Investigation Results of State of Art Algorithms in Denoising Text Images

Nastaran Emamjomeh

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Electrical and Electronic Engineering

Eastern Mediterranean University
September 2021
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Ali Hakan Ulusoy
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science in Electrical and Electronic Engineering.

Assoc. Prof. Dr. Rasime Uyguroğlu
Chair, Department of Electrical and Electronic Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Electrical and Electronic Engineering.

Prof. Dr. Hüseyin Özkaramanlı
Supervisor

Examining Committee

1. Prof. Dr. Hasan Amca
2. Prof. Dr. Hüseyin Özkaramanlı
3. Asst. Prof. Dr. Burcin Özmen

ABSTRACT

One of the crucial issues in acquired images is the corruption due to various kinds of noise. Recovering vital information of the original image from the noisy image is the purpose of image denoising. Denoising text images is of special interest since they convey important information. They are, however, corrupted by noise due to scanning, transfer or digitization. Old document images which date back to many years are affected by noise and deteriorations more seriously. Some of the deteriorations include ink-leakage from the black page, brown depigmentation of paper, fading text, background points. Reduction of noise, or denoising, is a vital step in the image processing of documents. Various methods for decreasing noise have been proposed by researchers already. Weighted Nuclear Norm Minimization (WNNM) will employ the low-rank standard. This algorithm uses low-rank models and produces good denoising results.

This thesis applies WNNM method to different text images and compares this method with some other traditional denoising methods, such as Block Matching and 3D Filtering (BM3D), Non-local Centralized Sparse Representation (NCSR), and Expected Patch Log Likelihood (EPLL). A preliminary investigation of the various methods is also carried out for some ancient document images. The results show that using the WNNM method and denoised images have a higher peak signal-to-noise ratio (PSNR) than other methods. For example, for $\sigma = 10$, WNNM denoised images in average have PSNR=32.21, while BM3D denoised images in average have PSNR=31.84, EPLL denoised images have PSNR=31.47, and NCSR denoised images have PSNR=31.93.

Keywords: Image Processing, Image Denoising, Text Image Denoising, Weighted Nuclear Norm Minimization Algorithm, Low-rank Minimization.

ÖZ

Edinilen görüntülerdeki önemli konulardan biri, çeşitli gürültüler nedeniyle oluşan bozulmadır. Orijinal görüntünün hayati bilgilerini gürültülü görüntüden kurtarmak, görüntü denoising'in amacıdır. Metin görüntülerinin devaziyesi, önemli bilgiler aktardıkları için özel ilgi çekicidir. Bununla birlikte, tarama, aktarım veya dijitalleştirme nedeniyle gürültü nedeniyle bozulurlar. Uzun yıllara dayanan eski belge görüntüleri gürültüden ve bozulmalardan daha ciddi şekilde etkilenir. Bozulmalardan bazıları siyah sayfadan mürekkep sızıntısı, kağıdın kahverengi depigmentasyonu, solgun metin, arka plan noktalarıdır. Gürültünün azaltılması veya denoising, belgelerin görüntü işlenmesinde hayati bir adımdır. Gürültüyü azaltmak için çeşitli yöntemler araştırmacılar tarafından zaten önerilmiştir. Ağırlıklı Nükleer Norm En Aza İndirme (WNNM) düşük dereceli standardı kullanacaktır. Bu algoritma düşük dereceli modeller kullanır ve iyi denoising sonuçları üretir.

Bu tez, farklı metin görüntülerine WNNM yöntemi uygular ve bu yöntemi Blok Eşleştirme ve 3B Filtreleme (BM3D), Yerel Olmayan Seyrek Gösterim (NCSR) ve Beklenen Düzeltme Eki Günlüğü Olasılığı (EPLL) gibi diğer bazı geleneksel denoising yöntemleriyle karşılaştırır. Bazı eski belge görüntüleri için çeşitli yöntemlerin ön araştırması da yapılmaktadır. Sonuçlar, WNNM yöntemini ve denoised görüntüleri kullanmanın diğer yöntemlere göre daha yüksek bir tepe sinyal-gürültü oranına (PSNR) sahip olduğunu göstermektedir. Örneğin, σ=10 için WNNM denoised görüntüler ortalama PSNR=32.21, BM3D denoised görüntüler ortalama PSNR =31.84, EPLL denoised görüntüler PSNR =31.47 ve NCSR denoised görüntüler PSNR = 31.93 vardır.

Anahtar Kelimeler: Görüntü İşleme, Görüntü Denoising, Metin Görüntü Denoising, Ağırlıklı Nükleer Norm En Aza indirme Algoritması, Düşük Dereceli En Aza İndirme.

ACKNOWLEDGEMENT

I would like to record my gratitude to Prof. Dr. Huseyin Ozkaramanli for his supervision, advice, and guidance from the very early stage of this thesis and for giving me extraordinary experiences throughout the work. Above all and the most needed, he provided me with constant encouragement and support in various ways. His ideas, experiences, and passions have truly inspired and enriched my growth as a student. I am indebted to him more than he knows.

My thanks goes to my family, who has always supported me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AWGN	Additive White Gaussian Noise
BM	Block Matching
BM3D	Block-Matching and 3D filtering method
DWT	Discrete Wavelet Transform
EPLL	Expected Patch Log Likelihood
FFT	Fast Fourier Transform
GMM	Gaussian Mixture Model
LRMF	Low Rank Matrix Factorization
MSE	Mean Square Error
NCSR	Non-local Centralized Sparse Representation
NL-means	Non-Local Means
NNM	Nuclear Norm Minimization
NSS	Non-Local Self-Similarity
OCR	Optical Character Recognition
PLE	Piecewise Linear Estimator
POSHE	Partially Overlapped Sub-Block Histogram Equalization
PSNR	Peak Signal-to-Noise Ratio
SAIST	Spatially Adaptive Iterative Singular value Thresholding
SPN	Stroke like Pattern Noise
SNR	Signal-To-Noise Ratio
SVD	Singular Value Decomposition
WNNM	Weighted Nuclear Norm Minimization

Chapter 1

INTRODUCTION

In modern life today, digital images play a vital role in daily uses like: satellite television, magnetic resonance imaging, digital cameras, and research and technology like Geographical Information System. Generally, datasets collected by image sensors are contaminated by noise. Imperfect instruments problems with the data acquisition process, and interfering natural phenomena can all corrupt the data of interest [1].

Image denoising means removing noise from a noisy image to recover a clean image. The image denoising problem is modelled like the formula below mathematically:

$$y = x + n \tag{1.1}$$

y is the noisy picture, *x* is the clear image which is unknown, and *n* shows Additive White Gaussian Noise (AWGN) with a standard deviation σ_n, which can be estimated in practical applications by various methods. The purpose of noise reduction is to decrease the noise in natural images while minimizing the loss of original features and improving the Signal-to-Noise Ratio (SNR) [2].

Recently, because of the increase in using computers in everybody's lives, converting documents into digital and comprehensible data is necessary. One way of changing printed documents into digital forms is by scanning documents. A usual problem faced when scanning documents is the noise that may happen in a picture due to the quality of the paper [4]. Any image processing method can have a few phases like (i) Pre-processing, (ii) Segmentation, (iii) Recognition and (iv) Post-processing. This pre-processing stage is a crucial stage, which primarily deals with noise removal [5].

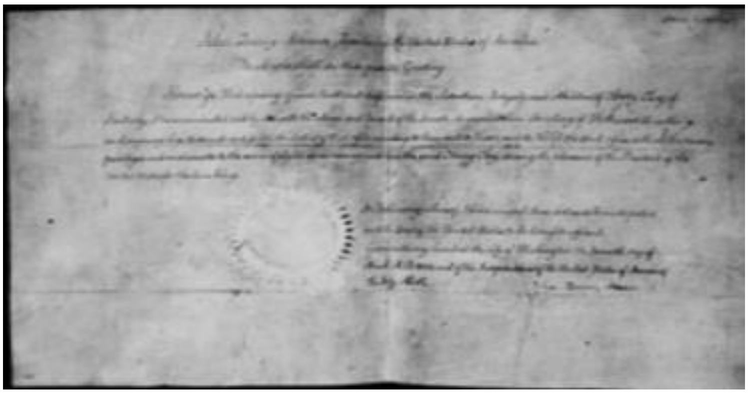
1.1 Problem Statement

In document images, noise reduces the accuracy of subsequent tasks of OCR (Optical Character Recognition) systems. It can appear in the foreground or background of an image and could be generated before or after scanning. The page rule line is a source of noise that interferes with text objects. The marginal noise usually appears in a large dark region around the document image and can be textual or non-textual. Some forms of clutter noise appear in an image because of document skew while scanning or are from holes punched in the document or background noise. This includes uneven contrast, show through effects, interfering strokes, background spots, etc. [4]. One example of this problem is the digital scan of books and documents in libraries. Recently, libraries wanted to keep their books in digital format, using an optical scanner to convert them into images. The marginal noise usually appears in a large and dark region around the margin of document images. It is one of the main problems found when scanning documents, especially thick books and old books. It is not only unpleasant to view on a display device but also cause a problem when library's users want to print this document image for reading, as shown in Figure1.1.

Figure 1.1: Vertical and horizontal marginal noise example [6].

From the study, some of the marginal noise was removed by using commercial programs like Adobe Photoshop by most librarians that scanned their books. This procedure makes librarians work inefficiently because all images are scanned at a high resolution of about 400 dpi. Additionally, if they want to edit this image it is time consuming. This problem was solved by cutting the thick book into separate pages and scanning it page by page, but this procedure damages the original text and marginal noise remains [6]. By filtering, we can remove this noise. The filters are two types: linear and nonlinear. Mean filters and wiener filters [30] are the linear filters that are utilized for noise reduction. They have specific disadvantages like blurring the edges and damaging lines, hence why nonlinear filters like median filters are used. This overcomes the limitations of linear filters to some degree, but it has disadvantages such as: the loss of corners and threads, blurring text content document records [5].

In addition, historical writings and scanned document photos usually have degradations like uneven contrast, show through effects, interfering strokes, background spots, humidity absorbed by paper in different areas, and uneven backgrounds. These problems cause challenges similar to those in an OCR system. Such degradations can destroy the blank spaces between lines and words. There are many methods to enhance background degradations in document images [4].



(a)

(b)

Figure 1.2: Examples of degradation of the background [4]

Pepper noise can appear in a document image during the conversion process and is also caused by dirt on the document. This noise can be composed of one or more pixels, but by definition, they are assumed to be much smaller than the size of the text objects. Simple filters can remove isolated pepper noise like median, but algorithms like k-fill or morphological operators will be more effective for noise removal if they are larger than that.

Printed documents come in many forms and infinite varieties of writing ink, and salt noise looks like a lack of ink in the document image. If the fragmentation is very high, it reduces segmentation and recognition accuracy. Simple filters can remove isolated salt noise like median. In 2007, a morphological-based method was proposed. This method solved one of the most critical problems of morphology-based approaches by using a learning phase for finding the parameters of a suitable structuring element. Shortly after, a dilation operator is used to fill places where there is a lack of ink. This method experiences several issues such as a high execution time as a result of the learning phase. It produced undesirable connections between some characters, particularly when the fonts were very thick [4].

1.2 Background of the Study

In earlier work, the image denoising is always assumed to be smooth or sparse under some prior knowledge, such as gradient. A classical model is a Total Variation (TV) model proposed by Rudin, Osher and Fatemi. It can effectively suppress the noise in an image and preserve the edges of the image. However, the recovered image usually suffers from the staircase effect [3].

Lately, patch-based image denoising with a non-local principle has led to several state-of-the-art algorithms. These algorithms exploit the self-similarity of natural images as prior knowledge. The Block-Matching and 3D filtering (BM3D) method has become the benchmark for denoising algorithms. More recently, another prior-named low rank has also been adopted for image denoising, such as Spatially Adaptive Iterative Singular Value Thresholding (SAIST), Low Rank Regularized Collaborative Filtering (LRCF), and Weighted Nuclear Norm Minimization (WNNM). This is because the matrix formed by stacking non-local similar patches from a noisy image will satisfy the low-rank criterion [3].

In-text documents, the noises exist in a document can be reduced by three approaches: (I) using of human to identify the noise type, and then applying the appropriate filters, (2) applying a bunch of filters directly to a document, and (3) presuming (without human identifier) that a document consists of a certain noise type. The first approach is not efficient since human-based noise identification is a time and resource-consuming task. The second approach can be categorized as a trial and error mechanism. Applying many filters to a noisy document can become a redundant task because there is no guarantee that the filters are applicable for the noises that exist in the documents. Even if they are applicable, their strength does not always suit the noise strength. The third approach will be successful only if the used noise model is accurate. Otherwise, the denoising procedure may degrade the documents. Considering various noise types inherently exist in ancient digital documents and the insufficiently of available noise reduction approaches to be directly applied to ancient noisy documents. An additional step, which is noise characterization, is required. Using this extra step means no need to employ a human identifier and trial and error mechanism. In addition, the assumption of a noise model is not required as well [7].

1.3 Thesis Contribution

This thesis studies different methods of image denoising, for Images which are corrupted by adding Additive White Gaussian Noise (AWGN) with a standard deviation σ_n. This noise type is a basic noise model which is used in information theory to simulate the effect of many random processes that happen in nature and elaborates about different methods of text and document denoising that has been done before. After adding noise, we then apply BM3D, EPLL, NCSR, and WNNM methods on 20 sample text images and some famous test images. Comparison of above methods is carried out to determine how effective they are in text image denoising. Finally, the performance of these methods is investigated on an ancient image. Furthermore, throughout the thesis, we have tested four traditional state of the art denoising algorithms on images that were contaminated with Additive White Gaussian Noise, which is assumed the worst possible noise for text images, not other types of noise. We then focused on WNNM denoising method for text images.

Chapter 2

LITERATURE REVIEW

2.1 Background

Image denoising is an applicable issue found in diverse image processing and computer vision problems. There are various existing methods for denoising images. The important property of a good image denoising model is that it should completely remove noise, as well as preserve edges. There have been numerous published algorithms and each approach has its assumptions, advantages and limitations. Image denoising has remained a fundamental problem in the field of image processing. Due to properties like sparsity and multiresolution structure, Wavelet Transform [31] have become an attractive and efficient tool in image denoising. Wavelet Transform is a signal representation similar to Fourier except that it uses basis functions (Wavelets), which have finite support (locally defined). DWT represents functions by dilation and shift of the basis Wavelets. With Wavelet Transform gaining popularity in the last two decades, various algorithms for denoising in Wavelet Domain were introduced. The focus was shifted from the Spatial and Fourier domain to the Wavelet Transform domain. Although Donoho's theory was not revolutionary, his methods did not need tracking or correlation of the Wavelet maxima and minima across the different scales as proposed by Mallat. Thus, there was a rehabilitated interest in Wavelet-based denoising techniques since Donoho demonstrated a simple approach to a difficult problem. Researchers published different ways to compute the parameters for the Wavelet thresholding. Data adaptive thresholds were introduced to attain an optimum value of the threshold. Later efforts found that considerable improvements in perceptual quality could be obtained by changing invariant methods based on the thresholding of an Undecimated Wavelet Transform. These thresholding techniques were applied to the non-orthogonal Wavelet coefficients to reduce artifacts. Multiwavelets were also used to get similar results. Probabilistic models using the statistical properties of the Wavelet coefficient seemed to outperform the thresholding techniques and gained ground. Recently, much effort has been devoted to Bayesian denoising in Wavelet domain, Hidden Markov Models and Gaussian Scale. Data adaptive transforms such as Independent Component Analysis (ICA) have been explored for sparse reduction. The development continues to focus on using different statistical models to model the statistical properties of the Wavelet coefficients and their neighbors.

There are various methods of image denoising. Aims of any of the approaches for filtering are:

- To suppress the noise effectively in uniform regions.
- To preserve edges and other similar image characteristics.
- To provide a visually natural appearance.

2.2 Transform Domain Filtering

Transformation or frequency domain techniques are based on manipulating the orthogonal transform of the image rather than the image itself. Transformation domain techniques are suited for processing the image according to the frequency content. The principle behind the frequency domain methods of image enhancement consists of computing a 2-D discrete unitary transform of the image, for instance, the 2-D DFT, manipulating the transform coefficients by an operator M and then performing the inverse transform. The orthogonal transform of the image has two components: magnitude and phase. The magnitude consists of the frequency content of the image. The phase is used to restore the image back to the spatial domain. The usual transform domain enables operation on the frequency content of the image; therefore high-frequency content such as edges and other subtle information can easily be enhanced [32]. The transform domain filtering can be separated based on choices of fundamental operations.

2.2.1 Spatial Frequency Filtering

It refers to the use of low pass filters using fast Fourier Transform. The noise is removed by deciding a cut-off frequency and adapting a frequency-domain filter when the noise components are decorrelated from the useful signal. The main disadvantage of Fast Fourier Transform is that the edge information is spread across frequencies because of FFT basis function. It is not localized in time or space, which means that time information is lost and hence low pass filtering results in the smearing of the edges. But the localized nature of Wavelet Transform both in time and space provides a particularly useful method for image denoising when the preservation of edges in the scene is of importance.

2.2.2 Wavelet Domain Filtering

Working in Wavelet domain is preferred because the Discrete Wavelet Transform (DWT) make the signal energy concentrate in a small number of coefficients. Hence, the DWT of the noisy image consists of a small number of coefficients having a high Signal to Noise Ratio (SNR) while a relatively large number of coefficients has low SNR. After removing the coefficients with low SNR (i.e., noisy coefficients), the image is reconstructed by using inverse DWT. As a result, noise is removed or filtered from the observations. A significant advantage of Wavelet methods is that it provides time and frequency localization simultaneously. Moreover, Wavelet methods characterize such signals much more efficiently than either the original domain or transforms with global basis elements such as the Fourier transform [1].

2.3 Non-local Regularization

While local denoising methods have low time complexities, the performances of these methods are limited when the noise level is high. The reason for this is that the correlations of neighborhood pixels are seriously disturbed by high-level noise. Lately, some methods have applied the NSS prior. This is because images contain extensive, similar patches at different locations. Pioneering work on Non-Local Means (NLM) used the weighted filtering of the NSS (Non-local Self-Similarity) prior to achieving image denoising, which is the most notable improvement for the problem of image denoising. Its basic idea is to build a pointwise estimation of the image, where each pixel is obtained as a weighted average of pixels centered at regions that are similar to the area centered at the estimated pixel. At present, most research on image denoising has shifted from local methods to non-local methods [2].

2.4 Sparse Representation

In sparse representation, one usually learns on an over complete dictionary using either a set of training data or noisy image patches. For obtaining the dictionary, we can use several optimization techniques. One of the most notable algorithms is KSVD algorithm [33]. Once the dictionary is learned, an image patch can be approximately represented with such a few atoms from the dictionary, rather than using all the dictionary atoms. Such a representation is referred to as sparse representation since patches are represented with few atoms. Many current image denoising methods exploit the scarcity prior of natural images. Methods in this category are all local, meaning they ignore the correlation between non-local information of the image. In the case of high noise, local information is seriously disturbed, and the result of denoising is not effective. Coupled with the NSS prior, the sparsity from self-similarity properties of natural images, which has received significant attention in the image processing community, is widely applied for image denoising. One representative work is the Non-local Centralized Sparse Representation (NCSR) model. The NCSR model naturally integrates NSS into the sparse representation framework, and it is one of the most commonly considered image denoising methods at present. NCSR is very effective in reconstructing both smooth and textured regions [2].

2.5 Different Kinds of Noises in Text Photos

Document images may be contaminated with noise during transmission, scanning or conversion to digital form. Different kinds of noises contaminate text images.

2.5.1 Ruled Line Noise

They usually write documents that are handwritten on lined paper that is pre-printed. Lines may result in the challenges:

- The ruled lines interfere with and connect to the text.
- Variable thicknesses in the ruled lines cause problems for the noise removal algorithms.
- Broken ruled lines cause problems for algorithms detecting them.
- Some letters, for example, 'z', which have horizontal lines, are removed by the algorithms as they are incapable of detecting differences between them and the ruled lines [4].

2.5.2 Marginal Noise

Marginal noise is another type of noise defined as dim shadows that emerge in upright or lateral borders of a photo. This type of noise is the result of thick scanning documents or the borders of pages in books; it can be textual or non-textual.

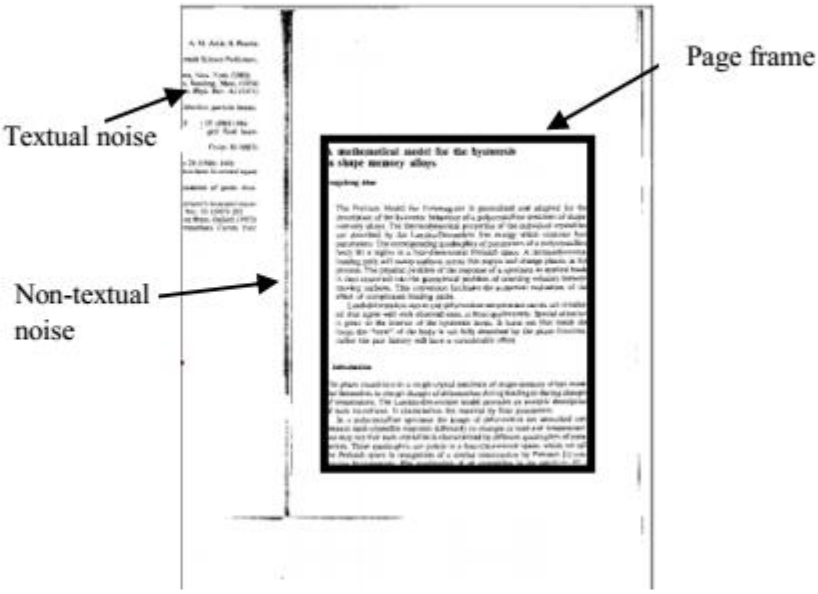


Figure 2.1: Example of marginal noise [4]

2.5.3 Clutter Noise

Another type of noise is clutter noise, which refers to unwanted foreground content that is typically larger than the text in binary images. This results from numerous sources such as punched holes, document image skew, or connecting vast amounts of pepper noise. The significant feature of clutter noise is larger than the text objects in the document image. One of the significant challenges facing clutter is its connectivity with text. Clutter often touches or overlaps some parts of the text, reducing segmentation and recognition accuracy in OCR systems.

2.5.4 Stroke Like Pattern Noise

Another type of noise is Stroke-like Pattern Noise (SPN), independent of the size or other properties of the text in the document image. SPN is similar to diacritics so its presence near textual components can change the meaning of a word. This noise is formed primarily due to the degradation or unsuccessful removal of underlying ruled lines that interfere with the foreground text, or it is formed by the remaining clutter noise after clutter removal approaches [4].

2.5.5 Background Noise

Today's OCR systems can only read black characters on the identical white background or in reverse with high recognition exactness. But when text is printed on a complicated background, colored background or background patterns (regular/not regular, periodic/not periodic), or just in case of noisy background, OCR systems do not operate well. Examples of such areas of overlapping text and background photos can be found in mail pieces where the address is written on pattern papers [8].

2.5.6 Salt and Pepper Noise

The first step in document image analysis is to capture a paper document into a binary image, the resulting image can be contaminated by salt-and-pepper noise during the conversion process and also from dirtiness in the document itself. This appears as randomly distributed black dots on a white background and white dots on a black document component. Each noise dot can be either an isolated pixel or composed of more than one pixel [9].

2.6 Ruled Line Noise Denoising Methods

Several methods have been proposed for ruled line removal. The methods can be divided into three major groups. First, there are mathematical morphology-based methods that depend on prior knowledge. The second group contains methods that employ Hough Transform to extract text features and find lines in every direction. The methods in the last group use Projection Profiles to estimate lines hence reducing the problem's dimensions, which then improves the accuracy of the first step in some methods of noise removal.

2.6.1 Mathematical Morphology Based Methods

The mathematical morphology-based methods are limited by the design and application of the structuring elements, which often require knowledge of the font size or use trial and error. Structuring elements are used to probe an image and draw conclusions on how they fit or miss the shapes in the image. Following the aforementioned step, some operations such as dilation are used to highlight the extracted features from the patterns in order to remove them more easily.

Methods in this group are based on tracing line like structures as candidates for rule lines for removal. In these methods, a structuring element is used to find the line patterns to facilitate the removal of the ruled lines by dilation and erosion. Because the structuring elements are designed for special purposes, these methods are incapable of handling large variations in the thickness of the ruled lines. On the other hand, with these methods no difference is perceived between the ruled lines and characters with horizontal strokes (such as 'z'), so removal of too many text pixels makes the recognition phase more difficult [4].

2.6.2 Hough Transform Based Methods

Hough Transform aims to find imperfect instances of objects within a certain class of shapes using a voting procedure. The voting procedure is carried out in parameter space. Object candidates are obtained as local maxima in a so-called accumulator space explicitly constructed by the algorithm to compute Hough Transform. It can be used to find straight lines, such as ruled lines, in an image. By extracting the dominant features of an image, Hough can find lines in every direction; this group of methods, therefore, is robust against document rotation as an earlier group. Methods using Hough Transform are computationally expensive but more robust against noise; they also cope better with broken lines than other methods.

A Hough Transform-based method was proposed to remove ruled lines in 1990 [10]. However, the method had problems that were mentioned earlier, so Random Hough

Transform was proposed, which performed better but because of the high computational cost, neither one is used.

2.6.3 Projection Profile-based Approaches

Projection Profile-based approaches operate by creating a horizontal histogram in which the hills of the histogram are the center locations of the horizontal ruled lines. Projection profiles ignore the line's thickness. Therefore, in the removal phase, the characters with horizontal strokes will be broken up. Another problem with this group of methods is sensitivity to rotation. However, compared to the methods mentioned before, reducing the problem's dimensions makes this group faster. The successful methods in this group have two phases: First, an image's projection profile helps estimate the ruled lines. Second, we make our estimation more accurate using some other methods such as searching vertical run lengths [11]. These groups of methods solve the third problem of ruled lines, as mentioned earlier.

2.7 Marginal Noise Denoising Methods

We can divide approaches of removing marginal noise into two sections. The first section recognizes and deletes noisy elements; the Second section identifies the actual content area or page frame of the document [4].

2.7.1 Identification of Noise Elements

The methods in this group search for the noise patterns in an image by extracting its features, then removing areas that contain those patterns. Zheng Zhang et al.'s [12] method employed vertical projection to recover document images that have marginal noise and decided whether the marginal noise was on the left or right side of the image based on the location of peaks in the profile. Then, by using extracted features, it detects the boundary between the shadows and cleans the area. However, this method suffers from the following problems:

- Because of using features like black pixels, there is no peak in projections to locate marginal noise in images that have marginal noise areas that are smaller than the text areas. Thus, it is not suitable for noises with variable regions.
- Because of ignoring the extraction of features in horizontal directions, this method cannot locate marginal noises in the horizontal margins of a page.

To overcome these problems, another algorithm was proposed. This algorithm has three steps:

1. Resolution reduction.
2. Block splitting to find possible local boundaries between connected blocks.
3. Block identification to determine which blocks contain marginal noise [13].

In 2004, Peerawit employed Sobel edge detection and identified noises to be removed by comparing the edge density of marginal noise and text. This method uses density as the selected feature because edge density is higher in noise than text. If the document has only non-textual marginal noise, this method cannot find significant differences between edge densities and, hence, cannot detect marginal noise. Moreover, this method is not suitable for detecting marginal noise in a small area [6].

2.7.2 Identifying Text Components

Another group of methods finds the page frame of the document, which it defines as the smallest rectangle that encloses all the foreground elements of the document image. This group performs better than the previous one because searching for text patterns is more straightforward than searching for noise features in a document.

In 2008 Shafait proposed a method that works in two steps. First, a geometric model is built for the page frame. Then a geometric matching method is employed in finding the globally optimal page frame with respect to a defined quality function. Although the method works well in practice, it requires prior text line extraction, which increases the computational cost and is hard to implement [14]. To overcome the shortcomings of this method, another algorithm was proposed that works in three steps:

1. A black filter is used; it selects them if the black regions are bigger than a pre-defined threshold area.
2. Connected component removal is used; first, all connected components are extracted from the image after applying a black filter. All components close to the image border are considered noise and, hence, removed from the image. Selecting an appropriate value for the threshold is dependent on prior knowledge.
3. A white filter is used; it extracts features similar to the black filter and removes everything up to the border if it finds a big white block.

2.8 Stroke like Pattern Noise Denoising Methods

The situation is challenging where the ruled lines are broken and degraded, as they cannot be perceived in straight lines even by the human eye. Thus, techniques like Hough Transform and projection profiles are inappropriate in such cases. Furthermore, because of their similarity in shape and size to smaller text components, morphology-based removal approaches are unsuitable because the successive erosion and dilation steps tend to degrade the text.

In 2011, Agrawal described the difference between SPN and ruled lines for the first time and proposed a solution. The method works in two steps. First, independent prominent text component features are extracted with a supervised classifier, then it uses their cohesiveness and stroke-width properties to filter smaller text components using an unsupervised classification technique [15].

2.9 Salt and Pepper Noise Denoising Methods

Simple filters like Median can remove isolated pepper noise. However, if it becomes bigger than it, rules such as k-fill or morphological operators would be more helpful in removing the noise. Documents that are printed publish in different shapes and with different kinds of ink and salt noise seems similar to ink shortage in the photo. It reduces segmentation and recognition accuracy whenever the fragmentation is very high.

Simple filters can remove isolated salt noise like median filter. In 2007, a morphological-based method was proposed. This method solved one of the most critical problems of morphology-based approaches by using a learning phase for finding the parameters of a suitable structuring element. Additionally, a dilation operator is used to fill places where there is a lack of ink. This method experienced some problems, such as a high execution time because of the learning phase. It produced undesirable connections between some characters, particularly when the fonts were very thick [4].

2.10 Background Noise Denoising

The denoising methods are divided into four main groups:

2.10.1 Binarization and Thresholding Based Methods

One of the methods to enhance the background quality of grayscale images employs thresholding and binarization techniques. Some resources divide thresholding techniques into two major groups. The methods in the first group use global algorithms which employ global image features to determine appropriate thresholds to divide image pixels into an object or background classes. The second group uses local image information to calculate thresholds, similar to the locally adaptive thresholding method that uses neighbourhood features such as the mean and standard deviation of pixels [16]. However, the methods of the second group are much slower than the first, but their accuracy is higher.

2.10.2 Histogram Based Methods

An image histogram acts as a graphical representation of the intensity distribution in an image. It plots the number of pixels for each intensity value. The histogram for a very dark image will have most of its data points on the left side and center of the graph. Conversely, the histogram for a very bright image with a few dark areas will have most of its data points on the right side and center of the graph, so the contrast in an image will be improved by using histogram equalization. Histogram-based methods solve most of the fuzzy logic-based method's problems.

Partially Overlapped Sub-Block Histogram Equalization or POSHE had been designed in the year 2001. The photo is separated into blocks. After that, in each of the blocks, histogram homology is performed. This method shows better implementation than previous approaches due to the extraction of local elements [17], [18].

2.10.3 Approaches Based On Morphology

Mathematical morphology is a robust methodology for enhancing uneven backgrounds. The operators are powerful tools for processing and analyzing shapes with structural features like borders, area etc. Methods in this group search for noise patterns, which appear as shadows in the background, with defined structuring elements. Then, in one or more steps, morphological operators like thickening and pruning remove shadows. Some algorithms in this group start with a pre-processing stage [4].

2.10.4 Genetic Algorithm Based Methods

The majority of difficulties arise during the separation of characters from the background. Backgrounds can have complex variations and a variety of degradations. To improve quality, well-known filters such as Fourier transform, Gabor filters, and wavelet transforms can be used. However, it is difficult for a single filtering technique to deal with a variety of degradations. Nagao et al. [19, 20] used GAs to construct an optimal sequence of image processing filters to extract characters from different sources to solve similar problems. In 2006, Kohmura extended previous work and used the algorithm for colour images. A filter bank of 17 well-known filters (mean, min, max, Sobel, etc.) was created in this approach to search for an optimal filtering sequence [21].

There are some problems, however, in using a genetic algorithm. The first is that the optimization procedure is relatively slow, as every fitness evaluation requires comparing two images. The second problem is the algorithm's inability to select appropriate filters for the optimization procedure automatically. In 2010 genetic algorithms were used to estimate the degradation function of an image. A degradation model has a degradation function that, together with an additive noise term, operates on an input image to produce a degraded image. In general, the more we know about the degradation function and the additive noise term, the better we are able to restore the image [22].

Chapter 3

METHODOLOGY

3.1 Introduction

The following chapter elaborates on the most successful traditional methods and algorithms of denoising, including the BM3D, EPLL, NCSR, and the WNNM, which have been applied in this thesis for comparison of sample text images chosen from the Internet. Several famous sample images include: Lena and Barbara, clarifying the WNNM denoising method. Finally, PSNR, which is used to measure quality of images, is explained.

3.2 BM3D Denoising Method

Collaborative filtering is the name of the BM3D procedure of filtering and grouping. It is realized in four stages:

- Finding image patches similar to a given image patch and applying them in a group in a 3D block;
- 3D linear transform of the 3D block;
- Shrinkage of the transform spectrum coefficients;
- Inverse 3D transformation. This 3D filter, therefore, filters out simultaneously all 2D image patches in the 3D block.

By attenuating the noise, collaborative filtering reveals even the finest details shared by the grouped patches. The filtered patches are then returned to their original positions. Since these patches overlap, many estimates are obtained which need to be combined for each pixel. Aggregation is a particular averaging procedure used to take advantage of this redundancy.

The first collaborative filtering step is much improved by a second step using Wiener filtering. This second step mimics the first step, with two differences. The first difference is that it compares the filtered patches instead of the original patches. The second difference is that the new 3D group (built with the unprocessed image samples but using the patch distances of the filtered image) is processed by Wiener filtering instead of a mere threshold. The final aggregation step is identical to those of the first step.

The proposed method improved on the NL-means method, which denoises jointly similar patches, but only by performing a patch average, which amounts to a 1D filter in the 3D block. The 3D filter in BM3D is performed on the three dimensions simultaneously.

The algorithms work in the case of Additive White Gaussian Noise where σ^2 denotes the variance. The algorithm is divided into two major steps:

1) The first stage provides an estimation of the denoised photo by the use of hard thresholding within the collaborative filtering.
2) The second step is based on the original noisy image, and the basic estimate received in the first step by using Wiener filtering. This step is defined by the demonstrative Wiener [23].

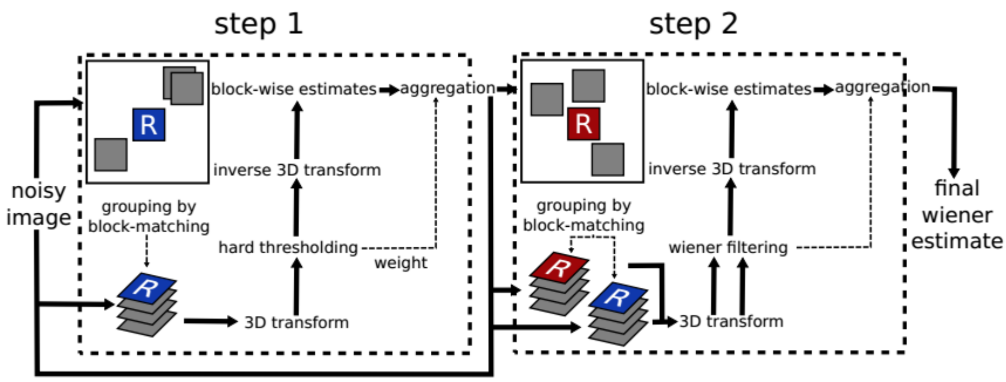


Figure 3.1: Scheme of the BM3D algorithm [24]

3.2.1 Block Matching

The Block-Matching (BM) is a particular matching approach that has been extensively used for motion estimation in video compression as a specific way of grouping. It is used to find similar blocks, then stacked together in a 3-D array (i.e., a group). An illustrative example of grouping by block-matching for images is given in Figure 3.2, where it shows a few reference blocks and the ones matched as similar to them [24].

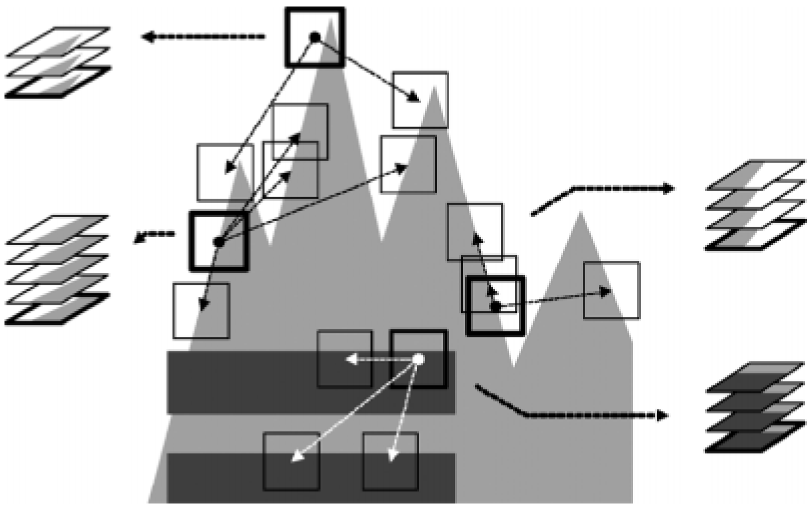


Figure 3.2: Simple example of grouping in an artificial image, where for each reference block (with thick borders) there exist perfectly identical ones [24].

3.2.2 Collaborative Filtering

Given a set of *n* pieces, the collaborative filtering of the set manufactures *n* guesses for each of the gathered pieces. Generally, these guesses can be dissimilar. The word "collaborative" is taken, which indicates that every one of the collected pieces cooperates to filter all the other fragments and reverse.

Constructive, collaborative filtering can be perceived as a reduction in the transform domain. Assuming d+1-dimensional groups of similar signal fragments are already formed, the collaborative shrinkage comprises of the following steps:

- Put in a d+1 -dimensional linear transform to the set.
- Reduce the transform coefficients to make the noise weak by using soft-thresholding, hard-thresholding, or Wiener filtering.
- Reverse linear transform to construct guesses of all gathered pieces.

The 3-D transform may take advantage of both kinds of correlation and, thus, produce a sparse representation of the true signal in the group. This sparsity makes the shrinkage very effective in attenuating the noise while preserving the features of the signal.

3.2.3 Algorithm

The algorithm is illustrated in figure 3.1 and proceeds like the following:

Step 1) Primary guess.

a) *Block-wise guesses*. For every one block in the photo which is noisy, the following steps are done:

I. *Grouping*. Find the blocks that are identical to the previously prepared ones and accumulate the blocks together in a 3-D group.

II. *Collaborative hard-thresholding*. Make use of a 3-D transform to the organized set, which weakens the noise by hard-thresholding of the transform coefficients, reverses the 3-D transform to make guesses of all gathered blocks, and turn back the guesses of the blocks to their original places.

b) *Aggregation*. Calculate the main guess of the actual photo by using weighted averaging of all overlapping gained block-wise guesses.

Step 2) *Final Estimate*: Intuitively executing enhanced gathering and collaborative Wiener filtering.

a) *Block-wise estimates*. For every one of the blocks, do the following procedures.

I. *Grouping*. Utilize block matching inside the main guess to find the blocks' places identical to the one which was formerly prepared. By utilizing these places, make two sets, one group from the noisy photo and one group from the main guess.

II. *Collaborative Wiener filtering*. On both groups, put in a 3-D transform. Execute Wiener filtering on the noisy set by utilizing the main guess's energy spectrum as the sample energy spectrum. By applying this in the inverse 3-D transform on the filtered coefficients this produces estimates of all grouped blocks and give back the guesses of the blocks to their main places.

b) *Aggregation*. Calculate a finishing guess of the true-photo by gathering all of the gained regional guessed by utilizing a weighted average [24].

3.3 EPLL Denoising Method

Image priors are a popular tool for image restoration tasks. However, the high dimension of images makes learning or optimization with such priors tedious. This is why state-of-the-art methods comprehend priors over small image patches.

Early approaches like Non-Local means (NL-means) were designed to search for identical patches of an image and average them. Instead of using similar patches of the same image, EPLL uses a prior (namely a Gaussian mixture model) cultivated from a huge set of patches taken from several images. Thus, can be considered as an external denoising method (the target image is denoised using other images), otherwise known as a "shotgun method". EPLL can be seen as an external version of the Piecewise Linear Estimator (PLE) method: PLE learns a GMM specific to each noisy image, whereas EPLL uses a fixed GMM learned once from a collection of clean patches.

Thus, the first step in EPLL will be to extract patches from a dataset of clean images and learn a GMM prior to them by maximizing the likelihood. Once a prior is set, given a noisy image “y”, the first approach to denoising could be to decompose it into overlapping patches, denoise every patch separately, and aggregate the results by simple averaging. This aggregation of overlapping patch estimates is common in patch-based algorithms; it improves the estimation as it averages a set of different estimates for any pixel.

However, applying the prior only on patches without any control on the whole image is not optimal. Indeed, averaging the values obtained for each pixel from the patches that contain it creates new patches in the constructed image, which might be unlikely under our prior. The EPLL (Expected Patch Log Likelihood) method by Zoran and Weiss [26] addresses this very problem. The aim of the method is simple: suppose we take a random patch from our reconstructed image, we wish this patch to be likely under our prior. In other words, we wish to find a reconstructed image in which every patch is likely under our prior while keeping the reconstructed image still close to the corrupted one. This variational approach is still popular as shown by recent denoising papers.

A flaw in patch-based modelling is the enforced locality of the model. Even if EPLL endeavors to work globally on the image, we neglect the long-range interactions present in the image [25].

3.4 NCSR Denoising Method

The Nonlocally Centralized Sparse Representation (NCSR) algorithm aims to reduce sparse coding noise and utilizes the Non-local Self-Similarity of the image to improve the effectiveness and accuracy of the image denoising algorithm, which has achieved a good denoising effect [27].

3.4.1 NCSR model

For a photo x $\in R^N$, dictionary $\Phi \in R^{n \times M}$ is stated ,n $\leq M$,we can show each patch as:

$$x_i \approx \Phi\alpha_{x,I} \tag{3.1}$$

And by solving a l_1-norm minimization problem, we can obtain the scattered decomposition:

$$\alpha_{x,i}=\arg_{\alpha_i}\min\left(\|x_i - \Phi\alpha_i\|_2^2 + \lambda\|\alpha_i\|_1\right) \tag{3.2}$$

The whole photo can be depicted as an encrypted collection $\{\alpha_{x,i}\}$, which can be reconstructed by least-square solution:

$$x \approx \Phi \circ \alpha_x=(\sum_{i=1}^{N} R_i^T R_i)^{-1} \sum_{i=1}^{N}\left(R_i^T \Phi\, \alpha_{x,i}\right) \tag{3.3}$$

The approach is based on sparse representation, and we should solve the minimization problem:

$$\alpha_y=\arg_\alpha \min\left(\|y - H\Phi \circ \alpha\|_2^2 + \lambda\|\alpha\|_1\right) \tag{3.4}$$

photo *x* can be restructured to:

$$\hat{x} = \Phi \circ \alpha_y \tag{3.5}$$

Since the image *x* is disturbed by noise and other factors, α_x and α_y are varied. By considering this, Dong explains the differentiation among the sparse coding coefficient:

$$v_\alpha = \alpha_{y-} \alpha_x \tag{3.6}$$

He called it sparse coding noise. It is evident that for reducing sparse coding noise, the key is to reconstruct high-quality images. Though α_x is not known, we put β instead of α_x. As a result, the model of sparse coding will become as:

$$\alpha_y = \arg_\alpha \min\left(\|y - H\Phi \circ \alpha\|_2^2 + \lambda \sum_i \|\alpha_i - \beta_i\|_p\right) \tag{3.7}$$

Getting the dictionary and the estimate is the purpose for solving this model. For obtaining the dictionary, we can use K- means and PCA method. At first, the photo is separated into patches, then all of the patches are divided into K classes using the K-means approach. Subsequently, a sub-dictionary of PCA is studied for each of the classes. β_i is used to employ a photo's Non-Local Self-Similarity. By considering a set of identical patches for photo, x_i is $\Omega_{I,}$ and also $\alpha_{I,\,q}$ defines sparse coding of the photo patch $x_{i,q}$ in Ω_i , β_i will be computed as follow by weighted approximation of $\alpha_{i,,q}$:

$$\beta_I = \sum_{q \in \Omega_i} \omega_{i,q} \, . \, \alpha_{i,q} \tag{3.8}$$

The weight is $\omega_{i\,q}$

$$\omega_{i,q} = \frac{1}{W} \exp\left(-\| \hat{x}_i - \hat{x}_{i,q} \|_2^2 \;/h\right) \tag{3.9}$$

W denotes normalization factor and the default constant is h:

$$\hat{x}_i = \Phi\alpha_i \tag{3.10}$$

and $\hat{x}_{I,q} = \Phi\hat{\alpha}_{i,q}$ (3.11)

are estimated for x_i and $x_{i,q}$ patches .

The algorithm of NCSR acquires the ultimate result by using iterative performance. After that formula could be changed to:

$$\alpha_y = \arg_\alpha \min\left(\|y - H\Phi \circ \alpha\|_2^2 + \sum_i \sum_j \lambda_{i,j} \left|\alpha_i(j) - \beta_j(j)\right|\right) \tag{3.12}$$

This problem is a distinguished optimization subject .The superseded method would be utilized to solve the problem's efficiently.

$$\alpha_i{}^{(i+1)}(j) = S_\tau\left(v_{i,j}^{(l)} - \beta_j(j)\right) + \beta_j(j) \tag{3.13}$$

Where $S\tau$ is the classic soft threshold operator.

$$v^{(l)} = K^T\left(y - K \circ \alpha^{(l)}\right)/c + \alpha^{(l)} \tag{3.14}$$

and $K = H\,\Phi$, $K^T = \Phi^T \circ \mathrm{H}^T$, $\tau = \lambda_{i,j}$ [28]. (3.15)

3.5 WNNM Denoising Method

Low-rank matrix approximation methods could be generally categorized into the Low-Rank Matrix Factorization (LRMF) methods and the nuclear norm minimization (NNM) methods. Given a matrix ***Y***, the aim of LRMF is to find a matrix ***X*** as close to ***Y*** as possible under specific data fidelity functions, at the time of being able to be factorized into the product of two low-rank matrices. The LRMF problem is fundamentally a nonconvex optimization problem. $\|\boldsymbol{X}\|_*$ is the nuclear norm [35] of a matrix ***X***, and it is described as the summation of its singular values.

$$\|\boldsymbol{X}\|_* = \Sigma_i\, |\sigma_i(\boldsymbol{X})\,|_1 \tag{3.16}$$

$\sigma i(\boldsymbol{X})$ is singular value of ***X***. Nuclear norm minimization approach wants to approximately calculate ***Y*** with calculating ***X*,** when minimizing nuclear norm of ***X***. The NNM low rank matrix approximation with *F*-norm is solvable.

$$\mathrm{X} = \mathrm{argmin}_X\, \|Y - X\|_F^2 + \lambda\|X\|_*, \tag{3.17}$$

λ is a constant which is positive, and we can calculate it by:

$$\mathrm{X} = \mathrm{US}_\lambda\,(\Sigma)V^T, \tag{3.18}$$

SVD [34] of ***Y*** is $\boldsymbol{Y} = \boldsymbol{U\Sigma V}^T$, and the soft thresholding function on diagonal matrix $\boldsymbol{\Sigma}$ with parameter λ is $S\lambda(\boldsymbol{\Sigma})$. We have:

$$S_\lambda\,(\boldsymbol{\Sigma})_{ii} = \mathrm{Max}\,(\boldsymbol{\Sigma}_{ii} - \lambda\,, 0). \tag{3.19}$$

for each diagonal element $\boldsymbol{\Sigma}ii$ in $\boldsymbol{\Sigma}$.

Although NNM has been widely used for low-rank matrix approximation, it still has some problems. To pursue the convex property, the standard nuclear norm treats each singular value equally and shrinks each singular value with the same amount λ. They designed a weighted nuclear norm and studied its minimization to improve the nuclear norm's flexibility. The weighted nuclear norm of a matrix $\boldsymbol{X}$ is described:

$$\|X\|_{w,*} = \sum_i |w_i \sigma_i (X)|_1, \tag{3.20}$$

$w_i \geq 0$ is a weight that is positive and it is allocated to $\sigma_i(\boldsymbol{X})$, and $\boldsymbol{w} = [w_1, \ldots, w_n]$

Weighted Nuclear Norm Minimization (WNNM) problem is defined as follows [29]:

$$\min_X \|Y - X\|_F^2 + \|X\|_{w,*} \tag{3.21}$$

3.5.1 WNNM for Image Denoising

For a patch, $\boldsymbol{y}_j$, in photo $\boldsymbol{y}$, we look for non-local identical patches among photo with the help of approaches like block matching. We have $\boldsymbol{Y}_j = \boldsymbol{X}_j + \boldsymbol{N}_j$, by collecting non-local similar patches in a matrix $\boldsymbol{Y}_j$, and both of $\boldsymbol{N}_j$ and $\boldsymbol{X}_j$ are matrices of the patch of original photo and noise, and $\boldsymbol{X}_j$ must be a low-rank matrix. So we can use the low-rank matrix approximation methods to estimate $\boldsymbol{X}_j$ from $\boldsymbol{Y}_j$. The entire image can be estimated by gathering all the denoised patches. WNNM model is applied to $\boldsymbol{Y}_j$ for estimating $\boldsymbol{X}_j$ for denoising the photos. With usage of σ_n^2 , which is the variance of the noise, for normalization of $\|\boldsymbol{Yj} - \boldsymbol{Xj}\|^2_{\mathrm{F}}$, following energy function, is proposed:

$$\hat{X}_j = \arg\min_{X_j} \frac{1}{\sigma_n^2} \|Y_j - X_j\|_F^2 + \|X_j\|_{w,*}. \tag{3.22}$$

The main subject is calculating the weight vector $\boldsymbol{w}$. we know that greater singular values of $\boldsymbol{X}_j$ are much more significant than the tinier singular value. In denoising, as singular values become larger, they must be minimized. As a result, the i-th singular value of $\boldsymbol{X}_j$ must be oppositely symmetrical with σ_i ($\boldsymbol{X}_j$).

$$w_i = c\sqrt{n} \,/\, \left(\sigma_i \left(X_j\right) + \varepsilon\right), \tag{3.23}$$

Where $c > 0$ is a constant, n is the number of identical patches in $\boldsymbol{Y_j}$ and $\varepsilon = 10^{-16}$is to avoid splitting by zero. One problem here is that σ_i ($\boldsymbol{X}_j$) is not available. Pressuming that the energy of noise is equally spread over each of the subspaces, and is measured by set of $\boldsymbol{U}$ and $\boldsymbol{V}$ and σ_i ($\boldsymbol{X}_j$) could be calculated as:

$$\hat{\sigma}_i(\boldsymbol{X}j) = \sqrt{\max(\sigma_i^2(\boldsymbol{Y}j) - n\sigma_n^2, 0)}, \tag{3.24}$$

σ_i ($\boldsymbol{Y}_j$) is the $\boldsymbol{i}$-th singular amount of $\boldsymbol{Y}_j$. We ensured that the calculated weights wi=1,...,n are in a non-descending arrangement because $\hat{\sigma}_i$ ($\boldsymbol{X}_j$) are arranged in a non-ascending arrangement.By using the described approaches for each patch and gathering them all, we can reproduce photo $\boldsymbol{x}$. [29]

3.6 Peak Signal to Noise Ratio (PSNR)

Peak Signal-to-Noise Ratio (PSNR) is a statement for the ratio between the maximum possible amount (power) of a signal and the power of false noise that influences the quality of its display. The PSNR is often represented as the logarithmic decibel scale since many signals have an extensive dynamic ratio between the largest and smallest feasible amounts of a changeable quantity.

PSNR is usually used for comparing the quality of images, and it is easily defined through the Mean Square Error (MSE). By giving a noise-free m×n image I and its noisy approximation K, MSE is described as:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i\,,j) - K(i\,,j)]^2 \tag{3.25}$$

So the PSNR is defined as:

$$PSNR = 10.\log_{10}\left(\frac{MAX_I^2}{MSE}\right) \tag{3.26}$$

$$= 10.\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right)$$

$$= 10.\log_{10}(MAX_I) - 10.\log_{10}(MSE)$$

If we consider O as the original image and X as a noisy image (σ_n^2) then we have:

$$PSNR=\frac{255^2}{\sum_{ij}(O-X)^2} \tag{3.27}$$

Chapter 4

EXPERIMENTS AND RESULTS

4.1 Introduction

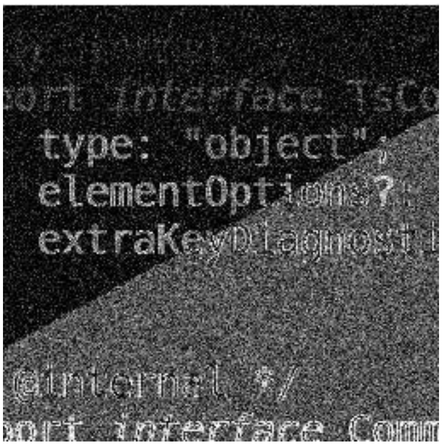
This chapter discusses the results we got by the implementation of BM3D, EPLL, NCSR, and the WNNM denoising methods on 20 text images chosen from the internet and some famous test images, which were contaminated by adding Additive White Gaussian Noise with variance σ_n^2 . Additionally, Peak Signal-to-Noise Ratio (PSNR) results of the chosen denoised images are compared and the price paid for 1dB improvement in SNR in the WNNM method is calculated. Lastly, we did the experiment on a picture of an ancient document and compared its PSNR value.

4.2 Results on Text Images

The following chapter discusses the results of the implementation of BM3D, EPLL, NCSR, and the WNNM denoising methods on text images and Barbara, Lena, Straw, and Monarch test images. Zero mean the Additive White Gaussian Noises with variance δ_n^2 *are* added to those test images to generate the noisy observations with δ=10, 20, 30, 40, respectively. All test images are 8 bit with the size 256×256, and Matlab did the programming.

At first choosing a text image, Added White Gaussian Noise on it with σ=40, and denoised it by BM3D, EPLL, NCSR, and WNNM method (Figure4.1), and compared their PSNR.

a. Original image



b. Noisy image(σ=40)

c. BM3D denoised image (PSNR=24.47)

d. EPLL denoised image (PSNR=24.54)

e. NCSR denoised image (PSNR=24.05)

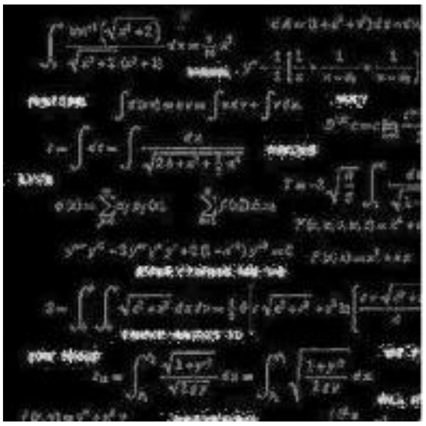
f. WNNM denoised image(PSNR=25.28)

Figure 4.1: Original, noisy (σ=40) and denoised text image (text16)

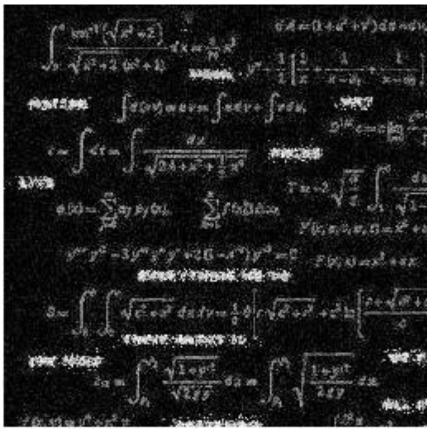
As evident in Figure 4.1, the WNNM denoised image has a higher PSNR value, so it has a higher quality.

In the text image shown in figure 4.1, we note that σ=40, and this corresponds to an initial PSNR of 16.06 dB. BM3D increases PSNR by 8.41dB (24.47-16.06), EPLL increases it by 8.48dB, NCSR increases it by 7.99 dB, and the WNNM increases it by 9.22 dB, which is the most amongst all used denoising methods.

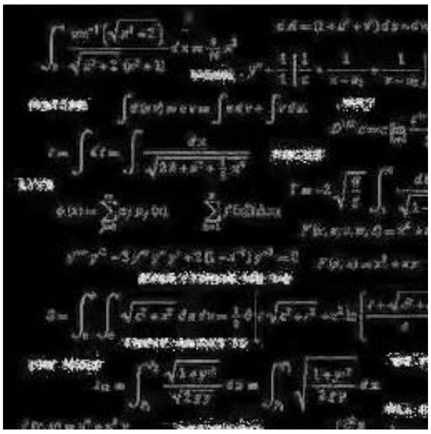
Also, the computational complexity of the BM3D method is 1.2 sec, for EPLL it is 36.58 sec, for NCSR it is 319.4 sec, and finally, for the WNNM it is 132.34 sec.



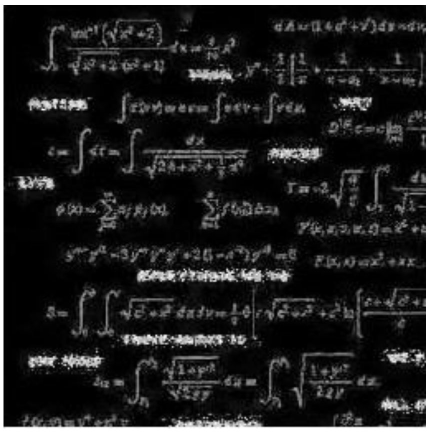
a. Original image



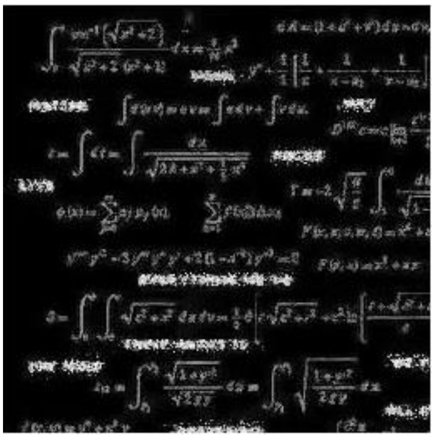
b. Noisy image(σ=20)

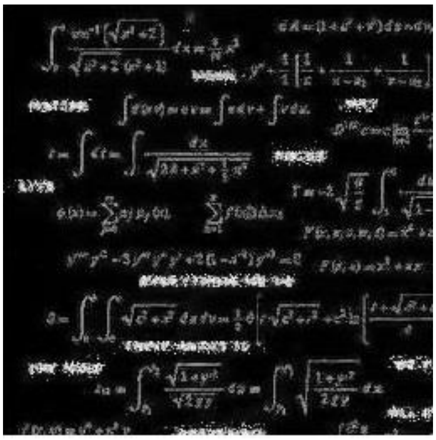


c. BM3D denoised image (PSNR=26.25)



d. EPLL denoised image (PSNR=26.48)





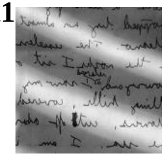
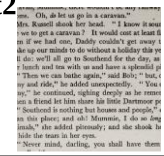
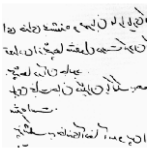
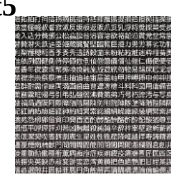
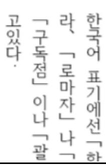
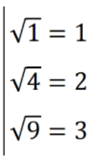
e. NCSR denoised image (PSNR=26.67) f. WNNM denoised image (PSNR=26.94)

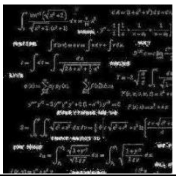
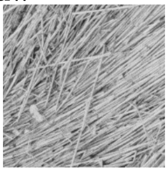
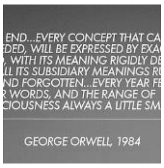
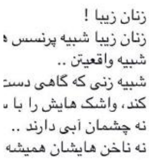
Figure 4.2: Original, noisy (σ=20) and denoised Mathematic text image (text8)

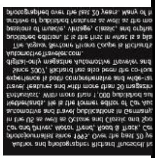
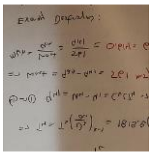
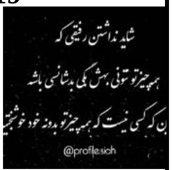
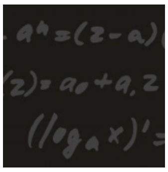
4.3 Test Images and Their PSNR Values

For testing BM3D, NCSR, EPLL, and the WNNM method, as well as Lena, Barbara, Straw, and monarch test images, 20 text images were chosen from the internet. All of the images are 8 bit, 256×256.

Table 4.1: PSNR results of the denoised images

	σ = 10				σ = 20			
	BM3D	**EPLL**	**NCSR**	**WNNM**	**BM3D**	**EPLL**	**NCSR**	**WNNM**
text1 	28.49	28.24	28.48	**28.51**	22.98	22.65	22.95	**23.19**
text2 	28.34	28.04	28.42	**28.48**	22.78	22.30	22.99	**23.10**
text3 	31.81	30.64	31.77	**32.15**	26.15	24.95	26.02	**26.53**
text4 	34.84	36.51	36.06	**36.71**	29.85	31.24	30.80	**31.45**
Lena 	30.52	30.63	30.68	**30.71**	26.87	27.11	27.10	**27.14**
Barbara 	34.77	33.76	34.87	**35.39**	31.24	29.97	31.17	**31.61**
text5 	28.35	28.12	28.32	**28.38**	22.68	22.50	22.88	**22.93**
text6 	36.68	36.30	36.52	**36.74**	30.85	30.76	31.13	**31.39**
text7 	41.46	40.47	40.86	**41.22**	35.34	34.82	35.34	**35.82**

text8 	31.35	31.42	31.36	**31.54**	26.25	26.48	26.67	**26.94**
Straw 	30.91	30.83	31.45	**31.72**	27.07	27.03	27.46	**27.62**
text9 	28.33	27.60	28.23	**28.26**	22.58	21.44	22.48	**22.55**
text10 	31.17	30.51	30.66	**30.98**	25.33	24.47	25.06	**25.49**
text11 	28.99	28.65	29.02	**29.11**	23.43	22.95	23.58	**23.74**
text12 	31.69	31.49	31.66	**31.83**	26.36	25.89	26.41	**26.63**
Monarch 	34.12	34.27	34.51	**35.03**	30.35	30.54	30.62	**31.11**
text13 	28.40	28.18	28.42	**28.43**	22.84	22.70	23.03	**23.08**
text14 	34.69	33.65	35.44	**35.82**	29.80	28.99	29.93	**30.74**

text15 	29.10	29.08	29.24	**29.39**	24.38	24.10	24.32	**24.74**
text16 	34.66	34.22	34.91	**35.76**	29.15	28.95	29.49	**30.54**
text17 	29.58	29.17	29.66	**30.19**	24.17	23.56	24.38	**24.86**
text18 	29.23	29.18	29.26	**29.49**	23.87	23.94	23.86	**24.27**
text19 	33.81	33.62	33.82	**34.45**	28.65	28.25	28.78	**29.88**
text20 	32.90	32.77	32.80	**32.96**	27.68	27.47	27.72	**27.95**
AVG.	31.84	31.47	31.93	**32.21**	26,69	26.37	26.84	**27.22**
	σ = 30				**σ = 40**			
text1	19.99	19.95	20.13	**20.16**	18.10	18.17	17.99	**18.28**
text2	19.66	19.42	20.13	**20.17**	17.59	17.63	18.12	**18.28**
text3	22.84	21.96	23.01	**23.25**	20.60	20.24	20.52	**21.15**
text4	28.49	29.08	28.67	**29.17**	27.81	28.07	27.62	**28.09**
Lena	25.44	25.51	25.62	**25.70**	24.69	24.64	24.73	**24.84**
Barbara	29.07	27.58	28.76	**29.48**	27.25	26.04	27.28	**27.86**
text5	19.49	19.54	20.02	**20.05**	17.49	17.81	17.93	**18.14**
text6	27.71	28.09	28.31	**28.16**	25.91	26.36	26.15	**26.27**
text7	32.12	31.81	32.41	**32.50**	29.99	29.58	29.84	**30.49**
text8	23.55	24.01	24.20	**24.28**	21.87	22.51	22.27	**22.64**
straw	24.94	24.64	25.12	**25.47**	23.16	23.26	23.64	**23.94**
text9	19.26	18.28	19.36	**19.38**	16.92	16.47	17.22	**17.30**
text10	21.87	21.28	22.03	**22.17**	19.54	19.36	19.61	**20.06**
text11	20.29	19.93	20.66	**20.69**	18.17	18.14	18.46	**18.72**
text12	23.38	23.12	23.71	**23.72**	21.45	21.40	21.72	**21.87**
Monarch	28.36	28.35	28.46	**28.91**	26.72	26.83	26.84	**27.47**
text13	19.90	19.94	20.18	**20.33**	18.16	18.32	18.39	**18.67**
text14	26.74	26.40	26.93	**27.61**	24.78	24.81	24.82	**25.47**

text15	21.92	21.79	22.16	**22.30**	20.30	20.39	20.44	**20.76**
text16	26.25	26.27	26.63	**27.23**	24.47	24.54	24.05	**25.28**
text17	20.96	20.65	21.48	**21.62**	18.74	18.82	19.41	**19.57**
text18	21.07	21.29	21.21	**21.35**	19.36	19.64	19.28	**19.67**
text19	25.37	25.36	25.92	**26.77**	23.32	23.60	23.26	**24.67**
text20	24.70	24.88	25.04	**24.95**	22.96	23.03	23.04	**23.16**
AVG.	23.89	23.71	24.17	**24.39**	22.05	22.06	22.19	**22.61**

Table 4.1 compares PSNR of 20 denoised test images with BM3D, EPLL, NCSR, and The WNNM methods, and shows that the PSNR of the WNNM method is higher than other methods, so the pictures denoised with the WNNM method are higher quality.

4.4 Results on an Old Text Image

The denoising methods have been tested on an old text image. A 256×256 sample old image was chosen.

a. Original old image

b. Noisy old image ($\sigma = 40$)

c. BM3D denoised image (PSNR=21.94)

d. EPLL denoised image (PSNR=22.42)

e. NCSR denoised image (PSNR=21.92)

f. WNNM denoised image (PSNR=22.48)

Figure 4.3: Original, noisy (σ=40) and denoised old text images

Figures 4.3 compares PSNR of a noisy ($\boldsymbol{\sigma}$ **= 40**) old document image with different image denoising methods. Because the PSNR value of the WNNM method is the highest, so this method showed better results.

Table 4.2: PSNR results of the denoised old image by different methods.

	σ = 10				**σ = 20**			
	BM3D	**EPLL**	**NCSR**	**WNNM**	**BM3D**	**EPLL**	**NCSR**	**WNNM**
history	29.88	30.03	29.93	**30.07**	22.52	25.93	25.58	**25.87**
	σ = 30				**σ = 40**			
history	23.39	23.87	23.48	**23.79**	21.94	22.42	21.92	**22.48**

Table 4.2 compares the PSNR of different denoising methods and indicates the WNNM method has the highest PSNR in all cases.

Chapter 5

CONCLUSION

Image denoising is an applicable issue found in diverse image processing and computer vision problems. There are various existing methods for the denoise image. The critical property of a good image denoising model is that it should completely remove noise as far as possible and preserve edges.

Ancient and historical manuscripts are usually kept in a museum for a long time. This causes the manuscripts to contain some noises, not only substantive but also additive.

In this thesis, BM3D, EPLL, NCSR, and the WNNM denoising methods are applied to 20, 8bit text images from the internet, including historical text images, which were contaminated by Additive White Gaussian Noise with variance σ_n^2, and PSNR of the denoised images were measured. The results in all different noise levels ($\sigma = 10,20,30,40$) showed that images denoised with the WNNM method have higher PSNR values, so they have a higher quality than the images denoised with other methods. Moreover focusing on the WNNM denoising method on text images contaminated with Additive White Gaussian Noise, in comparison to other types of noise in in-text images. By comparing this method with other traditional denoising methods, we have concluded that the WNNM works better in denoising text images than other methods.

REFERENCES

[1] Rajni, Anutam. (2014).Image denoising techniques-An overview. *International Journal of Computer Applications*, 86(16). doi:10.5120/15069-3436

[2] Fan, L., Zhang, F., Fan, H. et al. (2019). Brief review of image denoising techniques. Vis. Comput. Ind. Biomed. doi.:10.1186/s42492-019-0016-7.

[3] Xiaohua Liu, Xiao-Yuan Jing, Guijin Tang, Fei Wu, Qi Ge. (2017). Image denoising using weighted nuclear norm minimization with multiple strategies. *Signal Processing*, 135, 239-252. doi: 10.1016/j.sigpro.2017.01.003.

[4] Farahmand Atena, Sarrafzadeh Abdolhossein & Shanbezadeh, Jamshid. (2013). Document Image Noises and Removal Methods. *Lecture Notes in Engineering and Computer Science*. 2202. 436-440.

[5] S. S. Thakare, S. N. Kale. (2019). Noise reduction technique in scanned documents using Cuckoo optimization algorithm. *International Journal Of Engineering and Advanced Technology (IJEAT)*, 8(6S). doi:10.35940/ijeat.F1069.0886S19.

[6] Peerawit, Worapoj. (2004). Marginal noise removal from document images using edge density. *Proceedings of Fourth Information and Computer Eng. Postgraduate Workshop.*

[7] F. Arnia, Fardian, S. Muchallil and K. Munadi. (2015). Noise characterization in ancient document images based on DCT coefficient distribution. *13th*

International Conference on Document Analysis and Recognition (ICDAR) (pp. 971-975), doi: 10.1109/ICDAR.2015.7333906.

[8] M. Ali. Background Noise Detection and Cleaning in Document Images.(1996). *Pattern Recognition International Conference* (pp. 758). Vienna, Austria. doi: 10.1109/ICPR.1996.547270.

[9] K. Chinnasarn, Y. Rangsanseri and P. Thitimajshima. (1998). Removing salt-and-pepper noise in text/graphics images. IEEE. APCCAS 1998. 1998 IEEE Asia-Pacific Conference on Circuits and Systems. Microelectronics and Integrating Systems Proceedings (pp. 459-462). doi: 10.1109/APCCAS.1998.743809.

[10] Lei Xu, Erkki Oja, Pekka Kultanen. (1990). A new curve detection method: Randomized Hough transform (RHT). *Pattern Recognition Letters*, 11(5), 331-338 .doi: 10.1016/0167-8655(90)90042-Z.

[11] Z. Shi, S. Setlur and V. Govindaraju. (2010). Removing Rule-Lines from Binary Handwritten Arabic Document Images Using Directional Local Profile. *2010 20th International Conference on Pattern Recognition* (pp. 1916-1919). doi: 10.1109/ICPR.2010.472.

[12] Zheng Zhang and Chew Lim Tan. (2001). Recovery of distorted document images from bound volumes. *Sixth International Conference on Document Analysis and Recognition,* (pp. 429-433). doi: 10.1109/ICDAR.2001.953826.

[13] Kuo-Chin Fan, Yuan-Kai Wang, and Tsann Ran Lay. (2002). *Marginal noise*

*removal of document images.Pattern Recognition Society, Elsevier Science (*pp. 2593-2611).

[14] Faisal Shafait, Joost van Beusekom, Daniel Keysers, and Thomas M. Breuel. (2008). International Journal on Document Analysis and Recognition (IJDAR), 11(2), 81-96. doi: 10.1007/s10032-008-0071-7.

[15] M. Agrawal and D. Doermann. (2011). Stroke-Like Pattern Noise Removal in Binary Document Images. *2011 International Conference on Document Analysis and Recognition (*pp. 17-21). doi: 10.1109/ICDAR.2011.13.

[16] J. Sauvola and M. Pietikainen. (2000). Adaptive document image binarization. *Pattern Recognition International journal*,33(2),225-236. doi: 10.1016/S0031-3203(99)00055-2.

[17] Joung-Youn Kim, Lee-Sup Kim and Seung-Ho Hwang. (2001). an advanced contrast enhancement using partially overlapped sub-block histogram equalization. *IEEE Transactions on Circuits and Systems for Video Technology*, 11 (4), 475-484. doi: 10.1109/76.915354.

[18] Leung, Chung-Chu & Chan, Ka-Shing & Chan, Hoi-Mei & Tsui, Wai-Kin. (2005). A new approach for image enhancement applied to low-contrast-low-illumination IC and document images. *Pattern Recognition Letters*, 26, 769-778. doi:10.1016/j.patrec.2004.09.032.

[19] T. Nagao and S. Masunaga. (1996). Automatic construction of image

transformation processes using genetic algorithm. *3rd IEEE International Conference on Image Processing*, 3, 731-734. doi: 10.1109/ICIP.1996.560795.

[20] S. Aoki and T. Nagao. (1999). Automatic construction of tree-structural image transformations using genetic programming. *10th International Conference on Image Analysis and Processing*, 136-141. doi: 10.1109/ICIAP.1999.797584.

[21] Kohmura, H. & Wakahara, Toru. (2006). Determining Optimal Filters for Binarization of Degraded Characters in Color Using Genetic Algorithms. 18th International Conference on Pattern Recognition (ICPR'06), 3, 661 - 664. doi:10.1109/ICPR.2006.446.

[22] H. Deborah and A. M. Arymurthy. (2010). Image Enhancement and Image Restoration for Old Document Image Using Genetic Algorithm. *Second International Conference on Advances in Computing, Control, and Telecommunication Technologies* (pp. 108-112). doi: 10.1109/ACT.2010.24.

[23] Lebrun, Marc. (2012). An Analysis and Implementation of the BM3D Image Denoising Method. *Image Processing On Line*, 2, 175-213. doi:10.5201/ipol.2012.l-bm3d.

[24] Dabov, Kostadin & Foi, Alessandro & Katkovnik, Vladimir & Egiazarian, Karen. (2007). Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society*,16, 2080-95. doi:10.1109/TIP.2007.901238.

[25] Hurault, Samuel & Ehret, Thibaud & Arias, Pablo. (2018). EPLL: An Image Denoising Method Using a Gaussian Mixture Model Learned on a Large Set of Patches. *Image Processing On Line*, 8, 465-489. doi: 10.5201/ipol.2018.242.

[26] Zoran, Daniel & Weiss, Yair. (2011). From Learning Models of Natural Image Patches to Whole Image Restoration. *IEEE International Conference on Computer Vision*, 479-486. doi:10.1109/ICCV.2011.6126278.

[27] Lee, XiDa & Wu, JunFang. (2019). Image Denoising Algorithm based on Improved NCSR Model. *Journal of Physics: Conference Series*. doi: 10.1088/1742-6596/1314/1/012209.

[28] Dong, Weisheng & Zhang, Lei & Shi, Guangming & li, Xin. (2012). Nonlocally Centralized Sparse Representation for Image Restoration. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 22. doi: 10.1109/TIP.2012.2235847.

[29] Gu, Shuhang & Zhang, Lei & Zuo, Wangmeng & Feng, Xiangchu. (2014). Weighted Nuclear Norm Minimization with Application to Image Denoising. 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2862-2869. doi: 10.1109/CVPR.2014.36

[30] Akash Kethwas, Dr. Bhavana Jharia. (2014). Comparison Study on Image Denoising Through Wiener Filter. INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT),8(3). doi: IJERTV3IS080812

[31] Bouden, Toufik & Nibouche, Mokhtar. (2012). The Wavelet Transform for Image Processing Applications. doi: 10.5772/35982.

[32] Sharma, Sudhanshi & Kumar, Umesh. (2013). Review of Transform Domain Techniques for Image Steganography. International Journal of Science and Research (IJSR). doi: 10.13140/RG.2.1.4797.1928.

[33] Aharon, Michal & Elad, Michael & Bruckstein, Alfred. (2006). K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. Signal Processing, IEEE Transactions on. 54. 4311 - 4322. doi:10.1109/TSP.2006.881199.

[34] *Brady Mathews.*(2014). Image Compression using Singular Value Decomposition (SVD). The University of Utah.

[35] Nuclear norm Encyclopedia *of Mathematics. Retrieved from:*http://encyclopediaofmath.org/index.php?title=Nuclear_norm&oldid=4