiopoulos, P. (2010) *Correction of clipped pixels in color images*, IEEE Transactions on Visualization and Computer Graphics, Vol. 17(3), pp. 333–344.

Yamanaka, J., Kuwashima, S., and Kurita, T. *Fast and Accurate Image Super Resolution by Deep CNN with Skip Connection and Network in Network*. *International Conference on Neural Information Processing Systems.* Long Beach, CA, USA, 4-9 December 2017.

Yang, X., Xu, K., Song, Y., Zhang, Q., Wei, X. and Lau, R. W. H. *Image Correction via Deep Reciprocating HDR Transformation*. *IEEE Conference on Computer Vision and Pattern Recognition.* Salt Lake City, Utah, USA, 18-22 June 2018.

Appendix A: HDR Compatible Devices

Company	Samsung	Sony	Nikon	Fujifilm	Canon	Kodak	Casio
Digital Camera	NX2000	SLT A99	D7500	GFX100	EOS 80D	AZ528	EX-ZR700
	NX30	PXW-Z150	D7100	X-Pro3	6D DSLR	FZ152	EX-ZR400
	WB2200F	ZV-1	D800	X-T200	EOS Rebel T4i	EasyShare Z990	EX-ZR1000

Company	Samsung	Sony	Beko	Acer	Asus	Dell	BenQ
Display	The Frame QLED TV	Master Series Z9G	B55 OLED 9890 5B	ET322QK	ROG PG27UQ	S2418HX	EX3501R
	LC27HG70QQ MXZN	BVM-HX310	B65L 9785 5S Quatro	XB271HK bmiprz	Adaptive-Sync HDR IPS	AW2521HF	EW3270OU
	RU7300 Curved	LMD-A240 v3.0	Crystal Pro X-B49 A 950 A	Predator X27	ROG SWIFT PG35VQ	U2718Q	EL2870U

Appendix B: The Algorithm Proposed in this Thesis

Algorithm 1 One-channel saturation correction

Input: $\mathcal{I}$
Output: $\mathcal{I}_1$

Define one-channel saturation mask M as $\{M \mid R_\alpha, G_\alpha, B_\alpha \in \mathcal{I}\}$
for all $\alpha \in M$ **do**
    Calculate the gradient angle of α
    Obtain neighbors $\mathcal{N}_1$ and $\mathcal{N}_2$ in the gradient direction
    Solve the optimization in Eqn.21 accordingly
    Estimate the saturated pixel α value in Eqn. 22 to form $\mathcal{I}'$
end for
Obtain $\mathcal{I}_S$ from $\mathcal{I}'$ via Eqn. 23
Calculate intensity map $\mathcal{G}$ of $\mathcal{I}$
$\mathcal{I}_{SW} = \text{SlidingWindow}(\mathcal{I}, \mathcal{I}_S)$
Calculate the final output $\mathcal{I}_1$ via Eqn.24

Algorithm 2 Two-channel saturation correction

Input: $\mathcal{I}_1$
Output: $\mathcal{I}_2$

Define two-channel saturation mask M as $\{M \mid \{R_\alpha, G_\alpha\}, \{R_\alpha, B_\alpha\}, \{G_\alpha, B_\alpha\} \in \mathcal{I}_1\}$
Determine connected components in $\mathcal{I}_1$ through M
for all Connected components **do**
    for all α with the highest priority on the contour **do**
        Calculate the gradient angle of α
        Obtain neighbors $\mathcal{N}_1$ and $\mathcal{N}_2$ in the gradient direction
        Estimate saturated pixels via Eqn. 26 accordingly to form $\mathcal{I}_2$
    end for
end for

Algorithm 3 Three-channel saturation correction

Input: $\mathcal{I}_2$
Output: $\mathcal{I}_3$

Define three-channel saturation mask M as $\{M \mid \{R_\alpha, G_\alpha, B_\alpha\} \in \mathcal{I}_2\}$
Determine connected components in $\mathcal{I}_2$ through M
for all Connected components **do**
    for all α with the highest priority on the contour **do**
        Calculate the gradient angle of α
        Obtain 4 neighbors in the gradient direction
        Form the search-region
        Estimate saturated pixels via template matching to form $\mathcal{I}_3$
    end for
end for

Appendix B: The Algorithm Proposed in this Thesis (cont'd)

Algorithm 4 Post-processing algorithm

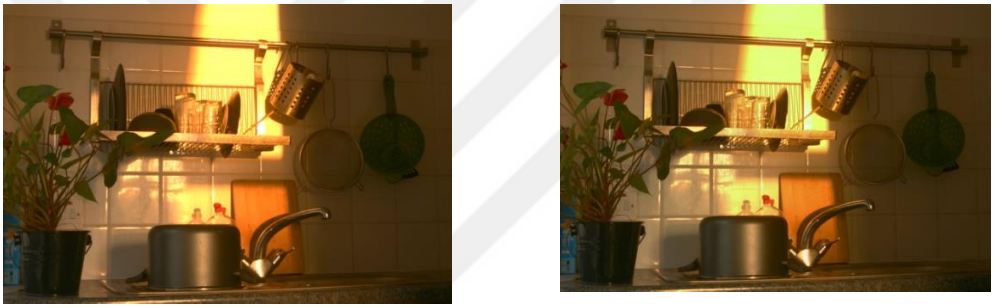
Input: $\mathcal{I}_3$, M
Output: $\mathcal{I}_O$

1: $\mathcal{I}_L = \text{CIELab}(\mathcal{I}_3)$
2: $\mathcal{I}_R = M(\mathcal{I}_L)$
3: $\mathcal{I}_R^B = \text{BilateralFilter}(\mathcal{I}_R)$
4: $\mathcal{I}_O = M(\mathcal{I}_R^B) + \neg M(\mathcal{I}_L)$
5: $\mathcal{I}_O \leftarrow \text{RGB}(\mathcal{I}_O)$

Appendix C: Algorithm Outcomes with Unsatisfying Performance

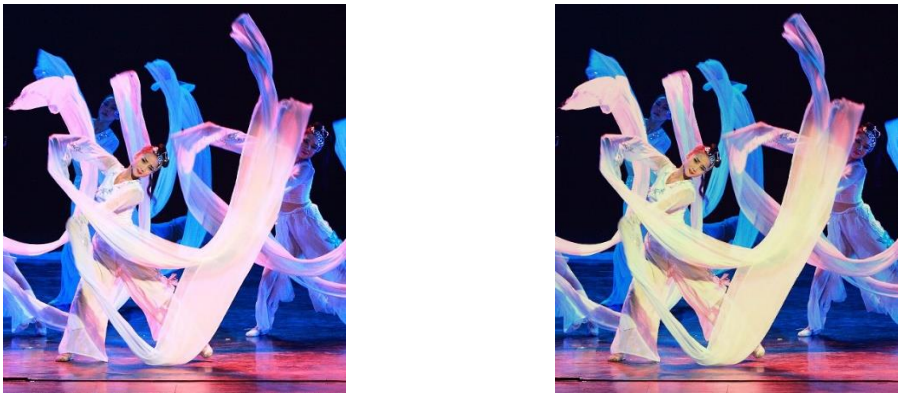
a) Input b) Proposed

Figure 23. The Result of the Proposed Method for *Dino*.



a) Input b) Proposed

Figure 24. The Result of Proposed Method for *Sink*.



a) Input b) Proposed

Figure 25. The Result of Proposed Method for *Dancer* (Source: Guo Algorithm, 2010).

Appendix C: Algorithm Outcomes with Unsatisfying Performance (cont'd)

a) Input b) Proposed

Figure 26. The Result of the Proposed Method for *House* (Source: Guo Algorithm, 2010).

Appendix D: The Publications during Master's Thesis

Published in or submitted to journals:

- D.Karakaya, O. Ulucan and M. Turkan (2020) *Image declipping: saturation correction in single images*. Computer Graphics Forum, (Submitted).
- O.Ulucan, D.Karakaya and M. Turkan. (2020) *Multi-exposure image fusion based on linear embeddings and watershed masking.* Signal Processing, (Accepted).
- D.Karakaya, O. Ulucan and M. Turkan (2020) *Electronic nose and its applications: a survey*. International Journal of Automation and Computing, Vol. 17(2), pp. 177 – 209.

Published in or submitted to indexed conference proceedings:

- D. Karakaya, O. Ulucan and M. Turkan. *A Comparative Study on Electronic Nose Data Analysis Tools*. *The Innovations in Intelligent Systems and Applications Conference.* Marmara University, Istanbul. 15 – 17 October (Accepted).
- O. Ulucan, D. Karakaya and M. Turkan. *A Large-Scale Dataset for Fish Segmentation and Classification*. *The Innovations in Intelligent Systems and Applications Conference.* Marmara University, Istanbul. 15 – 17 October 2020 (Accepted).
- Y. Oktar, O. Ulucan, D. Karakaya, E. O. Ersoy and M. Turkan. *Binocular Vision Based Convolutional Networks. IEEE Conference on Signal Processing and Communications Applications.* Gaziantep. 5 – 7 October 2020 (Accepted).
- O. Ulucan, D. Karakaya and M. Turkan. *Meat Quality Assessment based on Deep Learning. The Innovations in Intelligent Systems and Applications Conference.* Yasar University, Izmir. 31 October – 2 November 2019.
- D. Karakaya, O. Ulucan and M. Turkan. *A Comparative Analysis on Fruit Freshness Classification. The Innovations in Intelligent Systems and Applications Conference.* Yasar University, Izmir. 31 October – 2 November 2019.