UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

MATHEUS ROCHA BARBOSA

CONTRAST ENHANCEMENT AND EXPOSURE CORRECTION USING A STRUCTURE-AWARE DISTRIBUTION FITTING

Porto Alegre
2023

MATHEUS ROCHA BARBOSA

CONTRAST ENHANCEMENT AND EXPOSURE CORRECTION USING A STRUCTURE-AWARE DISTRIBUTION FITTING

Thesis presented to Programa de Pós-Graduação em Engenharia Elétrica of Universidade Federal do Rio Grande do Sul in partial fulfillment of the requirements for the degree of Master in Electrical Engineering.
Minor: Eletrônica – Instrumentação e Processamento de Sinais

ADVISOR: Prof. Ph.D. Jacob Scharcanski

Porto Alegre
2023

MATHEUS ROCHA BARBOSA

CONTRAST ENHANCEMENT AND EXPOSURE CORRECTION USING A STRUCTURE-AWARE DISTRIBUTION FITTING

This thesis was considered adequate for obtaining the degree of Master in Electrical Engineering and approved in its final form by the Advisor and the Examination Committee.

Advisor: _____
Prof. Ph.D. Jacob Scharcanski, UFRGS
Ph.D. at University of Waterloo – Waterloo, Canada

Examination Committee:

Prof. D.Sc. Eduardo Antonio Barros da Silva, UFRJ
D.Sc. at University of Essex – Essex, England

Prof. Ph.D. Ricardo Lopes de Queiroz, UNB
Ph.D. at The University of Texas – Arlington, United States of America

Prof. Dr. Valner João Brusamarello, UFRGS
Dr. at Universidade Federal de Santa Catarina – Florianópolis, Brazil

Coordinator of PPGEE: _____
Prof. PhD Sérgio Luís Haffner

Porto Alegre, January 2023.

ACKNOWLEDGMENT

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

ABSTRACT

Contrast enhancement and exposure correction are useful in domestic and technical applications, the latter as a preprocessing step for other techniques or for aiding human observation. Often, a locally adaptive transformation is more suitable for the task than a global transformation. For example, objects and regions may have very different levels of illumination, physical phenomena may compromise the contrast at some regions but not at others, or it may be desired to have high visibility of details in all parts of the image. For such cases, local image enhancement methods are preferable. Although there are many contrast enhancement and exposure correction methods available in the literature, there is no definitive solution that provides a satisfactory result in all situations, and new methods emerge each year. In special, traditional adaptive histogram equalization-based methods suffer from checkerboard and staircase effects and from over enhancement. This dissertation proposes a method for contrast enhancement and exposure correction in images named Structure-Aware Distribution Stretching (SADS). The method fits a parametric model of probability distribution to the image regionally while respecting the image structure and edges between regions. This is done using regional versions of the classical expressions for estimating the parameters of the distribution, which are obtained by replacing the sample mean present in the original expressions by an edge-preserving smoothing filter. After fitting the distribution, the cumulative distribution function (CDF) of the adjusted model and the inverse of the CDF of the desired distribution are applied. A structure-aware heuristic to indicate smooth regions is proposed and used to attenuate the transformations in flat regions. SADS was compared with other methods from the literature using objective no-reference and full-reference image quality assessment (IQA) metrics in the tasks of simultaneous contrast enhancement and exposure correction and in the task of defogging/dehazing. The experiments indicate a superior overall performance of SADS with respect to the compared methods for the image sets used, according to the IQA metrics adopted.

Keywords: Contrast enhancement, exposure correction, local histogram equalization, beta distribution.

RESUMO

Realce de contraste e correção de exposição são úteis em aplicações domésticas e técnicas, no segundo caso como uma etapa de pré-processamento para outras técnicas ou para ajudar a observação humana. Frequentemente, uma transformação localmente adaptativa é mais adequada para a tarefa do que uma transformação global. Por exemplo, objetos e regiões podem ter níveis de iluminação muito diferentes, fenômenos físicos podem comprometer o contraste em algumas regiões mas não em outras, ou pode ser desejável ter alta visibilidade de detalhes em todas as partes da imagem. Para esses casos, métodos de realce de imagem locais são preferíveis. Embora existam muitos métodos de realce de contraste e correção de exposição disponíveis na literatura, não há uma solução definitiva que forneça um resultado satisfatório em todas as situações, e novos métodos surgem a cada ano. Em especial, os métodos tradicionais baseados em equalização adaptativa de histograma sofrem dos efeitos checkerboard e staircase e de excesso de realce. Esta dissertação propõe um método para realce de contraste e correção de exposição em imagens chamado Structure-Aware Distribution Stretching (SADS). O método ajusta regionalmente à imagem um modelo paramétrico de distribuição de probabilidade, respeitando a estrutura da imagem e as bordas entre as regiões. Isso é feito usando versões regionais das expressões clássicas de estimativa dos parâmetros da distribuição, que são obtidas substituindo a média amostral presente nas expressões originais por um filtro de suavização que preserva as bordas. Após ajustar a distribuição, a função de distribuição acumulada (CDF) do modelo ajustado e a inversa da CDF da distribuição desejada são aplicadas. Uma heurística ciente de estrutura que detecta regiões suaves é proposta e usada para atenuar as transformações em regiões planas. SADS foi comparado a outros métodos da literatura usando métricas objetivas de avaliação de qualidade de imagem (IQA) sem referência e com referência completa nas tarefas de realce de contraste e correção de exposição simultâneos e na tarefa de defogging/dehazing. Os experimentos indicam um desempenho geral superior do SADS em relação aos métodos comparados para os conjuntos de imagens usados, de acordo com as métricas IQA adotadas.

Palavras-chave: Realce de contraste, correção de exposição, processamento local de histogram, distribuição beta.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AHE	adaptive histogram equalization
AHIQ	attention-based hybrid image quality assessment
AMEF	artificial multiple-exposure fusion
BRISQUE	blind/referenceless image spatial quality evaluator
CDF	cumulative distribution function
CEIQ	contrast-changed image quality measure
CLAHE	contrast limited adaptive histogram equalization
CNN	convolutional neural network
DB-CNN	deep bilinear CNN
EPS	edge-preserving smoothing
FADE	fog aware density evaluator
FCN	fully convolutional network
FFA-Net	feature fusion attention network
FGS-GU	fast global smoothing based guided upsampling
FGS	fast global smoothing
FR	full-reference
FSIM	feature similarity
GAN	generative adversarial network
GHE	global histogram equalization
GMSD	gradient magnitude similarity deviation
GSS	golden section search
HDR	high dynamic range
HE	histogram equalization
HSL	hue saturation lightness color system
HSV	hue saturation value color system
iHSV	inverted HSV color system

IL-NIQE	integrated local NIQE
IQA	image quality assessment
IRLS	iteratively reweighted least squares
LCDPNet	local color distributions prior network
LCD	local color distributions
LCT	lightness channel transformation
LDR	low dynamic range
LHE	local histogram equalization
LIME	low-light image enhancement
LPIPS	learned perceptual image patch similarity
LUT	look-up table
MAE	mean absolute error
MICF	MIC factorization
MIC	modified incomplete Cholesky
MLE	maximum likelihood estimation
MMT	maximum and minimum transformation
MoM	method of moments
Musiq	multi-scale image quality transformer
NIQE	natural image quality evaluator
NR	nor-reference
PCG	preconditioned conjugate gradient
PDF	probability density function
PSNR	peak signal-to-noise ratio
RGB	red, green, blue
SADS	Structure-Aware Distribution Stretching
SAMD	symmetric approximate minimum degree permutation
SSIM	structural similarity
VIF	visual information fidelity
WLS	weighted least squares

LIST OF SYMBOLS

$\mathbf{M}$	a matrix (bold Latin letter or bold Greek letter)
M	a scalar (regular letter)
$\mathcal{M}$	a tensor (calligraphic letter)
$\boldsymbol{\alpha}$	alpha parameter map
α	alpha parameter of beta distribution
$\tilde{\boldsymbol{\alpha}}$	attenuated alpha parameter map
$\tilde{\boldsymbol{\beta}}$	attenuated beta parameter map
$\mathrm{B}(a, b)$	beta function for inputs a, b
$\Phi(x; \alpha, \beta)$	beta distribution CDF for input x
$\phi(x; \alpha, \beta)$	beta distribution PDF for input x
$\boldsymbol{\beta}$	beta parameter map
β	beta parameter of beta distribution
X_Y	channel Y of a given color when discussing color systems, where Y is an element from the set of labels of the color systems
T	coefficient from the combination of transformations in the MMT color image extension
$\psi(x)$	digamma function for input x
r	downscale ratio
r	downscale ratio
$M_{i,j}$	element from row i and column j of matrix $\mathbf{M}$
$M_{i,j}^{(c)}$	element from row i, column j and slice c of tensor $\mathcal{M}$
$\Gamma(x)$	gamma function for input x
C	global smoothness influence
K	global transformation level
$\mathbf{I}$	$N \times N$ identity matrix
$\mathbf{1}$	$H \times W$ matrix of 1s
H	image height

W	image width
$\mathcal{F}$	input image (color)
$\mathbf{F}$	input image (grayscale)
$\mathscr{L}$	five-point spatially inhomogeneous Laplacian matrix
$\mathbf{Q}$	local smoothness influence
$l_{i,j}$	local transformation level for pixel (i, j)
$\mathbf{M}$	maximum channel
τ	MICF tolerance
$\mathbf{m}$	minimum channel
N	number of pixels
$\acute{\mathcal{F}}$	output image (color)
$\acute{\mathbf{F}}$	output image (grayscale)
p	power parameter of smoothness map generalized mean
γ	power parameter of smoothness map input
η	preservation level
η	preservation level of detail-preserving downscale
$\mathbf{G}$	regional geometric mean
$\bar{\mathbf{G}}$	regional geometric mean of inverted inputs
$\boldsymbol{\mu}$	regional mean
$\mathbf{s}$	regional variance
I_x	regularized incomplete beta function
h	scale parameter of MMT color image extension weights
σ	scale parameter of WLS weights Gaussian blur
z	scale parameter of smoothness map
$\mathbf{S}$	smoothness map
λ_{FGS}	smoothness parameter of FGS algorithm
$\breve{\alpha}$	target alpha parameter
$\breve{\beta}$	target beta parameter
λ	WLS smoothness parameter (trade-off factor of WLS objective function)
$\mathcal{W}$	WLS weights tensor

CONTENTS

1 INTRODUCTION

Contrast enhancement and exposure correction are useful in domestic and technical applications, the latter as a preprocessing step for other technique (e.g., segmentation, object recognition, object tracking etc.) or helping human observations (e.g., medical images, satellite photographs, micrographs etc.) (VIJAYALAKSHMI; NATH; ACHARYA, 2020). For some tasks, a global transformation is sufficient for a satisfactory enhancement. More often, however, a locally adaptive transformation suits better for the task. For example, objects and regions may have very different levels of illumination, physical phenomena may compromise the contrast at some regions but not at others, or it may be desired to have a high detail visibility in all parts of the image. For such cases, local image enhancement methods are preferable.

Although there are many contrast enhancement and exposure correction methods available in the literature, there is no definitive solution to provide a satisfactory result in all situations, and new methods emerge each year. In special, traditional adaptive histogram equalization based methods suffer from checkerboard and staircase effects and from over enhancement.

This dissertation proposes a regional algorithm for contrast enhancement and exposure correction, named Structure-Aware Distribution Stretching (SADS). SADS obtains smooth transformations adapted to the segments in a structure-aware way. Similarly to histogram equalization or histogram specification algorithms, SADS is based on the idea of spreading the distribution of an image channel through the application of an estimate of the cumulative distribution function (CDF) of the channel, followed by the application of the inverse of the CDF of the desired distribution. The central difference from conventional algorithms such as Contrast Limited Adaptive Histogram Equalization (CLAHE) (PIZER et al., 1990) lies in replacing the histogram with a distribution model, namely the beta distribution, and in fitting the model parameters regionally in a structure-aware way using an edge-preserving smoothing (EPS) filter, instead of subdividing the image into tiles and using bilinear interpolation (approach that produces the checkerboard effect).

An advantage of this proposed strategy is that the transformation obtained for a pixel of a given segment depends mostly on the distribution of pixels of that segment, with little influence from neighboring segments, and thus the transformations are better adapted to each segment. Another advantage is the absence of the checkerboard effect resulting from the subdivision into tiles adopted in traditional methods. Yet another advantage is that the transformations obtained are continuous and smooth, avoiding the staircase effect. Figure 1.1 shows the outputs of CLAHE and SADS, demonstrating the presence and absence of checkerboard effect for CLAHE and SADS, respectively.

EPS filters have already been used in methods to correct under exposure (GUO; LI; LING, 2016; YING et al., 2017) and under exposure mixed with over exposure (ZHANG;

Figure 1.1: Outputs of CLAHE and SADS, demonstrating the presence and absence of checkerboard effect for CLAHE and SADS, respectively.

(a) Input (b) CLAHE (c) SADS

Source: TM-DIED dataset (VONIKAKIS, 2021) (1.1a) and author (other images).

NIE; ZHENG, 2019), through the estimation of illumination maps. The proposed method differs from them in the use of the EPS filter to estimate parameter maps of a distribution model, generating a transformation that corrects under and over exposure and/or low contrast.

SADS was experimentally compared with relevant methods from the literature using objective no-reference (NR) and full-reference (FR) image quality assessment (IQA) metrics in the task of simultaneous contrast enhancement and exposure correction and in the task of defogging/dehazing. The experiments indicate a superior overall performance of SADS with respect to the compared methods for the image sets used, according to the IQA metrics adopted.

This dissertation is organized as follows. Chapter 2 (page 17) presents a review of related works. Chapter 3 (page 21) covers necessary theoretical background for this work. Chapter 4 (page 25) describes the proposed method. Chapter 5 (page 47) presents and discusses the experimental results. Chapter 6 (page 58) contains the conclusions and directions for future works. Finally, Appendix A (page 63) states the credits of all input images used in the experiments.

1.1 Objectives

The general objective of this work is to propose a structure-aware image enhancement algorithm based on the distribution of the image values. The specific objectives are:

- propose a method to fit the beta distribution to an image in a structure-aware way;
- propose an algorithm to enhance contrast and correct exposure locally based on such distribution fitting method;
- propose a heuristic to detect smooth or flat segments, which is used in a method to reduce enhancement in such regions;
- propose a color extension method that avoids loss of perceptual colorfulness.

1.2 Notations and Definitions

Bold Latin or bold Greek letters denote matrices. Calligraphic characters denote three-dimensional tensors or sets. Regular characters denote scalars. The elements of a matrix $\mathbf{X}$ are denoted by $X_{i,j}$, where i is the row and j is the column. The frontal (mode-1) slices of a three-dimensional tensor $\mathcal{X}$ are denoted by $\mathbf{X}^{(c)}$, where c is the label of the slice (e.g., r, g and b for an RGB image). The elements of the tensor $\mathcal{X}$ are denoted by $X_{i,j}^{(c)}$, where c is the label of the slice and i and j are the row and column, respectively.

Besides denoting tensor slices, a superscript inside parenthesis, as c in $\mathbf{X}^{(c)}$, is also used to label different but related matrices or scalars.

A single subscript, as i in x_i, may be used to denote ordered variables (e.g., a sequence of samples from a random variable). In this case, the subscript is an index which assumes numeric values. When discussing color systems, X_Y denotes the channel Y of a given color. In this case, the subscript is a non-numeric symbol equal to the label of a channel from a given color system (e.g., B for the blue channel from RGB color system).

In this work, a matrix correspond to a single channel image or to a channel from a multiple-channel image. Tensors correspond to multiple-channel images. H and W are used as the height and width of images, respectively, except when referring to downscaled images in Section 4.8 (page 39). To illustrate, if $\mathcal{F}$ is an RGB image, $\mathbf{F}^{(r)}$, $\mathbf{F}^{(g)}$ and $\mathbf{F}^{(b)}$ are its RGB channels, each channel having height H and width W.

Additionally, an equation containing i, j subscripts denotes a set of $H \times W$ equations, one for each combination of i and j, with $i \in \{1, \ldots, H\}$ and $j \in \{1, \ldots, W\}$, unless specified otherwise. For example, due to this convention, the equation $X_{i,j} = Y_{i,j}$ means that the matrix $\mathbf{X}$ is equal to the matrix $\mathbf{Y}$.

2 RELATED WORKS

This section presents a short review of the state of the art of contrast enhancement and exposure correction on images. The methods have been organized in three categories, namely (1) histogram based methods, (2) learning based methods and (3) other methods, presented next.

2.1 Histogram Based Methods

The level of contrast of a given region of an image is related to the histogram of the intensities of the pixels within this region. If the histogram is narrow, then there is little variability and the contrast is low. If the histogram spreads well over the full intensity range, then the contrast is high. Hence, by stretching the histogram of an image, the contrast is expected to improve. The most "spread" configuration that the histogram can have, which by the principle above would correspond to an excellent contrast, is a uniform (flat) shape, where all levels occur with the same frequency. It can be shown that this is the configuration with maximum entropy of the image levels (WANG; CHEN; ZHANG, 1999). Entropy here means the quantity $-\sum_{k=1}^{N_L} p_k \log_2(p_k)$, where p_k is the ratio between the frequency of occurrence of the discrete level k in the image and the number of pixels of the image, and N_L is the number of discrete image levels.

Viewing the image intensities as samples of a random variable, the normalized histogram of intensities of an image gives the discrete probability density function (PDF) of the intensities. The cumulative sum of the normalized histogram gives the discrete cumulative distribution function (CDF) of the intensities. From statistics theory, applying the CDF of a continuous random variable to the random variable will map it into a new variable with uniform distribution. This is approximately true for the discrete case. Hence, applying to the image the CDF of its intensities (obtained from its histogram) will spread the histogram of the image across the full pixel value range, thus increasing contrast. This method is known as Global Histogram Equalization (GHE), or simply Histogram Equalization (HE).

GHE is able to improve contrast adequately for some images, but it has some limitations. First, the transformation does not adapt to the needs of each region; it is the same for every pixel. Second, the output image histogram will not be exactly uniform, containing in many cases a "spiky" aspect that is accompanied by the production of artificial edges (staircase effect). Third, uniform is not necessarily the best target shape for the histogram of an image, and force it frequently will over enhance the contrast. Despite these limitations, GHE has a very low cost and is fully automatic, not requiring parameter adjustments. Besides, it is the starting point for more sophisticated algorithms.

Local Histogram Equalization (LHE) attempts to overcome the first limitation of GHE

mentioned above. This method moves a window through the image and executes histogram equalization for each position, using in the calculation only the pixels inside the window, and updating only the central pixel (LEE; PANT; LEE, 2015).

Since a histogram equalization is performed for each pixel, LHE is very costly. Adaptive Histogram Equalization (AHE), another local version of GHE, avoids this cost by dividing the image into rectangular tiles, obtaining the HE transformation for each tile separately and computing the output values with a bilinear interpolation of the values resulting from the transformation functions obtained for the neighbor tiles.

One of the main drawbacks of LHE and AHE is that they over enhance the contrast, specially in regions that naturally have little contrast (e.g., a clean sky region). Pizer (PIZER et al., 1987) proposed to overcome this problem for the AHE case by clipping the obtained histograms, which limits the maximum slope of the transformations computed for each tile, originating the method named Contrast Limited Adaptive Histogram Equalization (CLAHE).

The idea of histogram equalization leads to the idea of histogram specification. Histogram specification is a technique for changing the original distribution of an image to a desired distribution by mapping its original pixel values to new values. This is done by applying the CDF of the original distribution and then applying the inverse CDF of the target distribution. The first transformation uniformize the histogram, and the second transformation converts the uniform histogram into the target distribution.

Huang, Cheng and Chiu (2012) proposed to enhance the image in the following way. The histogram of the image is computed and normalized to obtain the discrete PDF $f : \{0, 1, \ldots, 255\} \to [0, \infty)$. A smoothed CDF ρ_s is used as an estimate for the CDF of the pixel values; it is defined by

$$\rho_s : \{0, 1, \ldots, 255\} \to [0, 1], \quad \rho_s(n) = \frac{1}{S} \sum_{k=0}^{n} (f(k))^a, \quad S = \sum_{k=0}^{255} (f(k))^a, \tag{2.1}$$

where $a \in (0, 1)$ is a user-defined parameter. The power transformation applied to the PDF values $f(k)$ reduces the impact of small fluctuations of the PDF in the smoothed CDF, thereby improving the stability and accuracy of this CDF estimate. Let $X_{i,j} \in \{0, 1, \ldots, 255\}$ be the integer pixel values of the image and $x_{i,j}$ the normalized values $x_{i,j} = X_{i,j}/255$. The output normalized image values $y_{i,j}$ are given by

$$y_{i,j} = x_{i,j}^{\gamma(X_{i,j})}, \quad \gamma(X_{i,j}) = 1 - \rho_s(X_{i,j}). \tag{2.2}$$

The results from the work show a smaller degree of over enhancement compared to previous image enhancement algorithms.

J. Lee, Pant and H. Lee (2015) combined CLAHE method with dynamic range compression (MONOBE et al., 2005) and used the local edge density to control the contrast gain. The strategy to control locally the contrast gain provides a significant improvement in comparison to standard CLAHE, avoiding over and under enhancement. The method was also able to boost detail information more than standard CLAHE for medical images.

Chang et al. (2018) combined CLAHE with a dual gamma transformation in order to enhance both contrast and luminance. The authors proposed a heuristic to estimate the optimal CLAHE clipping point based on textureness and dynamic range information, thus avoiding under and over enhancement by CLAHE. The algorithm first redistributes the block histogram in CLAHE using the clip limit points. Then, it enlarges the luminance of image blocks by applying the first gamma correction. Finally, when the image block

contains a large dynamic range, a second gamma correction is applied to compensate for dark regions and avoid over enhancement at bright regions. The results show that, in comparison with CLAHE and others algorithms, the method is more apt to lit up dark regions while preserving naturalness. However, we can observe that some of the results have too little contrast. This can be due to the fact that their algorithm lights up all the dark tones, including those inside reasonably exposed segments, or due to small results for the clipping points provided by their heuristic (if this is the case, however, the clipping points can be increased by changing the parameters of their heuristic).

2.2 Learning Based Methods

Supervised machine learning techniques can be used to learn complex nonlinear transformations in image enhancement algorithms. In this category of methods, the models are usually trained using datasets with examples of non-enhanced images and their enhanced counterparts (QI et al., 2021). Examples of supervised image enhancement methods that emerged in recent years are methods based on fully convolutional networks (FCN), reinforcement learning and U-Net (QI et al., 2021).

Eilertsen et al. (2017) designed a deep convolutional neural network (CNN) model to predict a high dynamic range (HDR) image corresponding to a low dynamic range (LDR) image. Their model is able to recover details missing due to the saturation of camera sensors. Promising results for a wide range of images were achieved. While the objective of their method is to predict HDR images that can be viewed in HDR displays, when the HDR output is tone mapped to a LDR image to be used in regular LDR displays, the resulting LDR image can be seen as an enhanced version of the input image.

Yang et al. (2018) proposed a method named Deep Reciprocating HDR Transformation (DRHT) to enhance images. Their method uses a system with two fully convolutional encoder-decoder networks. The first network receives the LDR image to be enhanced and predicts an HDR image, recovering missing details due to camera saturation in this process (similarly as made in (EILERTSEN et al., 2017)). The second network executes a tone mapping of the HDR output of the first network. The two networks are integrated end-to-end for joint training and prediction. Experiments on standard benchmarks suggest that their method tends to perform favorably against state-of-the-art image correction methods.

Liu et al. (2020) also proposed a supervised system to reconstruct an HDR image from a LDR input. Differently from previous learning based methods, their system uses domain knowledge of the LDR image formation process. More specifically, the authors modeled the HDR to LDR image formation process as (1) clipping, (2) nonlinear mapping from a camera response function and (3) quantization, and designed three specialized CNNs to reverse these steps. They also jointly fine-tuned the entire model end-to-end to reduce error accumulation. They reported that their method obtained satisfactory results in comparison with other state-of-the-art single image HDR reconstruction algorithms.

Wang, Xu and Lau (2022) exploits the local color distributions (LCD) as a prior for locating and enhancing under and over exposed regions. The LCD are used to represent these regions, and a local color distribution embedded module is used to formulate LCD in multi-scales and model the correlations across different regions. A dual-illumination learning mechanism is proposed to enhance the under and over exposed regions.

Unsupervised learning techniques were also used for image enhancement. For example, Chen et al. (2018) proposed an unpaired learning method for image enhancement

based on the two-way generative adversarial networks (GANs) framework. Given a set of images with the desired characteristics, their method learns a photo enhancer which transforms an input image into an enhanced image with those characteristics.

2.3 Other Methods

Yuan and Sun (2012) proposed a method to correct exposure in images. Their algorithm obtains S-shaped transformations adapted to the image needs. They proposed a zone-based region-level optimal exposure evaluation and a method to compensate for the compression of details in the middle tones potentially produced when dark or bright ares are stretched. Their results indicated that their method tends to obtain better performance than popular image editing tools and other previous automatic exposure correction methods.

Guo, Li and Ling (2016) proposed a method based on retinex theory to correct low-light images named Low-Light Image Enhancement (LIME). LIME estimates an illumination map by applying an edge-preserving smoothing filter to the maximum of the R, G and B channels. The algorithm then applies a gamma transformation with $\gamma \in (0, 1)$ to the estimated illumination map, and divides the original image by the result. The algorithm may be interpreted as a multiplication of the original image by a factor that is considerably larger than 1 for dark regions and slightly larger than or equal to 1 for adequately illuminated regions. The gamma transformation controls the degree of enhancement. LIME is very effective at correcting low-light images. Limitations of LIME are the fact that it can not correct over exposure and the fact that the output image can surpass the maximum image level, needing to be clipped, which causes loss of some details.

Ying et al. (2017) proposed a method to correct low-light images with less color and lightness distortion compared to previous methods. Their method uses a camera response model and an estimated exposure map to transform the input image in a retinex theory based way. In the results displayed in their paper, we can observe that their method indeed produced more natural results than the other methods from the comparison.

Zhang, Nie and Zheng (2019) observed that LIME can also be used to correct over exposed images by simply inverting the input image (i.e., applying $f(x) = 1 - x$ to each channel), applying LIME to the inverted image and inverting back the result. The authors then proposed to apply (separately) both original and inverted LIME to an input image, and fuse the results and the input image with an algorithm that detects which image is more appropriate in each location. This way, their method corrects under and over exposure. Their method often generates well-balanced images with corrected exposure, but we can observe that in some situations their fusion method may not select the image corrected for under exposure in locations where it should. Another limitation is the potential image level violation problem, inherited from LIME.

3 THEORETICAL BACKGROUND

The proposed Structure-Aware Distribution Stretching (SADS) method uses the beta distribution and the Weighted Least Squares (WLS) filter. Thus, these two subjects are briefly presented in this chapter. Posteriorly, after the full explanation of SADS method on Chapter 4 (page 25), section 4.11 (page 43) explains why the beta distribution was chosen for SADS. When presenting the SADS color image extension methods, a knowledge of the HSV color system and a definition of it are required, and therefore this topic is covered on section 3.3 (page 24). Additionally, section 3.3 (page 24) defines a color system called iHSV which will also be important in the presentation of the color image extension methods.

3.1 Beta Distribution

The beta distribution is a flexible family of distributions bounded in the interval $[0, 1]$, parameterized by the positive parameters α and β. The PDF of this distribution (family) is

$$\phi(x; \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathrm{B}(\alpha, \beta)}, \tag{3.1}$$

while its CDF is

$$\Phi(x; \alpha, \beta) = I_x(\alpha, \beta), \tag{3.2}$$

where B is the beta function, defined as

$$\mathrm{B}(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1}\, dt, \quad a, b \in \mathbb{R},\ a, b > 0, \tag{3.3}$$

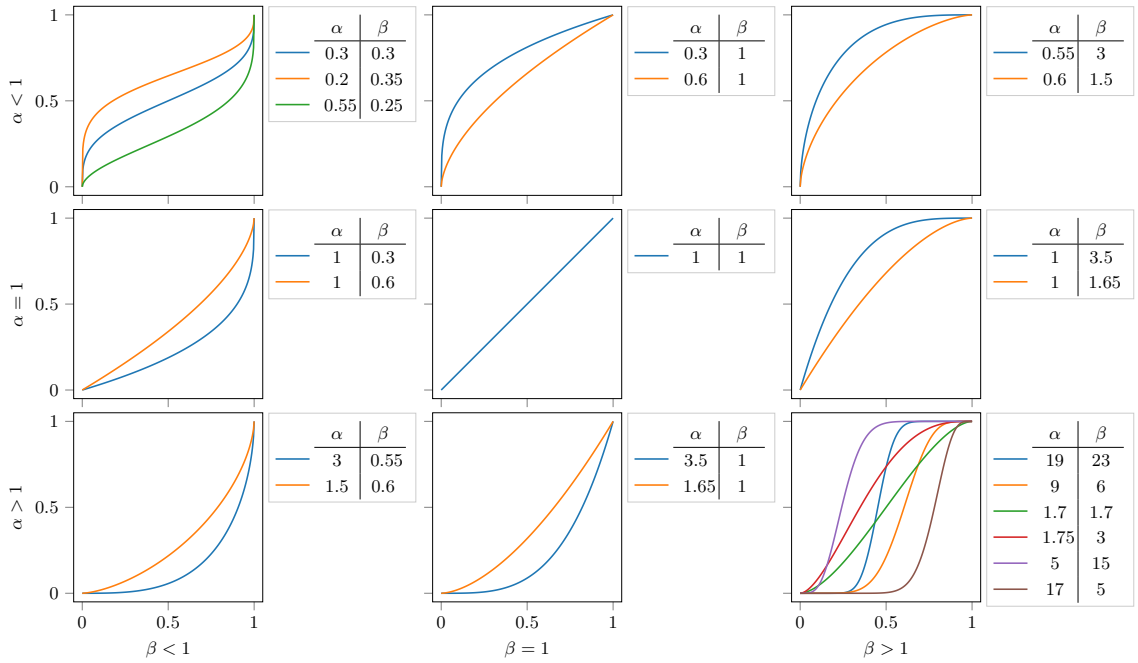
and I_x is the regularized incomplete beta function (GUPTA; NADARAJAH, 2004), defined as

$$I_x(a, b) = \frac{1}{\mathrm{B}(a, b)} \int_0^x t^{a-1}(1-t)^{b-1}\, dt, x \in [0, 1], \quad a, b \in \mathbb{R},\ a, b > 0. \tag{3.4}$$

Figure 3.1 shows graphics of the CDF $\Phi(x; \alpha, \beta)$ for different combinations of α, β. It can be observed that the beta distribution CDF is equal to the identity function for $\alpha = \beta = 1$ and assumes several nonlinear shapes for other combinations.

α and β can be estimated using method of moments (MoM) or maximum likelihood estimation (MLE). Let $\{x_i\}_{i=1}^n$ be n observations of a random variable X. The MoM

Figure 3.1: CDF of members of the beta distribution family.



Source: Author.

expressions for the parameters are (OWEN, 2008)

$$\begin{aligned} \alpha &= \left[\frac{\mu(1-\mu)}{s} - 1\right]\mu, \\ \beta &= \left[\frac{\mu(1-\mu)}{s} - 1\right](1-\mu), \end{aligned} \tag{3.5}$$

where $\mu = \sum_{i=1}^{n} x_i/n$ is the sample mean and $s = \sum_{i=1}^{n}(x_i - \mu)^2/n$ is the sample variance. The MLE method estimates α and β by solving the system (GNANADESIKAN; PINKHAM; HUGHES, 1967)

$$\begin{cases} \psi(\alpha) - \psi(\alpha+\beta) = \log(G), \\ \psi(\beta) - \psi(\alpha+\beta) = \log(\bar{G}), \end{cases} \tag{3.6}$$

where $\log$ is the natural logarithm, ψ is the digamma function and

$$\log(G) = \frac{1}{n}\sum_{i=1}^{n}\log(x_i),\ \log(\bar{G}) = \frac{1}{n}\sum_{i=1}^{n}\log(1-x_i). \tag{3.7}$$

The digamma function is defined as the logarithmic derivative of the gamma function. It is given by:

$$\psi(x) = \frac{d}{dx}\log(\Gamma(x)), \quad x \in \mathcal{U}, \tag{3.8}$$

where $\mathcal{U}$ is the set of the reals less the non-positive integers and $\Gamma(x)$ is the gamma function, defined as

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt, \quad x \in \mathcal{U}. \tag{3.9}$$

3.2 Weighted Least Squares (WLS) Filter

Let $\mathbf{F}$ be a single channel image. The Weighted Least Squares (WLS) (FARBMAN et al., 2008) filter is an EPS technique defined by

$$\mathrm{WLS}(\mathbf{F}, \mathcal{W}) = \arg\min_{\mathbf{U}} J(\mathbf{F}, \mathbf{U}, \mathcal{W}), \tag{3.10}$$

where

$$J(\mathbf{F}, \mathbf{U}, \mathcal{W}) = \sum_{i,j} \{(U_{i,j} - F_{i,j})^2 + \lambda[W^{(x)}_{i,j}(\partial_x \mathbf{U})^2_{i,j} + W^{(y)}_{i,j}(\partial_y \mathbf{U})^2_{i,j}]\} \tag{3.11}$$

and $\mathcal{W}$ is the weights tensor, whose slices are $\mathbf{W}^{(x)}$ and $\mathbf{W}^{(y)}$, formed by the horizontal weights $W^{(x)}_{i,j}$ and the vertical weights $W^{(y)}_{i,j}$, respectively. In (3.10) and (3.11), $\mathbf{U}$ is the dummy variable from the minimization operation. ∂_x and ∂_y are the horizontal and vertical difference operators, respectively, defined by

$$\begin{aligned}
(\partial_x \mathbf{F})_{i,j} &= \begin{cases} F_{i,j+1} - F_{i,j}, & i \in \{1, \ldots, H\};\ j \in \{1, \ldots, W-1\}, \\ 0, & i \in \{1, \ldots, H\};\ j = H, \end{cases} \\
(\partial_y \mathbf{F})_{i,j} &= \begin{cases} F_{i+1,j} - F_{i,j}, & i \in \{1, \ldots, H-1\};\ j \in \{1, \ldots, W\}, \\ 0, & i = H;\ j \in \{1, \ldots, W\}. \end{cases}
\end{aligned} \tag{3.12}$$

To act as an EPS filter, the weights $W^{(d)}_{i,j}$ are designed to be small in edges and large in other regions, making use of the gradients of the input image $\mathbf{F}$. The smoothness degree is controlled globally by the factor λ and locally by the weights $W^{(d)}_{i,j}$.

The objective function (3.11) is convex and has a global minimum, defining a unique solution $\mathbf{O} = \mathrm{WLS}(\mathbf{F}, \mathcal{W})$. Let $\mathbf{o}$ and $\mathbf{f}$ be the column vectors resulting from stacking the columns of $\mathbf{O}$ and $\mathbf{F}$, respectively. $\mathbf{o}$ is the solution to the system (FARBMAN et al., 2008)

$$\mathbf{A}\mathbf{o} = \mathbf{f}, \tag{3.13}$$

$$\mathbf{A} = \mathbf{I} + \lambda \boldsymbol{\mathscr{L}}, \tag{3.14}$$

where $\mathbf{I}$ is an $N \times N$ identity matrix.

$\boldsymbol{\mathscr{L}}$ is the five-point spatially inhomogeneous Laplacian matrix from the WLS weights, defined by

$$\boldsymbol{\mathscr{L}}_{p,q} = \begin{cases} \sum_{k \in \mathcal{N}(p)} w(p,k), & p = q, \\ -w(p,q), & q \in \mathcal{N}(p), \\ 0, & \text{otherwise.} \end{cases} \tag{3.15}$$

In (3.15), $\mathcal{N}(p)$ is the set of indexes of the 4-neighborhood of pixel p and $w(p,q)$ is the WLS weight from pixel p to pixel q. $w(p,q)$ satisfies $w(p,q) = w(q,p)$ and is defined only for $q \in \mathcal{N}(p)$ (or, equivalently, $p \in \mathcal{N}(q)$). Here, "pixel p" means the pixel with row p_i and column p_j given by

$$\begin{aligned}
p_i &= [(p-1) \mod H] + 1, \\
p_j &= \lfloor (p-1)/H \rfloor + 1,
\end{aligned} \tag{3.16}$$

where, as stated before, H is the image height. If $p_i < H$, then $w(p, p+1) = W^{(y)}_{p_i,p_j}$; if $p_j < W$, then $w(p, p+H) = W^{(x)}_{p_i,p_j}$. Therefore, $\boldsymbol{\mathscr{L}}$ is a symmetric sparse matrix formed by five diagonals, with an upper bandwidth of H.

Since $\mathbf{A}$ is sparse, the system in (3.13) is solved using methods for solving sparse systems, such as preconditioned conjugate gradient (PCG).

3.3 The HSV and iHSV Color Systems

HSV is a color system that represents a color in terms of three channels, namely hue, saturation and value. It is useful in many applications that require the manipulation of qualities of a color that are more aligned to human perception, compared to RGB channels (although HSV is not a perceptually uniform color system).

Let X_R, X_G and X_B be the RGB channels of a given color. Let X_M, X_m and X_C be the maximum, minimum and color range channels, defined as

$$X_M = \max(X_R, X_G, X_B), \tag{3.17}$$
$$X_m = \min(X_R, X_G, X_B), \tag{3.18}$$
$$X_C = X_M - X_m. \tag{3.19}$$

Then, the hue X_H, saturation X_S and value X_V of the color are defined by

$$X_V = X_M, \tag{3.20}$$
$$X_S = \begin{cases} 0, & X_M = 0, \\ \frac{X_C}{X_M} = 1 - \frac{X_m}{X_M}, & \text{otherwise}, \end{cases} \tag{3.21}$$
$$X_H = \begin{cases} \text{undefined}, & \text{if } X_C = 0, \\ \frac{X_G - X_B}{X_C} \bmod 6, & \text{if } X_M = X_R, \\ \frac{X_B - X_R}{X_C} + 2, & \text{if } X_M = X_G, \\ \frac{X_R - X_G}{X_C} + 4, & \text{if } X_M = X_B. \end{cases} \tag{3.22}$$

The inverted HSV (iHSV) is defined here as the HSV representation of the inverted color. More precisely, let $X_{\overline{H}}$, $X_{\overline{S}}$ and $X_{\overline{V}}$ be the iHSV channels of the color whose RGB channels are X_R, X_G and X_B. Then, $X_{\overline{H}}$, $X_{\overline{S}}$ and $X_{\overline{V}}$ are the HSV channels of the color with RGB channels $1 - X_R$, $1 - X_G$ and $1 - X_B$. Let $X_{\overline{M}}$, $X_{\overline{m}}$ and $X_{\overline{C}}$ be

$$X_{\overline{M}} = \max(1 - X_R, 1 - X_G, 1 - X_B), \tag{3.23}$$
$$X_{\overline{m}} = \min(1 - X_R, 1 - X_G, 1 - X_B), \tag{3.24}$$
$$X_{\overline{C}} = X_{\overline{M}} - X_{\overline{m}}. \tag{3.25}$$

The following relations hold:

$$X_{\overline{V}} = X_{\overline{M}} = \max_{u \in \{R,G,B\}} \{1 - X_u\} = 1 - X_m, \tag{3.26}$$
$$X_{\overline{m}} = \min_{u \in \{R,G,B\}} \{1 - X_u\} = 1 - X_M, \tag{3.27}$$
$$X_{\overline{C}} = X_{\overline{M}} - X_{\overline{m}} = X_M - X_m = X_C, \tag{3.28}$$
$$X_{\overline{S}} = \begin{cases} 0, & X_m = 1, \\ X_{\overline{C}} / X_{\overline{M}} = X_C / (1 - X_m), & \text{otherwise}, \end{cases} \tag{3.29}$$
$$X_{\overline{H}} = X_H. \tag{3.30}$$

4 PROPOSED METHOD

In this chapter we describe the proposed Structure-Aware Distribution Stretching (SADS) in detail. The central idea of SADS is to fit the beta distribution to image channels and use the fitted distribution to transform the image in a way inspired by the well known theory of histogram specification. Figure 4.1 shows a block diagram for SADS algorithm. The variables and steps present in the diagram will be described throughout the chapter.

The distribution parameters are estimated for the processed channels in a regional structure-aware way. After the estimation, each pixel of each processed channel is assigned to a pair of parameters of the beta distribution. An image or matrix containing the parameters calculated for all pixels is called a parameter map. Here, "regional structure-aware" means that the distribution parameters at a given pixel depend mostly on the content of the segment or region in which the pixel is located, where the segments and regions are separated by the image edges. This regional structure-aware adjust is made by replacing each sample mean in the expressions for the beta distribution parameters estimation with a WLS filter, which operates as a regional structure-aware mean.

As WLS is a global smoothing technique based on the minimization of an objective function, mathematically the estimated parameters at a given pixel depends on the values of the channel for the entire image. However, the term "regional" is justified by the fact that the estimated parameters at a given pixel are more strongly influenced by the content of the segment or region of the pixel, having little or computationally negligible influence from the content of other segments.

When naming variables obtained from regional structure-aware processes, we will use the term "regional" to indicate this property. For example, "regional variance" (which will be proposed later) is a variance obtained from a regional structure-aware process.

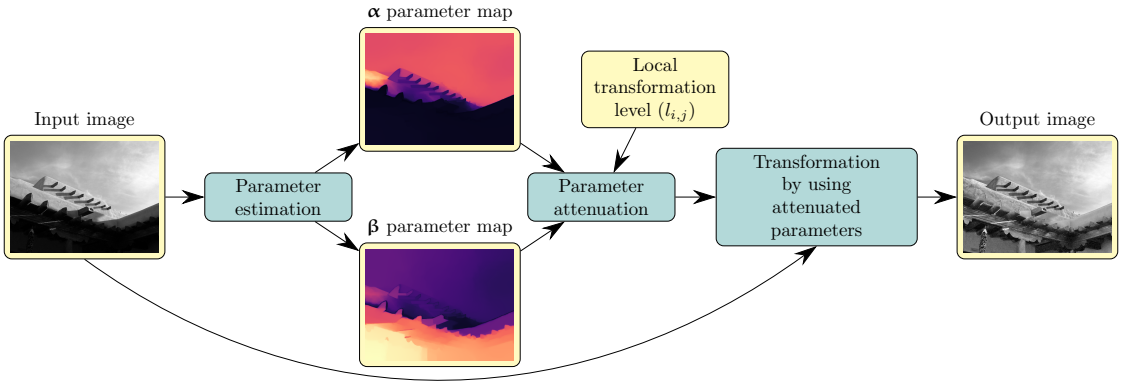
4.1 Designed Weighted Least Squares Filter

If the input image is a single channel grayscale image $\mathbf{F}$, the WLS weights $W_{i,j}^{(d)}$ are computed with

$$\begin{aligned} W_{i,j}^{(x)} &= \frac{1}{\max(\mid [\mathbf{K} * (\partial_x \mathbf{F})]_{i,j} \, (\partial_x \mathbf{F})_{i,j} \mid, \epsilon)}, \\ W_{i,j}^{(y)} &= \frac{1}{\max(\mid [\mathbf{K} * (\partial_y \mathbf{F})]_{i,j} \, (\partial_y \mathbf{F})_{i,j} \mid, \epsilon)}. \end{aligned} \tag{4.1}$$

In (4.1), $\mathbf{K}$ denotes a 2-d Gaussian kernel and $*$ denotes a convolution. ϵ is a small (set to $1 \cdot 10^{-6}$ in our experiments) tolerance term used to prevent division by zero. This formula is similar to the one used in (ZHANG; NIE; ZHENG, 2019), except that we have

Figure 4.1: Block diagram for SADS algorithm.



Source: Author.

removed the outer convolution and used a different strategy for handling the division by zero (in (ZHANG; NIE; ZHENG, 2019) and other works, the denominator of the factors are summed with the tolerance term). The scale parameter σ of the Gaussian kernel $\mathbf{K}$ is a parameter that can be chosen by the user; this parameter controls the degree of smoothing of textures within segments.

If the input image is an RGB color image $\mathcal{F}$, the WLS weights are computed with

$$\begin{aligned} W_{i,j}^{(d)} &= \frac{1}{\max(\sqrt{C_1 C_2}/2, \epsilon)}, \\ C_1 &= [\mathbf{K} * (\partial_j \mathbf{M})]_{i,j}^2 + [\mathbf{K} * (\partial_j \mathbf{m})]_{i,j}^2, \\ C_2 &= (\partial_j \mathbf{M})_{i,j}^2 + (\partial_j \mathbf{m})_{i,j}^2, \\ d &= \{x, y\}, \end{aligned} \tag{4.2}$$

where $\mathbf{M}$ and $\mathbf{m}$ are the maximum and minimum channels, given by

$$M_{i,j} = \max_{c \in \{r,g,b\}} \left\{ F_{i,j}^{(c)} \right\}, \quad m_{i,j} = \min_{c \in \{r,g,b\}} \left\{ F_{i,j}^{(c)} \right\}. \tag{4.3}$$

The factors $\sqrt{C_1/2}$, $\sqrt{C_2/2}$ from (4.2) are extensions of the factors $|[\mathbf{K} * (\partial_j \mathbf{F})]_{i,j}|$ and $|(\partial_j \mathbf{F})_{i,j}|$ from (4.1) that jointly consider the variations along the channels $\mathbf{M}$ and $\mathbf{m}$ (note the use of the properties $\sqrt{ab} = \sqrt{a}\sqrt{b}$ and $|ab| = |a||b|$ to separate the factors from (4.2) and (4.1)). The intention in jointly using the variations of $\mathbf{M}$ and $\mathbf{m}$, instead of using the variations of a single channel extracted from the image, is to improve the capability of the filter to discriminate where it should and should not smooth.

In the implementation, we can replace the division of $\sqrt{C_1 C_2}$ by 2 in (4.2) with a multiplication of ϵ and λ by 2, which will give the same result and will avoid the $2N$ divisions (one for each pixel and weight direction).

As in (GUO; LI; LING, 2016) and (ZHANG; NIE; ZHENG, 2019), the designed weights tend to be large inside segments and small along the edges between segments. The result of the WLS is thus a smoothed version of the filtered image, where the smoothing across the structural image edges is avoided.

4.2 Regionally Fitting the Beta Distribution to an Image

It is well known that the sample mean $\mu = \sum_{i=1}^{n} x_i/n$ satisfies

$$\mu = \arg\min_{\mu^*} \sum_{i=1}^{n} (x_i - \mu^*)^2, \tag{4.4}$$

where μ^* is the dummy variable from the minimization operation (similar to a dummy index in a summation notated with the $\sum$ operator). By replacing x_i in (4.4) with the pixel values of a grayscale image $\mathbf{F}$, allowing the mean to change from pixel to pixel and adding certain smoothness constraints in this variable mean, one arrives at an expression in the form of (3.10) (page 23). This suggests that the WLS filter is a species of regional structure-aware mean, at least in a heuristic sense. This motivates the definition of a (heuristic) regional mean $\boldsymbol{\mu}$ and a regional variance $\mathbf{s}$ by

$$\boldsymbol{\mu} = \mathrm{WLS}(\mathbf{F}, \mathcal{W}), \quad \mathbf{s} = \mathrm{WLS}((\mathbf{F} - \boldsymbol{\mu})^2, \mathcal{W}), \tag{4.5}$$

where the square in the expression for $\boldsymbol{\mu}$ is applied element by element. Note that $\boldsymbol{\mu}$ and $\mathbf{s}$ given by (4.5) are images or matrices with the size of the input image $\mathbf{F}$. As described before, the weights tensor $\mathcal{W}$ is computed with (4.1) (page 25), and is the same for the calculation of $\boldsymbol{\mu}$ and $\mathbf{s}$.

Expressions in (3.5) (page 22) can be converted into regional expressions by replacing the mean and variance with its regional versions, giving

$$\begin{aligned} \alpha_{i,j} &= \left[\frac{\mu_{i,j}(1-\mu_{i,j})}{s_{i,j}} - 1\right]\mu_{i,j}, \\ \beta_{i,j} &= \left[\frac{\mu_{i,j}(1-\mu_{i,j})}{s_{i,j}} - 1\right](1-\mu_{i,j}). \end{aligned} \tag{4.6}$$

This provides a regional, structure-aware, method of moments (MoM) based estimate of α and β.

The maximum likelihood estimation (MLE) method can also be converted into a regional form. For convenience, consider the modification of the system in (3.6) (page 22) by applying an exponential on each side, written as

$$\begin{cases} \exp(\psi(\alpha) - \psi(\alpha+\beta)) = G = \exp(\frac{1}{n}\sum_{i=1}^{n}\log(x_i)), \\ \exp(\psi(\beta) - \psi(\alpha+\beta)) = \bar{G} = \exp(\frac{1}{n}\sum_{i=1}^{n}\log(1-x_i)). \end{cases} \tag{4.7}$$

Using again the idea of replacing the sample mean with a WLS filter, we obtain the regional version of this system

$$\begin{cases} \exp(\psi(\alpha_{i,j}) - \psi(\alpha_{i,j}+\beta_{i,j})) = G_{i,j}, \\ \exp(\psi(\beta_{i,j}) - \psi(\alpha_{i,j}+\beta_{i,j})) = \bar{G}_{i,j}, \end{cases} \tag{4.8}$$

where

$$\begin{aligned} \mathbf{G} &= \exp(\mathrm{WLS}(\log(\mathbf{F}), \mathcal{W})), \\ \bar{\mathbf{G}} &= \exp(\mathrm{WLS}(\log(\mathbf{1}-\mathbf{F}), \mathcal{W})), \end{aligned} \tag{4.9}$$

being $\mathbf{1}$ a $H \times W$ matrix of 1s. As G and $\bar{G}$ in (3.6) (page 22) and (4.7) are the sample geometric mean of samples $\{x_i\}_{i=1}^n$ and $\{1-x_i\}_{i=1}^n$, respectively, $\mathbf{G}$ and $\bar{\mathbf{G}}$ will be called the regional geometric mean of $\mathbf{F}$ and $\mathbf{1}-\mathbf{F}$, respectively.

The solution to the systems (3.6) (page 22) or (4.7) for α and β requires a numerical method, e.g., Newton–Raphson method. To remove the need to use one of such methods for each pixel in (4.8), a look-up table (LUT) was built having G and $\bar{G}$ as inputs and α as output. This was facilitated by the fact that $G, \bar{G} \in [0, 1]$ (it would be less viable if $G, \bar{G}$ were unbounded). The entries of the LUT were computed with Newton–Raphson method and saved to be loaded by the proposed algorithm. The chosen grid size of the LUT was 1024×1024, where only the entries at the triangular region $G + \bar{G} \leq 1, G, \bar{G} > 0$ are calculated (because $G + \bar{G} \leq 1$ is always valid). For the application of the LUT, its entries are interpolated with spline 2-d interpolation.

Having a LUT for α, there is no need to build one for β, for the following reason. If α and β are interchanged in (4.7), then the new system is equivalent to the original, except that G and $\bar{G}$ switch places. As a consequence, if $\alpha = \mathrm{LUT}(G, \bar{G})$, then $\beta = \mathrm{LUT}(\bar{G}, G)$, where LUT denotes the LUT mapping, and a LUT designed to compute one of the parameter can compute the other by switching its indexes. The explicit expressions for the parameter maps are

$$\boldsymbol{\alpha} = \mathrm{LUT}(\mathbf{G}, \bar{\mathbf{G}}), \quad \boldsymbol{\beta} = \mathrm{LUT}(\bar{\mathbf{G}}, \mathbf{G}). \tag{4.10}$$

It is also possible to improve the results given by the LUT by using an iterative method to solve the MLE equations (4.8) for each pixel, having the LUT results as starting points. However, the approximation given by the designed LUT with spline interpolation is already sufficiently accurate for the algorithm.

Figure 4.2 shows results of fitting the beta distribution by using the proposed regional MoM and MLE techniques for a grayscale image. Figure 4.2a shows the graphics of the CDF at selected points. Figure 4.2b shows the full parameter maps in the same experiment, for the MLE case.

4.3 Transforming the Image

SADS is based on the well known histogram specification principle, which consists into apply the estimated CDF to spread the PDF and then apply the inverse of the CDF of the target (desired) distribution. First, consider the case in which the target distribution is the uniform distribution in the interval $[0, 1]$ $U(0, 1)$ (for which the second transformation is simply the identity function, that we do not need to apply). By the principle, the transformation would be simply $F_{i,j} \mapsto \Phi(F_{i,j}; \alpha_{i,j}, \beta_{i,j})$ (see (3.2) (page 21)), where $\alpha_{i,j}, \beta_{i,j}$ are the parameters estimated using the regional MoM or MLE described in the previous section. However, in order to provide a control over the degree of transformation, we propose the following modification. The output image $\acute{\mathbf{F}}$ is calculated with

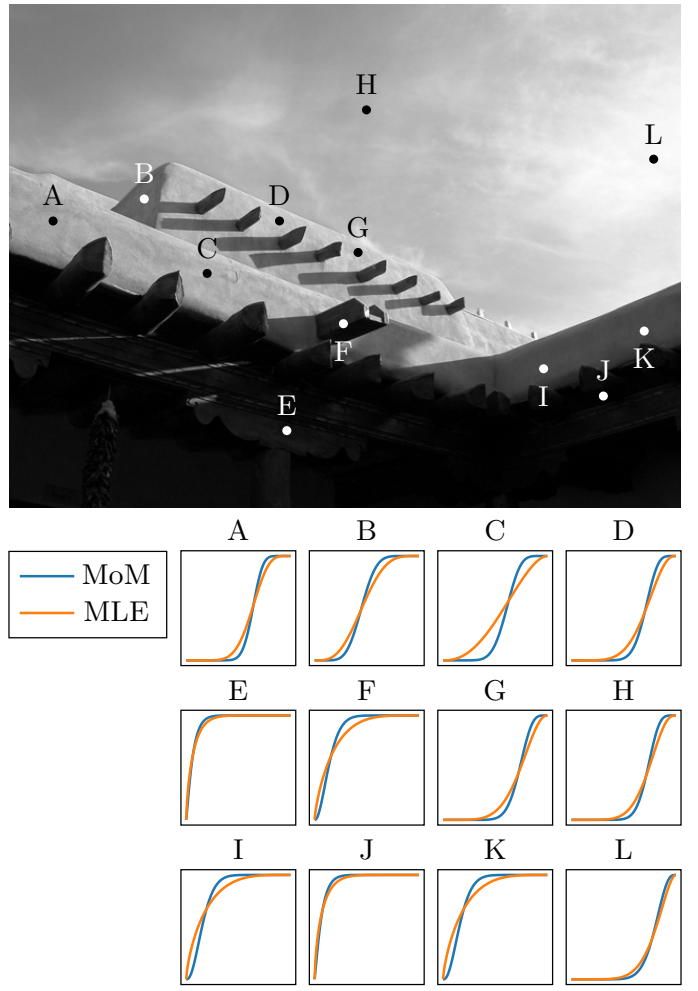
$$\acute{F}_{i,j} = \Phi(F_{i,j}; \tilde{\alpha}_{i,j}, \tilde{\beta}_{i,j}), \tag{4.11}$$

where $\tilde{\alpha}_{i,j}, \tilde{\beta}_{i,j}$, called attenuated parameters, are defined by

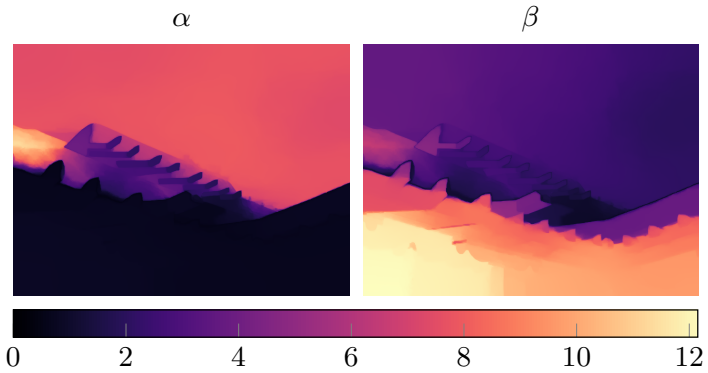
$$\tilde{\alpha}_{i,j} = \alpha_{i,j}^{l_{i,j}}, \quad \tilde{\beta}_{i,j} = \beta_{i,j}^{l_{i,j}}. \tag{4.12}$$

Following the conventions in this work, the matrices with elements $\alpha_{i,j}, \beta_{i,j}$ are denoted by $\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\beta}}$. $l_{i,j} \in [0, 1]$ in (4.12) is the local transformation level at pixel (i, j), which will be described in section 4.5 (page 32). By raising the beta distribution parameters to $l_{i,j}$, they are compressed in direction to 1. This reduces the degree of transformation, because the beta distribution CDF gets closer to the identity function if its parameters become

Figure 4.2: Results of fitting the beta distribution by using the proposed regional MoM and MLE techniques for a grayscale image.



(a) CDF at selected points.



(b) Full parameter maps for the regional MLE case.

Source: Author.

closer to 1. If $l_{i,j} = 0$, the parameters are equal to 1 and the transformation is an identity function. The closer $l_{i,j}$ is to 1 inside $\in [0, 1]$, the higher is the degree of transformation. This motivates the name used for $l_{i,j}$.

Now consider the case in which the target distribution is not $U(0, 1)$. We will restrict ourselves to the case in which the target distribution is a member of the beta distribution family (which is probably wide enough for most applications). Let $\breve{\alpha}, \breve{\beta}$ be the parameters of the target beta distribution. Again, let $\acute{\mathbf{F}}$ denote the output image. For this case, we propose to map the image by the transformations

$$\bar{F}_{i,j} = \Phi(F_{i,j}; \tilde{\alpha}_{i,j}, \tilde{\beta}_{i,j}), \tag{4.13}$$

$$\acute{F}_{i,j} = \Phi^{-1}(\bar{F}_{i,j}; \breve{\alpha}, \breve{\beta}), \tag{4.14}$$

where

$$\tilde{\alpha}_{i,j} = (\alpha_{i,j}/\breve{\alpha})^{l_{i,j}}\breve{\alpha}, \quad \tilde{\beta}_{i,j} = (\beta_{i,j}/\breve{\beta})^{l_{i,j}}\breve{\beta}. \tag{4.15}$$

Again, $l_{i,j}$ is used to control the transformation degree. The working principle of (4.15) is the following. If $l_{i,j} = 1$, then $(\tilde{\alpha}_{i,j}, \tilde{\beta}_{i,j}) = (\alpha_{i,j}, \beta_{i,j})$ and the transformation has full degree. If $l_{i,j} = 0$, then $(\tilde{\alpha}_{i,j}, \tilde{\beta}_{i,j}) = (\breve{\alpha}, \breve{\beta})$ and the transformation in (4.14) is the inverse of the transformation in (4.13), and hence the transformation has zero degree ($\acute{F}_{i,j} = F_{i,j}$). If $l_{i,j} \in (0, 1)$, then the resulting transformation is between these two extremes, which corresponds to an attenuated transformation (with respect to the full degree transformation). $\acute{\mathbf{F}}$ given by the transformations (4.13) and (4.14) is equivalent to $\acute{\mathbf{F}}$ given by (4.11) (page 28) if $\breve{\alpha} = \breve{\beta} = 1$.

The transformation (4.14) requires the calculation of Φ^{-1}, which is made with numerical methods and is more computationally intensive than the calculation of Φ. However, as in (4.14) $\breve{\alpha}, \breve{\beta}$ do not vary with i, we can build an 1-d LUT having x as input and $\Phi^{-1}(x; \breve{\alpha}, \breve{\beta})$ as output and apply this LUT to the values $\bar{F}_{i,j}$.

4.4 Smoothness Map

The smoothness map $\mathbf{S}$ is an image with values in the interval $[0, 1]$ set to have a large value at smooth regions in the image and low value at non-smooth regions. For a grayscale input $\mathbf{F}$, $\mathbf{S}$ is defined as

$$\begin{aligned} \mathbf{S} &= \exp(-\mathbf{D}/z), \\ \mathbf{D} &= \mathrm{WLS}(\mathbf{T}^p, \mathcal{W})^{1/p}, \end{aligned} \tag{4.16}$$

where $T_{i,j}$ is the squared magnitude of the gradient of F^γ at pixel i and p, z and γ are user-defined parameters. p, z and γ are positive. In our experiments, p is set to 0.025 and z is set to $1 \cdot 10^{-5}$ for defogging/dehazing and to $8 \cdot 10^{-4}$ for the other tasks (except in the calculation of Figure 4.3b, in which $z = 1.5 \cdot 10^{-3}$ was adopted). The powers and the exponential are applied element by element. $\mathbf{T}$ can be computed using standard well known gradient operators; the chosen one was the Sobel operator, because of its simplicity and effectiveness. The idea behind (4.16) is the following. $\mathbf{D}$ is a regional version of the generalized or power mean, defined as

$$M_p(x_1, \ldots, x_n) = \left(\frac{1}{n} \sum_{i=1}^{n} x_i^p \right)^{1/p}. \tag{4.17}$$

Figure 4.3: Smoothness map $\mathbf{S}$ for a color image.

(a) Input (b) Smoothness map

Source: pexels.com (4.3a) and author (4.3b).

As M_p satisfy $M_p < M_q \iff p < q$, M_1 is the arithmetic mean and $\lim_{t\to 0} M_t$ is the geometric mean, if p is very small and positive, then M_p will be between the arithmetic and geometric means, and close to the geometric mean. With such choice, the smallest gradients will have a larger influence on $\mathbf{D}$. The practical consequence is that $\mathbf{D}$ will be very small for a smooth segment, even if it contains large isolated gradient perturbations (e.g., a bird flying in the sky). $p < 0$ or the limit case $p \to 0$ would not be suited for the heuristic, because in these cases a single sample equal to zero makes M_p be zero, and $\mathbf{D}$ would inherit this characteristic. The exponential that maps $\mathbf{D}$ into $\mathbf{S}$ is a fuzzy indicator of how small is $\mathbf{D}$. $\mathbf{F}$ is raised to the power γ before the gradient calculation for the following reason. If $\gamma \in (0, 1)$, the gradients in the dark (bright) areas become larger (smaller); this is useful to avoid a large smoothness map $\mathbf{S}$ at areas in the image that are not flat but are so underexposed that its gradients are small. If $\gamma > 1$, the gradients in the bright (dark) areas become larger (smaller); this is useful to avoid a large smoothness map $\mathbf{S}$ at areas in the image that are not flat but are so overexposed that its gradients are small.

For a color input image $\mathcal{F}$, the smoothness map $\mathbf{S}$ is still computed using (4.16), except that the $T_{i,j}$ in now given by

$$T_{i,j} = (T_{i,j}^{(r)} + T_{i,j}^{(g)} + T_{i,j}^{(b)})/3, \tag{4.18}$$

where $T_{i,j}^{(c)}$ is the squared magnitude of the gradient of the channel c of $\mathcal{F}^\gamma$ at pixel (i, j). If we divide and multiply $T_{i,j}$ by an arbitrary number x in the expression for $\mathbf{D}$ in (4.16), we obtain

$$\begin{aligned} \mathbf{D} &= \mathrm{WLS}((x\mathbf{T}/x)^p, \mathcal{W})^{1/p} \\ &= \mathrm{WLS}((x\mathbf{T})^p/x^p, \mathcal{W})^{1/p} \\ &= [\mathrm{WLS}((x\mathbf{T})^p, \mathcal{W})/x^p]^{1/p} \\ &= \mathrm{WLS}((x\mathbf{T})^p, \mathcal{W})^{1/p}/x. \end{aligned} \tag{4.19}$$

As a consequence, the division by 3 in (4.18) can be replaced with a multiplication of the scale parameter z with 3, which is more efficient, because a single multiplication will be made instead of N divisions (one for each pixel). Figure 4.3 shows the smoothness map obtained for a color image with the described method.

The smoothness map is used to control the enhancement level according to the smoothness level. The higher the smoothness map, the higher the reduction of the local enhancement level relative to the non attenuated enhancement level. In practice, very dark regions may produce high values in the smoothness map, which would undesirably reduce the correction degree at such regions if the smoothness map was directly used to attenuate the transformation. To solve this conflict, the influence of the smoothness map is reduced at dark regions. This is made by using a factor called local smoothness influence, with symbol $\mathbf{Q}$. $\mathbf{Q}$ is given by

$$\begin{aligned} \mathbf{Q} &= \mathrm{WLS}(\mathbf{Q}', \mathcal{W}), \\ Q'_{i,j} &= \begin{cases} 1, & \text{if } L_{i,j} = \epsilon_l \text{ or } L_{i,j} = \epsilon_h, \\ \frac{1}{1+\exp(-10(L_{i,j}-0.5))}, & \text{otherwise,} \end{cases} \\ & i = 1, \ldots, H,\ j = 1, \ldots, W, \end{aligned} \tag{4.20}$$

where $\mathbf{L}$ is the gray level if the SADS input is a grayscale image and the lightness channel, defined by $\mathbf{L} = (\mathbf{M} + \mathbf{m})/2$ (see (4.3) (page 26)), if the input is a color image. ϵ_l and ϵ_h are the lowest and highest gray level, respectively, which are the values to which the image levels 0 and 255 (or 65535, for images with 16 bit by channel) are mapped when the image values are converted from integers to decimals in the interval $[0, 1]$. The expression that results in $\mathbf{Q}'$ is a logistic function on the lightness that results in low values (near 0) for dark pixels and high values (near 1) for bright pixels, except when the lightness has the lower or higher level possible, case in which $\mathbf{Q}'$ is 1. The exception is made to ensure that the local smoothness influence $\mathbf{Q}$ do not reduce the influence of the smoothness map in flat black or white areas. While the $\mathbf{Q}'$ design presented was the used in the experiments in this work, other designs of $\mathbf{Q}'$ are possible; the design must produce a predominantly low $\mathbf{Q}'$ (high $\mathbf{Q}'$) in the regions where it is desirable to reduce (keep) the influence from the smoothness map.

4.5 Local Transformation Level

The local transformation level $l_{i,j}$ is given by

$$l_{i,j} = K(1 - CQ_{i,j}S_{i,j}), \tag{4.21}$$

where K and C are user-defined parameters. K and C will be called global transformation level and global smoothness influence, respectively. As the names suggest, K controls the level of transformation globally, while C controls how much $Q_{i,j}S_{i,j}$ reduces $l_{i,j}$ (i.e., how much the transformation is reduced in smooth regions with high $\mathbf{Q}$). The higher K is, the higher the transformation level, and the higher C is, the higher the smoothness influence. Both K and C are primarily in the interval $[0, 1]$. However, it is also possible, and perhaps useful, to use a K larger than 1. If $K > 1$, then $l_{i,j}$ can be larger than 1, and if so, the transformation is intensified instead of attenuated. Normally, however, the full degree transformation (corresponding to $K = 1$, $C = 0$) is sufficiently enhanced or over enhanced, and then normally $K \in (0, 1]$. K near 0.8 provides a strong correction and K near 0.5 provides a lighter correction; the choice depends on the user interests.

4.6 Extension of SADS to Color Images

When the image to be enhanced is a color image $\mathcal{F}$, we must choose a color system to represent $\mathcal{F}$ and choose the channels of this color system that will be enhanced. The

Figure 4.4: Results of SADS using LCT extension with different color systems and MMT extension.

(a) Input (b) HSV (c) iHSV (d) HSL (e) CIELAB (f) MMT

Source: TM-DIED dataset (VONIKAKIS, 2021) (4.4a, row 1), pexels.com (4.4a, row 2) and author (other images).

steps of the grayscale SADS are then made for each channel to be enhanced. As many color systems and many color image enhancement strategies exist, there are many ways to extend SADS to color images. Two methods will be presented. The first is a simple luminance enhancement, and the second is a combination of strategies.

4.6.1 Luminance Channel Transformation (LCT)

The more direct way to extend SADS to color images consists in:

1. represent the image in a color system with a luminance[1] channel and two color channels, such as HSV, HSL or CIELAB;

2. apply to the luminance channel of the color system the steps of the SADS algorithm proposed for grayscale images;

3. convert back to RGB the image with transformed luminance.

This approach will be referred to as luminance channel transformation (LCT). The only steps of the grayscale SADS algorithm that are modified in LCT extension are the calculation of the WLS weights and the smoothness map, which follow the procedures described before for the calculation of $\mathcal{W}$ and $\mathbf{S}$ in the color image case.

Figure 4.4 shows the outputs of SADS for different color extension possibilities. Figures 4.4b to 4.4e show the outputs for the LCT extension using multiple color systems. It can be observed that most tested color systems produce a grayish aspect on the enhanced dark or bright regions of the LCT result. This problem will be expanded on in subsection 4.6.2. Among the LCT extension possibilities, the HSL-based provided the most sactisfatory results in these examples. However, as the other color systems, HSL can also produce grayish tones.

4.6.2 Maximum and Minimum Transformation (MMT)

When SADS or conventional histogram equalization methods are applied to color images by mapping the value X_V of HSV system, the corrected results for overexposed

[1]Here the term "luminance" is referring to the non-chromatic channel of a color system (HSV value, HSL lightness etc.), and is *not* being used with its original physical definition.

regions tend to have an unsaturated or grayish aspect. Before the explanation for why this happens, first consider the following lemma and corollary.

Lemma 1. *Let X_R, X_G and X_B be the RGB channels of a given color. Let X_H, X_S and X_V be the HSV channels of the same color. Let X'_R, X'_G and X'_B be the RGB channels of the color whose HSV channels are X_H, X_S and X'_V. Let $X_M = \max(X_R, X_G, X_B)$, $X_m = \min(X_R, X_G, X_B)$, $X'_M = \max(X'_R, X'_G, X'_B)$ and $X'_m = \min(X'_R, X'_G, X'_B)$. Let $X_C = X_M - X_m$ and $X'_C = X'_M - X'_m$. Then, if $X_V \neq 0$, it is the case that*

$$X'_C = \frac{X'_V}{X_V} X_C. \tag{4.22}$$

It is also the case that

$$X'_m = \frac{X'_M}{X_M} X_m. \tag{4.23}$$

Proof. As the HSV saturation is the same for both colors, by using its definition (see section 3.3 (page 24)), we obtain

$$\frac{X_C}{X_V} = \frac{X'_C}{X'_V}, \tag{4.24}$$

which gives

$$X'_C = \frac{X'_V}{X_V} X_C. \tag{4.25}$$

As $X_V = X_M$ and $X'_V = X'_M$ (see section 3.3 (page 24)), we obtain

$$\begin{aligned} X'_M - X'_m &= \frac{X'_M}{X_M}(X_M - X_m) \\ X'_M - X'_m &= X'_M - \frac{X'_M}{X_M} X_m \\ \therefore X'_m &= \frac{X'_M}{X_M} X_m. \end{aligned} \tag{4.26}$$

□

Corollary 1. *Let X_R, X_G and X_B be the RGB channels of a given color. Let $X_{\overline{H}}$, $X_{\overline{S}}$ and $X_{\overline{V}}$ be the iHSV channels of the same color. Let X'_R, X'_G and X'_B be the RGB channels of the color whose HSV channels are $X_{\overline{H}}$, $X_{\overline{S}}$ and $X'_{\overline{V}}$. Let $X_M = \max(X_R, X_G, X_B)$, $X_m = \min(X_R, X_G, X_B)$, $X'_M = \max(X'_R, X'_G, X'_B)$ and $X'_m = \min(X'_R, X'_G, X'_B)$. Then, if $X_m \neq 1$, it is the case that*

$$X'_M = 1 - \frac{1 - X'_m}{1 - X_m}(1 - X_M). \tag{4.27}$$

Proof. Let $X_{\overline{M}} = \max(1 - X_R, 1 - X_G, 1 - X_B)$, $X_{\overline{m}} = \min(1 - X_R, 1 - X_G, 1 - X_B)$, $X'_{\overline{M}} = \max(1 - X'_R, 1 - X'_G, 1 - X'_B)$ and $X'_{\overline{m}} = \min(1 - X'_R, 1 - X'_G, 1 - X'_B)$. As the iHSV channels are the HSV channels of the inverted color (see section 3.3 (page 24)), from the corollary statement, $X_{\overline{H}}$, $X_{\overline{S}}$ and $X_{\overline{V}}$ are the HSV channels of the color with RGB channels $1 - X_R$, $1 - X_G$ and $1 - X_B$. Similarly, $X_{\overline{H}}$, $X_{\overline{S}}$ and $X'_{\overline{V}}$ are the HSV channels of the color with RGB channels $1 - X'_R$, $1 - X'_G$ and $1 - X'_B$. As only the HSV value channel is different for the two inverted colors, Lemma 1 can be applied, giving

$$X'_{\overline{m}} = \frac{X'_{\overline{M}}}{X_{\overline{M}}} X_{\overline{m}}. \tag{4.28}$$

It is easy to show that

$$\begin{aligned} X_{\overline{M}} &= 1 - X_m, \\ X_{\overline{m}} &= 1 - X_M, \\ X'_{\overline{M}} &= 1 - X'_m, \\ X'_{\overline{m}} &= 1 - X'_M. \end{aligned} \tag{4.29}$$

Replace (4.29) in (4.28) yields

$$\begin{aligned} 1 - X'_M &= \frac{1 - X'_m}{1 - X_m}(1 - X_M) \\ \therefore X'_M &= 1 - \frac{1 - X'_m}{1 - X_m}(1 - X_M). \end{aligned} \tag{4.30}$$

□

The color range $X_C = X_M - X_m$ is a good measure for the color's intensity or colorfulness. This is the case because, when X_C is small, the color is close to the gray axis $R = G = B$, tending to have a grayish tone. As shown in Lemma 1, when a transformation is applied to the HSV value while keeping the HSV saturation and hue unchanged, the color range is multiplied by X'_V/X_V, where X_V and X'_V are the value and the new value, respectively. Hence, if X_V is reduced to correct an overexposed pixel, its color range is also reduced, which can cause the sensation of lack of colorfulness.

This problem of grayish aspect also happens if the transformation is made in the luminance channel of other conventional color systems, such as CIELAB and HSL, where for each color system it happens in specific situations. For example, if we use the iHSV color system described on section 3.3 (page 24), the problem will occur for the enhanced underexposed regions, which is the opposite situation with respect to the HSV system. For the HSL system, the color range is reduced when $\min(X'_L, 1 - X'_L)/\min(X_L, 1 - X_L) < 1$, where X_L and X'_L are the original and new HSL luminances, and when this happens blackish or whitish tones may be perceived. The CIELAB color system is approximately perceptually uniform; as dark and bright areas are perceived as having low colorfulness, when the CIELAB luminance channel is corrected this perceptual lack of colorfulness will be conserved, and hence the CIELAB enhancement result tend to look grayish in the originally dark and bright areas. This phenomenon can be seen in Figure 4.4 (page 33).

The second method of extension of SADS to color images was designed to overcome this problem. It consists in combining three hue-preserving transformations: the HSV and iHSV-based transformations and an additional transformation that will be detailed later. The HSV transformation spreads the distribution of the maximum channel $\mathbf{M}$ while preserving the HSV saturation, the iHSV transformation spreads the distribution of the minimum channel $\mathbf{m}$ while preserving the iHSV saturation and the third transformation spreads both $\mathbf{M}$ and $\mathbf{m}$ distributions by using a joint distribution model for these channels. This approach will be referred to as maximum and minimum transformation (MMT), and is described next.

First, $\mathbf{M}$ and $\mathbf{m}$ channels (see (4.3) (page 26)) are extracted from $\mathcal{F}$. Then, the regional estimation of the beta distribution parameters is made in parallel for $\mathbf{M}$ and $\mathbf{m}$. Let $\boldsymbol{\alpha}^{(M)}, \boldsymbol{\beta}^{(M)}$ and $\boldsymbol{\alpha}^{(m)}, \boldsymbol{\beta}^{(m)}$ be the parameter maps estimated for $\mathbf{M}$ and $\mathbf{m}$, respectively.

If MoM is chosen, the expressions for the parameters are

$$\begin{aligned}\alpha_{i,j} &= \left[\frac{\mu_{i,j}(1-\mu_{i,j})}{s_{i,j}} - 1\right]\mu_{i,j},\\ \beta_{i,j} &= \left[\frac{\mu_{i,j}(1-\mu_{i,j})}{s_{i,j}} - 1\right](1-\mu_{i,j}),\end{aligned} \tag{4.31}$$

where

$$\begin{aligned}\boldsymbol{\mu}^{(M)} &= \mathrm{WLS}(\mathbf{M}, \mathcal{W}),\\ \boldsymbol{\mu}^{(m)} &= \mathrm{WLS}(\mathbf{m}, \mathcal{W}),\\ \mathbf{s}^{(M)} &= \mathrm{WLS}((\mathbf{M}-\boldsymbol{\mu}^{(M)})^2, \mathcal{W}),\\ \mathbf{s}^{(m)} &= \mathrm{WLS}((\mathbf{m}-\boldsymbol{\mu}^{(m)})^2, \mathcal{W}).\end{aligned} \tag{4.32}$$

If MLE is chosen, the expressions are

$$\begin{aligned}\alpha_{i,j}^{(c)} &= \mathrm{LUT}(G_{i,j}^{(c)}, \bar{G}_{i,j}^{(c)}),\\ \beta_{i,j}^{(c)} &= \mathrm{LUT}(\bar{G}_{i,j}^{(c)}, G_{i,j}^{(c)}),\\ c &= \{M, m\},\end{aligned} \tag{4.33}$$

where

$$\begin{aligned}\mathbf{G}^{(M)} &= \exp(\mathrm{WLS}(\log(\mathbf{M}), \mathcal{W})),\\ \bar{\mathbf{G}}^{(M)} &= \exp(\mathrm{WLS}(\log(\mathbf{1}-\mathbf{M}), \mathcal{W})),\\ \mathbf{G}^{(m)} &= \exp(\mathrm{WLS}(\log(\mathbf{m}), \mathcal{W})),\\ \bar{\mathbf{G}}^{(m)} &= \exp(\mathrm{WLS}(\log(\mathbf{1}-\mathbf{m}), \mathcal{W})).\end{aligned} \tag{4.34}$$

The attenuated parameters are

$$\tilde{\alpha}_{i,j} = \alpha_{i,j}^{l_{i,j}}, \quad \tilde{\beta}_{i,j} = \beta_{i,j}^{l_{i,j}}, \tag{4.35}$$

where $l_{i,j}$ is given by (4.21) (page 32).

Now we define the transformation. For compactness, the symbol $\Phi_{i,j}^{X,Y}$ is defined as

$$\Phi_{i,j}^{X,Y} = \Phi(X_{i,j}, \tilde{\alpha}_{i,j}^{(Y)}, \tilde{\beta}_{i,j}^{(Y)}), \quad X, Y \in \{M, m\} \tag{4.36}$$

(see (3.2) (page 21) and (4.11) (page 28)). $\Phi_{i,j}^{X,Y}$ is the result of the application of the beta CDF to X using the parameters estimated for Y, $X, Y \in \{M, m\}$.

The results $\acute{M}_{i,j}, \acute{m}_{i,j}$ of the transformations of $M_{i,j}, m_{i,j}$ need to satisfy two conditions: (i) $\acute{M}_{i,j} = \acute{m}_{i,j} \iff M_{i,j} = m_{i,j}$ and (ii) $\acute{M}_{i,j} \geq \acute{m}_{i,j}$. Condition (i) ensures that a color with nonzero color range is mapped into a color with nonzero color range, while a color with zero color range is mapped into a pixel with zero color range. Condition (ii) ensures that the new maximum is larger than the new minimum, which is necessary for the validity of the new color. A natural pair of transformations to apply would be $\acute{M}_{i,j} = \Phi_{i,j}^{M,M}$ and $\acute{m}_{i,j} = \Phi_{i,j}^{m,m}$, which consists in applying the beta CDF to $M_{i,j}$ and $m_{i,j}$ using the parameters estimated for $\mathbf{M}$ and $\mathbf{m}$, respectively. However, this choice would not meet the conditions (i) and (ii), and therefore cannot be used. Instead, the method here described combines three transformations satisfying the conditions. The outputs of these transformations will be denoted by $\acute{M}_{i,j}^{(u)}$ and $\acute{m}_{i,j}^{(u)}$, for $u \in \{1, 2, 3\}$.

The first transformation operates in the HSV system. The HSV value, which is equal to $M_{i,j}$, is transformed to $\Phi_{i,j}^{M,M}$, while the HSV saturation and hue are preserved. Thus, from Lemma 1 (page 34), $M_{i,j}$ and $m_{i,j}$ are mapped to

$$\begin{aligned} \acute{M}_{i,j}^{(1)} &= \Phi_{i,j}^{M,M}, \\ \acute{m}_{i,j}^{(1)} &= \frac{\Phi_{i,j}^{M,M}}{M_{i,j}} m_{i,j}. \end{aligned} \tag{4.37}$$

The second transformation operates in the iHSV system (see section 3.3 (page 24)). The iHSV value, which is equal to $1 - m_{i,j}$, is transformed to $1 - \Phi_{i,j}^{m,m}$, while the iHSV saturation and hue are preserved. From the Lemma 1 (page 34), $M_{i,j}$ and $m_{i,j}$ are mapped to

$$\begin{aligned} \acute{M}_{i,j}^{(2)} &= 1 - \frac{1 - \Phi_{i,j}^{m,m}}{1 - m_{i,j}}(1 - M_{i,j}), \\ \acute{m}_{i,j}^{(2)} &= \Phi_{i,j}^{m,m}. \end{aligned} \tag{4.38}$$

The third transformation is based on viewing $M_{i,j}, m_{i,j}$ as if they were samples of a single random variable X. The distribution of X is the mixture distribution of the individual distributions of $\mathbf{M}$ and $\mathbf{m}$, with equal mixture weights. Equivalently, the CDF of X is the mean between the CDF of M and m. A valid model for the CDF of X is then $g(x) = (\Phi_{i,j}^{x,M} + \Phi_{i,j}^{x,m})/2$. The third transformation consists in applying g to $M_{i,j}$ and $m_{i,j}$, mapping them to

$$\begin{aligned} \acute{M}_{i,j}^{(3)} &= (\Phi_{i,j}^{M,M} + \Phi_{i,j}^{M,m})/2, \\ \acute{m}_{i,j}^{(3)} &= (\Phi_{i,j}^{m,M} + \Phi_{i,j}^{m,m})/2. \end{aligned} \tag{4.39}$$

It is easy to verify that the three proposed transformations satisfy the conditions (i) and (ii).

As discussed before, the HSV (iHSV) based transformation causes lack of colorfulness on overexposed (underexposed) regions. The third transformation (4.39) tends to increase the colorfulness in all areas, but such increase may be excessive. With this in mind, the final transformed maximum and minimum channels, denoted by $\acute{M}_{i,j}$ and $\acute{m}_{i,j}$, are defined as

$$\begin{aligned} \acute{M}_{i,j} &= (1 - T)[v_{i,j}\acute{M}_{i,j}^{(1)} + (1 - v_{i,j})\acute{M}_{i,j}^{(2)}] + T\acute{M}_{i,j}^{(3)}, \\ \acute{m}_{i,j} &= (1 - T)[v_{i,j}\acute{m}_{i,j}^{(1)} + (1 - v_{i,j})\acute{m}_{i,j}^{(2)}] + T\acute{m}_{i,j}^{(3)}, \end{aligned} \tag{4.40}$$

where T is a user-defined parameter in the interval $[0, 1]$, $v_{i,j}$ is given by

$$\begin{aligned} v_{i,j} &= \frac{v_{i,j}^{(1)}}{v_{i,j}^{(1)} + v_{i,j}^{(2)}}, \\ v_{i,j}^{(1)} &= \exp(-h(M_{i,j}^2 + m_{i,j}^2)), \\ v_{i,j}^{(2)} &= \exp(-h[(1 - M_{i,j})^2 + (1 - m_{i,j})^2]) \end{aligned} \tag{4.41}$$

and h is a scale parameter for the exponential.

It is the case that $v_{i,j}^{(1)}, v_{i,j}^{(2)} \in (0, 1]$. Moreover, $v_{i,j}^{(1)}$ is near 1 if the color is dark and near 0 if the pixel is bright. $v_{i,j}^{(2)}$ is near 1 or 0 in the opposite situations. $v_{i,j}^{(1)}$ and $v_{i,j}^{(2)}$ can be

interpreted as fuzzy indicators that the pixel is dark and bright, respectively. $v_{i,j}$ is defined in a way that makes the coefficients $v_{i,j}, 1 - v_{i,j} \in (0, 1]$ from the convex combination in (4.40) be proportional to the quantities $v_{i,j}^{(1)}, v_{i,j}^{(2)}$. The idea behind (4.40) is to combine the HSV and iHSV-based transformations assigning a larger contribution to the HSV (iHSV) based transformation in dark (bright) pixels, and then combine the result with the third transformation (4.39) using a convex combination with T as weight. The larger $v_{i,j}$ is, the larger is the contribution of the first transformation relative to the second transformation. The larger T is, the larger is the contribution of the third transformation relative to the combination of the first two transformations.

The final output image $\acute{\mathcal{F}}$ is computed in the following way. The maximum and minimum channels of $\acute{\mathcal{F}}$ must be $\acute{\mathbf{M}}$ and $\acute{\mathbf{m}}$, respectively, and the hue of $\acute{\mathcal{F}}$ must be equal to the hue of the original image $\mathcal{F}$. This leads to the expression

$$\acute{F}_{i,j} = \frac{\acute{M}_{i,j} - \acute{m}_{i,j}}{M_{i,j} - m_{i,j}}(F_{i,j} - m_{i,j}) + \acute{m}_{i,j}, \tag{4.42}$$

which is a hue-preserving transformation (UEDA et al., 2018) that maps $M_{i,j}$ and $m_{i,j}$ into $\acute{M}_{i,j}$ and $\acute{m}_{i,j}$, respectively.

Outputs of SADS using MMT extension to color images are shown in Figure 4.4f (page 33).

4.7 WLS Calculation

The proposed algorithm demands multiple WLS filter calculations. The weights are the same in all cases, and hence only the right-hand side of the sparse system in (3.13) (page 23) change, being always equal to the image to be filtered reshaped as a vector. To take advantage of this, a decomposition based approach is used to approximate the filter. The procedure is the following.

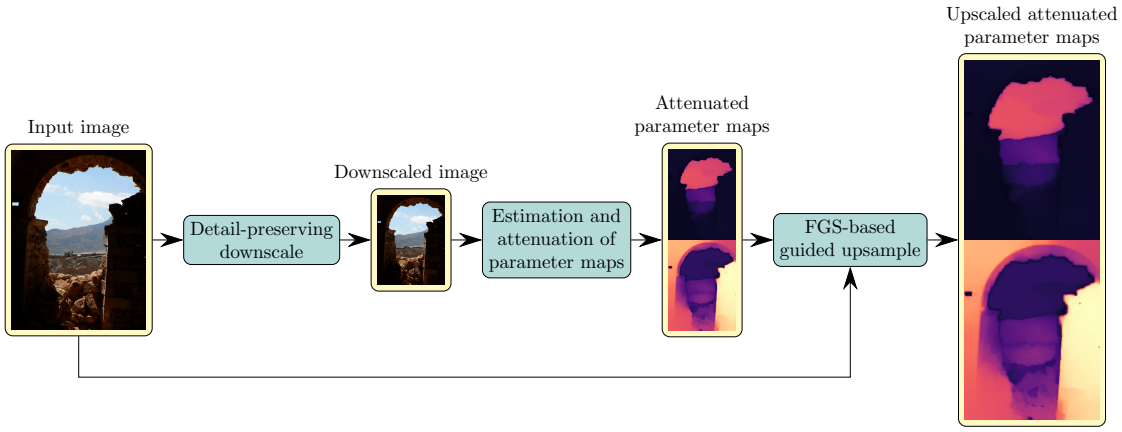
First, a symmetric approximate minimum degree permutation (SAMD) is obtained for $\mathbf{A}$. Then, the modified incomplete Cholesky (MIC) factor $\mathbf{R}$ of the SAMD permutation of $\mathbf{A}$ is calculated. Let $\mathbf{O} = \mathrm{WLS}(\mathbf{F}, \mathcal{W})$ and $\mathbf{o}$ and $\mathbf{f}$ be the column-major vectorization of $\mathbf{O}$ and $\mathbf{F}$, respectively. Consider $\hat{\mathbf{o}}$ defined as the solution to the system

$$\mathbf{R}\mathbf{R}^T\mathbf{P}\hat{\mathbf{o}} = \mathbf{P}\mathbf{f}, \tag{4.43}$$

where $\mathbf{P}$ is the permutation matrix that permutes the elements of a vector using the SAMD reordering obtained for $\mathbf{A}$. The system in (4.43) can be rewritten as $\mathbf{P}^T\mathbf{R}\mathbf{R}^T\mathbf{P}\hat{\mathbf{o}} = \mathbf{f}$ by multiplying both its sides by $\mathbf{P}^T$ and using the property $\mathbf{P}^T\mathbf{P} = \mathbf{I}$ (which is valid because $\mathbf{P}$ is a permutation matrix). Since $\mathbf{A} \approx \mathbf{P}^T\mathbf{R}\mathbf{R}^T\mathbf{P}$, it follows that $\hat{\mathbf{o}}$ is an approximation for $\mathbf{o}$. The system in (4.43) can be broken in two systems with coefficient matrices $\mathbf{R}$ and $\mathbf{R}^T$; the first gives the vector $\mathbf{R}^T\mathbf{P}\hat{\mathbf{o}}$, while the second gives the vector $\mathbf{P}\hat{\mathbf{o}}$. These two systems are sparse and triangular, and thus they can be efficiently solved with forward and backward substitution.

The accuracy of this approximation for the WLS filter depends on the drop tolerance parameter τ of the MIC factorization (MICF) used to decompose the SAMD of $\mathbf{A}$. In practice, for a sufficiently small τ, the result still behaves as an EPS filter. In comparison with the original WLS filter, it was observed that the approximated method leads to less diffusion across prominent edges, which can be advantageous for the algorithm. If a τ too small is chosen, the approximation to the original WLS filter is very exact but the

Figure 4.5: Block diagram for the SADS speed-up scheme with a downscale and a guided upsampling algorithm.



Source: Author.

Cholesky factor $\mathbf{R}$ may have too many nonzero elements, and the computational needs could become very expansive for large resolutions; if a τ too large is chosen, there are less computational needs, but the approximation to the original WLS filter may display noticeable artifacts. τ is set to $1 \cdot 10^{-5}$ in our experiments.

If a better approximation for the true WLS filter with more diffusion is desired, it is proposed to use a permuted preconditioned conjugate gradient (PCG), where the preconditioner is $\mathbf{R}$ and the permutation is given by the previously computed SAMD. The PCG is initialized with the approximated solution in (4.43), and hence can be viewed as a refinement to it. It is up to the user to decide if PCG is used to refine the solution or not. In practice, the PCG refinement will be strictly necessary only if the desired smoothing degree is too large to be suitably achieved with the approximation (4.43) by only increasing the smoothing parameter λ.

4.8 Speed-up with Downscale and Guided Upsampling

The computational cost of the algorithm may be prohibitively large for a large image. For this reason, it is proposed to downscale the original image to a tractable resolution if the original image size is larger than a threshold value. The algorithm is then executed using the downscaled image as input, until the step in which the attenuated parameter maps $\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\beta}}$ or $\tilde{\boldsymbol{\alpha}}^{(c)}, \tilde{\boldsymbol{\beta}}^{(c)}$, $c \in \{M, m\}$ are obtained. The attenuated parameter maps are then upscaled to the original resolution using a guided upsampling algorithm. Figure 4.5 shows a block diagram representing this procedure.

The downscale is made using a detail-preserving downscale algorithm inspired in (WEBER et al., 2016). Algorithm 1 explains the steps of the downscale algorithm used. Let r be the downscale ratio. In the algorithm, $\mathrm{downscale}(\cdot)$ denotes a conventional downscale with ratio r using a box kernel and $\mathrm{upscale}(\cdot)$ denotes a conventional upscale to the original resolution using bicubic interpolation. As before, $\mathbf{K}*$ denotes a convolution with a Gaussian kernel. $\hat{\mathcal{F}}$ is a coarsened version of $\mathcal{F}$. $\mathbf{V}$ is an image of weights that measure the distance between each pixel of $\mathcal{F}$ and $\hat{\mathcal{F}}$, having the same size as a channel of $\mathcal{F}$. Symbols $\circ$ and $\oslash$ denote the element wise multiplication and division, respectively. η controls the degree of preservation of details of the downscale algorithm.

Algorithm 1 Detail-preserving downscaling.

Input: input image $\mathcal{F}$ with n_{ch} channels, downscale ratio r, detail preservation level η, weights tolerance ϵ_d

Output: downscaled image $\mathcal{Z}$

1: $\hat{\mathbf{F}}^{(c)} = \text{upscale}(\mathbf{K} * (\text{downscale}(\mathbf{F}^{(c)})))$,
$c = 1, \ldots, n_{ch}$

2: $V_{i,j} = \sum_{c=1}^{n_{ch}} \left(F_{i,j}^{(c)} - \hat{F}_{i,j}^{(c)} \right)^2$,
$i = 1, \ldots, H$, $j = 1, \ldots, W$

3: $\bar{V} = \frac{1}{N} \sum_{i=1}^{H} \sum_{j=1}^{W} V_{i,j}$

4: $V_{i,j} \leftarrow (V_{i,j} / \bar{V} + \epsilon_d)^{\eta/2}$

5: $\mathbf{Z}^{(c)} = \text{downscale}(\mathbf{V} \circ \mathbf{F}^{(c)}) \oslash \text{downscale}(\mathbf{V})$,
$c = 1, \ldots, n_{ch}$

Algorithm 1 is derived from the algorithm proposed in (WEBER et al., 2016) by simplifying some of its steps and making them more straightforwardly implementable. The output of the Algorithm 1 has similar characteristics to the output of the original algorithm in (WEBER et al., 2016), although the two algorithms are not equivalent.

As mentioned, after obtaining the attenuated parameter maps, they are upscaled to the original resolution using a guided upsampling algorithm. As visible in Figure 4.2b (page 29), the parameter maps are piecewise smooth — i.e., they are smooth inside segments and all the significant gradients are in the structural edges. This is also valid for the attenuated parameter maps. The objective in using a guided upsampling algorithm is to conserve this property. If a traditional interpolation technique (e.g., bicubic interpolation) is used, the upscaled attenuated parameter maps would have blurred edges, and the output image would have large halos beside structural edges.

The guided upsampling algorithm is based on the sparse data interpolation scheme described in subsection IV.C of (MIN et al., 2014). Such sparse data interpolation scheme uses the fast global smoothing (FGS) EPS filter proposed in (MIN et al., 2014). The guided upsampling algorithm will be referred as FGS-based guided upsampling (FGS-GU). Algorithm 2 describes the FGS-GU.

4.9 Model to Select FGS-GU Smoothness Parameter

The choice of the smoothness parameter λ_{FGS} from FGS-GU is not trivial. A value too small do not provide the smoothing required for the sparse data interpolation, and a value too big causes over smoothing of the upsampled parameter map, potentially producing halos on SADS result (FGS is less edge-preserving than original WLS). The best value of λ_{FGS} depends on the downscale ratio r (the ratio between the downscaled and original dimensions). The smaller r is, the larger the λ_{FGS} needed, because more smoothing is required for the sparse data interpolation. For this reason, a model was created to provide a suitable value of λ_{FGS} given r. Here, the model and the procedures to obtain this model are described.

Two square (width equal height) images were selected (Figure 4.6 (page 42)). For each image, the α parameter map was computed and attenuated. The parameter map was then downscaled for resolutions ranging from 8×8 to the original image size minus 1 in each dimension. For each downscale resolution, Golden Section Search (GSS) was used to find the value of λ_{FGS} that minimizes the mean of absolute differences between the

Algorithm 2 FGS-GU

Input: image to be upscaled $\mathbf{E}$ of size $H' \times W'$, guidance image $\mathcal{F}$ of size $H \times W$
Output: upscaled image $\mathbf{U}$

1: Initialize the auxiliary images $\mathbf{H}$ and $\tilde{\mathbf{E}}$ as matrices of 0s with the target resolution $H \times W$
2: Let

$$\begin{aligned} i^*(i) &= \left\lceil \left(i - \frac{1}{2}\right) \frac{H}{H'} \right\rceil, \\ j^*(j) &= \left\lceil \left(j - \frac{1}{2}\right) \frac{W}{W'} \right\rceil \end{aligned} \tag{4.44}$$

Compute the nonzero elements of $\mathbf{H}$ and $\tilde{\mathbf{E}}$ with

$$\begin{gathered} H_{i^*(i),j^*(j)} = 1, \\ \tilde{E}_{i^*(i),j^*(j)} = E_{i,j}, \\ i = 1, \ldots, H';\ j = 1, \ldots, W'. \end{gathered} \tag{4.45}$$

3: Compute the output image $\mathbf{U}$ with

$$\mathbf{U} = \mathrm{FGS}(\hat{\mathbf{E}}, \mathcal{F}, \lambda_{\text{FGS}}) \oslash \mathrm{FGS}(\mathbf{H}, \mathcal{F}, \lambda_{\text{FGS}}), \tag{4.46}$$

where $\mathrm{FGS}(\mathbf{X}, \mathcal{Y}, \lambda_{\text{FGS}})$ denotes the FGS filter proposed in (MIN et al., 2014) applied to the image $\mathbf{X}$ using $\mathcal{Y}$ as guidance image and λ_{FGS} as smoothness parameter.

Figure 4.6: Images used to generate data for the λ_{FGS} model.

(a) (b)

Source: shutterstock.com (4.6a) and idealista.pt (4.6b).

original parameter map and the result of upsampling the downscaled parameter map with FGS-GU. GSS was applicable in this case because it was verified graphically that the objective function minimized was convex. From the results of this procedure applied to the two images, data with $(r, \text{optimal } \lambda_{\text{FGS}})$ pairs was gathered. Such data was then used to fit a model to predict the optimal λ_{FGS} given the downscale ratio r.

The model was chosen by analyzing the shape outlined by the gathered data. Let (x_d, y_d) be the $(r, \text{optimal } \lambda_{\text{FGS}})$ pairs in the data. Let a (k) indicate the k-th data sample. y_d appears to grow boundless when x_d approaches 0, and approaches 0 when x_d approaches 1 (both behaviors were expected). By multiplying the y_d with x_d^2, the singularity of y_d when x_d approaches 0 was removed. Hence, a rational function in the form $f(x) = \text{pol}\, x/x^2$ was chosen as the model, where pol is a polynomial of degree 6. The polynomial was fitted to the pairs (x_d, y'_d), where $y'_d = x_d^2 y_d$. To this end, Iteratively Reweighted Least Squares (IRLS) with L-2 regularization was used. The weights of the IRLS were computed in each iteration with

$$\begin{aligned} w(k) &= 1/\sqrt{0.2 + d(k)^2}, \\ d(k) &= \text{pol}(x_d(k)) - y'_d(k). \end{aligned} \tag{4.47}$$

The regularization matrix was $1 \cdot 10^{-6}\, \text{diag}(0, \ldots, 6)$, where each diagonal element correspond to a polynomial term, by degree order (the constant polynomial term received no regularization, and the other polynomial terms received a regularization proportional to their degree). 6 iterations of the IRLS were run. The choice (4.47) for the IRLS weights corresponds to a minimization of the objective function $\sum_k \sqrt{0.2 + d(k)^2}$, which is more robust to outliers than the simple sum of squares of the residuals $d(k)$. Robustness to outliers was required because the observed data has deviations from the smooth apparent relationship between r and optimal λ_{FGS}, notably an attraction of λ_{FGS} to 0 when r approaches 0.5. Regularization was required because the observed data is noisy.

In practice, although the obtained model fitted the data adequately, it was observed that it produced values of λ_{FGS} not sufficiently large for the upsample. For this reason, a

Figure 4.7: Effect of the downscale/guided upsample speed-up scheme on the output of SADS.

(a) Input	(b) $r = 0.2$	(c) $r = 0.4$	(d) $r = 0.6$	(e) $r = 0.8$	(f) No downscale
-	MAE $= 0.0463$	MAE $= 0.0276$	MAE $= 0.0179$	MAE $= 0.0093$	-

Source: TM-DIED dataset (VONIKAKIS, 2021) (4.7a) and author (other images).

factor of 10 was multiplied to the fitted model. The final model is

$$\begin{aligned} f(x) &= 10 \sum_{k=0}^{6} \theta_k x^{k-2}, \\ \theta &= (0.25, -0.34, 5.34, -18.64, \\ &\quad 15.22, 11.98, -13.72), \end{aligned} \tag{4.48}$$

where here θ was rounded to two decimal places for compactness. Tests showed that the final model provided suitable values of λ_{FGS} for all values of r.

Figure 4.7 shows outputs of SADS with and without the downscale/guided upsample scheme, using varying downscale ratios. The mean absolute error (MAE) between the SADS output without the scheme and each SADS output with the scheme is given. The original resolution is 867×1155 (nearly 1 mega pixel). The results show that when donwsample is used there is some loss of local contrast gain, specially in the dark areas inside the building. The lower r is, the higher the loss of local contrast gain. Nevertheless, there is a high similarity between the results. The effect of using the downsample scheme is somewhat similar to increasing the smoothness parameter λ from the WLS filter.

4.10 Summary of SADS Main Algorithms

In this Section the main SADS algorithms are summarized. Algorithm 3 is the SADS algorithm for grayscale images. Algorithm 4 (page 45) is the SADS algorithm for color images using MMT extesion (Section 4.6.2 (page 33)). The SADS algorithm for color images using LCT extension (Section 4.6.1 (page 33)) is simply Algorithm 3 applied to the luminance channel, using the expressions for the WLS weights and smoothness map designed for color images.

4.11 Reasons for Using the Beta Distribution

The beta distribution is used in the proposed method for the following reasons. First, the shapes assumed by the beta distribution CDF are adapted to the enhancement tasks. It is the case that:

- A transformation that corrects dark areas should be concave downwards, similarly to the shape of $f(x) = x^a$, $x \in (0, 1]$, $a \in (0, 1)$.
- A transformation that corrects bright areas should be concave upwards, similarly to the shape of $f(x) = x^a$, $x \in (0, 1]$, $a \in (1, \infty)$.

Algorithm 3 SADS for grayscale images.

Input: Image to be enhanced $\mathbf{F}$
Output: Enhanced image $\acute{\mathbf{F}}$

if $N > n_{max}$ **then**
  Downscale $\mathbf{F}$ using Algorithm 1 (page 40), with downscale ratio $r = \sqrt{n_{max}/N}$ and preservation level η. From here, symbol $\mathbf{F}$ denotes the downscaled image, except at the upscale step
end if
Compute the WLS weights $\mathcal{W}$ with (4.1) (page 25), the Laplacian matrix $\mathscr{L}$ using (3.15) (page 23) and the matrix $\mathbf{A}$ with (3.14) (page 23)
Compute the SAMD of $\mathbf{A}$ and store the permutation indexes
Compute the MIC factor $\mathbf{R}$ of the SAMD permutation of $\mathbf{A}$, using τ as drop tolerance
if $C > 0$ **then**
  Compute the smoothness map $\mathbf{S}$ with (4.16) (page 30)
  Compute the local smoothness influence $\mathbf{Q}$ with (4.20) (page 32)
end if
if Use MLE **then**
  Compute $\mathbf{G}$ and $\bar{\mathbf{G}}$ with (4.9) (page 27)
  Compute $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ with (4.10) (page 28)
else
  Compute $\boldsymbol{\mu}$ and $\mathbf{s}$ with (4.5) (page 27)
  Compute $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ with (4.6) (page 27)
end if
if $C > 0$ **then**
  Compute $l_{i,j}$ with (4.21) (page 32)
else
  Set $l_{i,j} = K$
end if
Compute $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\boldsymbol{\beta}}$ with (4.15) (page 30).
if Donwscale was made **then**
  Upscale $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\boldsymbol{\beta}}$ with Algorithm 2 (page 41) using the original image (not downscaled) $\mathbf{F}$ as guidance image and $H \times W$ as target resolution
end if
Apply transformations (4.13) (page 30) and (4.14) (page 30). If $\breve{\alpha} = \breve{\beta} = 1$, transformation (4.14) (page 30) is the identity function and does not need to be calculated

Algorithm 4 SADS for color images with MMT extension.

Input: Image to be enhanced $\mathcal{F}$
Output: Enhanced image $\acute{\mathcal{F}}$

if $N > n_{max}$ **then**
  Downscale $\mathcal{F}$ using Algorithm 1 (page 40), with downscale ratio $r = \sqrt{n_{max}/N}$ and preservation level η. From here, symbol $\mathcal{F}$ denotes the downscaled image, except at the upscale step
end if
Compute the WLS weights $\mathcal{W}$ with (4.2) (page 26), the Laplacian matrix $\boldsymbol{\mathscr{L}}$ using (3.15) (page 23) and the matrix $\mathbf{A}$ with (3.14) (page 23)
Compute the SAMD of $\mathbf{A}$ and store the permutation indexes
Compute the MIC factor $\mathbf{R}$ of the SAMD permutation of $\mathbf{A}$, using τ as drop tolerance

if $C > 0$ **then**
  Compute the smoothness map $\mathbf{S}$ with (4.16) (page 30), using the $T_{i,j}$ definition in (4.18) (page 31).
  Compute the local smoothness influence $\mathbf{Q}$ with (4.20) (page 32)
end if
if Use MLE **then**
  Compute $\mathbf{G}^{(c)}$ and $\bar{\mathbf{G}}^{(c)}$ with (4.34) (page 36)
  Compute $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ with (4.33) (page 36)
else
  Compute $\boldsymbol{\mu}$ and $\mathbf{s}$ with (4.32) (page 36)
  Compute $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ with (4.31) (page 36)
end if
if $C > 0$ **then**
  Compute $l_{i,j}$ with (4.21) (page 32)
else
  Set $l_{i,j} = K$
end if
Compute $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\boldsymbol{\beta}}$ with (4.35) (page 36).
if Donwscale was made **then**
  Upscale $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\boldsymbol{\beta}}$ with Algorithm 2 (page 41) using the original image (not downscaled) $\mathcal{F}$ as guidance image and $H \times W$ as target resolution
end if
Use (4.36) (page 36) to (4.42) (page 38) to obtain the output image $\acute{\mathcal{F}}$.

- A transformation that enhances middle tone contrast should have a sigmoid shape.
- A transformation that enhances contrast simultaneously in dark and bright areas should have a logit shape (a shape similar to the shape of the logit function $f(x) = \log(x/(1-x))$).
- The transformation that conserves the input is the identity function.

Figure 3.1 (page 22) shows that the beta distribution CDF can have all these useful shapes, allowing SADS to correct under and over enhancement and enhance contrast, and also conserve the input if the values are already evenly distributed.

Second, the support of the beta distribution CDF is the interval $(0, 1)$, allowing a direct correspondence to the interval of the normalized image values.

Third, the MLE expressions can be written in terms of sample means of all the image values, allowing the conversion of these expressions into a heuristic regional version by replacing the sample means with WLS filters.

Fourth, the beta distribution CDF, which is the regularized incomplete beta function $I_x(\alpha, \beta)$, can be quickly and efficiently computed with partial fraction expansions.

4.12 Reasons for not Using Segmentation

It would be possible to use a segmentation algorithm to segment the image and for each segment compute the parameters of the beta distribution, instead of using WLS filters. However, this approach would have the following drawbacks.

Segmentation techniques are not always reliable and may fail to correctly separate segments, making SADS more unpredictable. Additionally, segmentation requires considerably more computational power than the WLS filter, and is also a much more complex task than it.

Moreover, if segmentation is used, the parameter maps would be exactly piecewise constant. Thus, the parameter maps would always change abruptly, even in the cases in which a smooth variation is required. For example, if an area in the image changes smoothly, or if there is a smooth or diffuse transition between two segments, a smooth variation of the parameter maps is required. Segmentation would make the parameter maps change abruptly in these cases, creating artificial edges in the output image. Furthermore, the parameter maps should not be always constant inside segments, because they should adapt to variations within each segment.

It should also be noted that recent algorithms, such as LIME (GUO; LI; LING, 2016) and Zhang's method (ZHANG; NIE; ZHENG, 2019), do not use segmentation to calculate illumination maps; they use the WLS filter instead.

5 EXPERIMENTS

The evaluation of the SADS method was conducted through a series of experiments that assess SADS contrast enhancement, exposure correction and dehazing capabilities using no-reference (NR) and full-reference (FR) metrics, compared to other methods in the literature. The results are presented in Sections 5.1, 5.2 (page 49), 5.3 (page 51), 5.4 (page 53), 5.5 (page 55) and 5.6 (page 56). SADS was executed on Matlab R2015a, except in the experiment of Section 5.5 (page 55), which was conducted on Matlab Online.

5.1 Contrast Enhancement and Exposure Correction, Dataset without References

In this experiment, 50 natural images with contrast or exposure defects were enhanced using CLAHE (PIZER et al., 1987) on HSL color space, Ying's method (YING et al., 2017), Zhang's method (ZHANG; NIE; ZHENG, 2019), Local Color Distributions Prior Network (LCDPNet) (WANG; XU; LAU, 2022) and the proposed SADS. For convenience, Ying's and Zhang's methods will be referred to as Ying17 and Zhang19, respectively. The images are from TM-DIED dataset (VONIKAKIS, 2021). The results of the methods were quantitatively evaluated using the NR IQA metrics Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) (MITTAL; MOORTHY; BOVIK, 2012), Natural Image Quality Evaluator (NIQE) (MITTAL; SOUNDARARAJAN; BOVIK, 2012), Integrated Local NIQE (IL-NIQE) (ZHANG; ZHANG; BOVIK, 2015), Deep Bilinear CNN (DB-CNN) (ZHANG et al., 2018), Contrast-Changed Image Quality Measure (CEIQ) (YAN; LI; FU, 2019) and Multi-Scale Image Quality Transformer (Musiq) (KE et al., 2021).

CLAHE was executed applying Matlab's 'adapthisteq' built-in function to the lightness of HSL color space. For the other concurrent algorithms, the official implementations were used. The implementations of Ying17 and Zhang19 are in Matlab, and the implementation of LCDPNet is in Python with Pytorch.

Table 5.1 shows the parameters fixed for SADS in this experiment. These parameters were chosen empirically, visually evaluating the output images compared to the input images. For CLAHE, a clipping limit of 0.02 and an 8×8 grid of tiles were adopted. For Zhang19, the γ parameter (defined in (ZHANG; NIE; ZHENG, 2019)) was set to 0.7, and the other parameters were the same as in the original code. For the other algorithms, the original parameter choices were kept.

Table 5.2 (page 49) shows the median of each metric for the 50 images, for each enhancement algorithm. An upwards (downwards) arrow after the metric name means the higher the better (the lower the better). The best and second best results of each row

Table 5.1: Parameters of SADS for the contrast and exposure experiment.

Symbol	Description	Value
σ	Scale parameter of WLS weights Gaussian blur	2
λ	Trade-off factor of WLS objective function	0.25
-	Estimation method	MLE
-	Use PCG for WLS filter?	No
τ	ICF tolerance	$1 \cdot 10^{-5}$
z	Scale parameter of smoothness map	$8 \cdot 10^{-4}$
p	Power parameter of smoothness map's generalized mean	0.025
γ	Power parameter of smoothness map's input	0.025
K	Global transformation level	0.8
C	Local transformation level factor	0.5
-	Color image extension method	MMT
T	Coefficient from the combination of transformations in the MMT color image extension	0.7
h	Scale parameter of MMT color image extension weights	0.5
η	Preservation level of Algorithm 1 (page 40)	0.75

Table 5.2: Median of the IQA metrics in the contrast and exposure experiment without references.

	CLAHE	Ying17	Zhang19	LCDPNet	SADS
BRISQUE ↓	<u>15.802</u>	16.215	16.807	18.281	**15.395**
NIQE ↓	**2.300**	2.475	2.414	2.648	<u>2.326</u>
IL-NIQE ↓	**18.483**	19.996	19.329	21.169	<u>18.627</u>
DB-CNN ↑	58.022	57.678	<u>59.756</u>	53.860	**59.785**
CEIQ ↑	<u>3.585</u>	3.407	3.572	3.409	**3.592**
Musiq ↑	66.892	67.186	<u>68.088</u>	65.649	**68.652**

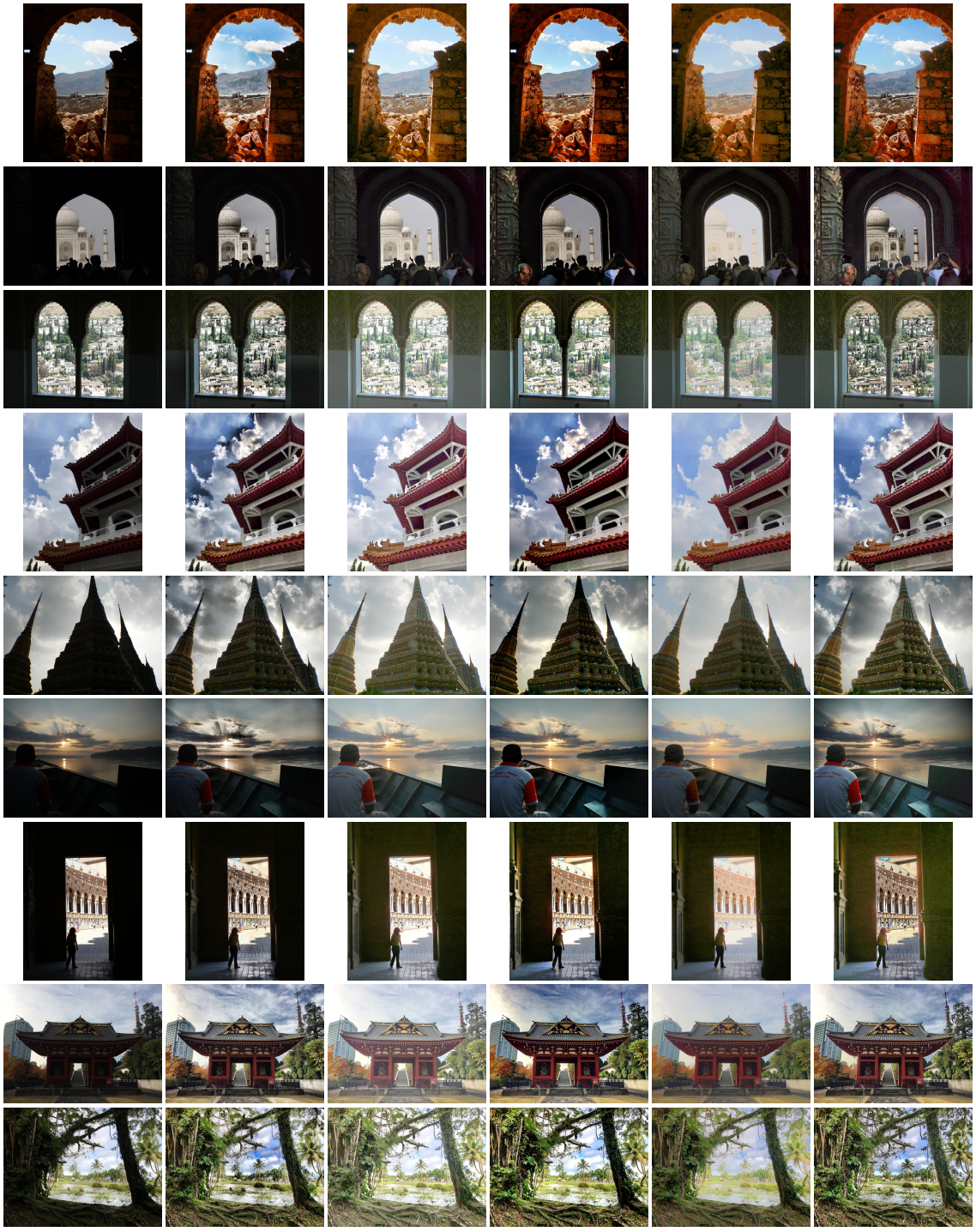
are highlighted with bold and underline, respectively. The proposed SADS had the best result for four of the IQA, and the second best for two of the IQA.

Figure 5.1 shows some of the image outputs in this experiment. The outputs verify the high capability of SADS to enhance contrast and correct exposure. We can observe that, when CLAHE performs well for an image, SADS output is similar to CLAHE's output, but when CLAHE performs badly, SADS achieve a much better result. In special, CLAHE was unable to correct some of the dark regions in the inputs, and over enhanced sky regions. Zhang19 method also had similar outputs to SADS in some cases, but in other cases over enhanced sky areas or enhanced dark regions less than SADS. A problem observed with Zhang19 method is that apparently the fusion method adopted by it do not select the LIME result in some situations where it should, thus leaving some dark regions without enhancement (e.g., in the third image in Figure 5.1, at the center of the window). Ying17 method performed well for enhancing low light regions, but generated results too bright in many cases, and could not correct overexposed regions (which was expected, since it is an algorithm for low light correction only). LCDPNet generated images with well-balanced exposure, but in some cases do not enhanced contrast satisfactorily or produced slightly artificial colors.

5.2 Contrast Enhancement and Exposure Correction, Dataset with References

In this experiment, 17 natural images with contrast or exposure defects were enhanced with the same algorithms from the experiment of Section 5.1 (page 47), and the outputs were quantitatively compared to reference images. The image pairs are from MEF dataset (MA; ZENG; WANG, 2015), which contains 17 image sequences with multiple exposure levels (≥ 3) for each scene, and the results of 8 multiple-exposure fusion algorithms applied to each sequence. One image from each image sequence was selected as the input image. The fusion results generated with the algorithm from (MERTENS; KAUTZ; VAN REETH, 2009), labeled 'Mertens07' in the dataset, were used as reference images (i.e., the ground truth). The quantitative comparison between the outputs and the reference images was made using some of the most relevant FR IQA metrics in the literature, namely Peak Signal-to-Noise Ratio (PSNR), Structural Similarity (SSIM) Index (WANG et al., 2004), Visual Information Fidelity (VIF) (SHEIKH; BOVIK, 2005, 2006), Feature Similarity (FSIM) (ZHANG et al., 2011), Gradient Magnitude Similarity Deviation (GMSD) (XUE et al., 2013), Learned Perceptual Image Patch Similarity (LPIPS) (ZHANG et al., 2018) and Attention-Based Hybrid Image Quality Assessment (AHIQ) Network (LAO et al., 2022). The outputs were also evaluated with the NR IQA metrics

Figure 5.1: Sample results of the contrast and exposure experiment without references



(a) Input (b) CLAHE (c) Ying17 (d) Zhang19 (e) LCDPNet (f) SADS

Source: TM-DIED dataset (VONIKAKIS, 2021) (5.1a) and author (other images).

Table 5.3: Median results of the FR IQA metrics in the contrast and exposure experiment with references.

	CLAHE	Ying17	Zhang19	LCDPNet	SADS
PSNR ↑	13.047	**19.705**	16.070	17.117	16.531
SSIM ↑	0.515	0.745	0.651	0.696	**0.763**
VIF ↑	0.400	**0.460**	0.412	0.406	0.437
FSIM ↑	0.812	**0.890**	0.861	0.862	0.876
GMSD ↓	0.147	0.105	0.116	0.111	**0.101**
LPIPS ↓	0.267	**0.221**	0.229	0.233	0.229
AHIQ ↑	0.208	**0.464**	0.262	0.317	0.358

Table 5.4: Median results of the NR IQA metrics in the contrast and exposure experiment with references.

	CLAHE	Ying17	Zhang19	LCDPNet	SADS
BRISQUE ↓	17.593	16.658	18.796	20.537	**14.986**
NIQE ↓	2.821	2.929	2.778	2.996	**2.548**
IL-NIQE ↓	22.946	23.496	23.355	26.677	**21.869**
DB-CNN ↑	53.678	53.193	54.143	48.892	**55.600**
CEIQ ↑	3.021	3.362	3.269	3.295	**3.395**
Musiq ↑	64.782	65.671	66.074	64.472	**66.374**

used in Section 5.1 (page 47). The parameters used for SADS and the other enhancement algorithms were the same as in the experiment of Section 5.1 (page 47).

Table 5.3 and Table 5.4 show the median of the results of the FR and NR metrics, respectively. Table 5.3 shows that Ying17 and the proposed SADS achieved results perceptually closer to the reference images. The good FR scores of Ying17, which is design to correct only low light images, is due to the fact that all selected input images were low light images (the high exposure images in the MEF dataset sequences have large saturated areas or are too close to the selected ground truth images, and thus none of them was selected as input image). Table 5.4 shows that SADS achieved the highest quality according to all NR metrics.

Figure 5.2 shows some of the image outputs in this experiment. Some of the observations made in Section 5.1 (page 47) regarding the performance of SADS with respect to the other methods also apply to the outputs in this experiment.

5.3 Image Defogging/Dehazing

In this experiment, 306 natural foggy/hazy images were enhanced using DehazeNet (CAI et al., 2016), Artificial Multiple-Exposure Fusion (AMEF) (GALDRAN, 2018), Feature Fusion Attention Network (FFA-Net) (QIN et al., 2020), DehazeFormer (SONG et al., 2022) and the proposed SADS. The images were selected from the LIVE Image Defogging Database (CHOI; YOU; BOVIK, 2015). The results of the methods were quantitatively evaluated using Fog Aware Density Evaluator (FADE) NR metric (CHOI; YOU; BOVIK, 2015). The FADE metric estimates the fogginess/haziness density of an image, and thus a lower FADE result indicates a better performance of the algorithm. The parameters used for SADS where the same used in the contrast and exposure experiments, except for a change in z and T for the values indicated in Table 5.5 (page 53). A smaller

Figure 5.2: Sample results of the contrast and exposure experiment with references.

(a) In/GT (b) CLAHE (c) Ying17 (d) Zhang19 (e) LCDPNet (f) SADS

Source: MEF dataset (MA; ZENG; WANG, 2015) (5.2a) and author (other images).

Table 5.5: Parameters of SADS changed for the defogging/dehazing experiment.

Symbol	Description	Value
z	Scale parameter of smoothness map	$1 \cdot 10^{-5}$
T	Coefficient from the combination of transformations in the MMT color image extension	0.5

Table 5.6: Median of the FADE metric in the defogging/dehazing experiment.

DehazeNet	AMEF	FFA-Net	DehazeFormer	SADS
0.677	<u>0.647</u>	1.376	1.078	**0.619**

z is used to have higher values in the smoothness map only for virtually flat regions, helping the transformation to perform better in hazy areas. A smaller T is used to increase the contribution of the third transformation in MMT extension for color images, which increases the colorfulness, as required to attenuate the foggy/hazy aspect.

The official implementation of the defogging/dehazing algorithms was used. DehazeNet and AMEF implementations are in Matlab, and FFA-Net and DehazeFormer implementations are in Python with Pytorch. The default parameters were used in case the model had parameters. For FFA-Net, the 'ots' pre-trained weights were used. For DehazeFormer, the 'outdoor/dehazeformer-s' pre-trained weights were used.

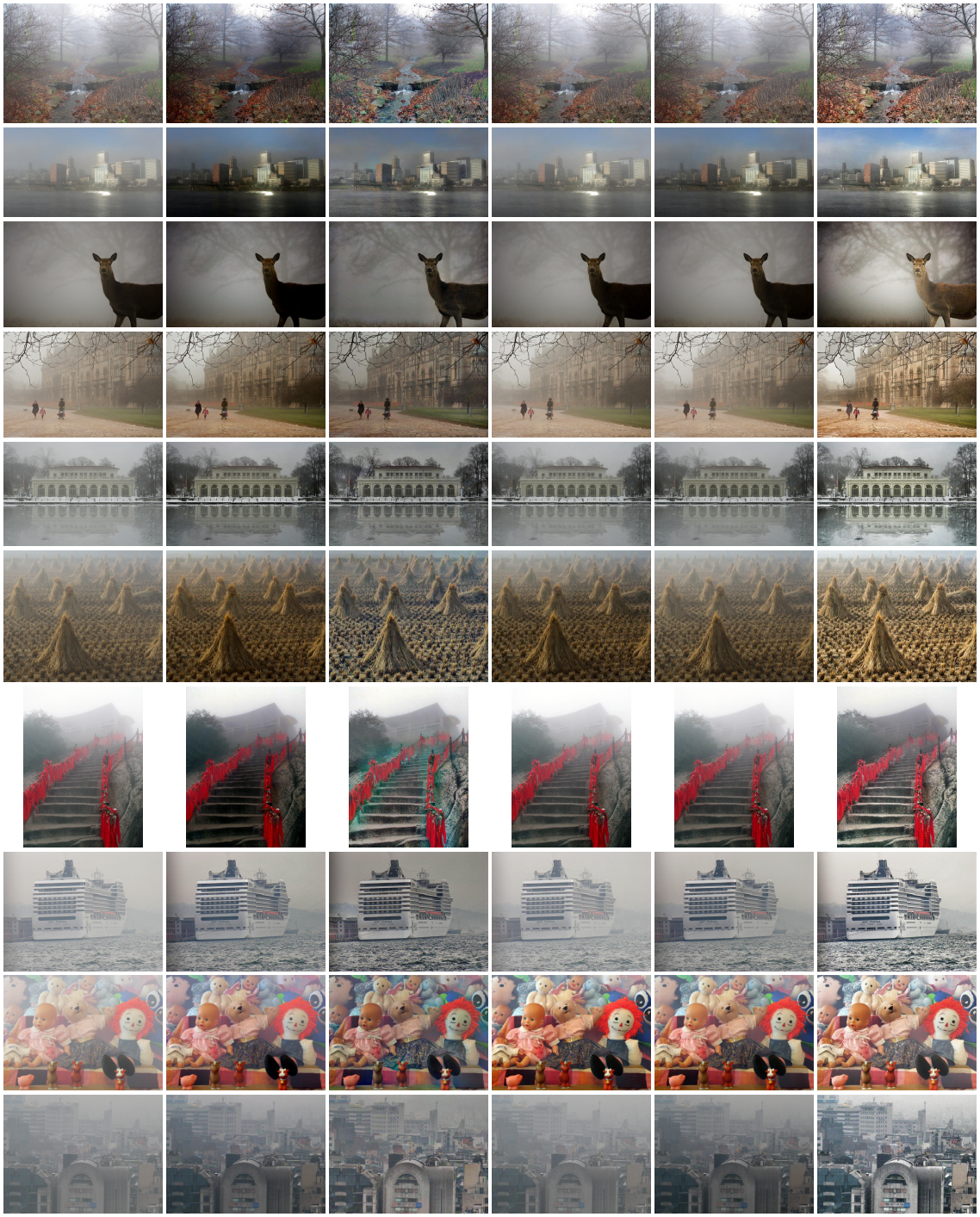
Table 5.6 shows the median of the FADE metric for each algorithm. The proposed SADS achieved the lower median FADE, verifying that it is competitive in the defogging/dehazing task.

Figure 5.3 shows some of the image outputs obtained in this experiment. Generally, AMEF and SADS generated results visually better than the other algorithms, showing concordance with Table 5.6. DehazeNet tended to produce results too dark in foggy areas. In a few cases, AMEF produced unnatural blue/green tones, e.g., in the image with stair. In some cases, FFA-Net and DehazeFormer produced only slight corrections to foggy areas. SADS was able to correct fogginess/haziness at least partially and at the same time improve contrast and correct exposure.

5.4 Effect of Smoothing Parameter λ and Global Transformation Level K on the Output

In this experiment, SADS was executed for varying combinations of the parameters λ and K. MMT color extension method was used. Figure 5.4 (page 55) shows the results, where each column corresponds to a value of K and each row corresponds to a value of λ. For the effect of K, it can be observed that the larger K is, the larger is the degree of enhancement with respect to the input image. For the effect of λ, it can be observed that, when λ is small, finer details are boosted and there is less preservation of the relative contrast between segments, characterizing a more local intra-segment contrast gain. When λ is large, more of the contrast between segments is preserved, and the contrast gain becomes less local. In conclusion, λ controls the spatial scale in which the enhancement occurs, while K, as expected, controls the global enhancement degree.

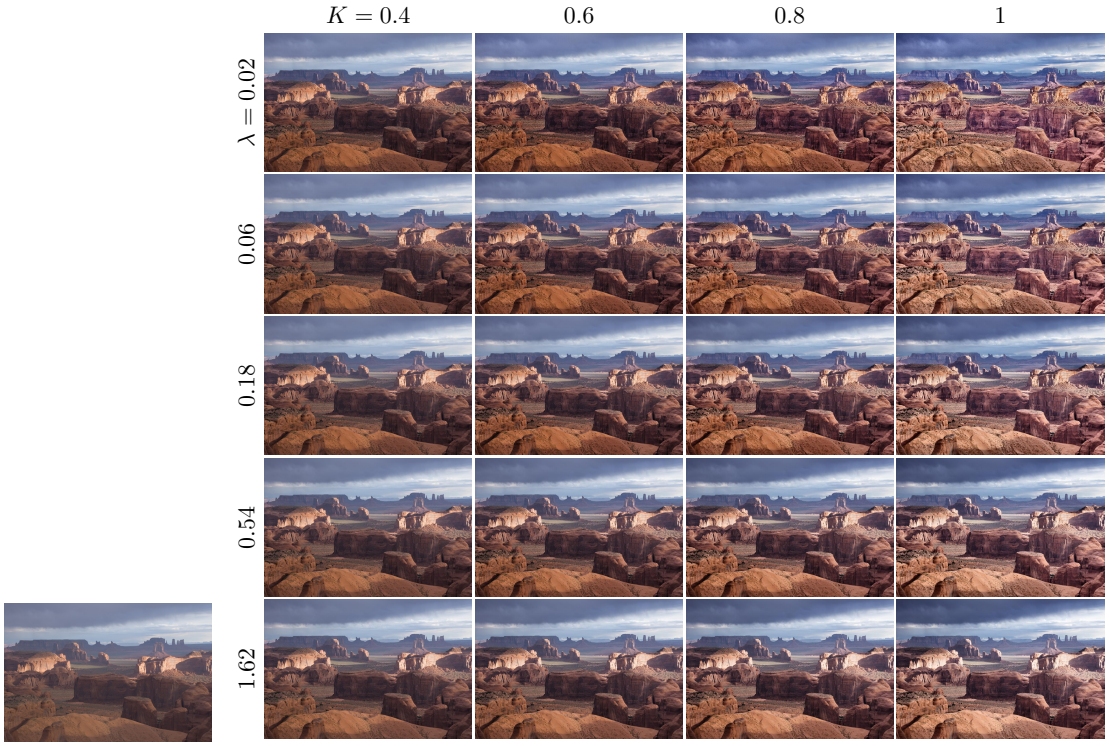
Figure 5.3: Sample results of the defogging/dehazing experiment.



(a) Input (b) DehazeNet (c) AMEF (d) FFA-Net (e) Dehaze-Former (f) SADS

Source: LIVE Image Defogging Database (CHOI; YOU; BOVIK, 2015) (5.3a) and author (other images).

Figure 5.4: Effect of λ and K on the SADS output.



(a) Input (b) SADS outputs

Source: (FARBMAN et al., 2008) (5.4a (page 55)) and author (5.4b (page 55)).

5.5 Runtime Analysis

The following runtime analysis was conducted for SADS. Each one of the 50 input images used in the experiment of Section 5.1 (page 47) was resized to a set of 10 different resolutions. The set of resolutions was the same for each image, except for the interchange between the first and second dimensions when the image is tall relative to when the image is wide. The largest resolution is 1 mega pixel, with an aspect ratio of 3:4 for wide images and 4:3 for tall images. The other resolutions are given approximately by the largest multiplied by $k/10$, with $k = 1, \ldots, 9$. SADS was applied to each image and each resolution, and the execution time was measured in each case. This procedure was made for SADS with both LCT and MMT color extension methods (Section 4.6 (page 32)); in the LCT case, HSL was the color system used to extract the luminance. The CPU used by Matlab online at the experiment was an Intel Xeon Platinum 8375C @ 2.90GHz. Table 5.7 (page 56) shows the median of the execution times for each resolution and choice of color extension method.

The downscale/upscale scheme proposed on Section 4.8 (page 39) to speed up the algorithm was not used in this experiment. By using such scheme, lower execution times can be achieved, and SADS can be applied to larger resolutions without increasing significantly its execution time. In practice, if the downscale/upscale speed up scheme is used, the total execution time is the execution time of SADS on the resolution to which the image is downscaled added to the execution times of the downscale and the guided upsample (which are not very high).

Table 5.7: Median of execution times of SADS by resolution and color extension method (LCT with HSL color system and MMT)

Resolution	Time (s)	
($W \times H$ or $H \times W$)	LCT	MMT
115×86	0.0540	0.0553
231×173	0.1134	0.1189
346×260	0.2251	0.2412
462×347	0.4061	0.4369
577×433	0.6875	0.7319
693×520	1.0864	1.1637
808×606	1.4992	1.5979
924×693	2.0813	2.2314
1039×779	2.6915	2.9014
1155×866	3.4545	3.6927

5.6 Limitations of the Proposed SADS Method

One of the main limitations of the proposed SADS is that the transformation obtained for disconnected parts of a segment can be inconsistent, causing unnatural results. This occurs because the WLS diffusion is attenuated by prominent edges, and thus the WLS based regional mean potentially changes stepwisely from a disconnected part of a segment to a neighbor part. For example, in Figure 5.5 (page 57), the sky segment is divided in two sides by the antenna, which causes a jump discontinuity in the SADS transformation for the sky, producing an unnatural aspect.

Other problem occurs when the WLS diffusion inside a segment 'leaks' across the edges of a small segment or a small disconnected part of segment contained inside or overlapping the former larger segment. An example occurs in Figure 5.6. The tree branches enclose parts of the sky segment, and the transformation applied to this enclosed sky areas is essentially the same transformation applied to the tree segment, generating the bright sky areas between the tree branches. In its turn, the larger sky segment partially encloses some of the tree branches in the top of the tree, and the transformation applied to this enclosed branches is essentially the same transformation applied to the larger sky segment, generating the dark branches in the top of the tree.

These problems also occurs for other WLS based methods, such as LIME (GUO; LI; LING, 2016) and Zhang19, but they were observed in a larger degree for SADS.

Another limitation of SADS is the noise amplification on noisy segments that receive very large contrast gains. This problem can be mitigated by the following strategy, inspired in (GUO; LI; LING, 2016). A denoising filter such as BM3D (DABOV et al., 2007) would be applied to the SADS output. The original and filtered SADS output would then be combined by a convex combination, where the coefficient of the convex combination would depend on the local contrast gain, which can be measured by a heuristic depending on certain variables, such as the maximum slope of the beta CDF fitted for each pixel. The objective in combining the original and filtered SADS outputs according to the contrast gain would be to reduce the filtering degree in areas that do not need to be filtered, similarly to what is done in (GUO; LI; LING, 2016).

Figure 5.5: Example of inconsistent transformations obtained for different parts of a disconnected segment.

(a) Input (b) Zhang19 (c) SADS

Source: TM-DIED dataset (VONIKAKIS, 2021) (5.5a) and author (other images).

Figure 5.6: Example of leakage of the parameter maps through edges of small semi enclosed parts of segments.

(a) Input (b) SADS (c) SADS (highlight)

Source: TM-DIED dataset (VONIKAKIS, 2021) (5.6a) and author (other images).

6 CONCLUSION AND FUTURE WORK

A method for contrast enhancement and exposure correction, namely Structure-Aware Distribution Stretching (SADS), has been proposed for grayscale and color images. As traditional histogram equalization based methods, SADS relies on the idea of stretching the probability distribution of channels from the image. However, different from these methods, SADS models the distributions of the channels with the beta distribution, fitting such distribution regionally, with a structure-aware scheme based on replacing the sample mean operation with an EPS filter. SADS showed promising results for the tasks of contrast enhancement, exposure correction and image defogging/dehazing.

In comparison to the well known retinex based algorithms, SADS has the advantage of not suffering from the interval $[0, 1]$ violation problem. Besides, SADS can enhance the contrast of middle tone image regions, differently from most retinex methods that tend to enhance the contrast only in dark or bright (by inverting the image) areas. In comparison to typical adaptive histogram equalization algorithms, SADS has the following advantages: do not produce the checkerboard effect; do not produce or produces in a much smaller degree the staircase effect; applies a transformation better adapted to the segments; applies a transformation with a simple smooth regular shape. In comparison to learning based methods, SADS has the advantage of providing consistent outputs for all classes of images, and not only images similar to the images in a training dataset, and also the advantage of having multiple parameters that can be adjusted to better adapt to each task (some learning methods also provide parameters enabling user adjustments, however).

It is important to note that, in each experiment, the user-defined parameters for SADS were kept constant for all input images. However, superior results can be achieved by adjusting the SADS parameters individually for each image.

Possible directions for future works are: (i) the use of other distributions besides the beta distribution as probability distribution model; (ii) adaptive ways of selecting or modifying the SADS parameters; (iii) a solution or mitigation for the limitation involving disconnected segments described in Section 5.6 (page 56); and (iv) the possibility of a multiscale version of SADS that combines the transformations or parameter maps obtained for different smoothness degrees. In the direction (i) in special, there is the possibility of using a constrained beta distribution where α, β satisfy $\alpha\beta = 1$; this may create a variant of SADS focused on exposure correction only, without middle range contrast enhancement; it would also be possible to combine the parameter maps of the original SADS and such constrained SADS, thus enabling a control of the amount of middle range contrast enhancement.

REFERENCES

CAI, B. et al. Dehazenet: an end-to-end system for single image haze removal. **IEEE Transactions on Image Processing**, [S.l.], v.25, n.11, p.5187–5198, 2016.

CHANG, Y. et al. Automatic contrast-limited adaptive histogram equalization with dual gamma correction. **Ieee Access**, [S.l.], v.6, p.11782–11792, 2018.

CHEN, Y.-S. et al. Deep photo enhancer: unpaired learning for image enhancement from photographs with gans. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2018. **Proceedings...** [S.l.: s.n.], 2018. p.6306–6314.

CHOI, L. K.; YOU, J.; BOVIK, A. C. Referenceless prediction of perceptual fog density and perceptual image defogging. **IEEE Transactions on Image Processing**, [S.l.], v.24, n.11, p.3888–3901, 2015.

DABOV, K. et al. Image denoising by sparse 3-D transform-domain collaborative filtering. **IEEE Transactions on image processing**, [S.l.], v.16, n.8, p.2080–2095, 2007.

EILERTSEN, G. et al. HDR image reconstruction from a single exposure using deep CNNs. **ACM transactions on graphics (TOG)**, [S.l.], v.36, n.6, p.1–15, 2017.

FARBMAN, Z. et al. Edge-preserving decompositions for multi-scale tone and detail manipulation. **ACM Transactions on Graphics (TOG)**, [S.l.], v.27, n.3, p.1–10, 2008.

GALDRAN, A. Image dehazing by artificial multiple-exposure image fusion. **Signal Processing**, [S.l.], v.149, p.135–147, 2018.

GNANADESIKAN, R.; PINKHAM, R.; HUGHES, L. P. Maximum likelihood estimation of the parameters of the beta distribution from smallest order statistics. **Technometrics**, [S.l.], v.9, n.4, p.607–620, 1967.

GUO, X.; LI, Y.; LING, H. LIME: low-light image enhancement via illumination map estimation. **IEEE Transactions on image processing**, [S.l.], v.26, n.2, p.982–993, 2016.

GUPTA, A. K.; NADARAJAH, S. **Handbook of beta distribution and its applications**. [S.l.]: CRC press, 2004.

HUANG, S.-C.; CHENG, F.-C.; CHIU, Y.-S. Efficient contrast enhancement using adaptive gamma correction with weighting distribution. **IEEE transactions on image processing**, [S.l.], v.22, n.3, p.1032–1041, 2012.

KE, J. et al. Musiq: multi-scale image quality transformer. In: IEEE/CVF INTERNATIONAL CONFERENCE ON COMPUTER VISION, 2021. **Proceedings...** [S.l.: s.n.], 2021. p.5148–5157.

LAO, S. et al. Attentions Help CNNs See Better: attention-based hybrid image quality assessment network. In: IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2022. **Proceedings...** [S.l.: s.n.], 2022. p.1140–1149.

LEE, C.; LEE, C.; KIM, C.-S. Contrast enhancement based on layered difference representation of 2D histograms. **IEEE transactions on image processing**, [S.l.], v.22, n.12, p.5372–5384, 2013.

LEE, J.; PANT, S. R.; LEE, H.-S. An adaptive histogram equalization based local technique for contrast preserving image enhancement. **International Journal of Fuzzy Logic and Intelligent Systems**, [S.l.], v.15, n.1, p.35–44, 2015.

LIU, Y.-L. et al. Single-image HDR reconstruction by learning to reverse the camera pipeline. In: IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2020. **Proceedings...** [S.l.: s.n.], 2020. p.1651–1660.

MA, K.; ZENG, K.; WANG, Z. Perceptual quality assessment for multi-exposure image fusion. **IEEE Transactions on Image Processing**, [S.l.], v.24, n.11, p.3345–3356, 2015.

MERTENS, T.; KAUTZ, J.; VAN REETH, F. Exposure fusion: a simple and practical alternative to high dynamic range photography. In: COMPUTER GRAPHICS FORUM, 2009. **Anais...** [S.l.: s.n.], 2009. v.28, n.1, p.161–171.

MIN, D. et al. Fast global image smoothing based on weighted least squares. **IEEE Transactions on Image Processing**, [S.l.], v.23, n.12, p.5638–5653, 2014.

MITTAL, A.; MOORTHY, A. K.; BOVIK, A. C. No-reference image quality assessment in the spatial domain. **IEEE Transactions on image processing**, [S.l.], v.21, n.12, p.4695–4708, 2012.

MITTAL, A.; SOUNDARARAJAN, R.; BOVIK, A. C. Making a “completely blind” image quality analyzer. **IEEE Signal processing letters**, [S.l.], v.20, n.3, p.209–212, 2012.

MONOBE, Y. et al. Dynamic range compression preserving local image contrast for digital video camera. **IEEE Transactions on Consumer Electronics**, [S.l.], v.51, n.1, p.1–10, 2005.

OWEN, C. B. **Parameter estimation for the beta distribution**. [S.l.]: Brigham Young University, 2008.

PIZER, S. M. et al. Adaptive histogram equalization and its variations. **Computer vision, graphics, and image processing**, [S.l.], v.39, n.3, p.355–368, 1987.

PIZER, S. M. et al. Contrast-limited adaptive histogram equalization: speed and effectiveness. In: PROCEEDINGS OF THE FIRST CONFERENCE ON VISUALIZATION IN BIOMEDICAL COMPUTING, 1990., 1990. **Anais...** [S.l.: s.n.], 1990. p.337–338.

QI, Y. et al. A comprehensive overview of image enhancement techniques. **Archives of Computational Methods in Engineering**, [S.l.], p.1–25, 2021.

QIN, X. et al. FFA-Net: feature fusion attention network for single image dehazing. In: AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE, 2020. **Proceedings...** [S.l.: s.n.], 2020. v.34, n.07, p.11908–11915.

SHEIKH, H. R.; BOVIK, A. C. A visual information fidelity approach to video quality assessment. In: THE FIRST INTERNATIONAL WORKSHOP ON VIDEO PROCESSING AND QUALITY METRICS FOR CONSUMER ELECTRONICS, 2005. **Anais...** [S.l.: s.n.], 2005. v.7, n.2, p.2117–2128.

SHEIKH, H. R.; BOVIK, A. C. Image information and visual quality. **IEEE Transactions on image processing**, [S.l.], v.15, n.2, p.430–444, 2006.

SONG, Y. et al. Vision Transformers for Single Image Dehazing. **arXiv preprint arXiv:2204.03883**, [S.l.], 2022.

UEDA, Y. et al. Hue-preserving color contrast enhancement method without gamut problem by using histogram specification. In: IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP), 2018., 2018. **Anais...** [S.l.: s.n.], 2018. p.1123–1127.

VIJAYALAKSHMI, D.; NATH, M. K.; ACHARYA, O. P. A comprehensive survey on image contrast enhancement techniques in spatial domain. **Sensing and Imaging**, [S.l.], v.21, n.1, p.1–40, 2020.

VONIKAKIS, V. **TM-DIED**: the most difficult image enhancement dataset. Accessed: 2022-06-25, `https://sites.google.com/site/vonikakis/datasets/tm-died`.

WANG, H.; XU, K.; LAU, R. W. Local Color Distributions Prior for Image Enhancement. In: EUROPEAN CONFERENCE ON COMPUTER VISION, 2022. **Anais...** [S.l.: s.n.], 2022. p.343–359.

WANG, Y.; CHEN, Q.; ZHANG, B. Image enhancement based on equal area dualistic sub-image histogram equalization method. **IEEE transactions on Consumer Electronics**, [S.l.], v.45, n.1, p.68–75, 1999.

WANG, Z. et al. Image quality assessment: from error visibility to structural similarity. **IEEE transactions on image processing**, [S.l.], v.13, n.4, p.600–612, 2004.

WEBER, N. et al. Rapid, detail-preserving image downscaling. **ACM Transactions on Graphics (TOG)**, [S.l.], v.35, n.6, p.1–6, 2016.

XUE, W. et al. Gradient magnitude similarity deviation: a highly efficient perceptual image quality index. **IEEE transactions on image processing**, [S.l.], v.23, n.2, p.684–695, 2013.

YAN, J.; LI, J.; FU, X. No-reference quality assessment of contrast-distorted images using contrast enhancement. **arXiv preprint arXiv:1904.08879**, [S.l.], 2019.

YANG, X. et al. Image correction via deep reciprocating HDR transformation. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2018. **Proceedings...** [S.l.: s.n.], 2018. p.1798–1807.

YING, Z. et al. A new low-light image enhancement algorithm using camera response model. In: IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION WORKSHOPS, 2017. **Proceedings...** [S.l.: s.n.], 2017. p.3015–3022.

YUAN, L.; SUN, J. Automatic exposure correction of consumer photographs. In: EUROPEAN CONFERENCE ON COMPUTER VISION, 2012. **Anais...** [S.l.: s.n.], 2012. p.771–785.

ZHANG, L. et al. FSIM: a feature similarity index for image quality assessment. **IEEE transactions on Image Processing**, [S.l.], v.20, n.8, p.2378–2386, 2011.

ZHANG, L.; ZHANG, L.; BOVIK, A. C. A feature-enriched completely blind image quality evaluator. **IEEE Transactions on Image Processing**, [S.l.], v.24, n.8, p.2579–2591, 2015.

ZHANG, Q.; NIE, Y.; ZHENG, W.-S. Dual illumination estimation for robust exposure correction. In: COMPUTER GRAPHICS FORUM, 2019. **Anais...** [S.l.: s.n.], 2019. v.38, n.7, p.243–252.

ZHANG, R. et al. The unreasonable effectiveness of deep features as a perceptual metric. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2018. **Proceedings...** [S.l.: s.n.], 2018. p.586–595.

ZHANG, W. et al. Blind image quality assessment using a deep bilinear convolutional neural network. **IEEE Transactions on Circuits and Systems for Video Technology**, [S.l.], v.30, n.1, p.36–47, 2018.

APPENDIX A CREDITS TO INPUT IMAGES

The input images in Figures 1.1 (page 15), 4.4 (page 33) (first row), 5.5 (page 57) and 5.6 (page 57) are from TM-DIED dataset (VONIKAKIS, 2021). The input image in Figure 4.2 (page 29) is from DICM dataset (LEE; LEE; KIM, 2013), where a conversion for grayscale was realized. The input images in Figures 4.3 (page 31) and 4.4 (page 33) (second row) are from pexels.com, by Julius Silver and Ron Lach, respectively. The input images in Figures 5.1 (page 50), 5.2 (page 52) and 5.3 (page 54) are from the datasets used in their respective experiments. The input image in Figure 5.4 (page 55) is from (FARBMAN et al., 2008), a courtesy of Norman Koren in that paper.