A MATRIX-BASED IMPLICIT PRIOR FOR JOINT DEMOSAICING AND SUPER-RESOLUTION

A PROJECT REPORT

Submitted by

SNEHA C (312213106095)

SREENITHY C (312213106097)

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

IN

ELECTRONICS AND COMMUNICATION ENGINEERING

SRI SIVASUBRAMANIYA NADAR COLLEGE OF ENGINEERING, KALAVAKKAM

ANNA UNIVERSITY: CHENNAI 600 025

APRIL 2017

ANNA UNIVERSITY: CHENNAI 600 025

BONAFIDE CERTIFICATE

Certified that this project report **"A MATRIX-BASED IMPLICIT PRIOR FOR JOINT DEMOSAICING AND SUPER-RESOLUTION"** is the bonafide work of "**SNEHA C (312213106095) and SREENITHY C (312213106097)"** who carried out the project work under my supervision.

SIGNATURE

Dr. S. Radha
HEAD OF THE DEPARTMENT
Department of ECE
SSN College of Engineering
Old Mahabalipuram Road,
Kalavakkam – 603 110

SIGNATURE

Mr. Jino Hans W.
SUPERVISOR
Assistant Professor
Department of ECE
SSN College of Engineering
Old Mahabalipuram Road,
Kalavakkam – 603 110

Submitted for the project viva-voce examination held on _____

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ACKNOWLEDGEMENT

We express our sincere gratitude and heartfelt thanks to:

Dr. Shiv Nadar, esteemed Chairman of SSN Institutions, for extending a generous hand in providing the best of resources to the college;

Dr. S. Salivahanan, our respected Principal, who has been a source of motivation to all the staff and student members of our college;

Dr. S. Radha, the Head of the Department, for her continuous support and encouragement throughout our project;

Our project coordinator and the panel members, for their timely reviews and erudite inputs that helped us improve our project;

Mr. W. Jino Hans, our beloved project guide, for his invaluable guidance, innovative suggestions, patience and unflinching support throughout the course of this project;

All the staff members of the Department of Electronics and Communication Engineering for their assistance at all times;

And finally, our families, classmates and friends, for helping us throughout this journey.

ABSTRACT

This project aims at implementing two image processing techniques, namely demosaicing and super-resolution. Most of the digital cameras capture only one of the three primary colour components – red, green or blue at each pixel point using colour filter arrays. Demosaicing algorithms focus on interpolating the missing colour components using the available pixel information. The proposed algorithm exploits the inherent relationships in the colour difference planes, i.e. red-green and blue-green planes. This method also takes edges into account using novel edge classification methods, thus giving a better output. The result of the demosaicing process is a fully reconstructed RGB colour image obtained from the incomplete colour filter mosaic from an image sensor.

Super-resolution (SR) techniques are used to improve the resolution of images after they are captured by extracting relationships between high and low resolution images. This relation is obtained either from patches of different scales or by using a training set to build an operator through supervised learning. We propose to use a matrix based operator to reduce structural incongruities, which are otherwise prevalent in vector-based methods. The concept of a matrix based regression operator is extended to a low level feature SR algorithm. This approach achieves competitive performance efficiency and effectiveness which is checked by experimenting for various use cases.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Image processing refers to the field of signal processing which works on visual or image data. It is a continually evolving field today with rapid developments. In the last five years, there has been a remarkable growth in the level of interest in image morphology, image enhancement, neural networks, full-colour image processing, image compression standards, facial recognition, and knowledge-based image analysis systems. Image processing methods are motivated from two major applications: improvement of image-level information for human interpretation, and processing of pictorial data for machine perception. Humans have the ability to perceive information as an image better than any other form. Eye-sight and vision allows humans to observe and understand the surroundings. Image analysis and computer vision aim to duplicate the effect of human vision by electronically perceiving and understanding images.

A major part in dynamic advancement in this field is credited to digital systems. Modern-day digital technology has made it possible to manipulate multi-dimensional signals for use in image systems. The first step in the process is image acquisition and requires an imaging sensor as well as the capability to interpret the signal captured by the sensor. If the output of the sensor is not already in digital form, an analog to digital converter is additionally used.

After a digital image is obtained, the next step is to perform pre-processing. The key function of pre-processing is to improve the image for use in other applications. It deals with methods to enhance contrast, remove noise, and isolate regions with specific textures. The next stage deals with segmentation which partitions an input image into its constituent parts or objects. Weak or

erratic segmentation algorithms almost always guarantee eventual failure, in terms of character recognition, wherein the key role of segmentation is to extract individual characters and words from the background.

1.1 IMAGE ACQUISITION IN DIGITAL CAMERAS

In most digital cameras, a Colour Filter Array (CFA) is placed on top of the monochrome image sensor to acquire the low-resolution colour information of the image scene. Each sensor cell has its own spectrally selective filter (R,G, or B) and thus, the acquired CFA data constitutes a mosaic-like monochrome image. The type of CFA determines the arrangement of recorded colour components and the most widely used CFA pattern is the Bayer pattern. This widely adopted solution keeps cost and size of digital cameras under control, because the image sensor is the most expensive component of the camera. If we want to capture a colour image, we must capture or reconstruct the data corresponding to the other two colour-channels of the image.

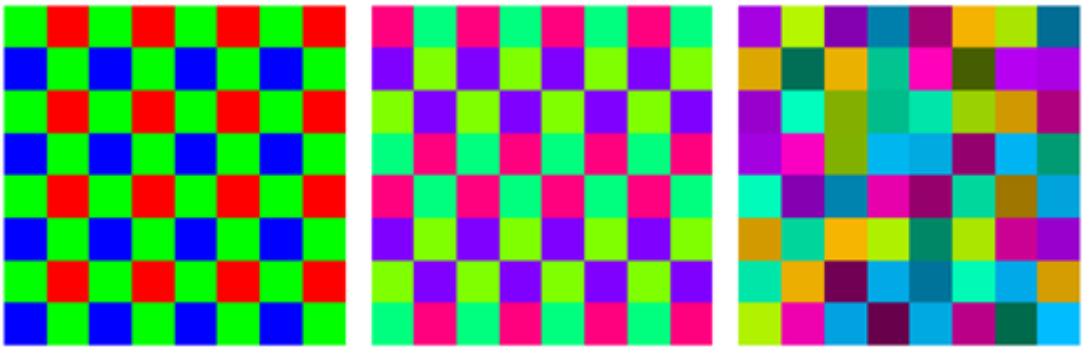


Fig 1.1: Type of colour filter array (CFA) patterns

1.2 BAYER CFA PATTERN

The Bayer CFA uses the three additive primary colours as the filter sensors to achieve and reconstruct red, blue and green data of a scene. In a Bayer Filter, for each 2×2 set of pixels, two diagonally opposed pixels have green filters, and the other two have red and blue filters. Because the spectral response of a green (G) channel corresponds to that of the human visual system's luminance

channel, the green components are sampled at twice the rate of the red (R) or blue (B) components. This technique is used in most digital cameras. Figure 1.2 shows the Bayer filter and arrangement of R, G and B sensors, the image is taken from Ye and Ma (2015).

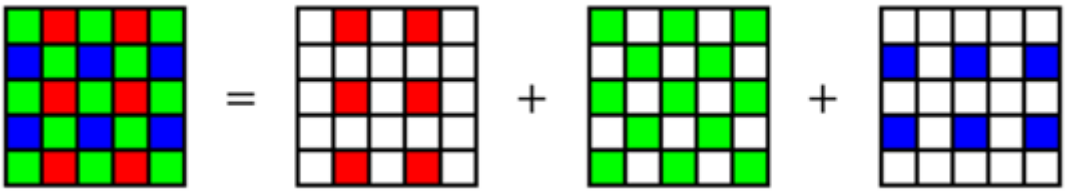


Fig 1.2: Bayer filter and arrangement of R, G and B sensors

1.3 COLOUR IMAGE RECONSTRUCTION

Demosaicing is the process that estimates the two missing colour components at each pixel location to restore a full-colour image from the mosaiced image. The most popular principle in the demosaicing literature is the green plane-first rule. The key motivation behind this principle is that the green component is less aliased than the other two. Thus, having a full-resolution green plane could facilitate the recovery of blue and red planes. If the demosaicing process is not properly carried out, severe colour artifacts will be incurred on the restored full-colour image. In demosaicing, it is commonly assumed that colour ratios and colour differences are constant over small regions. These are referred to as the colour ratio model and colour difference models, respectively. Interpolation along an object boundary is always preferable to interpolation across it because the discontinuity of the signal at the boundary contains high-frequency components that are difficult to estimate. Visually, two types of artifacts are generated in the demosaiced image: one is the pattern of alternating colours along the edge, called zipper effect, and the other is the noticeable colour errors called false colour.

Many methods have been proposed for interpolating the missing colour components in mosaiced image. Because of its simplicity and effectiveness,

many methods utilize the colour difference model, i.e., red-green and blue-green. In both cases, correct interpolation of the missing green components is crucial because interpolated green components are used to reconstruct the other colours. Improper interpolation of neighboring pixel values leads to demosaicing artifacts, such as false colours and the so called zipper effect.

One strategy for interpolation of neighboring pixels is the edge-directed method. In this method, two or more predictors are estimated along the candidate directions, and one of them is selected as the value of the missing pixel. The objective of this strategy is to perform interpolation along edges rather than across them. Another strategy is computing a weighted sum of predictors. After estimating predictors along the candidate directions, some weights are calculated on the basis of the edge directions. Each missing pixel is then interpolated by the weighted sum of predictors with the calculated weights.

Other strategies use various schemes such as pattern matching, median filtering, bilateral filtering, and optimization-based filtering. Demosaicing has been studied in the frequency domain as well. One promising class of algorithms is based on a frequency-domain explanation of the spatial multiplexing of red, green and blue components; it has been termed lumachroma de-multiplexing.

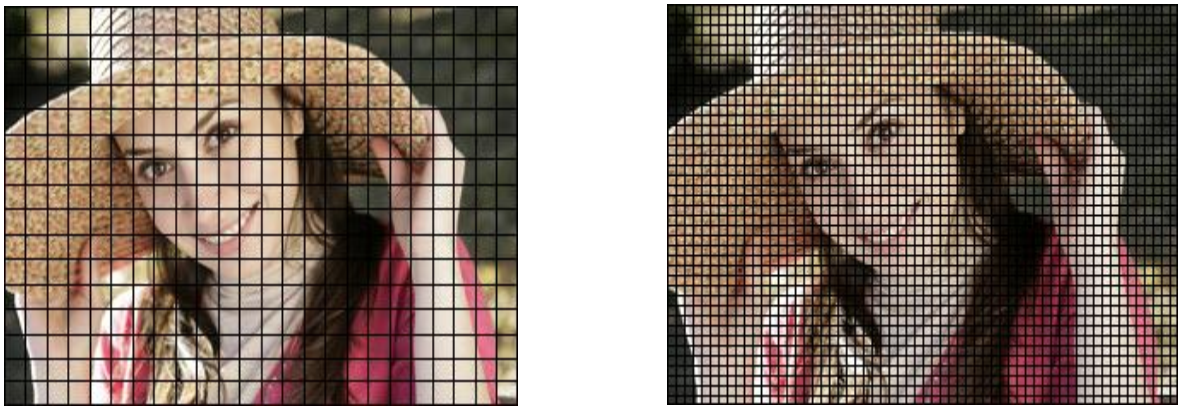
1.4 IMAGE RESOLUTION

Image resolution is the detail an image holds. Resolution quantifies how close lines can be to each other and still be visibly resolved. It is defined for all types of images: raster digital image, analog image captured by film camera and synthetic images. It is also the capability of the sensor to observe or measure the smallest object clearly with distinct boundaries. There is a difference between the resolution and a pixel. A pixel is actually a unit of the digital Resolution depends upon the size of the pixel.

Usually, with any given lens setting, the smaller the size of the pixel, the higher the resolution will be and the clearer the object in the image will be. Images having smaller pixel sizes might consist of more pixels. The number of pixels correlates to the amount of information within the image. Thus, an image that holds more detail is said to be of higher resolution than a corresponding one with less detail. Higher resolution images convey more complex structured data.

1.5 NEED FOR RESOLUTION ENHANCEMENT

When magnified, the surface of a CCD looks like a large, dot-filled grid. Each of these dots is a light receptor, called a photodiode. One dot equals one pixel, which is the smallest unit of an image. While the resolution of a film camera depends on the quality of the lens, the resolution of a digital camera depends on the number of pixels in the CCD. This is because a digital camera's CCD records an image using a fixed grid pattern. A higher spatial resolution may be attained by the image capturing system with a larger density of sensors. An imaging system with insufficient detectors will generate low resolution images with artifacts like blocky effects, ringing, checkerboard effect, etc. due to the aliasing from low spatial sampling frequency.



(a) (b)

Fig 1.3: An image scene captured with (a) small number of CCD sensors and (b) large number of CCD sensors

In digital image processing applications, images with high spatial resolution are desired for subsequent processing and analysis. As the resolution increases, the image becomes clearer. It becomes sharper, more defined, and more detailed as well. This is because there's more information in the same small space. Generally, the resolution of images obtained depends on the hardware employed. These days, images of high resolution are available by using modern high quality image sensing technologies such as high precision optics and charge-coupled devices (CCDs), but they are expensive.

This has led to the idea of resolution enhancing algorithms for obtaining higher resolution images in various applications. Normal camera sensors with low image resolution capability are used as such, but the resolution is improved using image processing techniques. This cuts down on the hardware and equipment expenses, since only software processing has to be done on the captured images. Therefore, resolution up-gradation techniques can improve the clarity of these images and help to solve the shortcomings of hardware improvisation.

1.6 SUPER-RESOLUTION

A class of techniques known as super-resolution algorithms has been developed to obtain high-resolution images from a collection of low resolution photographic images. Super-resolution (SR) algorithms are techniques that increase the high frequency components, replicating higher dimensional manifolds and removing the degradations caused by the imaging process of the low-resolution camera. These methods put together a collection of low-resolution images containing aliasing artifacts to restore a high-resolution image.

The general approach requires the re-sampling of a high resolution image of the training database to construct the low resolution image. The goal is then

to recover the high resolution image based on the input images and the imaging model that produces the low resolution observed images. Thus the accuracy of imaging model is vital for super-resolution. Incorrect modeling may result in further degradation of the image. The relative association between the high resolution and constructed low resolution images must then be learned. This represents the global correspondence between them, and the super-resolution procedure can then be applied to any test image of interest. In Figure 1.4, the image to the left is of lower spatial resolution and the image to the right is of higher spatial resolution obtained from the former image using SR techniques.

Figure 1.4: Super-resolution example

1.7 EXISTING SR ALGORITHMS

Super-resolution algorithms are classified based on the approach used to model the LR images. The two main categories are multi-frame and single frame super-resolution. Multi-frame image super-resolution (SR) aims to utilize information from a set of low-resolution (LR) images of the same scene to compose a high-resolution (HR) one. Image registration which is most important part of multi-frame super-resolution requires accurate alignment using the registration parameter. These images are then fused, which is the process of combing information of interest in two or more LR images into a single high-resolution image. Single-frame super-resolution generate high

resolution image from single degraded image or low resolution image. In single frame SR technique, the missing high frequency information in the LR image is estimated from large number of training set images and added to the LR image. This category of SR algorithms is also called as learning based super-resolution techniques as the LR-HR relation is learnt from a training database. On the other hand, single image super-resolution techniques use various levels of scale (size of the image) at different resolutions to extract fundamental characteristics between regions of low and high resolution. Single image super-resolution techniques are becoming more popular nowadays due to their simplicity and also there is no need for using a large training set to learn image pair relationship.

The organization of the thesis is presented in next section.

1.8 ORGANISATION OF THESIS

The rest of the thesis is arranged as follows:

Chapter 2 provides a literature overview of the various papers analyzed for this project,

Chapter 3 provides an introduction to traditional dictionary based algorithms and sparse representation,

Chapter 4 follows up with matrix valued regression and is extended to feature based regression analysis. The image pair operator and its properties are described. This chapter also explains in detail the training and testing phase of the algorithm,

Chapter 5 talks about the implementation of the above elucidated demosaicing and super-resolution algorithm in different applications,

And finally, Chapter 6 provides the future work in the domain and conclusion statements.

CHAPTER 2

LITERATURE SURVEY

There have been several outcomes of previous work in the field of super-resolution, each work proposing a novel methodology that has its own successes and limitations. Subsets of these papers which pertain to our field of research have been referred to as part of the project. The major paradigms of demosaicing algorithms include spatial domain methods or frequency domain methods. The paradigms of SR algorithms are: pixel interpolation-based algorithms, edge-based algorithms and regression based-algorithms. Very few works have been done regarding integrating both demosaicing and super-resolution.

In one of the seminal papers on demosaicing by Lukac et al. (2004) presented a normalized colour-ratio model for colour tiller array (CFA) interpolation schemes. The proposed normalized model enforces the underlying modelling assumption in both smooth and high frequency image regions. Using the proposed model, which represents a generalization of the conventional colour- ratio model, significantly boosts the performance of most well-known CFA interpolators, in both objective and subjective image quality measures.

Wang et al. (2012) has presented an adaptive demosaicing algorithm by exploiting both the non-local similarity and your regional correlation in along with filter array image. First, just about the flattest nonlocal image patches are searched within the searching window devoted to the estimated pixel. Second, the plot, which can be regarded the absolute most just like the current plot, is selected among the absolute easiest nonlocal patches. Third, in line with the similar degree and your regional correlation degree, the obtained nonlocal

image patches together with the current patch are adaptively chosen to estimate the missing colour sample.

Fang et al. (2012) has proposed an easy frequency domain analysis approach for joint demosaicing and sub pixel based down-sampling of single sensor Bayer images. From this, they integrated demosaicing into down-sampling by directly performing sub-pixel based down-sampling within the Bayer domain, which means that the computational complexity is reduced.

Ling Shao et al. (2014) proposed a content adaptive demosaicing strategy utilizing structure analysis and correlation involving the red, green and blue planes. Those two aspects were chosen in the classification associated with a block of pixels to produced trained filters. The planned technique aims to reconstruct a first class demosaiced image originating from a Bayer pattern in any colour filter array efficiently. Experimental results showed that the proposed strategy performs comparatively as higher end methods.

The work of Wu et.al (2016) titled “Bayer Demosaicing With Polynomial Interpolation”, introduced a polynomial interpolation-based demosaicing. This method makes three contributions: calculation of error predictors, edge classification based on colour differences, and a refinement stage using a weighted sum strategy. The predictors are generated based on the polynomial interpolation, and combined per an edge classifier. After populating three colour channels, a refinement stage is applied to enhance the image quality and reduce demosaicing artifacts.

Glasner et al. (2009) propose a unified framework for the classical multi-image super-resolution and example-based super-resolution. The proposed method further aims to recover new missing high-resolution details that are not explicitly found in any individual low-resolution image. It shows how this combined approach can be applied to obtain super-resolution from as little as a single image (with no database or prior examples). As shown in Figure 2.1, their

approach is based on the observation that patches in a natural image tend to redundantly recur many times inside the image, both within the same scale, as well as across different scales. Recurrence of patches within the same image scale (at sub-pixel misalignments) gives rise to the classical super-resolution, whereas recurrence of patches across different scales of the same image gives rise to example-based super-resolution. Their approach attempts to recover at each pixel its best possible resolution increase based on its patch redundancy within and across scales.

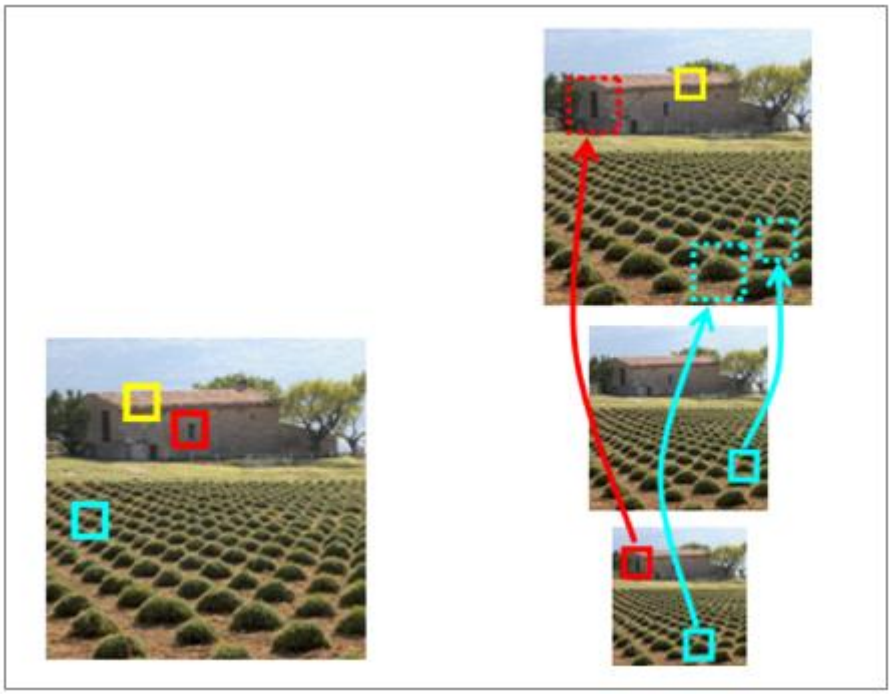


Figure 2.1: Recurrence of patches within and across scales of a single image.

Tang et al. (2013) suggested a regression approach on matrix scales for image SR. Single-image super-resolution is firstly treated as a problem of matrix-value regression. By using matrix-value regression techniques, some desired properties are found. Firstly, the matrix value regression technique greatly promotes the efficiency of learning from image pairs. As a result, the matrix-value regression based super-resolution algorithm can be smoothly applied to big data setting. Secondly, the matrix-value regression technique

makes it possible to design a patch-to-patch super-resolution algorithm. It is the first patch-to-patch algorithm in the field of single-image super-resolution.

Tang et al. (2015) have successively proposed an image pair analysis technique that provides significant image pair priori describing the dependency between training image pairs for various learning-based image processing. For avoiding the information loss caused by vectorising training images, a novel matrix-valued operator learning method is proposed for image pair analysis. Sample-dependent operators, named image pair operators (IPOs) by them, are employed to represent the local image-to-image dependency defined by each of the training image pairs. A linear combination of IPOs is learned via operator regression for representing the global dependency between input and output images defined by all of the training image pairs. The proposed operator learning method enjoys the image-level information of training image pairs because IPOs enable training images to be used without vectorising during the learning and testing process. The computational and memory complexities of the proposed algorithm heavily reduced.

Milanfar et al. (2005) have proposed a fast and robust method for joint multi-frame demosaicing and colour super-resolution. They have considered these problems in a joint setting because both refer to resolution limitations at the camera. They gave a multi-frame demosaicing technique based on a maximum a posteriori estimation technique by minimizing a multi-term cost function. The L1 norm is used for measuring the difference between the projected estimate of the high-resolution image and each low resolution image, removing outliers in the data and errors due to possibly inaccurate motion estimation. Bilateral regularization is used for spatially regularizing the luminance component, resulting in sharp edges and forcing interpolation along the edges and not across them.

Theodor Heinze et.al (2012) proposed an artificial neural network (ANN) based framework for joint demosaicing and super-resolution of colour filter array (CFA) raw image sequences. This is based on the idea that image processing steps are special cases of neural network processing and can be integrated into one single processing step. The proposed algorithm is self-adaptive to the number of frames; however it is too slow for real-time video data processing.

CHAPTER 3

DEMOSAICING ALGORITHM BASED ON EDGE INFORMATION

The reconstruction of full colour images from a CFA based detector necessitates a process of manipulating the values of any other colour separations at each pixel. The methods of these types are commonly referred as colour interpolation. The image below shown in Figure 3.1 shows the output from the Bayer layer image sensor each pixel has only Red, Green or Blue components.

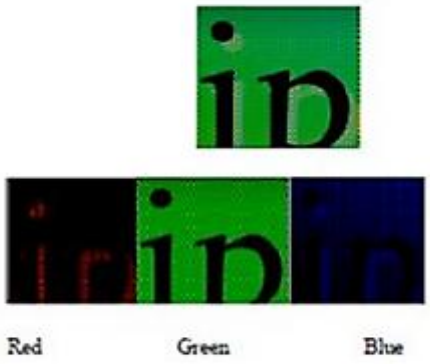


Fig 3.1: Bayer Filter Sample

Image reconstruction based on CFA may introduce colour artifacts and blur the image edge. Interpolation algorithm is simple without considering the edge information in image reconstruction process but the image reconstructed may be blurring and has low quality. So a digital camera is means to reconstruct a whole RGB image using all above information. The resulting image is like the shown image given in Figure 3.2.

Fig 3.2: Reconstructed Image

This reconstructed image will be accurate in uniform coloured areas but it has a loss of resolution and has edge artifacts. A reconstructed image from a CCD with a Bayer pattern CFA measures only 33% information of the original image. Two common types of artifacts are Zippering and False colouring.

1. Zippering or Blurring Artifacts

   Zippering and blurring effect is that when there is a one-side effect of CFA demosaicing, which occurs generally along edges. Therefore, edge blurring occurs along the edges in an on/off pattern.

2. False Colour Artifacts

   A regular artifact of CFA demosaicing is false colouring. It is that artifact which manifests along edges wherever quick or unpleasant adjustments in colour arise, as a result of mis-interpolation crosswise, rather than along the length of an edge.

Thus, it is important that a demosaicing algorithm must be chosen taking into consideration all these factors. Since the number of green pixel sensors is double the blue and red, reconstructing intensity image in green channel with full-resolution facilitates the unknown colour reconstruction of the scene, such as blue channel and red channel. There is a need for taking interpolation direction into consideration in order to ensure that interpolation is carried out effectively. Among Bayer mode, the Bayer method outputs the green square shaped grid and the rectangular grid on the blue and red array. Green channel contains more colour sensor data, so the first to carry out the interpolation of green components.

Abbreviations and Notations used in the proposed method are as follows

g(i,j) -	Interpolated green component at pixel location (i,j)
b(i,j) -	Interpolated blue component at pixel location (i,j)
r(i,j) -	Interpolated red component at pixel location (i,j)

G(i,j)-	Green Bayer component at pixel location (i,j)
R(i,j)-	Red Bayer component at pixel location (i,j)
B(i,j)-	Blue Bayer component at pixel location (i,j)
Lh -	Horizontal gradient at pixel position (i,j)
Lv -	Vertical gradient at pixel position (i,j)

The three steps used in this reconstruction are as given below:

(i) New calculation of gradients

(ii) Edge classification based on colour differences and gradient directions

(iii) Refinement stage using a weighted sum strategy for artifacts reduction

3.1 CALCULATING EDGE INFORMATION

The missing $g_{i,j}$ is calculated from the image filtered in both vertical direction and horizontal direction, and the optimally fusion can be done based on the two direction's estimations. Then the green intensity reconstruction full resolution image can be used to guide the interpolating of the unknown $r_{i,j}$ and unknown $b_{i,j}$. Lh and Lv can be used to determine how to interpolate the colour information of three different channels based on the interdependency of an image. Lh and Lv which indicate two different gradients of image information separately at pixel position (i, j) can be calculated by the below given formulas 3.1 and 3.2.

By comparing Lh and Lv, it can be estimated if there is a greater gradient changing or not at pixel (i,j) in the image. If Lh >Lv , then there are more great gray changing between the left parts and the right parts than between the above parts and the below parts at the pixel (i,j), so the unknown colour intensity at pixel (i,j) can be calculated by using the known colour information of its above parts and its below parts at the pixel (i,j). If Lv>Lh, then there are more great

gray changing between the above parts and the below parts than the left parts and the right parts at the pixel(i,j), so the unknown colour intensity at pixel (i,j) can be calculated by using the known colour information of the left and right parts to the pixel (i,j).

$$\begin{aligned} Lh = \sum_{m=0,\pm 2} & \left(\left|G_{i+m,j-1} - G_{i+m,j+1}\right|\right. \\ & \left. + \left|2R_{i+m,j} - R_{i+m,j-2} - R_{i+m,j+2}\right|\right) \\ & + \sum_{m=0,\pm 2} \left(\left|B_{i+m,j-1} - B_{i+m,j+1}\right|\right. \\ & \left. + \left|2G_{i+m,j} - G_{i+m,j-2} - G_{i+m,j+2}\right|\right) \end{aligned} \tag{3.1}$$

$$\begin{aligned} Lv = \sum_{m=0,\pm 2} & \left(\left|G_{i-1,j+m} - G_{i+1,j+m}\right|\right. \\ & \left. + \left|2R_{i,j+m} - R_{i-2,j+m} - R_{i+2,j+m}\right|\right) \\ & + \sum_{m=0,\pm 2} \left(\left|B_{i+m,j-1} - B_{i+m,j+1}\right|\right. \\ & \left. + \left|2G_{i,j+m} - G_{i-2,j+m} - G_{i+2,j+m}\right|\right) \end{aligned} \tag{3.2}$$

3.2 INTERPOLATING UNKNOWN GREEN INTENSITY

In order to achieve high quality image for HVS, the local edge information is used in the interpolating missing green component's procedure.

If Lh>2.Lv or Lv>2.Lh at pixel position (i,j), then there is a significantly gradient changing in image obviously, otherwise, there is a small gradient changing in image or there is a smooth local region in image. Here, according to different cases of the variable, $g_{i,j}$ can be interpolated by the following formula:

$$
g_{i,j} = \begin{cases} \left(\frac{G_{i,j-1} + G_{i,j+1}}{2} + \frac{2R_{i,j} - R_{i,j-2} - R_{i,j+2}}{4}\right) & Lv > 2Lh \\ \left(\frac{G_{i,j-1} + G_{i,j+1}}{2} + \frac{2R_{i,j} - R_{i,j-2} - R_{i,j+2}}{4}\right) & Lh > 2Lv \\ \frac{G_{i-1,j} + G_{i,j-1} + G_{i,j-1} + G_{i,j+1}}{4} & \\ \quad + \frac{4R_{i,j} - R_{i-2,j} - R_{i+2,j} - R_{i,j-2} - R_{i,j+2}}{8} & other \end{cases} \tag{3.3}
$$

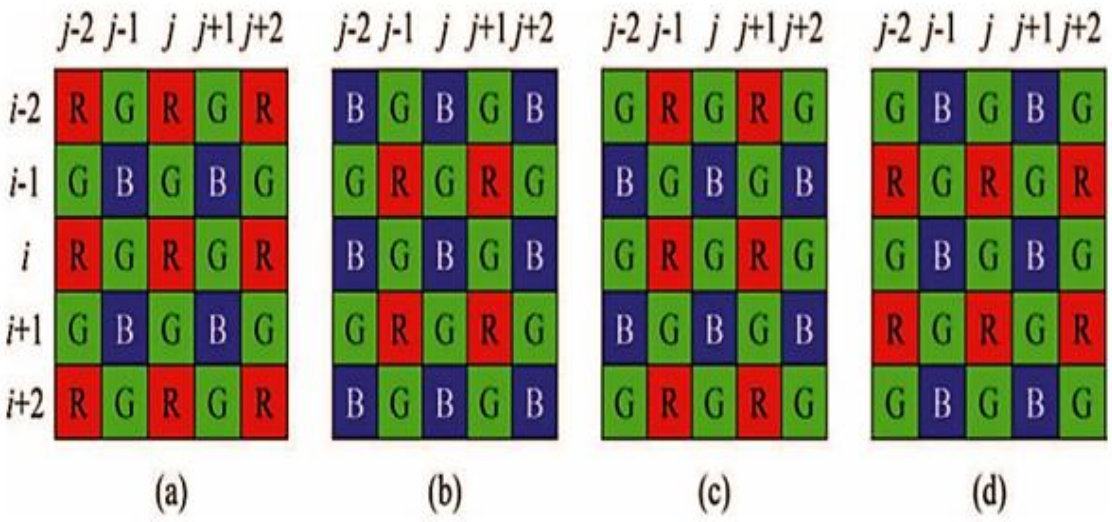


Fig.3.3: Bayer CFA Pattern with different center pixels

3.3 INTERPOLATING OTHER COLOUR PIXELS

3.3.1 Interpolating unknown red intensity at green pixel

Now using the interpolated green components from the previous step, the red and green intensities are estimated. The Figure 3.3 shows the Bayer CFA Pattern with different center pixels and taken from Wang(2012). The corresponding missing red sample, $r_{i,j}$ at a green pixel position in a RGRGRG as shown in Figure 3.3.(c) line is given by the formula below:

$$r_{i,j} = G_{i,j} + \frac{R_{i,j-1} + R_{i,j+1} - g_{i,j-1} - g_{i,j+1}}{2} \quad (3.4)$$

The corresponding missing red sample, $r_{i,j}$ at a green pixel position in a GBGBG line as shown in figure 3.3.(d) is given by the formula below:

$$r_{i,j} = G_{i,j} + \frac{R_{i-1,j} + R_{i+1,j} - g_{i-1,j} - g_{i+1,j}}{2} \quad (3.5)$$

3.3.2 Interpolating unknown blue intensity at green pixel

The unknown blue intensity $b_{i,j}$ in a RGRG and GBGB can be interpolated by the formulas given below:

$$b_{i,j} = G_{i,j} + \frac{B_{i-1,j} + B_{i+1,j} - g_{i-1,j} - g_{i+1,j}}{2} \quad (3.6)$$

$$b_{i,j} = G_{i,j} + \frac{B_{i,j-1} + B_{i,j+1} - g_{i,j-1} - g_{i,j+1}}{2} \quad (3.7)$$

3.3.3 Interpolating unknown blue intensity at red pixel

The unknown blue intensity $b_{i,j}$ can be interpolated as below:

$$b_{i,j} = g_{i,j} + \sum_{m=\pm1} \sum_{n=\pm1} \frac{1}{4}\left(B_{i+m,j+n} - g_{i+m,j+n}\right) \quad (3.8)$$

3.3.4 Interpolating unknown red intensity at blue pixel

The unknown red intensity $r_{i,j}$ can be interpolated as below:

$$r_{i,j} = g_{i,j} + \sum_{m=\pm1} \sum_{n=\pm1} \frac{1}{4}\left(R_{i+m,j+n} - g_{i+m,j+n}\right) \quad (3.9)$$

CHAPTER 4

LEARNING BASED SUPER-RESOLUTION

Various image super-resolution algorithms have been recommended with different assumptions and evaluation criteria based on the application. Super-resolution algorithms can be roughly classified into three groups. The first group generates HR images from LR inputs through a predefined mathematical formula without any training. The algorithms in this category include interpolation-based methods such as bilinear and bi-cubic, which generate HR pixel intensities by using weighted average of neighboring LR pixel values. But, interpolation-based methods tend to generate intensities that are locally similar to neighboring pixels and generate very smooth regions. The second group is the edge based methods in which priors are learned from edge features. Various edge features have been used such as the depth and width of an edge or the parameter of a gradient profile. Edge based methods have the advantage of sharpness and limited artifacts but is less effective for modeling high frequency structures like textures.

The third group is the learning based approach that uses a patch or feature-based method to learn the relationship between local image detail in low resolution and high resolution versions of the same scene. This learned knowledge is then incorporated into the priori term for reconstruction. The training database of learning based SR algorithms needs to have a proper generalization capability. Using a larger database does not necessarily generate better results, on the contrary, a larger number of irrelevant examples not only increase the computational time of searching, but it also disturbs the search. Some methods of learning the mapping functions are weighted average, sparse dictionary representation, kernel regression, etc.

4.1 REGRESSION BASED SR ALGORITHMS

Regression analysis is one of the most common methods among learning based super-resolution techniques. It is the process of estimating the relationship existing between variables. It formulates a correspondence between the dependant and the independent variable. With regression analysis, we get to know how the criterion or the dependant variable changes in accordance with one or more predictors, which are the independent variables. Applying this methodology to the case of super-resolution, it targets at establishing a universal relationship between low resolution and high resolution image pairs in the training set. This relation is obtained in the form of a function or an operator. It represents the global correspondence between the LR-HR image pairs and can be used on any LR image to improve its resolution. Some familiar methods of regression analysis are linear regression and nonlinear regression. In both these approaches, a least mean square error condition may be considered as a constraint to establish the best approximation. The data is fitted by a method of successive approximations where the sum of squares must be minimized by an iterative procedure.

Linear regression is an approach for characterizing the relationship between a scalar dependant variable *y* and one or more independent variables denoted *X*. Regression is done by fitting a predictive model to an observed data set of *y* and *X* values. The method is called a simple linear regression if there is only one independent variable involved. When more than one independent variable exists, the method is called multiple linear regression. The relationships are modeled using linear predictors whose unknown model parameters are estimated from the data. Commonly, line fitting and curve fitting techniques are used for this purpose. Nonlinear regression is a form of analysis in which data is modeled by a function which is a nonlinear combination of the model parameters and depends on one or more independent variables.

Regression models in SR, train the function or the operator to map the LR-HR relationship. This mapping can be either linear or non-linear. For ease of computation, linear operators are preferred whereas non-linear operators provide the advantage of minimal errors. This operator can later be used on low resolution test images to perform super-resolution. Thus, all regression based methods have a specific training phase and a testing phase.

4.2 IMAGE PATCH VECTORISATION

In usual regression based SR algorithms, the image pairs are converted into vectors during both the training and the testing phases. Image vectorisation is the process of converting images into a vector graphic. Computers understand images only as arrays of pixel values, limiting the operations that can be done to manipulate the pixels. Thus vectorisation transforms the problem of learning image-to-image dependency into the problem of learning vector-to-vector dependency. The vectorisation of images however suffers a drawback when it leads to the loss of image level information due to structural incongruity. In complex images where structural similarity is crucial, vector based priors prove to be inefficient. This can be overcome by avoiding the vectorisation of images and directly representing them as matrices.

4.3 MATRIX VALUE REGRESSION ANALYSIS

Successful image-pair analysis is the key point in performing learning based super-resolution algorithms. The image-pair regression operator defines the relation between training patch-pairs and should be effective enough to be applied as a global priori. Thus, the learning or training phase determines the preciseness of the operator and the quality of super-resolution output. As said earlier, vectorisation of images leads to the loss of image level image information and structural similarity. The matrix value regression technique

greatly promotes the efficiency of learning from image pairs and makes it possible to design a patch-to-patch super-resolution algorithm. Matrix based regression is also advantageous when compared to vector based methods since higher order differentiable components are more easily represented and identifiable in the matrix form rather than the vector form.

4.4 IMAGE PAIR OPERATOR

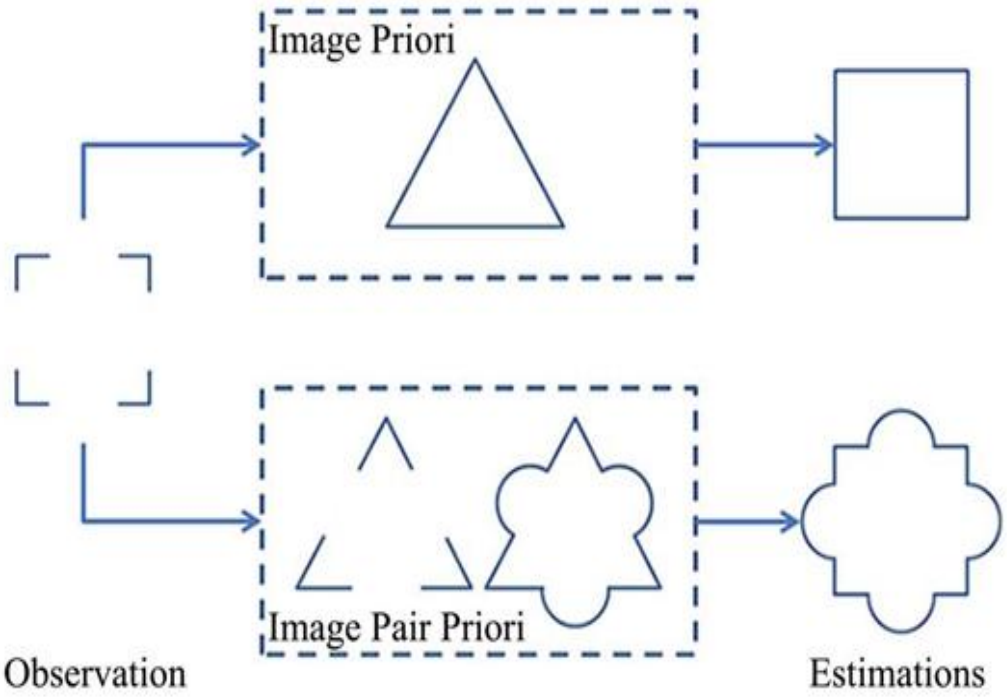


Figure 4.1: Image priori and pair-priori

Image patch-pair analysis is a method to learn the priori defined by a set of training image pairs. Image priors developed for specific tasks supply well targeted information. Training this image pair-priori by regression methods yields an operator that models the global image to image relationship. Thus, image pair analysis is a promising supervised learning technique to gain more information beyond the image priori for image processing. In figure 4.1, it could be found in the example that image pair priori makes the relation between

known training image pair preserved by the image pair of the observation and its estimation. Image pair analysis offers more chance to enhance the performance of learning-based image processing, such as offering larger magnification factor.

It is seen that image priori provides details specific to its own self and is not a global solution. Thus it does not provide targeted information to the algorithm. Alternatively, image pair-priori defines a mapping between the pair of images, and that is used to train the operator. Well learned pair priori provides more information than a normal priori, when used separately.

Tang et.al (2015) in the paper 'Image Pair Analysis With Matrix-Value Operator' defines an image pair operator (IPO) to represent the local dependency of image pairs using matrix-value operator. The performance of this IPO is based on the type of training set used. Thus, by adopting a simple constraint of least root mean square error based linear regression, the matrix value operator (MVO) can be well designed.

An image patch-pair denoted as $s = (x, y) \in R^{p*p}$ defines a linear image pair operator (IPO) $M \in R^{p*p}$. Mathematically, it is a natural choice to explicitly represent the local dependency of the image pairs via the matrix-valued operator. Each image pair $s = (x, y)$ potentially defines an operator M which satisfies,

$$y = M(x) \tag{4.1}$$

Restricting M to be a linear operator, it becomes

$$y = M \cdot x \tag{4.2}$$

where the linear operator M is represented by a matrix with the size $(p \text{ x } p)$, and ' $\cdot$ ' represents the matrix multiplication. For simplicity, the matrix multiplication sign ' $\cdot$ ' has be omitted in further use.

Assume that the image patch x is a full rank matrix. It is clear that a data-dependent linear operator *M* can be deduced from (4.2),

$$M = y\, x^{-1} \tag{4.3}$$

where x^{-1} means the inverse matrix of x. The data-dependent linear operator exactly models the dependency defined by the image pair $s = (x, y)$ because it follows from (4.3) that

$$y = Mx = (yx^{-1})x = y(x^{-1}x) = y \tag{4.4}$$

Thus, the local dependency of image pair can be equivalent to a data-dependent linear operator in some sense. Based on the above observations, an IPO is defined as follows:

Definition 1: For an image pair $\mathrm{s} = (\mathrm{x}, \mathrm{y})$, the IPO connected with the image pair is

$$M = yx^{\dagger} \tag{4.5}$$

where $\mathrm{x}^{\dagger}$ is the Moore–Penrose inverse of the matrix x.

The IPO operator *M* greatly depends on the full rank condition of its image patch-pairs to formulate the matrix inverse. For rank deficient matrices, computation of inverse is not stable. Generally in matrix-based image pair analysis methods the patch-pairs are assumed to be full rank matrices.

However, this places a huge constraint on the type of input images used. Also images with more smooth textures will intuitively have more rank deficient patches. Thus image patches extracted from natural images are intuitively rank deficient. To accommodate rank deficient patch-pairs to represent the image-pair priori, matrix inverse is computed by factoring the patch-pairs with singular value decomposition (SVD).

4.5 COMPUTATIONAL COMPLEXITY

The advantage of IPOs is that they employ 2-D similarity to represent the difference between training and test image patches. Therefore from (4.4), we get

$$M(x') = (y \cdot x^{-1}) \cdot x' = y \cdot (x^{-1} \cdot x') = y' \tag{4.6}$$

where, $z' = (x', y')$ is a test image pair. The term $x^{-1}x'$ records the difference between training image patch x and test image patch x' in 2-D style. From (4.6), the HR estimation y of the LR test image x can be found effectively using the MVR operator. A 2-D similarity measure will offer more information on the structure difference between two image patches.

The IPO operator significantly reduces the computational complexity by reducing the number of variables required to represent the operator. As the image patch-pairs are matrices of size $(p \times p)$, the image-pair regression operator will be a matrix of size $(p \times p)$. Therefore, it is required to have p^2 variables to represent the matrix-based regression operator. Nevertheless, in vector based approaches, as image patches are column vector of size $(p^2 \times 1)$, the regression operator that maps the two vectors should be a matrix of size $(p^2 \times p^2)$, requiring p^4 variables.

4.6 MATRIX VALUED OPERATOR

4.6.1 TRAINING PHASE

In this step, the relationship between some HR examples from a specific class like face images, finger-prints, etc. and their LR counterparts are learned. The training samples are used to learn an optimum matrix value regression operator. It is a learned image-pair priori which links the low-resolution and high-resolution patches. A few high quality images $\{Y_h^i \in \mathbb{R}^{m \times n}\}$ are first collected, captured by the high-resolution imaging device, and are considered as

high-resolution examples. These high-resolution example images are then down-scaled by a scale factor s. These down-scaled images form the corresponding low-resolution image set $\left\{Y_l^i \in \mathbb{R}^{\frac{m}{s}\times\frac{n}{s}}\right\}$. To avoid mismatch in matrix dimensions, the LR images are up-scaled to the size of the HR example images using an interpolation operator $Q : \mathbb{R}^{\frac{m}{s}\times\frac{n}{s}} \to \mathbb{R}^{m\times n}$ and is denoted by $\left\{X_l^i \in \mathbb{R}^{m\times n}\right\}$.

Thus, the set of images in $S = \left\{X_l^i, Y_h^i\right\}$ denote the training image pairs. Let x and y respectively denote image patches of size $(p \times p)$ extracted from X_l and Y_h respectively. We learn that for each image patch y obtained from the image Y_h centered at its origin (i, j), a self-similar example patch x around its origin (i_s, j_s) exists in the LR image X_l, where $i_s = [i/s + 0.5]$, and $j_s = [j/s + 0.5]$. Here, s is the above mentioned scale factor. This correspondence between y and x is learned by the matrix-value operator.

Let the training patch-pair set be denoted as $S_n = (x_i, y_i)_{i=1}^n \in \mathbb{R}^{p\times p}$, where (x_i, y_i) is the low-and high-resolution patch-pairs and n denotes the number of training patch-pairs. The matrix-value operator mapping the low-resolution image space to the high-resolution image space is then defined as $M : X \mapsto Y$. It represents the image level relation between the LR and HR training patches in the training sequence S_n. The optimal matrix-value operator M^* is subsequently learned from the training set S_n using the least square regression model,

$$M^* = \text{argmin}_M \sum_{i=1}^n \|y_i - Mx_i\|_F^2 \tag{4.7}$$

where $\|.\|_F$ is the Frobenius norm. This least square regression model can be thought of as in the Hilbert-Schmidt and the optimal matrix-value operator is given by;

$$M^* = \text{argmin}_M \sum_{i=1}^n \|y_i - Mx_i\|_F^2 \tag{4.8}$$

Denote $F_i(M) = \|y_i - Mx_i\|_F^2$

$= \|y_i\|_F^2 - 2\langle y_i, Mx_i\rangle_F + \langle Mx_i, Mx_i\rangle_F$

$= \|y_i\|_F^2 - 2\langle y_i\, x_i{}^T, M\rangle_F + \langle Mx_i x_i{}^T, M\rangle_F$

Therefore, Eq. (4.8) becomes

$$M^* = \underset{M}{\operatorname{argmin}} \sum_{i=1}^{n} F_i\,(M)$$

Thus the optimal target function is given by

$$F(M) = \sum_{i=1}^{n} F_i\,(M)$$

$$= \sum_{i=1}^{n} (\|y_i\|_F^2 - 2\langle y_i\, x_i{}^T, M\rangle_F + \langle Mx_i x_i{}^T, M\rangle_F)$$

$$= \left(\sum_{i=1}^{n} \|y_i\|_F^2 - 2\left\langle \sum_{i=1}^{n} y_i\, x_i{}^T, M\right\rangle_F + \left\langle M \sum_{i=1}^{n} x_i x_i{}^T, M\right\rangle_F\right)$$

$$= K_0 - 2\langle K_1, M\rangle_F + \langle MK_3, M\rangle_F \qquad (4.9)$$

where $K_0 = \sum_{i=1}^{n} \|y_i\|_F^2$, $K_1 = \sum_{i=1}^{n} y_i\, x_i{}^T$ and $K_2 = \sum_{i=1}^{n} x_i x_i{}^T$

The optimal matrix-value operator M^* should satisfy the local minima condition. Therefore,

$$\frac{\partial}{\partial M} F(M) = 0$$

$$\frac{\partial}{\partial M} (K_0 - 2\langle K_1, M\rangle_F + \langle M, M\rangle_F) = 0 \qquad (4.10)$$

Therefore the optimal Matrix-value operator is given by

$$M^* = K_1 K_2{}^{-1} \qquad (4.11)$$

As previously stated, due to the rank deficiency obstacle, we compute the inverse of the auxiliary matrix K_2 by factorizing it with singular value

decomposition, thus $K_2 = U\Sigma V^T$, where U and V are orthogonal matrices and Σ is a diagonal matrix with singular values.

$$M^* = K_1(V\Sigma^{-1}U^T) \tag{4.12}$$

The optimal matrix-value operator M^* shown in Eq. 4.12 explicitly represents the image-level correspondence between the low and high-resolution image patch-pairs. The matrix-value operator resulting from the training phase is expected to better reconstruct the fine details from the low-resolution images.

The procedure to obtain optimal MVR is summarized in Algorithm 1. Figure 4.2 provides an overview of the proposed algorithm.

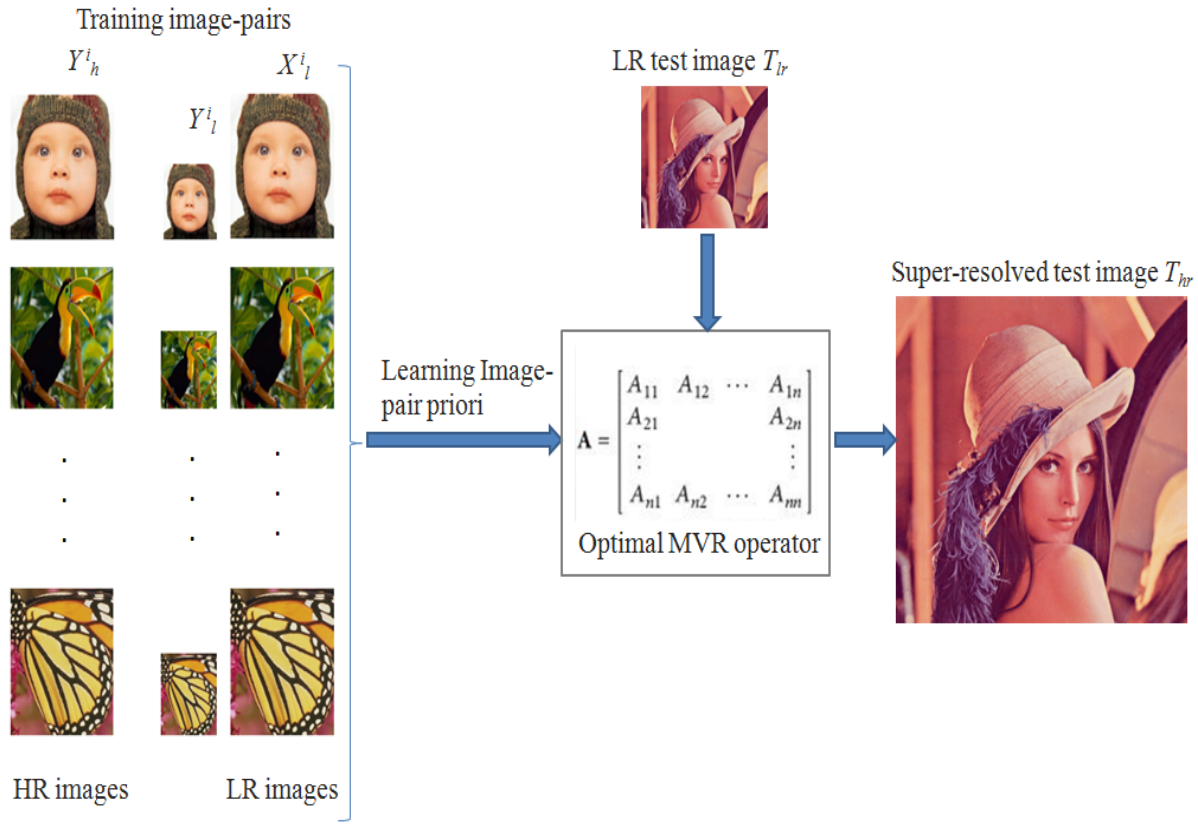


Figure 4.2: Proposed matrix based SR methodology

4.6.2 RECONSTRUCTION PHASE

This phase makes use of the matrix-value regression operator learned from the previous step to super-resolve a low-resolution image. The test image T_{lr} is interpolated by a factor s. Non-overlapping image patches of size $(p \times p)$ are extracted from the interpolated test image. This collection of low-resolution patches is represented as $T_s = \{t_{lr}^i\}_{i=1}^n$. Every test image patch in the set T_s is super-resolved using the matrix value regression operator:

$$t_{hr} = M^* t_{lr} \tag{4.13}$$

The super-resolved test image patches are merged together to form the super-resolved high-resolution image T_{hr}. The steps involved in the reconstruction phase are summarized in Algorithm 2.

Algorithm 1 Training phase

Input : Training image patch-pairs $S_n = (x_i, y_i)_{i=1}^n \in \mathbb{R}^{p \times p}$,

Output: Optimal Matrix-value operator M^*

Steps:

1. Calculate the auxiliary matrices K_1 and K_2

$$K_1 = \sum_{i=1}^{n} y_i \, {x_i}^T$$

$$K_2 = \sum_{i=1}^{n} x_i {x_i}^T$$

2. Factorize the auxiliary matrix K_2 using SVD

$$K_2 = U \Sigma V^T$$

3. Find the inverse of auxiliary matrix K_2

$${K_2}^{-1} = V \Sigma^{-1} U^T$$

4. Find the optimal Matrix-value operator

$$M^* = K_1 (V \Sigma^{-1} U^T)$$

7. **Result:** Optimal Matrix-value operator, M^*

Algorithm 2 Reconstruction phase

Input : Optimal Matrix-value operator M^*, low-resolution test image T_{lr}

Output: Super-resolved HR image T_{hr}

Steps:

1. Construct non-overlapping patches of size $p \times p$ from test image

$$T_{lr} \rightarrow T_s = \{t_{lr}^i\}_{i=1}^n$$

2. For every test image patch t_{lr}^i , find the super-resolved patch

$$t_{hr} = M^* t_{lr}$$

3. Merge the super-resolved patches $\{t_{hr}\}_{i=1}^n$

$$T_{hr} = \{t_{hr}\}_{i=1}^n$$

Result: Super-resolved HR image, T_{hr}

CHAPTER 5
RESULTS AND DISCUSSION

Image enhancement techniques are gaining importance nowadays due to the wide variety of applications which make use of image samples to process information related to their application. Image enhancement techniques used in many areas such as forensics, astro-photography, fingerprint matching, etc. It includes contrast enhancement, edge sharpening, blur reduction, removing noise are just some of the techniques used to make the images bright. Images obtained from fingerprint recognition, safety measures videos analysis and indulgence scene investigations are enhanced to help in identification of culprits and protection of victims. In this project demosaicing followed by super-resolution is applied to standard test images. The training of the data for super-resolution is also done using standard test images due to similarity in terms of presence of objects of interest present in the image.

Many works till now have separately focused on either of these techniques. The combined framework of the proposed method is studied using qualitative and quantitative measurements. The performance of the proposed method is evaluated and compared with some reference benchmarks. In the experiments the McM dataset was used. The McM dataset has been used in many recent survey papers. The McM dataset consists of 18 full colour images with a pixel resolution 500 × 500; however for experimental purpose we have used only 5 images.

(a)

(b)

(c)

(d)

(e)

Fig 5.1. Images #1, #5, #11, #17, #18 from the McM dataset correspond to (a) through (e) respectively

5.1 EVALUATION PARAMETERS

The performance is examined by the experimental results obtained by the proposed algorithm, both for demosaicing as well as super-resolution. The effectiveness of the proposed algorithm is a measure of visual experience obtained from the reconstructed image. The reconstructed image is evaluated both qualitatively and quantitatively to assess its effectiveness.

5.1.1 Qualitative Evaluation

Qualitative evaluation depends on a few attributes of the reconstructed image. The image is visually inspected for its naturalness and sharpness to assess the quality of the reconstructed image. The sharpness of an image is assessed based on the high-frequency details present in it. It is desired that the algorithm should not introduce any counterfeit HF details. Similarly, image naturalness is attributed to the distortions and artifacts present in the image. If the fine-details in the image are not preserved, it will introduce jaggy and ringing and staircase artifacts. These artifacts will severely affect the quality of the image. These attributes in the images can be evaluated by visual comparison of the images.

5.1.2 Quantitative Evaluation

For quantitatively evaluating the objective performance of the demosaicing process, the difference incurred between the original and the demosaiced images are measured based on the following image fidelity metrics,

1) Mean Squared Error(MSE)
2) Peak signal-to-noise ratio(PSNR)
3) Structural Similarity Index(SSIM)
4) Feature Similarity Index(FSIM)

Mean Squared Error (MSE) factor represents the average of the squares of the errors or deviations, i.e., the difference between the estimator and what is estimated. The Peak Signal Noise Ratio gives the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale. PSNR is most commonly used to measure the quality of reconstruction. A high PSNR score indicates that the magnified image is free from distortions and is more likely to carry HF details. The PSNR of an image is defined by,

$$PSNR = 10log10\left(\frac{255^2}{MSE_{xy}}\right) \tag{5.1}$$

where $MSE_{x,y} = \frac{||x-y||^2}{W*H}$, W is the width of the image patches x and y, H is height of both the image patches.

SSIM index is a method for measuring the similarity between two images based on an initial uncompressed or distortion-free image as reference. The SSIM of the reconstructed image is obtained using,

$$SSIM(x,y) = \left(\frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}\right) \tag{5.2}$$

where

$$\mu_x = \frac{1}{W * H\,\Sigma_{i=1}^{W*H} x_i}$$

$$\mu_y = \frac{1}{W * H\,\Sigma_{i=1}^{W*H} y_i}$$

$$\sigma_x = \frac{1}{W * H - 1\Sigma_{i=1}^{W*H}((x_i - \mu_x)^2)^{\frac{1}{2}}}$$

$$\sigma_y = \frac{1}{W * H - 1\Sigma_{i=1}^{W*H}((y_i - \mu_y)^2)^{\frac{1}{2}}}$$

where c_1 , c_2 are constants.

FSIM is a similarity measurement based on comparison between image features like gray levels, texture with the coefficients in the Fourier, structure, etc.

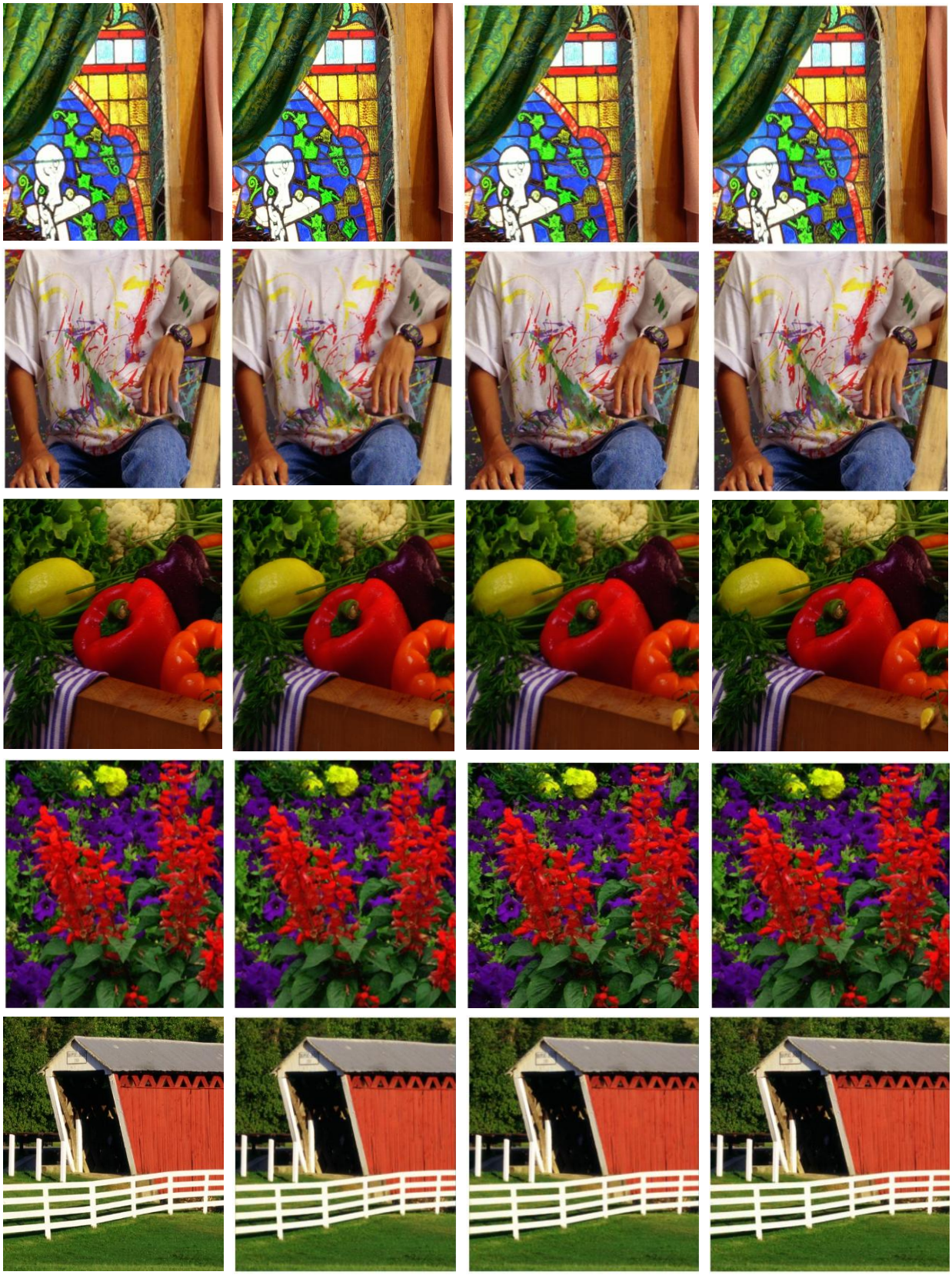
5.2 RESULTS OF PROPOSED DEMOSAICING ALGORITHM

The standard demosaicing algorithms against which the results are evaluated include bilinear intepolation and edge corrected bilinear interpolation. Bilinear Interpolation - They process each component plane separately and find the missing levels by applying linear interpolation on the available ones, in both main directions of the image plane. Considering the {GRG} structure, the missing blue and green values at the center pixel are respectively estimated by bilinear interpolation thanks to the following equations :

$$\hat{B} = \frac{1}{4}(B_{-1,-1} + B_{1,-1} + B_{-1,1} + B_{1,1}) \quad (5.3)$$

$$\hat{G} = \frac{1}{4}(G_{0,-1} + G_{-1,0} + G_{1,0} + G_{0,1}) \quad (5.4)$$

Edge corrected bilinear interpolation makes use of the edges which have much stronger luminance than chrominance components. The bilinear interpolation estimate is corrected by a measure of the gradient for the known colour at the pixel location. The gain factor α controls the intensity of such correction.



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

Fig. 5.2. Variation of test image output for (a) ground truth image (b) bilinear interpolation (c) edge corrected bilinear interpolation (d) proposed method

Table 5.1 and 5.2 summarizes the quantitative comparison of the proposed method with various demosaicing algorithms. Table 5.1 shows that the proposed demosaicing algorithm has the highest quantitative measures compared with other state-of-the-art algorithm. It is evident from Table 5.1, the PSNR index that the proposed algorithm reconstructs the image with minimum distortions and the high SSIM index validates that the image-level information is preserved by the proposed matrix-based implicit prior. From Table 5.2 where the demosaicing algorithms are compared based on similarity inidices, namely structural similarity index and feature simlarity index. The proposed method performs better compared to the other two considered algorithms for the images besides #11, where bilinear interpolation performs a better reconstruction. As shown in Figure. 5.3 qualitative evaluation between the reconstructed images from various algorithms clearly reveal that the edge corrected bilinear interpolation performs a better reconstruction. Thus, it is observed that the reconstruction varies with the image selected as well. Algorithm that works well for an image doesn't have to work efficiently for all images in the set.

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

Fig. 5.3. Variation of test image #11 output for (a) ground truth (b)bilinear interpolation (c) edge corrected bilinear interpolation (d) proposed method

Image	Bilinear Interpolation	Edge Corrected Bilinear Interpolation	Proposed method
#1	26.4881/145.9733	27.5163/115.0988	**27.5488/114.728**
#5	30.3223/60.3750	32.1644/42.5037	**33.9301/41.6937**
#11	37.8748/21.1639	**36.7924/10.8103**	37.1407/12.5605
#17	31.6560/44.4105	31.3591/47.5519	**32.8057/42.0009**
#18	28.7281/87.1511	33.1792/31.2726	**33.6016/30.7206**

Table 5.1: Comparison of output PSNR/MSE for different algorithms

Image	Bilinear Interpolation	Edge Corrected Bilinear Interpolation	Proposed method
#1	0.9463/0.9839	0.9640/0.9916	**0.9682/0.9923**
#5	0.9549/0.9908	0.9709/0.9961	**0.9793/0.9969**
#11	0.9832/0.9942	**0.9896/0.9983**	0.9883/**0.9983**
#17	0.9763/0.9923	0.9617/0.9960	**0.9773/0.9968**
#18	0.9573/0.9764	0.9802/0.9956	**0.9869/0.9958**

Table 5.2: Comparison of output SSIM/FSIM for different algorithms.

5.3 RESULTS OF PROPOSED SUPER-RESOLUTION ALGORITHM

Generally super-resolution techniques have been used to improve the visual appeal of an image, and to extract finer details from the image. It has been extensively applied for surveillance camera images, medical diagnosis using images and in other industrial applications. Nowadays, the multifarious advantages of super-resolution have given rise to a deluge of applications. Notably, super-resolution finds many applications in images obtained in

space/underground exploration, underwater imaging, remote-sensing and in social media.

Each application has to be chosen carefully, taking into account the general structure and information content in the image. For example, in selfie images taken from the front camera of the smartphone, a typical image contains a strong background in terms of information complexity and a solid, smooth front-ground of facial information. So training the algorithm using structurally similar images leads to an improvement in the results leading to sharper details, edges and facial features.

In the medical application, images are standard, defined and are taken for a particular purpose like X-Ray imagery, MRI imagery etc. Thus the images fall into the same class. For example, in X-Ray images, the region of interest constitutes a very small proportion of the complete image. In an X-ray of an arm fracture, the desired information is concentrated in the bright parallel section (the wounded arm) whereas the other regions are void of data. We train our operator for that particular class of images with a predefined structure, leading to better results.

In another application called CFA imaging, the training is tricky. The images obtained are, strictly speaking, not the exact ground truth image. They are demosaiced images derived from a CFA image matrix incomplete with each colour component. To counter this, we can impose priors on the available CFA sub-spaces to improve the interpolated demosaiced output.

More training images were obtained from the Caltech 101 database available online at http://www.vision.caltech.edu/Image_Datasets/Caltech101/ based from California Institute of Technology website. This dataset contains pictures of objects grouped into 101 categories, with about 40 to 800 images per category. The size of each image is roughly 300 x 200 pixels.

The training and testing colour images are first converted into the YCbCr channel. Being the only channel that is sensitive to the human eye, our target being the enhancement of human visual perception, only the luminance channel is considered during super-resolution. The LR test images are artificially generated by down sampling the images using the Bicubic decimator. These down sampled LR images are then blown up to the size of the target HR image and are subsequently blocked contiguously into non-overlapping patches of size 27*27 as default. The LR test images are then super-resolved with a scale factor ‘s’ of either 2, 3 or 4 – the same with which the LR-HR training image pairs were generated.

Figure 5.4: Sample training images

Qualitative and quantitative evaluations have been carried out to assess the effectiveness of the proposed algorithm. Qualitative evaluation of SR methods relies on a few attributes of the reconstructed image such as sharpness, naturalness, and granularity. The sharpness of an image is assessed based on the high frequency details it preserves. The naturalness of an image is affected by the artifacts present in it. Various artifacts such as ghosting, ringing, jagging, and staircase artifacts generally affect the quality of an image.

A visual comparison is made to assess the fidelity of the proposed algorithm qualitatively. The effectiveness of the proposed method is quantitatively evaluated based on a few objective performance metrics such as mean square error (MSE), peak signal-to-noise ratio (PSNR), and structural similarity indices (SSIM/FSIM). A high PSNR score indicates that the scaled-

up image is free from distortions and effectively reconstructs the HF details. Similarly, a high SSIM value (typically 1) implies that the scaled-up image has a very similar structure to its ground truth. The sample training images shown in Figures 5.4 are used to train the MVR operator and perform super-resolution and results are compared with the algorithms given by Yang (2013), Kim(2010) and bicubic interpolation. Table 5.3 and 5.4 summarizes the quantitative comparison of various SR algorithms on test images for 2x magnification.

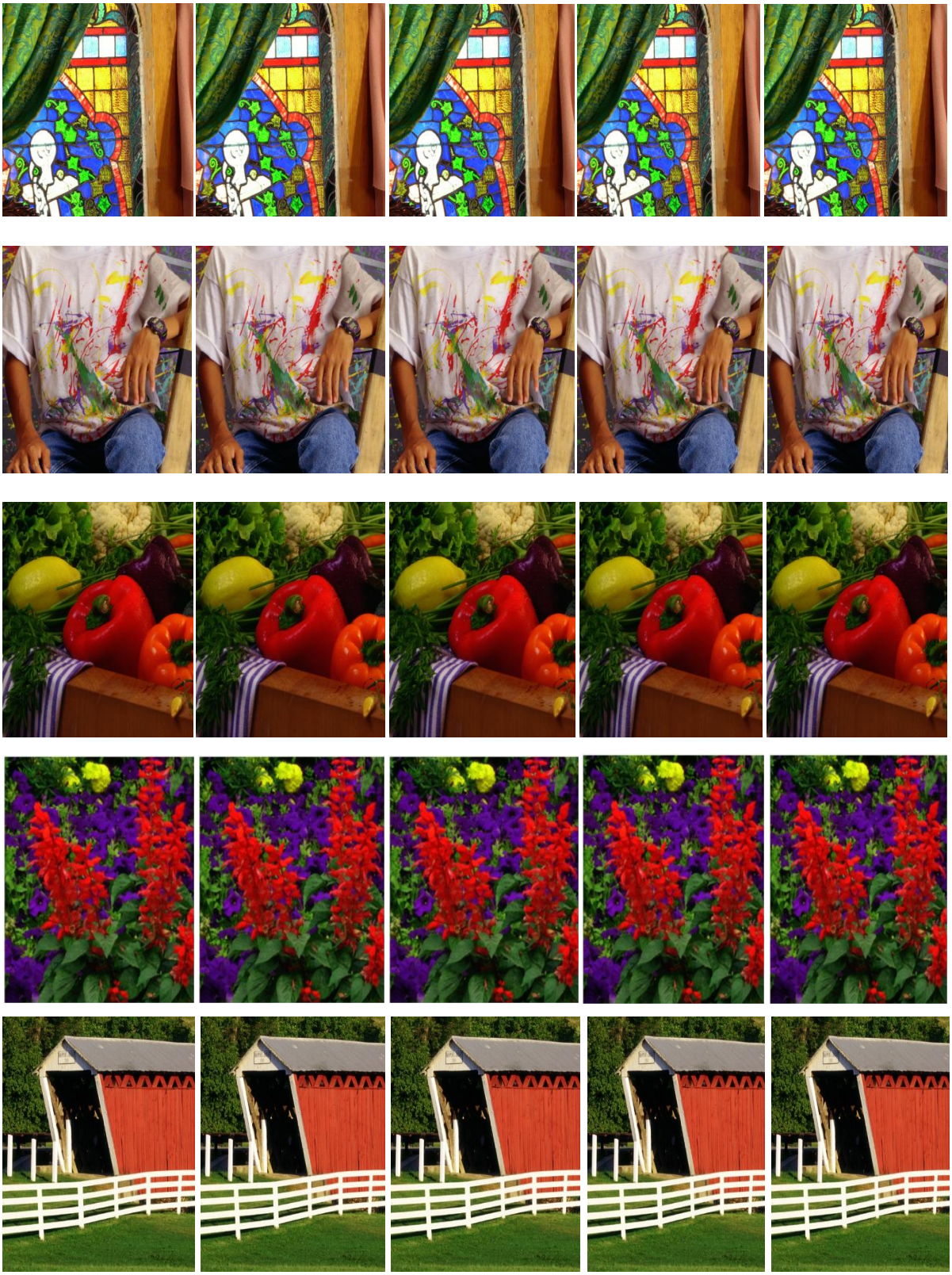
For more comparison, 2× magnification of images is carried out with 5 sample images from the McM dataset. Figure 5.5 depicts the qualitative visual comparison for these five images. Figure 5.5(a) depicts the test LR image. Figures 5.5(b) - 5.5(d) depict the SR images reconstructed by Yang et al.'s, Kim et al.'s, and bicubic interpolation. Figure 5.5(e) shows the proposed SR image.

Test Image	**Yang et al.**	**Kim et al.**	**Bicubic**	**Proposed**
#1	29.21/36.8405	30.08/38.4296	28.51/38.6784	**32.2320/36.1936**
#5	35.2037/15.3167	36.1034/15.7789	35.3966/15. 6333	**36.1903/15. 1459**
#11	36.0158/9.9365	36.9645/10.9005	35.3480/10.9985	**37.7509/9.9141**
#17	33.5583/33.8607	33.7947/34.3309	33.5891/34.4502	**34.8601/33.1241**
#18	31.4590/62.9132	**32.2815**/63.0309	30.21/62.5519	30.1246/**62.1864**

Table 5.3: Comparison of output PSNR/MSE for different SR algorithms.

Test Image	**Yang et al.**	**Kim et al.**	**Bicubic**	**Proposed**
#1	0.9766/0.9754	0.9801/0.9787	0.9825/0.9748	**0.9875/0.9817**
#5	0.9372/0.9701	0.9785/0.9882	0.9761/0.9833	**0.9884/0.9898**
#11	0.9365/0.9766	0.9847/0.9815	0.9893/0.9834	**0.9906/0.9875**
#17	0.9583/0.9730	0.9735/0.9757	0.9851/0.9769	**0.9863/0.9808**
#18	0.9653/0.9699	**0.9788/0.9712**	0.9457/0.9535	0.9685/0.9562

Table 5.4: Comparison of output SSIM/FSIM for different SR algorithms.



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

Fig 5.5: Variation of test image for different SR algorithms (a) Test LR image (b)Yang et.al (c) Kim et.al (d) Bicubic Interpolation (e)Proposed method

In Yang et al.'s SR based on sparse representation model, two coupled dictionaries are trained simultaneously from random raw image patches. Based on a dictionary pre-t-rained from thousands of natural images, Yang et al.'s method seems to produce natural-looking results. Though Yang's algorithm faithfully reconstructs natural-looking images, it can be observed from Table 5.3 that the objective measures are not the best among other comparative algorithms. This is because the fine details in the image are not well preserved due to the fact that a universal dictionary used in this method fails to represent complex structures accurately. Due to the fact that a natural image priori is used to post-process the SR image, Kim et al.'s method effectively reproduces more visually appealing images. The PSNR and SSIM value for Kim et al.'s method is better than other comparative algorithms, as a post-processing with an image edge priori is carried out on the reconstructed image.

On the contrary, the proposed method preserves the sharp details and fine textures in most of the images without affecting the naturalness of the image. Also it is observed that the proposed method provides more photo-realistic details as it does not introduce any counterfeit fine details. The proposed method achieves the best PSNR and SSIM value which indicates that the proposed algorithm reconstructs the LR image with minimal distortions and a high SSIM value corroborates the effectiveness of the structural similarity which has been preserved by the proposed matrix-based regression algorithm. The proposed method performs better than other state-of-the-art SR approaches as it avoids vectorisation of image patch-pairs during training phase of the MVR operator, which intuitively preserves structural similarity and image-level information within patch-pairs. Also, as the MVR operator is trained with HR images captured by the rear camera of the smart-phone it effectively corresponds to the relation between LR-HR patch-pairs, thereby improving the performance.

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

Fig 5.6: Variation of local image output from image #11 (a) LR local image, (b)Yang et.al (c) Kim et.al (d) Bicubic Interpolation (e) Proposed method

The image in 5.6 (a) represents the local image from LR test image #11.Though the stripes in the table cloth are sharp in Figure 5.6(b), it is not the same pattern as in the ground truth as the fine details in the table cloth are not well preserved. In Figure 5.6 (c) and (d), the texture on the table cloth is blurred when compared with (b). The result in 5.6 (e) from our proposed method shows good improvement in resolution and preserves sharp texture details.

The performance of the proposed algorithm can be influenced by the training dataset used to train the MVR operator. To validate this, a performance evaluation based on variation in dataset is carried out. It is observed that the training images similar in structure as the test image lead to better results than when the training and testing images depict different scenes. This self-similarity improves the interdependency between the images as well as the image patches and results in a more robust and efficient MVR operator that results in improved output.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 CONCLUSION

In this work, an image reconstruction algorithm is proposed to reconstruct mosaics based on an improved interpolating method and is followed with an image super-resolution algorithm. The demosaicing algorithm introduces green channel's edge information to interpolate green components firstly, then to reconstruct red and blue components by adopting fully reconstructed green plane. Our work proposes a demosaicing algorithm based on CFA edge information to remove the mosaic phenomenon, it reconstructs colour image by analysing local spatial gradient information. Because the method uses the horizontal and vertical gradient details, information present in the edge pixels is preserved and hence provides good results. The edges and textures of scene can be preserved in the output, and the reconstructed image has good visual quality. The results show that this algorithm has good performance, when quantitatively verified.

Further, we implemented a novel matrix-based regression methodology instead of the traditional vectorisation approach. We found that matrix-based operators better model the image structural information available spatially, while the vector operators fail to preserve this structural information. Due to better modelling, the results of our algorithm are competitive when compared to other state of the art methods. We proposed a novel patch variance based sampling method to balance the computational cost of feature regression. The algorithm is verified using standard test images.

6.2 FUTURE WORK

The demosaicing of images with weak spectral correlation can be done for further improvement of this work. An integrated colour demosaicing algorithm with colour correction algorithm may be proposed for better enhancement. Also some evolutionary optimization technique can also be considered to enhance the results. For the case of super-resolution, we believe that implementing higher order manifold learning in matrix domain will lead to better results, because we preserve the manifold structure in matrix domain. The learning can also be assisted by better optimization techniques such as gradient descent, conjugate gradient descent and evolutionary optimization techniques such as honey bee optimization, and genetic algorithms. We can also use back-propagation to control the parameter error iteratively. Another avenue of research is to better model the image matrix and the cost function, considering the class of images and its structure.

REFERENCES

[1] Chang, H., Yeung, D.Y., and Xiong, Y. (2004) 'Super-resolution through neighbor embedding', in Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. I–I.

[2] Cok D. R., 'Signal processing method and apparatus for producing interpolated chrominance values in a sampled colour image signal', Feb. 10 1987, US Patent 4,642,678.

[3] Fang et.al., (2012) 'Joint Demosaicing and Subpixel-Based Down-Sampling for Bayer Images: A Fast Frequency-Domain Analysis Approach', IEEE Transactions on Multimedia, Vol.14, No.4, pp. 1359- 1369.

[4] Freeman, W. T., Jones, T. R. and Pasztor, E. C. (2002) 'Example-based Super-resolution', IEEE Computer graphics and Applications, Vol. 22, No.2, pp. 56–65.

[5] Glasner, D., Bagon, S. and Irani, M. (2009) 'Super-resolution from a single image', in Proceedings of the International Conference on Computer Vision, pp. 349–356.

[6] He, H. and Siu, W.C. (2011) 'Single image super-resolution using Gaussian process regression', in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 449–456.

[7] Heinz, T., Lowis. M, and Polze, A. (2012) 'Joint multi-frame demosaicing and super-resolution with artificial neural networks', International Conference on Systems, Signals and Image Processing (IWSSIP'12), pp. 540-543.

[8] Kim, K. I. and Kwon, Y. (2010) 'Single-image super-resolution using sparse regression and natural image prior', IEEE transactions on pattern analysis and machine intelligence, Vol. 32, No. 6, pp. 1127–1133.

[9] Ling Shao and Amin Ur Rehman, (2014) 'Image Demosaicing using content and color-correlation analysis' ,Signal Processing, Elsevier, pp. 84-91.

[10] Lukac, R. and Plataniotis, K. N. (2004) ‘A normalized model for colour ratio based demosaicking schemes’, in Proceedings of the International Conference on Image Processing (ICIP’04), Vol. 3, pp. 1657–1660.

[11] Milanfar, P., Farsiu, S. and Elad, M. (2006) ‘Multiframe Demosaicing and Super-Resolution of Color Images’, in IEEE Transactions on Image Processing, Vol. 15, No. 1, pp 141-159.

[12] Tang, Y. and Chen, H. (July 2013) ‘Matrix-value regression for single image super-resolution’, in Proceedings of the International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR '13), pp. 215-220.

[13] Tang, Y. and Yuan, Y. (2015) ‘Image pair analysis with matrix-value operator’, IEEE Transactions on Cybernetics, Vol. 45, No. 10, pp. 2042-2050.

[14] Wang, G., Xiu-Chang, Z., Zong-Liang, G. (2012) ‘Image Demosaicing by Non-local Similarity and Local Correlation’, ICSP Proceedings, pp. 806–810.

[15] Wu, J. et.al. (2016) ‘Bayer Demosaicing With Polynomial Interpolation’, IEEE Transactions on Image Processing, Vol. 25, No. 11, pp. 5369-5382.

[16] Yang, J., Wright, J., Huang, T. S. and Ma, Y. (2010) ‘Image super-resolution via sparse representation’, IEEE Transactions on Image Processing, Vol. 19, No. 11, pp. 2861–2873.

[17] Yang, J., Lin, Z. and Cohen, S. (June 2013) ‘Fast image super-resolution based on in-place example regression’, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13), Portland, Ore, USA, pp. 1059-1066.

[18] Ye, W. and Ma, K.K. (2015) ‘Colour image demosaicing using iterative residual interpolation’, IEEE Transactions on Image Processing, Vol. 24, No. 12, pp. 5879–5891.