Agent Based Statistical Approach for Image Inpainting with High Fidelity

A thesis submitted to the University of Hyderabad
in partial fulfillment of the requirements for the award of

Doctor of Philosophy

in

Computer Science

by

Raghava Morusupalli

09MCPC04

School of Computer and Information Sciences

University of Hyderabad

Hyderabad

Telangana– 500 046

India

December 2016

CERTIFICATE

This is to certify that the thesis entitled "Agent Based Statistical Approach for Image Inpainting with High Fidelity" submitted by **RAGHAVA MORUSUPALLI** bearing registration number **09MCPC04** in partial fulfillment of the requirements for the award of **Doctor of Philosophy** in the **School of Computer and Information Sciences** is a bonafide work carried out by him under our supervision and guidance.

The thesis is free from plagiarism and has not been submitted previously in part or in full to this or any other University or Institution for the award of any degree or diploma.

Parts of this thesis have been presented in the following International Conferences:

1. Raghava,M.,Arun, Agarwal., Raghavendra, Rao C. Spatial Anisotropic Interpolation Approach for Text Removal from Images, MIWAI 2013, LNCS, Vol 8271,153-164, 2013. (ISBN: 978-3-642-44948-2)

2. Raghava,M.,Arun, Agarwal., Raghavendra, Rao C. A Scalable Spatial Anisotropic Interpolation Approach for Object Removal from Images using Elastic Net Regularization, MIWAI 2016, LNCS, Vol 10053,126-140, 2106. (ISBN: 978-3-319-49396-1)

Further, the student has passed the following courses towards fulfillment of coursework requirement for Ph.D.

	Course Code	Name	Credits	Result
1	CS900	Data Structures and Algorithms	4	Pass
2	CS901	Operating Systems and Programming	4	Pass
3	CS909	Digital Image Processing	4	Pass
4	CS918	Trends in Soft Computing	4	Pass

Supervisor	Supervisor	Dean
Prof. Arun Agarwal	**Prof. C Raghavendra Rao**	**Prof. Arun Agarwal**

DECLARATION

I, **Raghava Morusupalli** hereby declare that this thesis entitled **"Agent Based Statistical Approach for Image Inpainting with High Fidelity"** submitted by me under the guidance and supervision of **Prof. Arun Agarwal** and **Prof. C Raghavendra Rao** is a bonafide research work and is also free from plagiarism. I also declare that it has not been submitted previously in part or in full to this University or any other University or Institution for the award of any degree or diploma. I hereby agree that my thesis can be deposited in Shodganga/INFLIBNET.

A report on plagiarism statistics from the University Librarian is enclosed.

Signature of the Student
Date:
Name: **Raghava Morusupalli**
Regd No: 09MCPC04

Prof. Arun Agarwal
Supervisor

Prof. C Raghavendra Rao
Supervisor

School of Computer and Information Sciences

University of Hyderabad

Hyderabad

India

DEDICATION

I dedicate this thesis

to

My parents,

Smt C Kamala Garu & Sri C Madhusudhana Reddy Garu

and

Lord Lakshmi Narasimha Swamy

ACKNOWLEDGMENTS

I would like to thank my parents Sitapathi Rao and Shyamala, my wife (Jyothirmai), my daughter (L N Harini), all my brothers and their family members, my uncle PVK Srinivasa Rao and aunt Rukmini, my in-laws J Satyanarayana and Janaki for their love and constant encouragement towards achieving the goal-completing thesis writing which would not have been otherwise possible.

I owe my indebtedness to my supervisors **Dr. Arun Agarwal**, Dean and Professor, School of Computer and Information Sciences, University of Hyderabad, and **Dr. C. Raghavendra Rao**, Professor, School of Computer and Information Sciences, University of Hyderabad, whose constant encouragement, involved participation in subject discussions, patient listening, insightful criticism, motivating remarks, meticulous care, supervision and thorough support from the preliminary to the concluding level of my research work that helped me to cognize the intricacies involved in the mathematical models of this thesis area and made this thesis possible.

It is an honor for me to thank **Prof Atul Negi** and **Prof S Bapi Raju** for serving on my review committee. I thank them for their precious time in reviewing my research progress, steering the work in a right direction and providing thoughtful comments and suggestions.

I express my sincere and heartfelt thanks to the Administrative Office personnel for their valuable and timely help at every stage of my research endeavor.

I must talk about a noble personalty who show eternal affection, caring nature and also, is capable of prescribing simple solutions to complex problems through professional counsellings and instilling confidence in research scholars during their difficult situations. She is none other than *Smt C Vijaya Garu,* better half of Prof C Raghavendra Rao.

I also thank Dr C V Raghava - Chairman CVRCE, Prof L C Siva Reddy - Head, Department of CSE and my mentor, Dr K V Chalapathi Rao Sir, Dr PSVS Sai

Prasad, Dr Ch S Sastry (IITH), Dr R Sita Ramaiah, and Dr K Hima Bindu for their constant support in many forms. I shall express my special thanks to Dr D. Sudhir Reddy, Scientist-E, ADRIN, who instilled confidence in me at the early stages of my research work, stood beside me during tough times and helped me to bring up the thesis into a meaningful form.

I would like to acknowledge the Study Leave for 13 months with full pay provided by the Management of CVR College of Engineering, Ibrahimpatan, Ranga Reddy(Dt), Telangana, which facilitated me to participate in discussions with my supervisors, to visit various institutions and to have interactions with research community.

I would like to express my sincere thanks to my colleagues and friends at CVRCE-B Vikranth, CVS Satyamurthy, B Ram Babu, V Dattatreya, C Ramesh, Ch Ram Mohan, L Venkateswarlu, Dr H N Lakshmi, Dr A Vani Vathsala, G Ravi Shankar Reddy, Dr Uma Reddy, and C Samuel. I, also express my special thanks to all my childhood friends Bh Lakshmi Kanth, Dr P Vijaya Bhaskar, A Srinivas, S Ravikanth, M Purna Chandra Rao, K Hari Babu, M Shaym Babai, Dr D Narayana, K Srinivas, Sarath Chandra, N Narasimha Rao, Mr Navaid Alam, Dr JVS Srinivas, my well wishers K Appa Rao Mastaru, Late K Kamala Mohana Sarma, and my seniors at UoH Dr N Vijaya Lakshmi, Dr N Naveen and Dr P Pavan Kumar for their moral and technical support to accomplish this thesis.

I shall acknowledge the support I received from all my students at CVRCE while delivering the duties.

I wish to apologize the people who really helped me in various forms but missing a special mention over here. ...

... Thank you all!

Raghava M

ABSTRACT

In the realm of digital images, understanding the nuances of manual approaches of *image preservation* and implementing them through a computer program is a fascinating and challenging problem. The process of recovering the original image from a damaged image such that the resultant filled in image provides a holistic and consistent information with necessary smoothness is termed as *inpainting*. It is quite common that a binary mask is utilized to specify the target portions (inpaint region) of the image that must be recovered by utilizing the known portions of the image (source region). The inpaint region and the source region together constitute the input image and are complement to each other. The inpainting is an inverse problem which is ill-posed and has attracted many researchers for addressing diversified applications through various *control of regularity* strategies aimed at increasing the fidelity of the solutions. The perceptions of different researchers towards this problem led to agglomeration of various domain specific methodologies. The diffusion-based and self-similarity based algorithms available in the literature explicitly model either the structures or textures but not the both. The concurrent and hybrid algorithms are good at solving small instances of the problem but, they tend to produce low fidelity results and turn out to be intractable, if the relative size of the inpiant region to that of the input image increases.

The current work addresses the identified challenges of inpainting by redefining the prior associated with the existing models and proposes new *control of regularity* models in two stages; i. cast the problem in statistical domain and ii. design a machine learning based agent. For a better visualization of the solution space, this thesis introduces a qualitative terminology referring to category, classes and their association. The primary focus of this thesis is to build the statistics for understanding the characteristics of the input image and to suggest various *control of regularity* measures in terms of *spatial anisotropy*, for achieving high fidelity solutions in case of large scale inpainting problems. Subsequently, it develops a machine learning based agent which selects the suitable control of regularity by

analyzing the nature of the content in the image and the semantics of the gaps present in the image. Overall, this thesis evolves a statistical methodology which pitches the inpainting problem into the context of *stationary and non-stationary* behavior of data present in the image and alleviates the texture-structure confusion.

The first contribution, being the simple and an elegant improvement to the Best-Exemplar framework, is *Granular Approach for Inpainting* (GAI). This model redefines the boundary between the inpaint region and source region, which offers a mechanism to reduce the size of search space by utilizing the granules generated through histograms. It also proposes a new priority term which harnesses the selection of target patch and an anisotropic self-similarity measure for selecting the Best-Exemplar from the reduced search space. The appealing feature of this contribution is, it tends to solve large scale text removal *inpainting category* problem by defining the patch with smallest size, in contrast to larger size patches required by the state of the art exemplar models. The data structure support for the implementation of this framework appears simple to realize, yet brings down the time complexity substantially.

The second contribution is to model the inpaint problem as a *Spatial Anisotropic Interpolation* (SAI) over a random field, which develops a metaphoric representation of the knowledge extracted from the image using kriging. SAI model building starts with an effective design space construction step through the proposed Systematic Sampling scheme. It solves the statistical convolution of polynomial regression with Gaussian regression of the error random field. In SAI, the model building is implemented in Design and Analysis of Computer Experiments (DACE) framework and assumes that the data present in the image exhibits a piecewise constant trend. The least square solution of the polynomial regression associated with trend model is regularized through a correlation prior which is mathematically defined as a basis function. In parameter space, the optimal initialization of parameter description of spatial anisotropy is realized by Maximum Likelihood principle. Finally, the inferences are achieved through developing a moving win-

dow based model to solve large scale text removal *inpainting category* problems. Experiments demonstrate that the SAI can handle both texture synthesis and structure propagation aspects concurrently through a unified approach.

The third contribution is to extend the SAI model to *Scalable SAI* (SSAI) for solving the object removal *inpainting category* problems. SAI model is unable to solve the large scale object removal problem due to the following reasons. Systematic Sampling of SAI does not preserve the interpolation feature, kriging is a poor candidate model for extrapolation and naive basis selection in model building. Thus, it is imperative to develop a sampling scheme which can retain the interpolation feature, while showing an apposite concern to the non-stationary features present in the image. SSAI is a natural extension of SAI in three aspects. It extends the Systematic Sampling to 'Not Only Symmetric Hierarchical Sampling' (NoSHS) which inherently exploits the anisotropic correlation feature of the design space. Next, it proposes Hurst classifier that assesses the trend model and selects the basis in an intelligent way, in contrast to the fixed basis utilization in SAI. The third and novel extension of the SSAI model lies in employing a proactive approach for determining the order of the basis, based on a feature selection algorithm, in the light of *Elastic Net* regularization of Gaussian regression error. The adaptivity features of SSAI increase the fidelity of the results that are demonstrated to be superior than hybrid approaches. SSAI model is able to handle large scale structure interpolation and texture synthesis seamlessly without modeling them explicitly.

Finally, this thesis evolves an intelligent *Inpaint Agent* which considers the morphological feature and stationarity property of the image and invokes an appropriate inpainting model proposed in this thesis. The overall disciplined attempt of this thesis is, to activate an effort for developing a *content dependent delivery agent* that implements three different *control of regularity* features that alleviate the requirement of texture and structure awareness to launch an appropriate inpainting model on the problem at hand. Also, the agent-based methodology removes the human intervention and offers a comprehensive solution with theoretical

underpinning for diversified *inpainting category* problems. The experiments conducted on the benchmark images produce visually pleasing results when compared to hybrid methods and judge the capabilities of the *Inpaint Agent* to be better, based on PSNR and SSIM measures. Future directions for the thesis are to incorporate a fuzzy classification system and to realize content aware local anisotropy models aimed at tunable coherence in the results by utilizing the higher order representations.

List of Figures

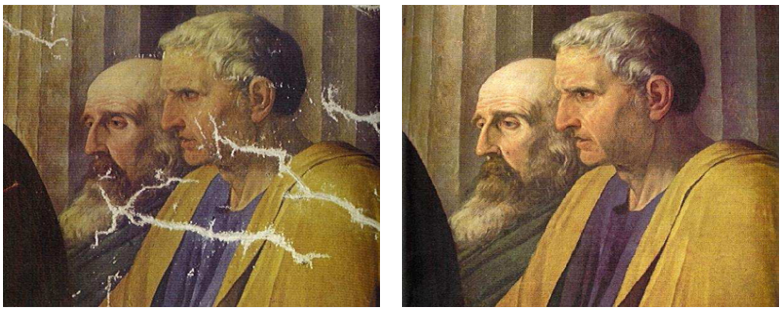
List of Tables

Contents

Chapter 1

Introduction

One of the data capturing and preservation techniques, which is good old one, is photography, taken either on paper or in digital form. Due to non-resistant nature of the photography paper over humidity, some portions of the image become sticky and start losing the informative character. These damages were patched up manually through touchups by the artists. However, the quality of the restored portion will not be the same as the other regions of the photograph. Inpainting is the process in which restoration of the digital image of this nature is performed through reconstructing the damaged portion. In the modern technology-driven world, the digital image is becoming more handy, in various walks of life: medical, fashion technology, modeling, gaming, etc. The identification of the lump of a tumor and treating the patient for cancer, plastic surgery, beautification and hairdressing also fall into inpainting domain. The present study is confined to the inpainting and its related applications on digital images. Figure 1.1(a) presents a deteriorated wall painting in which scratches are prominently visible. It represents a local scenario which requires only propagation of neighboring information. Figure 1.1(b) shows the effort made by an artist to restore it. This problem can also be generalized, and posed artificially, by designating certain objects present in the image as unwanted and are subjects of interest for removal. The real interest in this object removal is to reveal the information occluded by such unwanted features. In Figure 1.2(a), the person to the right, is considered as unwanted object and has



(a) (b)

Figure 1.1: Scratch Removal Example, 'Cornelia, A Mother of the Gracchi' by J. Suvee

(a) (b)

Figure 1.2: Object Removal Example, 'The Commissar vanishes' from D King

to be removed. Figure 1.2(b) presents the resultant image wherein the unwanted object is removed while preserving the unity of the content by reconstructing the wall and retouching the water surface.

Further, the image quality suffers due to the forward transformation and inverse operations involved while transmitting the image from one location to another over a transmission channel. The transform coefficients, Fourier and Wavelet coefficients for example, over the transmission line often get corrupted or missed altogether resulting in a noticeable negative result on the quality of the received image. The enhancement of such a distorted image, involves fine tuning the recovery process of missing coefficients based on the representation model of the image. This scenario, being in the transform domain, needs a robust mathematical framework. Figure 1.3(a) depicts the situation where the missing wavelet coefficients led to the distortion of the entire image. The recovery of the missing Wavelet

coefficients and the corresponding reconstructed image is presented Figure 1.3(b). The illustrations in Figure 1.1 and Figure 1.2 are, classically considered as crack

(a) (b)

Figure 1.3: Missing Wavelets Reconstruction, My improved implementation of [Chan et al., 2002]

and object removal problems from an image, respectively. These image manipulation techniques are classical instances of *image inpainting* and given the modern challenges faced by the computer vision community, inpainting stands at the heart of many inference drawing applications- erasing the scratches or date stamp from images, removal of trees form the satellite images in order to get the perfect picture of roads, finding stationary objects in a parking area, reconstruction of missing quotients, retargeting and shuffling of objects in the images are a few to mention.

1.1 Definition of Inpainting

The rediscovery of the missing information of an image, through a computer program is called 'image inpainting'. It is a nontrivial task as it involves accommodating certain implicit information which characterizes the complex features present in the vicinity of the damaged/unwanted region, into the process. Though this task is impossible to achieve completely, it will suffice to solve the less severe version problem, obtaining a restored image wherein the modifications are not discernible to any user who is not aware of the original image. Despite the existence of the problem in different forms, as previously demonstrated, they are considered to be similar problems but with varying scales. Figure 1.4 presents the inpaint

region as an irregular hole and its complement area as the source region.

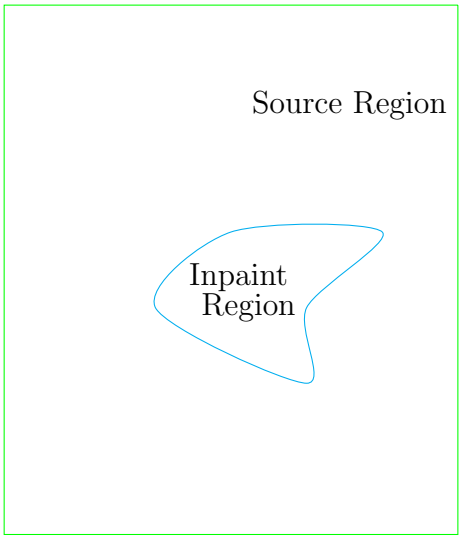


Figure 1.4: The Domain of Inpainting

The inpainting problem does not share many characteristics of the regular image processing tasks, like denoising, segmentation or restoration [Aubert and Kornprobst, 2006]. Figure 1.5, depicts the steps involved in solving a regular inpainting problem. It highlights the specification of the inpaint region by manual effort and encapsulating it into a binary mask, followed by application of some inpainting methodology to get the result. The real challenge lies in the selection of an appropriate inpainting model based on the nature of the image content and the morphological features of the gap inferred by the mask. The following sections

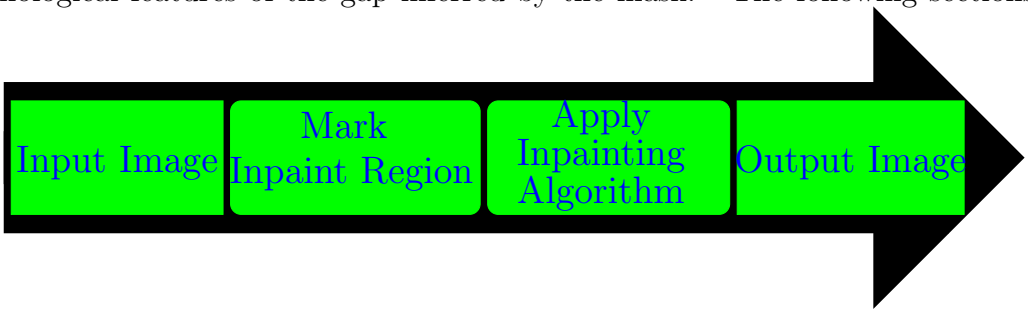


Figure 1.5: The inpainting procedure

provide the problem definition, organization, and highlights of the thesis.

1.2 Motivating Aspects and Problem Statement

The inpainting problem has been mostly addressed by different authors with perceptions to textures or structures by preconceiving various forms of control of regularity associated with the problem along with meta-data such as filling scratches and object removal. An arbitrary choice of these methods may result in inferior solutions. Moreover, the morphological features of real world images vary about scale and shape which require utilization of sophisticated algorithms to identify them. Hence, essentially it is another exciting and equally challenging problem to arbitrate the nature of the image. The present study focuses on developing a schematic hybrid methodology to address some of the identified issues associated with inpainting through pitching the problem into the context of stationary and non-stationary models. Statistical properties of the image, if properly modeled, are capable of increasing the fidelity of inference algorithms aimed at predicting the missing information by incorporating a new *control of regularity*. This observation is the driving factor and serves as the seed of contributions towards the present work.

Problem Statement: The primary goal of this thesis is to build the statistics for understanding the characteristics of the input image and to suggest various control of regularity measures in terms of spatial anisotropy. To achieve high fidelity solutions for large scale inpainting problems. To develop an intelligent agent which selects a suitable control of regularity to handle the non-stationary nature present in the real-world images. Finally, the contributions of this thesis led me to the development of ***Agent Based Statistical Approach for Image Inpainting with High Fidelity.***

1.3 Major Contributions

The present work makes an attempt to perceive the *inpainting classes* based on the analysis of the mask by extracting the statistical measures and semantics among

the gaps present in the mask. At the highest level, the statistical information accounts for the size of holes found in the mask and evolves two *inpaint classes*. The *first class* has only one hole and the second class has multiple holes. The *first class* can be further divided into two sub-classes, *small scale* and *large scale*, based on the size of the hole and the *second class* can be divided into three sub-classes based on the size distribution of holes as *all small*, *mixed* and *all large* classes. On the other hand, the meaningful semantics established among the holes suggest different *inpainting categories*- scratch removal, text removal, object removal, objects retargeting and objects shuffling. Table 1.1 associates the *inpainting class* with *inpainting category*. The given input image is analyzed

Table 1.1: The table showing the Inpainting categories and classes

Inpainting Category	Inpainting Class Morphological features of gaps in mask
Scratch Removal	Single and Small, Multiple and Small
Text Removal	Multiple and Small
Object Removal, Image Completion	Single and Large
Objects Retargeting	Multiple and Large
Objects Shuffling	Multiple and Large

to extract some inferential statistics that are utilized to choose an appropriate control of regularity methodology (also called as prior) to solve the inpainting problem at hand. The proposed control of regularity is abstracted in terms of spatial anisotropy at three levels; similarity measure, sampling and interpolation, which are elaborated in the upcoming chapters.

According to the literature, higher order Partial Differential Equations (PDE) and Total Variational (TV) based inpainting methods diffuse the information into the source region, available in the vicinity of inpaint region in a controlled and anisotropic manner to preserve structures. But, these methods suffer from the inherent slow convergence, over smoothing and short range propagation capability features. The interior point based Best-Exemplar methods fill the inpaint region by locating self-similar patches from other portions of the image and copy them into the inpaint region, while being capable of synthesizing, only the tex-

tures. Functionally, these algorithms address the ill-posed nature of the problem through well-defined control of regularity methods. While doing so, these algorithms perceive the image as a mixture of texture and structure components and could address only one aspect of them and fail to accommodate the other one. In the case of natural images, image characteristics are non-interceptive rendering automatic invocation of these methods difficult. For handling this clumsy situation, sophisticated methods are evolved, that jointly model them or separate texture from structures and process them independently. But these methods are highly intractable and could only fill-in small inpaint portions. The proposed agent-based model, perhaps for the first time in the domain of inpainting, solves the problem through a generic statistical model which inherently handles these aspects without making explicit attempts to capture them.

By nature, interior-point based Best-Exemplar approaches do not reveal any clue about exemplar patch and demand almost total exploration of the source region. This finding led to develop an intuitive algorithm GAI, as the first contribution to this thesis. GAI redefines the boundary of the inpaint region and exploits it to reduce the size of the search space by constructing granules from the image through statistical tools. As a novel contribution, GAI introduces a pseudo-similarity measure which is inherently structure-aware and involves less computational effort. The proposed model is employed on text removal and scratch removal problems. Those are typical inpaint category problems, in which the mask is highly cluttered with disconnected components that impair the existing methods to fail.

The literature is enriched with interpolation schemes for image representation based on Markov Random Fields and Radial Basis Functions. But there is no evidence in the literature about developing inpainting models on random fields that exploit spatial statistics. The proposed SAI model, as the second contribution, employs the Systematic Sampling scheme proposed in this work and operates the moving window based DACE approach for solving multiple regressions involved in the kriging. Inherently this model regularizes the least square solutions through

expressing the anisotropy in the form of correlation models. SAI could solve the Text Removal problem with higher fidelity than the GAI model, if the input image exhibits certain characteristics and possesses the desired statistical features.

The third contribution addresses the complicated situation induced by object removal from images by developing SSAI model which employs the *Not only Symmetric Hierarchical Sampling* strategy for retaining the interpolation feature and exploits the anisotropy property. Trend modeling is a critical issue in large-scale spatial interpolation. The sufficiencies of the polynomial basis selection from a well-defined collection of basis functions and finalizing its order turn out to be two different subjective issues. This thesis made the first aspect into a transparent feature by proposing Hurst classifier; a well-defined set of rules, based on the Hurst exponent extracted from the image. The other, an equally critical issue is determining the order of the polynomial. The proposed model addresses this issue as a considerable contribution which, even overcomes the stigma of Blind kriging through feature selection based on sparse representations induced by Elastic net regularization of Generalized Least Square solution of the polynomial regression.

Finally, this thesis investigates the possibilities of evolving **Inpaint Agent** which intelligently suggests, the invocation of appropriate inpainting methodology from the proposed algorithms. Inpaint Agent achieves this capability by considering two essential parameters- the morphological features of the mask and the assessed stationary property associated with the content of the image by the *Hurst classifier* without taking structure or texture aspects into consideration.

1.4 Organization of the Thesis

The remaining chapters of the thesis are organized as follows:

Chapter 2 provides the literature review of inpainting problem in detail. The first four sections offer the complete description of the two fundamental models related to structure propagation models. Subsequent two sections elaborate the texture synthesis models. The next two sections explain the latest developments

that make use of all basic models to realize both the aspects through domain decomposition and hybrid approaches that combine exemplar models with diffusions. The final two sections provide the introduction to interpolation models and quality metrics that are selected to assess the goodness of the results.

Chapter 3 presents the proposed *Granular Approach for Inpainting* (GAI). The first section elaborates the design issues of Best-Exemplar methods. The second section addresses the ambiguity in *target front* construction and target patch selection issues followed by granules-based search space reduction technique. The third part details the proposed similarity measure which accounts for the local anisotropy while selecting the Best-Exemplar patch. Fourth section deals with the implementation aspects of GAI. The fifth section presents the algorithm of GAI, followed by experimental results and summary.

Chapter 4 details about the *Spatial Anisotropic Interpolation* (SAI) model. The first two sections present an introduction to random fields and the characterization of anisotropy. The next section discusses the anisotropic interpolation followed by the *Systematic Sampling* and *spatial correlation* as a measure of *spatial anisotropy*. The subsequent section introduces the moving window scheme which models the trend with a constant basis and presents the detailed SAI algorithm. Final section presents the experimental results and summary.

Chapter 5 presents a *Scalable SAI* (SSAI) model for object removal problem. The first section demonstrates the low fidelity solutions offered by SAI while solving the object removal problem. The subsequent two sections highlight the issues of sampling followed by *Not only Symmetric Hierarchical Sampling* scheme to adapt the spatial anisotropy. Section four discusses various important design methodologies related to object removal problem with elastic net regularization as a effective prior to the spatial interpolation. Section five details the overall SSAI algorithm and the last section illustrates the capabilities of SSAI by presenting the results of the experiments conducted on a handful of images.

Chapter 6 presents the proposed *Inpaint Agent* (IA) framework which analyzes the morphological features of the mask and nature of the image content regarding

Hurst Exponent and prudently invokes an appropriate inpainting model from the collection of algorithms developed in Chapters 3, 4 and 5. This transparent nature of the agent methodology could make the framework as a decision system. Subsequently, the validation of the proposed framework is demonstrated by presenting the quality measures of the results acquired from the experiments conducted in previous three chapters.

Chapter 7 presents overall conclusions and future directions.

Chapter 2

Literature Review

This chapter presents the formal definition to Image Inpainting problem and offers an in-depth perspective on various existing models to solve it, that involve diversified spatial philosophies. These models include diffusion of information for solving an inverse problem, imputing the missing values by a synthesizing mechanism, extrapolating the missing data, representing the data in a sparse manner and some hybrid versions. The functional aspect of these models is to implement some tools to capture meaningful patterns from the images and utilize them to solve the inpainting problem. The common feature of these models is that they apply a variational philosophy, anisotropy in various forms.

2.1 Formal Definition

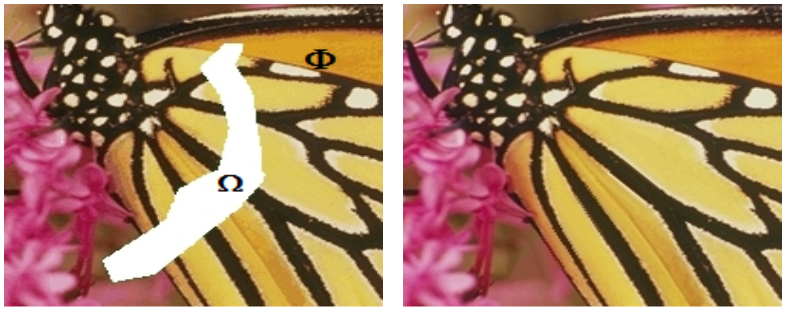
Achieving the rebuilding of missing information of damaged or hidden regions through an intelligent computer algorithm is the goal of *image inpainting*. However, this task is not trivial as the scale and the complexity of the problem increases exponentially with (i) the proportion of missing pixels to the size of the image (ii) the topological and large structural aspects of the missing region(s). Further, the assessment of the quality of a restored image is equally challenging when the original image is not available for comparison.

Let F be the image which is mathematically defined as $F : \wp \subset \Re^2 \rightarrow \Re^n$

where $m = 2$ and $n = 3$ for color images denoting the visible reflectance channels Red, Green and Blue or $n = 1$ for grayscale images. Let Ω and Φ be partitions of $\wp$, F_Ω and F_Φ be functions of the type F. Let I be the input image which is derived from F as a composition of (F_Ω, F_Φ) where Ω represents the set of pixels whose values need estimation- inpaint region and Φ represents the set of pixels whose information is available- the source region. The inpainting algorithm essentially makes use of a binary Mask M which is a characteristic function over $\wp$- with 1 representing the inpaint region location and 0 representing the source region pixel, passed as one of the arguments to it. Inpainting is a task intended to build a schematic process P for obtaining $\hat{F}$ as an estimate of F

$$\hat{F} = P(I, F_\Omega, F_\Phi)$$

therein the modifications are indiscernible to new observers, see Figure 2.1(a) and 2.1(b). The similarity of the $\hat{F}$ with F assesses the goodness of the process P. The objective of this exercise is to ensure that the resulting image is visually plausible and retains the unity or global uniformity of the entire image. But, the construction of P is an example of Inverse problem as it involves the determination of source image from the input image for which there is no unique solution. Hence, inpainting is an example of ill-posed problem and even the solution procedure is unstable [Hadamard, 1923]. The image processing community addresses the ill-posed feature of many problems through applying some form of **control of regularity** [Black et al., 1998]. In real world situations, an expert artist renovates a damaged image, manually either by continuing the surrounding information or copying the coherent portions that are present in the vicinity of the target region. Translation of this original art form into a computer program to process the digital images is a challenge for image processing community. Researchers from various domains e.g. computational fluid dynamics, mathematics, computer graphics, and signal processing have been attempting this problem by utilizing the domain-specific tools. The scale of the problem ranges from removal of small



(a) (b)

Figure 2.1: (a) Input image with inpaint region Ω and source region Φ (b) Result of inpainting

scratches to larger objects from images. It gains a tougher form by introducing man-made constraints as discussed later on. Adding another dimension, the nature of the image content significantly influences the data structure selection and the representation aspects. Throughout this work, the content of the input image is assumed to exhibit some stationary behavior, which is very rational in natural images. Alongside, there was another domain of study 'texture synthesis'- which produces large scale similar textures from the given smaller texel element for getting some artistic effects. Upcoming sections trace the evolution of the inpainting algorithms and the involved regularization aspects that address different physical phenomena associated with the image properties in chronological order.

2.2 Taxonomy of Image Inpainting

While the human is capable of interpreting the missing areas or occluded portions utilizing the 'global view' of the image, realizing this ability through a computer program is a challenge in front of the computer vision community. Over the past decade, this problem has been addressed by researchers through several techniques that encompass many representation models for the image features such as statistics about pixel values, structures, texture patterns and exhibition of certain spatial properties like self-similarity. Figure 2.2 gives the taxonomy and the complete spectrum of models developed by authors hailing from different

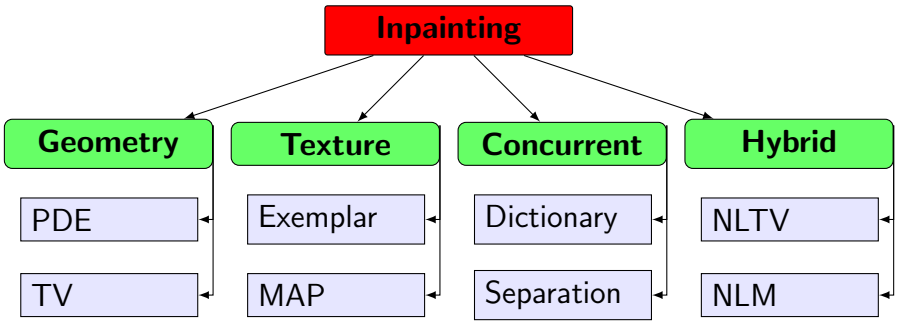


Figure 2.2: Taxonomy of image inpainting solution space

domains of science and engineering. The early solutions of inpainting attempt to solve the problem by adapting the image restoration algorithms [Nitzberg et al., 1993a] [Masnou and Morel, 1998]. Therein the inpainting problem is modeled based on the low-level image features, extracted from the image. Thus inpainting is one of the instances of inverse problems. The functional nature of any inverse problem is about inferring some features of the image which are either not available directly or lost due to the adopted initial representation model. For example, depth information is not encoded directly in 2-D images, but can be derived from multiple images. Hence, extracting such low-level features and using them effectively to solve a vision related problem turns out to be a challenging task. This aspect is a more complicated concern of inpainting problem as it requires sophisticated data structures or descriptors that encode even middle-level information. Perhaps, the descriptor selection serves only half of the purpose. The other and equally important aspect are to run an inference algorithm on the selected data structure. The fundamental nature of these inference algorithms is that they are very sensitive to the initial conditions and result in unstable numerical methods that give rise to multiple solutions. This feature turns the inpainting as an ill-posed problem [Hadamard, 1923]. In image processing domain, to deal with the ill-posed problems, standard practice is to add some extra information based on some image priors to the data model and solve the resulting mathematical model. This aspect lies at the heart of many engineering solutions under a generic title-regularization. For example, at the highest level, the assumption on which all

inpainting algorithms work is that the pixels inside the inpaint region share the same properties, be it statistical or geometrical, with the source region pixels.

2.3 Transport Based Methods

The entire solution space of the inpainting problem is, logically divided into two classes. The first class of the solutions make use of either local or non-local features and work in a greedy manner. The other class of algorithms offer global optimal solutions and aim at maintaining the unity of the image. The capabilities of these methods also vary, based on the scale of the problem and the nature of the content present in the image. The initial attempts address the inpaint problem by modeling it as a transport problem by drawing the principles from fluid dynamics. This model mimics the manual inpainting process followed by the professional artists while repairing the old wall paintings, therein the information available in the vicinity of the inpaint region is propagated into the inpaint region. If the content of the image is smooth, then the propagation aspect is similar to the diffusion of heat through a physical phenomenon by the same amount in all directions which is known as isotropic diffusion. The heat equation [Aubert and Kornprobst, 2006] can represent the diffusion process which is a Partial Differential Equation (PDE) of linear nature

$$\frac{\partial I(t,x)}{\partial t} = \Delta I(t,x) \geq 0; \quad x \in \Re^2 \qquad I(0,x) = F(x). \tag{2.1}$$

Equation 2.1 emphasizes that the Laplacian $\Delta I(t,x)$ which models the smoothness at the location x on the boundary $\partial\Omega$ of the inpaint region, shall be propagated along all directions. And the propagation phenomenon is parametrized as a continuous variable t. The Equation 2.1 is a second order linear partial differential equation. The functionality realized through this model is similar to a smoothing filter which diffuses the information along the lines of constant intensity [Bertalmio et al., 2001] that is modeled through isophotes. The isophotes are the directions

of constant intensity denoted as $\nabla^{\perp}$; normal to the direction of the local gradient vector ∇, available at p_x. The weakness of this gradient data structure is, it is local in nature and capable of speaking about only the direction of diffusion but not the magnitude. The isophotes work across the edges and hence distort them. This structure ignorance forms the second problem with heat equation based method.

2.3.1 Edge Preserving Capabilities

To address the first problem, Perona and Malik [Perona and Malik, 1990] extended the linear heat equation to a non-linear diffusion model by introducing a decreasing operator g that determines the amount of diffusion at a location x based on the square of the gradient value available as

$$\frac{\partial I(t,x)}{\partial t} = div(g((|\nabla I|)^2)\nabla I(t,x)) \geq 0; \qquad x \in \Re^2 \qquad I(0,x) = F(x). \tag{2.2}$$

The output of g is called the diffusion coefficient which is computed using one of the following two equations

$$g(\|\nabla I\|) = EXP(-\frac{|\nabla I|^2}{\alpha}) \tag{2.3}$$

or

$$g(|\nabla I|) = \frac{1}{1 + (\|\nabla I\|/\alpha)^2}. \tag{2.4}$$

The g annihilates to zero at edge locations which enable the model to preserve the edges. On the other hand, it assumes the value one at flat areas which mostly, mimics the behavior smoothing filter. In literature, such functions are referred to as edge stopping functions, and the associated structure-aware diffusion process is called anisotropic diffusion. It is worth noticing that the Equation 2.2 is a non-linear second order PDE involving local operator ∇. Authors in [Black et al., 1998] compared the similarity between these PDE regularizations with the robust

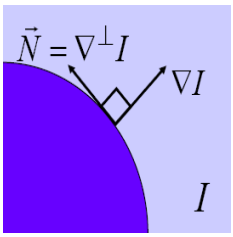


Figure 2.3: Gradient and Normal directions

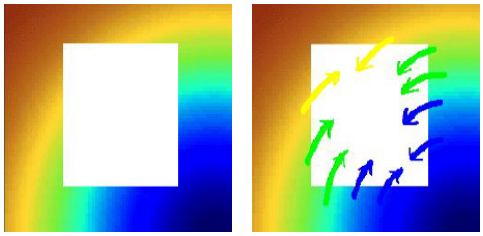
regression concepts of statistics. According to their observations the Laplacian, which is the divergence of the gradient-based regularization g in Equations 2.3 & 2.4, though generic, is inferior to solve large scale inpainting problems. The artifacts of this method are over smoothing and distortion of edges as the diffusion takes place across the edges. Bertalmio [Bertalmio et al., 2001] pioneered in this field by developing a non-linear PDE of order three, for controlled transportation of data into the inpaint region, with a title 'anisotropic diffusion'. Bertalmio empirically modeled the transport equation using a third order discrete PDE, which is analogous to the fluid dynamics based Navier-Stokes equation. The implementation relies on an exact forward time based finite differences scheme which diffuses different amounts of data along different directions. This phenomenon is referred as anisotropic diffusion and could prevent the edge distortions whose dynamics can be expressed as

$$\frac{\partial I}{\partial t} = \nabla(\Delta I)\nabla I^{\perp} \tag{2.5}$$

with the boundary condition

$$\nabla(\Delta I).\nabla I^{\perp} = 0. \tag{2.6}$$

The physical interpretation of this equation says, the Laplacian (smoothness) of image intensity (see Figure 2.3) available in the neighboring source region of the boundary ($\partial\Omega$) shall be transported into the inpaint region along the isophotes ($\nabla^{\perp} I$), see Figure 2.4. This transportation method simply prolongs the structures that hit the boundary of the inpaint region, inside the inpaint region to preserve the structures.



(a) (b)

Figure 2.4: (a) Input image (b) The direction of transportation along isophotes

Recent experiments proved that the superior quality results of inpainting are achievable by employing higher order PDEs. These, for instance, Cahn-Hillard equations-based methods [Baatz et al., 2008] of order four, are capable of interpreting the geometrical structures present in the image in a succinct manner and produce good results. But these methods are less preferred for interactive applications because of their high computational complexity and are best suited for filling in small patches only. The authors in [Chatterjee et al., 2009] solved the inpainting as an instance of extrapolation problem by working in time and frequency domains alternatively. Initially, the input image is treated as the band-limited version of the original signal to be reconstructed. Then the inpainting problem is solved by applying 'Popoulis-Gerchberg' theorem on the input image through 'alternate projections' method from spatial to frequency domains and the reverse of that in an iterative manner. Discrete Cosine Transform (DCT) is employed to transform the image into the frequency domain. This method can fill fine scratches through Harmonic extrapolation.

2.3.2 Utilization of higher order feature descriptors

The nature and quality of solutions to ill-posed problems like inpainting thoroughly depend on the mathematical operator employed to extract some latent information into a data structure and the nature of the algorithm which efficiently utilizes it. The methods discussed so far only make use of gradient so as to determine the direction in which the Laplacian of information, available in the vicinity of

the inpaint region, can be transported. But the quantum of diffusion cannot be spelled by the gradient operator and even the stopping criterion is not specified formally and hence the time parameter is incorporated.

2.3.2.1 Combinatorial Formulation

To overcome the drawbacks related to the direction of diffusion and the amount to the diffusion Grady [Grady and Schwartz, 2003] solved inpainting as an interpolation problem by mapping the diffusion of the Laplacian problem with Dirichlet boundary conditions onto a combinatorial Laplace equation. Therein they solved the combinatorial formulation of Dirichlet problem by exploiting the correspondence between the 'continuum differential' operators and the combinatorial differential operators. Initially, they launched the image onto a graph by employing the space-variant imaging technique and then specified the area to be removed from the image, manually. If the graph is uniformly weighted then the interpolation is isotropic; otherwise, it turns out to be an instance of anisotropic interpolation. In case of anisotropic interpolation the edge weights in graph are computed as the absolute difference of pixels intensities I_i and I_j at boundary locations i and j, as

$$w_{i,j} = EXP(-b|I_i - I_j|) \tag{2.7}$$

where b is a dimensionless parameter which controls the anisotropy induced by the pixel intensities. The authors [Grady and Schwartz, 2003] claim that in addition to achieving the steady state solution for stopping criterion, the sampling density can avoid the unwanted smoothing which is one of the prescriptions to the second ailment of diffusion based methods- the quantum of diffusion.

2.3.2.2 Tensor Fields

Authors in [Weickert, 1999] introduced structure tensor operator G at each point in the image, which is expressed in terms of ∇ and its transpose $\nabla^\top$, as

$$G = \nabla I \nabla I^T \tag{2.8}$$

which is a 2×2 matrix. The eigenvectors v_1 and v_2 of G determine the direction of the anisotropy and the corresponding eigenvalues λ_1 and λ_2 specify the quantum of diffusion. The authors [Benzarti and Amiri, 2013], [Jia and Tang, 2003] introduced the diffusion tensor as

$$D = f(\lambda_1, \lambda_2) v_1 {v_1}^\perp \tag{2.9}$$

where v_1 is the eigenvector along the major axis denoting the maximum change direction. Thus, they replaced the g function in Equation 2.2 with Equation 2.9 involving D, which is capable of steering the flow of structural information towards the privileged direction based on the content of the image. Overall, this method derives intermediate level information from the image field and effectively utilizes it for preserving the structures.

Note: It is easy to visualize that the transport based methods work across the edges and despite the deployment of anisotropic diffusion models, result in over smooth images and loss of structure information. Hence it is desirable to evolve a prudent diffusion operator that propagates the information along the edges which is possible by replacing the generic square measure of derivative with a linear order derivative.

2.3.3 A priori Knowledge and Fast Implementations

Before going into the linear order operators, we discuss the faster implementation of the diffusion kernels which are isotropic in nature. Oliveira [Oliveira et al.,

2001] developed two kernels, that could inpaint the finer regions. Their algorithm repeatedly convolves the inpaint region with the two kernels which is equivalent to replacing the pixel value with the weighted average of the neighboring pixels. Figure 2.5 shows the two isotropic filters. Though the proposed convolution

$$\begin{array}{ccc} \ulcorner a & b & a \urcorner \\ b & 0 & b \\ \llcorner a & b & a \lrcorner \end{array} \qquad \begin{array}{ccc} \ulcorner c & c & c \urcorner \\ c & 0 & c \\ \llcorner c & c & c \lrcorner \end{array}$$

(a) (b)

Figure 2.5: (a) First kernel with a=0.073235 and b= 0.176765 (b) Second kernel with c=0.125

mechanism is fast, it suffers from limitations with regard to the size of the inpaint region, poor texture generation capability and hence cannot be used to infill natural images with larger holes.

2.4 Diffusion Based Approaches

Another set of models represents the inpaint problem as an 'energy' function in 'variational' framework to solve an inverse problem. This model fundamentally minimizes the unbounded energy function which makes it ill-posed. A standard way out of this issue is to include some *aprior* information into the model along with the fidelity term under the regularization title [Black et al., 1998]. The quality of the inverse problem solution varies with the regularization term [Black et al., 1998]. The diffusion-based methods generally work based on Bayesian principle [Rudin et al., 1992]. These methods include regularization term as prior into the data model and expresses the energy term as

$$E(I) = \lambda R(I) + \frac{1}{2}\|I - F\|^2. \tag{2.10}$$

The first term R in the above equation acts as *prior* (regularization term) and coins the associated problem into Bayesian model and λ is the scaling factor which

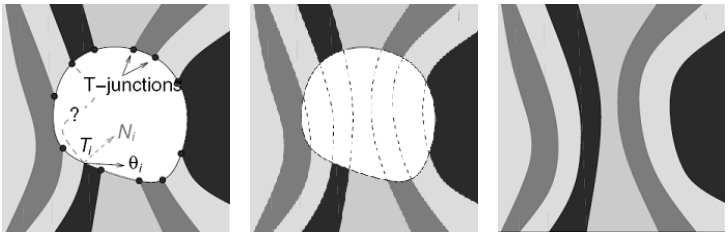
must be tuned to get better results. For example, the denoising problem is solved by solving the following convex problem

$$min \quad \frac{1}{2}\|I - F\|^2 + \lambda R(I). \tag{2.11}$$

The nature of the regularization term determines the quality of the solution and literature is enriched with many regularization terms. One of the *prior* terms is the *Total Variation* (TV) model introduced by Rudin, Osher and Fatemi [Rudin et al., 1992]. In case of discrete and smooth 2-D image I, the TV is defined as

$$\|I\| = \sum_x \sum_y |\nabla I_{x,y}| \tag{2.12}$$

where ∇ is the discrete gradient operator along x-axis and y-axis and is calculated as $\nabla I = (\partial_x I, \partial_y I)^\top$. These discrete derivatives are calculated in terms of either forward or backward difference method along the corresponding direction. Literature is enriched with lot of numerical algorithms to solve this TV minimization problem [Chambolle, 2004] [Wang et al., 2008]. The minimization effort of the above Equation 2.12 leads to a higher order PDE. In case of image segmentation Mumford [Nitzberg et al., 1993b] modeled the image as a piecewise-smooth function. Then the PDE evolves a functional, which penalizes the three functional elements- the missing smoothness within sub-regions, the length of the boundary between the sub regions and the deviation between the model and the input image. This approach was initially extended by Masnou and Morel [Masnou and Morel, 1998] to solve the disocclusion problem. The authors therein extended the idea of using T-junctions for solving disocclusion problem by continuing the level lines into the target region and joining them with the corresponding compatible level lines at the other end as shown in the Figure 2.6. We characterize two T-junctions as compatible if they belong to the same level set and the their ∇I be the same at both points. Then the regions enclosed by the edges are filled. The native assumption behind the success of this idea is that image is composed of many segments and each segment internally is smooth. This methodology closely



(a) (b) (c)

Figure 2.6: (a) T-junctions on boundary (b) Level curves continued (c) Disoccluded image

mimics human visual capability to connect the partially visible curves by extending them by utilizing the curvature details of the level lines. The mathematical model of this method involves finding the curve γ which connects two T-junctions on the level set and is characterized by θ_1, N_1 and θ_2, N_2 where θ stands for the normalized gradient and N refers to the normal at the T-junction, with minimum length as

$$\int_{\gamma} (\alpha + \beta\kappa^p) dH + \langle \theta_1, N_1 \rangle + \langle \theta_2, N_2 \rangle. \tag{2.13}$$

The second and third terms, enclosed in the pair of angular brackets denote the angle between the two arguments enclosed, and H is generalized functional space aimed at measuring the length of regular curves regarding Hausdorff measure and κ denotes the curvature. The disocclusion model proposed by Morel [Masnou and Morel, 1998] finds such minimum length curves for all level curves to connect the T-junctions first. Subsequently, the contours are filled with appropriate gray values. The existence of the solution for the model is available in [Aubert and Kornprobst, 2006]. The algorithmic feature of level lines is that they exhibit spatial causality, and hence they fit into dynamic programming model.

2.4.1 TV Minimization Based

Let us now, define the variational model formally. The image I is a function over the domain $\Re^2$. If I has partial derivatives at each location in the image domain with finite total variation, then we say I is a function with bounded variations

denoted as $BV(S)$. The total variation is represented, as an *energy* functional

$$E(I) = \int_I |\nabla I| dxdy. \tag{2.14}$$

The numerical solution for solving the minimization of this energy equation using steepest descent results in the following iterative steps

$$I_{t+1} = I_t + step \cdot \frac{\partial I}{\partial t} \tag{2.15}$$

$$\frac{\partial I}{\partial t} = \nabla \cdot (\frac{\nabla I}{|\nabla|}). \tag{2.16}$$

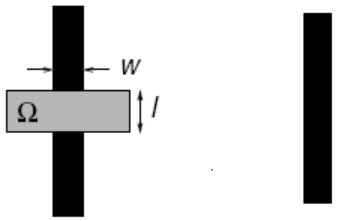
This solution diffuses the information along the normals at the boundary of the inpaint region but the diffusion strength only depends on the strength of the level lines. Hence they work across the isophotes and preserve the edges. The physical interpretation of the energy model is to minimize the length of the level lines. We can see the resemblance between TV model and the Bayesian model, but the former deal only with deterministic variables as follows

$$I_t = \nabla \cdot (\frac{\nabla I}{|\nabla I|}) + \lambda(F - I) \tag{2.17}$$

where the first term on the right hand side, stands for the prior and the second term represents the data term. Thus the optimal solution for this problem involving the penalization of length term, results in a straight line inside the inpaint region and hence can connect, only short and linear structures as shown in the Figure 2.7 and

$$\arg\min_{\gamma_\lambda} \int length(\gamma_\lambda) d\lambda \tag{2.18}$$

where $\gamma_\lambda = \{x \in \Omega / I(x) = \lambda\}$ is the level line for the gray value λ. In other words, the anisotropic diffusion of the total variation model diffuses with conductivity coefficient $\frac{1}{|\nabla|}$ which do not consider the geometry of the level line. Hence, the diffusion is done only along the linear interpolated lines, and the direction



(a) (b)

Figure 2.7: (a) Input image (b) TV inpainted image

of diffusion is not modeled correctly causing it to limit to give good results for narrow holes only. The chronological developments in variational models [Chan et al., 2002] range from simple Total Variation (TV) to Curvature Driven Diffusion (CDD) of information into the inpaint region. These models also come under the local category as they don't involve global features such as large scale textures. To overcome this connectivity problem Chan and Shen [Chan et al., 2002] proposed Curvature Driven Diffusion to introduce the curvature term- a second order derivative, into the model as follows

$$\frac{\partial I}{\partial t} = \nabla.(\frac{g(\kappa)\nabla I}{|\nabla|}). \tag{2.19}$$

The g in the above equation is a non-decreasing function with κ referring to the curvature. In the CDD model the conductivity coefficient is therefore changed to $\frac{g(\kappa)\nabla I}{|\nabla|}$, where g annihilates large curvatures and stabilizes small ones. The underlying interest to introduce curvature in the initial TV model is to increase the order of the associated Euler-Lagrange equation from second order to third order. Overall, the operational feature of variational inpainting algorithms is, they treat the image as a collection of regions, separated by sharp edges and the content inside each region is smooth. The solutions of the minimization of variational functional results in a higher order PDE.

2.4.2 Euler's Elastica Model

The original idea behind Morel's [Masnou and Morel, 1998] approach to solving the disocclusion problem that arises if one objects obscure another and is an extension of Elastica curve model proposed by Euler [Chan et al., 2002]. The automatic selection of T-junctions on the level curve is also possible. The functional behavior of Morel's model is expressed, along level curves as follows

$$\epsilon(C) = \int_c (u + v\kappa^2) ds \tag{2.20}$$

where C is the curve along which the integration is carried out on the curve segment ds and the curvature κ is expressed as the Euclidean distance. The underlying mathematical principle of this minimization is to estimate the geometrical walk between two T-junctions in Maximum Likelihood sense. The step length is a random variable drawn from i.i.d. and determines the rotation at each step point. Tony Chan along with his team members identified the functional difference between the Betarlmio's transport phenomenon and the Morel's diffusion based TV minimization mechanism. Therein authors observed, the transportation takes place along the isophotes which are perpendicular to the gradient direction and diffusion is carried out along the normal which is perpendicular to the gradient. This behavior endows the TV model to have edge preserving capability when compared to PDE model. Hence authors in [Chan et al., 2002] combined the two models into a single model and called it as Euler Elastica model. Their idea was to reformulate the Morel [Masnou and Morel, 1998] idea to work with surface integral directly instead of finding individual level curves, by dealing with Coarea that gave rise to the functional

$$E(I, \lambda) = \int_\Omega (\alpha + \beta\kappa^2) |\nabla I| dx \; + \; \frac{k}{2} \int_\Phi (F - I)^2. \tag{2.21}$$

The first term of in the Equation 2.21 is the regularization term and the second one is the data term. The minimization attempt of the associated Euler-Lagrange

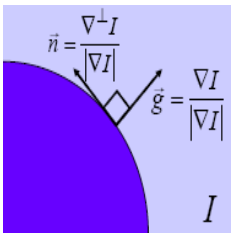
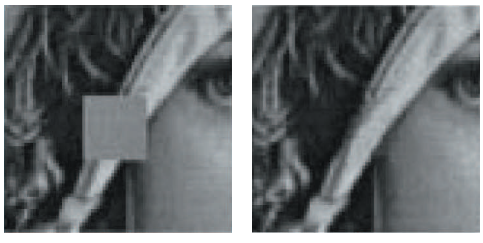


Figure 2.8: Normalized Gradient and Normal directions

equation of the functional results in a 4^{th} order PDE

$$\frac{\partial I}{\partial t} = \nabla(\alpha + \beta\kappa^2)g - \frac{2\beta}{|\nabla I|}\frac{\partial(\kappa|\nabla I|)}{\partial n}n \tag{2.22}$$

where n and g are normalized Normal and Gradient operators as shown in the Figure 2.8. In this formulation, the regularization term is a random walk which generates the smooth curves. Inherently, this model possesses two diffusion processes that are orthogonal to each other. The first one works along the isophotes i.e. across the edges, to penalize the length of the level line which mimics the Bertalmio's [Bertalmio et al., 2001] model and the other one along the Normals i.e. along the edges, to regularize the curvature which is the component of TV model by Chan [Chan et al., 2002]. These two processes penalize the diffusion process based on the ratio $\frac{\alpha}{\beta}$. Inherently, this model accommodates both the transport and diffusion processes in it and works well even to fill slightly larger gaps as shown in Figure 2.9. The solution of the Equation 2.21 has two terms



(a) (b)

Figure 2.9: (a) Input image (b) Result of Euler Elastica Model

out of which one involves the L_1 norm and the other one involves the L_2 norm which results in unstable numerical solutions. To solve this problem Goldstein

and Osher [Goldstein and Osher, 2009] proposed Bregman's Methods which are applicable for image inpainting problem also. This aspect is, further elaborated in Chapter 4 (Section 4.2.3).

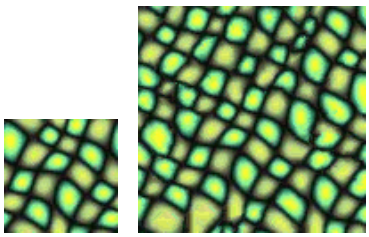
2.4.3 Limitations

The two models which are diffusion [Chan et al., 2002] and transport [Bertalmio et al., 2001] based phenomena, discussed earlier, involve PDE of higher orders which are either formulated directly [Bertalmio et al., 2001] or derived by solving an energy functional [Chan et al., 2002]. The former model involves the L_2 norm of the gradient operator and the latter involves only L_1 norm. It is a well-known fact that the TV based filter is capable of preserving the edges in the domain of image denoising and restoration. But, these models determine the pixel values in the inpaint domain only by using the information which is available in the vicinity of the inpaint region and capable of extending the structures smoothly. Overall, these methods come under the structure-aware category that solves the scratch removal *inpaint category* problems with a single or multiple smaller holes *inpaint class*. But, in some real world examples, the content away from the inpaint region Ω may resemble the inpaint region rendering the local methods to fail. In addition to this, geometry-based methods assure only the structure propagation to form contours and filling these contours to with smooth details. But they do not account for preserving the finer details- texture, inside each contour. These problems are addressed by different authors as discussed below.

2.5 Texture Synthesis Models

The second family of inpainting methods targets the texture synthesis. The texture is defined either as a regular and globally repeated pattern. For example, the tiles on the floor, the arrangement of bricks in a wall or non-regular like pile of pebbles with some global statistics. Efors and Leung [Efros and Leung, 1999] developed a 'locate and copy' strategy which exploits the self-similarity principle. The intuition

in this principle is that many natural images possess smaller repeated structures. If any algorithm is capable of reproducing these structures for many a time and arranges them in a systematic order, then larger textures can be synthesized from the given sample texture, see Figure 2.10. The operational feature of this algorithm

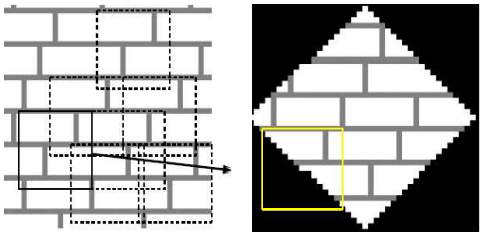


(a) (b)

Figure 2.10: (a) Sample Texture (b) Synthesized larger Texture

is as follows. Consider a point p on the boundary of the input image I and construct a patch Ψ_p of size $n \times n$ with p at center. Observe in Figure 2.11(b) that the patch is incomplete in the sense that it has some pixels with missing values. Then all possible adjacent overlapping patches $\Psi_q$of size $n \times n$ are extracted from the image and compared with Ψ_p. Based on some measure of similarity, the patch with maximum similarity referred to as Exemplar patch, $\Psi_{\tilde{p}}$ is located and used to assign values to the corresponding missing values in Ψ_p. Overall, the locate and copy based procedure is referred to as Exemplar approach. The coherence feature among the synthesized patches is possible by confining the subsequent searches for Exemplars to the top and left overlapping blocks of the previous Exemplars. The repeated application of this process results in bigger with consistent texture characteristics. The Figure 2.11(a), also shows multiple Exemplars highlighted in the rectangular box. In Figure 2.11(b) Ψ_p with M has finite informative pixels and remaining are unknown pixels.

Now locating the $M \times M$ Exemplar patch from the given image is a complex problem for larger images. For instance, if the size of the image is $N \times N$ then time complexity for locating the Exemplar patch $\Psi_{\tilde{q}}$ through scanning the entire image using a row-major scanning algorithm is $O(M^2)$. If the distance of between Ψ_p and Ψ_q is computed using Sum of the Squared Differences (SSD) on informative



(a) (b)

Figure 2.11: (a) Highlights multiple exemplars in square of dotted lines (b) Shows target patch in a square in solid Yellow lines

pixels then Figure 2.11(a) shows multiple patches with same SSD value. Thus the algorithm has to resolve the selection of Exemplar from multiple exemplars. Clearly, this is a replication of textures process rather than real synthesis. Successive application of this localized extension procedure on various pixels lying on the boundary could synthesize the new image of a larger size having repeated versions of consistent textures.

2.5.1 Best-Exemplar based approaches

Figure 2.12(a), shows that topology of the Ω may be irregular. The boundary points of the inpaint region Ω referred to as fill front ($\partial\Omega$) and they separate Φ from Ω. Now adopting the texture synthesis approach developed by Efros [Efros and Leung, 1999] to solve inpainting problem directly end up with a lot of artifacts [Criminisi et al., 2004]. There are related problems, due to the greedy nature of Efors [Efros and Leung, 1999] algorithm, such as either exploring the wrong part of the search space and piling up the texture garbage through recycling the verbatim copies of a texture sample. To handle this uncontrolled behavior of Exemplar approach researchers proposed few improvements. The crucial component of them is about determining the order in which the inpaint region patches are replaced. As a naive solution, the order can be determined by evolving an onion-peel procedure therein inpaint region is divided into concentric layers and are filled layer by layer. These layers can be constructed by utilizing the fill front of the boundary of the inpaint region and the pixels on it can be selected to for replacement. This heuristic

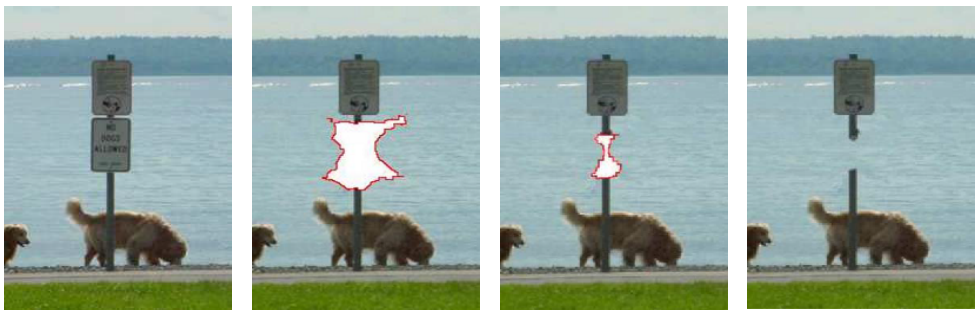
procedure resulted in poor results and suffered from texture garbage problem as shown in Figure 2.12(c) &(d). Another problem with the onion-peeling approach is

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

Figure 2.12: (a) Input Bungee (b) The Mask (c) Result of first run (d) Result of the second run

the random selection of target patch on the boundary also collapses the structures that approach the inpaint region, see Figure 2.13(d). The failure of the method is attributed to the uncertainty associated with the selection of a target patch from the fill front. Perhaps, this is one of many such issues based on which different variants of the Exemplar methods are developed for texture synthesis. These methods are broadly classified into two groups.

- Greedy or Local
- Energy based or Global Optimal



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

Figure 2.13: (a) Input Image (b) - (c) Partially filled inpaint region (d) Results with broken pole structure

It is trivial to observe that, these exemplar-based methods are, non-local in nature, but work in a greedy manner i.e. they inpaint the hole in a single pass. Though this model seems to be simple, it is necessarily required to answer many design issues that otherwise end up with texture garbage that have a significant bearing on the quality of the result. Even there is a proliferation of variants based on different design issues, the primary feature of these methods aim at large-scale texture synthesis but fail to propagate the structures inside the inpaint region [Buyssens et al., 2015]. The development of robust Exemplar algorithms [Criminisi et al., 2004] [Wexler et al., 2007] [Komodakis, 2006] are discussed in the next sub-sections.

2.5.2 Greedy Methods

The Greedy Exemplar methods are implemented by either using fragment or patch approaches. The difference between the fragment and patch can be best explained in terms of the shape of the probing element. A fragment refers to a circular probing element whereas the patch is a square. Drori implemented a fragment-based iterative algorithm for image completion [Drori et al., 2003]. This algorithm initially builds a confidence map which is based on the colors available in the source region around the boundary pixels. This map is used to select the target pixel with the largest confidence value. Then, the algorithm iteratively infers color of unknown pixels from visible parts of the image. To get consistent results, this was implemented in a multiscale model. This algorithm is capable of preserving both the structures and the textures, but the computation time runs into hours [Drori et al., 2003]. Criminisi [Criminisi et al., 2004] clearly groups the pixels into source and inpaint regions and demarks the boundary $\partial\Omega$, see Figure 2.14. Notice that the pixels lying on the boundary are considered as inpaint region pixels. If the boundary of inpaint region assumes convex shape and smaller in size then the quality of results from this algorithm would be superior. The behavior of the greedy method is as follows. For every pixel p on the boundary, a patch Ψ_p of size $n \times n$ having p at the center is constructed. Such patches have at least one and

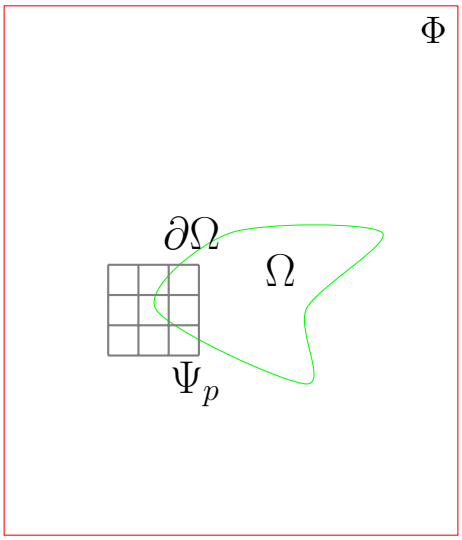


Figure 2.14: Fill front and a target patch on it

at most 8 source region pixels. A priority term Pr is assigned to every boundary pixel, by utilizing the intensity values of neighboring source region pixels spanned by the patch. The priority is computed by deriving two terms confidence C(p) and data term DT(p) as follows

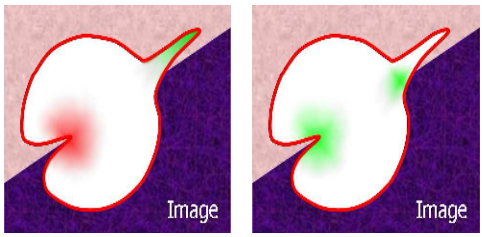
$$\mathrm{C}(p) = \frac{\sum_{p \in \Psi_p \bigcap \Omega} I_p}{Size(\Psi_p)} \quad \text{and} \quad \mathrm{DT}\ (p) = \frac{|\nabla {I_p}^{\perp} \cdot n_p|}{\gamma}. \tag{2.23}$$

Then the Priority term is computed as

$$\mathrm{Pr}(p) = \mathrm{C}(p) * \mathrm{DT}(p) \tag{2.24}$$

in which C encodes the confidence, in terms of the relative number known pixels present in the $\Phi(p)$ with respect to the size of the patch. This term favors the patches lying on the outward projecting features of inpaint region as they contain more number of source region pixels, see Figure 2.15(a). The DT term determines the strength of the data term present in the patch where γ is a normalizing factor whose value can be 255 for instance. If the data term is relatively smaller then the underlying area spanned by the patch is a smooth area. Otherwise a structural feature is expected in that patch. Hence the data term favors inward projection

portions of inpaint region, see Figure 2.15(b). Initially, the C term for each pixel



(a) (b)

Figure 2.15: (a) Confidence term (b) Data term

present in the image is fixed as - C = 0 for inpaint region pixels and C = 1 for the source region pixels. The priority term Pr, strikes the balance between the C and the DT terms. This priority term for a fill front pixel is not fixed and updated for the boundary pixels after each iteration as discussed later on. The inpaint algorithm selects the patch with highest priority as the target patch Ψ_p and invokes the Best-Exemplar method which searches for the most similar patch among the all candidate patches from Φ, while utilizing only the partial known information in Ψ_p. The similarity is measured between the known pixels in Ψ_p and the corresponding pixels in the candidate patch Ψ_q in various ways, for example in terms of a Sum of Squared Differences (SSD), Bhattacharya Distance [Bugeau et al., 2010] and normalized cross-correlation. Formally the Best-Exemplar is selected based on the following SSD measure

$$\Psi_{\tilde{p}} = \arg\min_{\Psi_q \in \Phi} \quad d(\Psi_p, \Psi_q).$$

The patch $\Psi_{\tilde{p}} \in \Phi$ is the most similar to the Ψ_p available in source region Φ. Basic nature of SSD based measure is, it works well for smooth areas. Then the target patch pixels are replaced with corresponding pixels of the Best-Exemplar patch while resulting in erosion of inpaint region. Then both the source region and inpaint region are updated and the confidence terms for the new fill front pixels are calculated. This step requires only to update the pixels abutting the Ψ_p. This procedure is executed in a iterative manner and gradually fills the entire Ω.

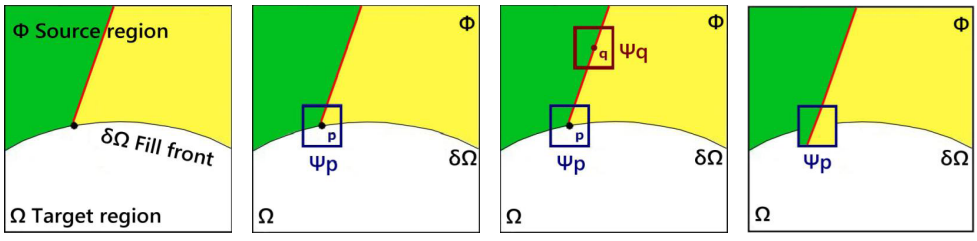
Algorithm 1 Best-Exemplar

Input: Input Image I and the associated Mask M

Output: Inpainted Image

1. Repeat steps 2 to 7
2. Extract the fill front $\partial\Omega$
3. Compute the priority for the fill front pixels and select the target patch Ψ_p
4. Select the Best-Exemplar $\Psi_{\tilde{p}}$ with minimum $d(\Psi_p, \Psi_q)$
5. Copy the Ψ_q into Ψ_p at locations r, $\forall r \in \Phi_p \cap \Omega$
6. Mark Ψ_p as inpainted and update Ω and Φ
7. Update $C(r)$, $\forall r \in \Phi_p \cap \Omega$
8. Until Ω becomes NULL set

Though this Best-Exemplar algorithm addresses two subjective issues of the Exemplar model and offers convincing results for inpainting problem, it has few limitations. To highlight a few of them, it can propagate only linear structures, see Figure 2.16, and smaller textures but not curves. Also, it is computationally expensive as it is required to search the entire Φ in each iteration. And, it considers only one Best-Exemplar and replaces target patch without ensuring the consistency with the neighboring patches. Literature is, enriched with



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

Figure 2.16: (a) The boundary (b) The target patch (c) The best Exemplar (d) Structure propagation after copying the best Exemplar into the target region

plenty of improvements to the Criminisi framework. The first one to be noted is Ashikhmin's [Ashikhmin, 2001] idea to restrict the search space, in step 3 of the Best-Exemplar algorithm to one of the neighbors of the best Exemplar located in the previous iteration. This idea improves the coherence between the pixels that are synthesized. The step 3 in Criminisi algorithm searches for the best Exemplar in each iteration in an exhaustive manner. There are many efficient data structures like *kd*-tree, *vp*-tree [Kumar et al., 2008] developed in literature that accelerate the nearest neighbor search operation that works well for fixed size patches Ψ wherein each pixel account for a dimension of the node in the tree. But in the case of inpainting, the number of known pixels varies from patch to patch. This aspect spoils the performance of these data structures and hence cannot be applied directly to optimize the Exemplar search operation and hence not discussed in this work. Averbuch [Averbuch et al., 2006] addressed this aspect by introducing a search structure which facilitates the Best-Exemplar search step. Their inpainting algorithm operates in two phases namely learning phase and synthesis phase. During learning phase, which is done offline even before the inpainting algorithm starts, the search structure $A[d, c]$ is constructed based on the patch size $c = m \star n$ and the dynamic range d of the pixel values. Learning step involves flattening of each patch from the source region and making entries into the search table. Suppose, Φ_l refers to the patch with size 3×3 whose index is l i.e. $c = 9$ then a raster scan of the pixels generates a flat record R_l with pixel values. The record number l is entered $A[d, R_l(i)] = l$ with $1 \leq i \leq c$. Before that, each patch is given an index number. The search table has multiple columns; $m \star n$ in the case of gray scale image and $3 * n * n$ for color images, and the number of rows equal to the range of pixel values. The flattening of a source region patch involves raster scanning the pixels in the patch and entering the index number of the patch in the corresponding cell of the search table. Once this data structure is ready then to inpaint a patch drawn from the fill front and the search table is queried to select the Best-Matching patch. For measuring the similarity L_∞ norm is utilized, which returns the maximum of pixel-wise differences between the two

patches. This data structure harnesses the Exemplar search operation which is essential in every step of the Criminisi algorithm and improves the time complexity by order two. Another bunch of works concentrated on actually determining the priority of the boundary pixels which is at the heart of the Exemplar approach. Being a greedy method, a small tinkering in the priority term leads to substantial improvement in the end results of this algorithm and also ensures that the texture is synthesized while preserving the structure. Keeping this point in mind, authors Xu [Xu and Sun, 2010] introduced a sparsity-based Data term for computing the priority. In Criminisi model the Data term involves computation of gradient and accounts for the structures. The authors in [Xu and Sun, 2010] model the Data term which ensures that the patches lying on the edges will be processed first by exploiting a fact that edges are rare phenomena than textures in images. Therein, the underlying idea is that the patch on the fill front which is most dissimilar with other patches in the source region is deemed to present on some structure. On the other hand, the most similar patches possess textural information. This aspect is expressed as

$$D(p_i) = \|w_{p_i}\|_2 \sqrt{\frac{N_s}{N}} \tag{2.25}$$

where N_s refers to the number of source region patches and N is the total number of patches. A large value of D infers that the patch is lying on some structure. This Data term pronounces the structural aspect more profoundly through sparsity than that of Criminisi model and leads to the improved results. The authors in [Le Meur et al., 2011] introduced the tensor-based Data term as

$$D(p_i) = \alpha + (1-\alpha) EXP(\frac{\eta}{(\lambda_1 - \lambda_2)^2}) \tag{2.26}$$

where λ_1, λ_2, are eigenvalues, as discussed in Equation 2.9, of structure tensor and η and α are hyper-parameters that can be fixed manually. The Figure 2.17, from [Buyssens et al., 2015] gives the effect of different Data terms on Bungee experiment and spells the ill-posed nature of the inpainting problem.

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

Figure 2.17: Comparison of different Data terms (a) The Bungee (b) Results of Data term in [Criminisi et al., 2004] (c) Data term in [Xu and Sun, 2010] (d) Data Term in [Le Meur et al., 2011]

2.5.3 Global Optimization

The methods so far discussed work based on either local information propagation or greedy in nature or follow the locate and copy approaches. These methods strive to achieve the right definitions for the image but may fail to ensure the global consistency of the overall results in the case of natural images. Levin [Levin et al., 2003] made an attempt to get the grasp of the global statistics of the image and included the same in the process of inpainting. Therein the proposed idea is to capture the look of the image through a probability distribution of secondary data extracted from it. The authors demonstrated, through histograms, that some two images with different structural features possess similar statistical properties. In contrast, marginal statistical properties of the extracted features, namely gradient magnitude and pairwise gradient angle are presented in the form of the histogram that started capturing different looks. This observation led to remodeling the inpaint problem with a constraint that the predicted marginal statistical feature must agree with the *empirical marginal* features available in data. The model starts with expressing the probability of an image I as

$$P(I; V_k) = \frac{1}{Z} EXP(\sum_k \sum_{x,y} V_k(f_k(x, y))) \tag{2.27}$$

where $f_k(x, y)$denotes the extracted feature k at location (x, y) and V_k is the data structure that holds such a feature. This way of defining features makes the representation of the probability function equivalent to Markov Random Field (MRF) [Li, 2009]

$$P(g(I)) = \frac{1}{Z} \prod_i \psi_i(g_i, g_i) \prod_{i,j} \psi_{(i,j)}(g_i, g_j). \tag{2.28}$$

The gradient g stands for the random variable at each location of MRF. The potential function ψ represents the node potential, if $i = j$ which is approximated by empirical marginal probability $\tilde{p}$. The term $\psi_{ij}(g_i, g_j)$ represents pairwise potential which measures the pairwise coherence on edges (the pairwise clique) $\langle i, j \rangle$ as $\tilde{p}(\theta_{i,j})$. Given this potential function the goal is to fill the inpaint region with the maximum marginal probability of the gradient at each (x, y) to generate the complete gradient field. This equation is solved by discretization of the 'gradient random' field and applying max-product Belief Propagation numerical algorithm [Li, 2009]. Finally, the integration step is applied on the gradient random field to get the overall result. This algorithm ensures to maintain the global look of the image. The patch based Exemplar methods make good use of non-local information during inpainting process. But the replacement policy of target patch with Best-Exemplar is accomplished in a greedy manner without showing any regard to the coherence with the neighboring patches resulting in a poor overall appearance of the image. Komodakis [Komodakis, 2006] worked on this design issue and enhanced the model by introducing an energy term $E(l_i)$which ensures the most probable Exemplar patch $l_i \in L$ a pixel n_i must be given which shows maximum coherence with neighboring patches on the region of overlap. This intuition is expressed by defining a Markov Random Field (MRF) over the rectangular arrangement of pixels each denoting a node n_i and by minimizing the energy term (E) in Maximum A Posteriori (MAP) framework which is defined as

$$E(l_m) = v_i(l_m) + v_{ij}(l_m, l_n) \tag{2.29}$$

where v_i term denotes the potential of node n_i that encodes quantitatively how well the patch l_m agrees with the source region around n_i. The term $v_{i,j}$ pairwise potential which encodes the coherence between the pair of patches about neighboring nodes n_i, n_j which is expressed in terms of SSD computed over the common locations of overlapping patches. It is equally important to learn the neighborhood structure in MRF [in Lee et al., 2006] [Roth and Black, 2005], which is out of the scope of this topic. In that proposed model the neighborhood is defined using 4-Neighbors and each edge correspond to a clique over which the entire image is factored. Solving this minimization problem is NP-hard as the number of possible patches that can be extracted from source region is very large. Literature is enriched with a lot of approximations to this labeling problem. Belief Propagation (BP) [Yedidia et al., 2003] is the most widely used message passing algorithm which infers the marginal probability with which each node n_i falling inside the inpaint region is assigned a possible label l as follows. Node n_i sends a message $m_{i,j}(l_q)$ to the neighboring node n_j indicating how likely node n_i believes that n_j must assume a label l_q which is defined as

$$m_{i,j}(l_q) = min_{l_p \in L}\{v_i(l_p) + v_{i,j}(l_p, l_q) + \sum_{k:k\neq j\langle k,i\rangle \in E} (m_{k,,i}l_p)\}. \tag{2.30}$$

These messages are passed across edges E iteratively until they converge. Similar to the Criminisi model the nodes that are assigned with priorities that are used to decide the order of message of passing. At the end belief of each node is computed as

$$b_i(l) = -v_i(l) - \sum_{k\langle k,i\rangle \in E} (m_{k,i}(l)). \tag{2.31}$$

This algorithm works well only for the graphs without loops. But it is evident that the MRF defined over the image has loops. To improve the performance of the BP therein authors introduced label pruning policy, naming it as priority-BP, to drop some labels from evaluation. This global optimal method could solve object removal problem while preserving the structure and synthesizing the texture which

demonstrates the global coherence. Alternative to BP, flow based algorithms are developed to solve the MRF model using graph cut algorithms. Karthik [Alahari et al., 2010] proposed very active label pruning algorithms that are tractable in nature and works for undirected graphs also. Recently, authors [He and Sun, 2012] observed that the histogram of the patch offsets of the Best-Exemplar patches is sparsely distributed over the image and the peaks in the histogram provide reliable information. The authors therein constructed the shifted versions of the image according to the peaks of the patch offset histogram and combined them using MRF approach. The models discussed in this section handle the texture synthesis problems under object removal *inpainting category* which falls into single and large hole *inpaint class*.

2.6 Simultaneous Inpainting

In this section, the methods that aspire to meet both the functional requirements of the inpaint problem through decomposition are discussed. In this class, there are two variants. The first category of algorithms separate the texture and structure components and inpaint each one by utilizing the component specific algorithms. For example, Bertalmio solved the inpainting problem through this model. The other category deals with inpainting through sparse representations. This model could inpaint both the components simultaneously which can also produce the individual component if wanted. This type of algorithms is referred to as sparse representation based algorithms.

2.6.1 Domain Decomposition Approaches

The first such attempt, made by Meyer [Meyer, 2001] expresses the image I as the sum of structure component (piecewise smooth) $u \in BV(R^2)$ and the texture component (oscillatory) v. This approach is an extension of Rudin [Rudin et al.,

1992] noise removal idea to model the smooth area with structures as

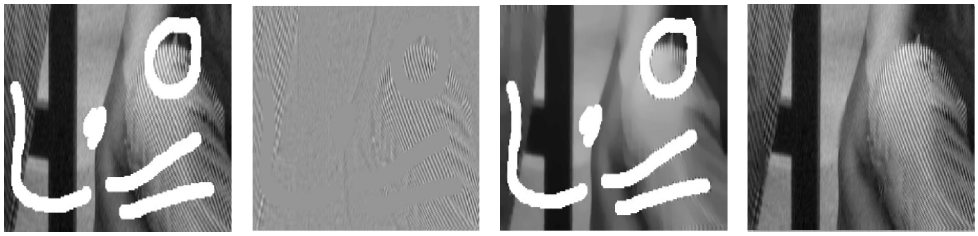
$$\inf \quad I(u) = \int |\nabla u| + \lambda \|v\|^2_{L2} \quad ; \qquad I = u + v. \tag{2.32}$$

Here λ is a tuning parameter. The idea incorporated in this equation is just about regularization of the noise removal solution using Bounded Variation functions. And, without loss of generality v, the fidelity term can represent the texture in place of noise. But the weakness of this model is for small values of λ the regularization term cannot adapt to the local oscillations because of the over smoothing nature of L_2 norm over v. Such a disability results in loss of texture and therein exhibits poor capability to separate the structure from the texture [Aubert and Kornprobst, 2006]. To overcome this problem Meyer [Meyer, 2001] has proposed a new fidelity term v involving a weaker norm instead of L_2 norm. The intuition in this proposed norm is to introduce the distance function for defining the fidelity term v over the space $G(R^2)$ which accounts for the data term in different amounts along different directions as

$$|v|_G = \inf\{ \quad |g|_\infty \quad / \quad v = div(g); \quad g = (g_1, g_2); \quad |v| = \sqrt{{g_1}^2 + {g_2}^2} \quad \} \tag{2.33}$$

$$g_1 \in L^\infty(R^2) \quad g_2 \in L^\infty(R^2); \text{where} \quad |g| = \quad \max_{1,2}|g_{1,2}|.$$

Observe that the involved space G defined over R^2 is endowed with L^∞ norm which ensures infinitely differential nature of the function g. Obviously, such a function is similar to sin or cos functions and g measures the anisotropic data components. Figure 2.18 shows the result of component *separation* algorithm. Upon separating the texture from structure, Figure 2.18(b) & (c), the inpainting problem is solved by working on both the components individually. Bertalmio [Bertalmio et al., 2001] applied the diffusion-based approach to inpaint the structure component and the Exemplar method on the texture component. Then the two inpainted components are combined to get the overall result Figure 2.18(d).



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

Figure 2.18: (a) The Barbara (b) Texture component (c) Cartoon component (d) Inpainted image

2.6.2 Dictionary Based Sparse Representations

In recent times the signal processing community developed a set of tools that are very impressive and capable of solving a wide range of computer vision problems. Sparse representation, the latest aspect in signal processing, invests on a principle that many physical characters in an image if either transformed or adequately represented possess a lot of sparsity. For example sin function in Fourier Domain is highly sparse and the gradient of a piecewise constant area of an image which is away from an edge is also sparse [Elad et al., 2005]. In this direction Elad [Elad et al., 2005] perceived a signal as a 'sum of a mixture of components' and each one is represented in sparse manner over a collection of transformations. The investigation behind this representation is any image, composed of both structures and textures cannot be effectively represented by a single kernel. Hence, such morphological aspects must be captured and represented in a collection of transformations that are arranged in the form of a matrix forming a dictionary D defined as

$$I = D\alpha \tag{2.34}$$

where α denotes the coefficients of the transformations. Elad modeled the problem of recovering the individual components present in the image by developing a data modeling framework 'morphological diversity'. This framework assumes that the input image I is a sum of the mixture and offers intuitions to construct a dictionary based on related transformation for each morphological feature. For

example, the curvelets can represent the structure component x_s and DCT is effective in handling the textures x_t. Each of these individual transforms are referred to as sub-dictionaries of the dictionary, and each sub-dictionary is composed of the corresponding basis functions, also called atoms. The significance of the dictionary is the prior knowledge represented in atoms of the dictionary offer a sparse representation for the related component and a dense representation for the other component as follows

$$\min\{\|\alpha_{DCT}\|_p^{\ p} + \|\alpha_{curv}\|_p^{\ p}\} \quad \text{s.t.} \qquad \|I - D_{DCT}\alpha_{DCT} - D_{curv}\alpha_{curv}\|_p \leq \sigma. \tag{2.35}$$

Here α denotes the coefficients of the corresponding transformation within the sub-dictionary D and σ models the noise parameter. To solve this constrained problem a stage wise hybrid algorithm was developed by Starck [Starck et al., 2005] which extracts the morphological components x_t and x_s (MCA). The overall idea underneath this model is to regularize the solution of the inverse problem by using sparsity as a prior. The MCA can be directly extended to solve inpainting problem by recasting the Equation 2.35 which includes the mask M which specifies the inpaint region as

$$\min \quad \{\|\alpha_{DCT}\|_2 + \|\alpha_{curv}\|_2\} \quad \text{s.t.} \quad \|I - MD_{DCT}\alpha_{DCT} - MD_{curv}\alpha_{curv}\|_2 \leq \sigma. \tag{2.36}$$

Fadili [Fadili and Starck, 2005] pursued the scope of turning this model into a well known framework Expectation-Maximization (EM) for estimating maximum likelihood when some data is missing while harnessing the sparsity by penalizing the solution

$$min_{\alpha,\lambda} \quad \{\|I - M * D_{DCT}\alpha_{DCT} - M * D_{curv}\alpha_{curv}\|_2 + \lambda(\|\alpha_{DCT} + \alpha_{curv}\|)\} \tag{2.37}$$

where the latter term in the Equation 2.37 denotes the penalization which necessarily ensures the sparse representation for α and is solved iteratively. These models could solve scratch removal *inpaint category* problem under simultaneous

taxonomy that falls in multiple but smaller holes *inpainting class* in the mask.

2.7 Hybrid Methods

This sub-section offers an overview of the methods that combine the useful features of so far discussed standard methods to realize global optimal solutions. All such methods express the global features as energy function and make attempts to minimize it optimally.

2.7.1 Global Coherence Methods

In the recent times, the image processing community explored possibilities to use non-local information to solve image denoising problem [Buades and Coll, 2005] In these methods, an image is denoised by convolving it with a local geometry aware pointwise mask [Tschumperle and Deriche, 2005]. Authors in [Buades and Coll, 2005] expressed a fact that natural images possess similar patches that might be spatially far away and introduced a non-local means approach. The authors therein applied the weighted average of the distances between a target patch and all the other patches in the image. Transform domain denoising methods deal with small image patches. These denoising algorithms project the image patches onto a content aware orthonormal basis such as a wavelet or DCT to produce a set of coefficients. Then the transform coefficients are modified by hard thresholding the smaller coefficients followed by inverse transformation. This process is extended to remove the seam by considering the average of overlapping patches. Rajwade [Rajwade et al., 2013] in his model HOSVD, combined the non-local means with transform-domain techniques to take the advantage of both the approaches. The Exemplar based methods are the candidate methods for improvisation in many ways [Buyssens et al., 2015]. The Criminisi model picks only one non-local Best-Exemplar in a greedy manner which induced texture garbage as shown in the Figure 2.16. Many improvements are suggested to this model that suggest to either consider either more than one exemplar and derive a new patch from that

or combine this method with another paradigm.

2.7.2 Nonlocal Means

Wong [Wong and Orchard, 2008] extended the idea of Variational framework in the context of Best-Exemplar model. Therein, the authors [Wong and Orchard, 2008] adapted the pixel based replacement policy while working with patches $\Phi(p)$ around the pixel $p \in \Omega$. The robustness of this algorithm is based on the simple idea to utilize multiple patches $\Phi(q_i), i = 0...n$ that are selected from the most similar patches. The similarity measure $d(\Phi(p), \Phi(q_i))$ is the sum of squared differences between the corresponding source region pixels in the two patches, which is same as that of [Criminisi et al., 2004] approach. Once the most n similar patches Φ_r, $r = 1, 2, ...n$ are selected then the weights for each patch is computed in the form of a Gaussian model

$$w(i) = e^{-\frac{d(\Phi(p),\Phi(q_i))}{h}} \tag{2.38}$$

which suggests that the contribution of the patch decreases exponentially with dissimilarity. Ultimately, the target patch $\Psi_{\hat{p}}$ is computed based on the convex combination of the finite number of n most similar patches, which is called Nonlocal Means (NLM)

$$\Phi_{\hat{p}} = \frac{\sum_i w(i)\Phi(q_i)}{\sum w(i)}. \tag{2.39}$$

2.7.3 Nonlocal TV Minimization

Based on the definition of inpainting, the original image F is assumed to be a transformed by T which results in incomplete one which is the input image I for the inpainting problem. Then the TV model solves the inpainting problem in Nonlocal TV minimization framework (NLTV) as

$$\hat{F} = T^{-1}I. \tag{2.40}$$

A suite of algorithms proposed by Giloba [Gilboa and Osher, 2008] utilize nonlocal information within variational framework which shares the similarity with the Exemplar-based model solves the inpainting inverse problem. Therein the authors modeled the input image as a weighted graph that links together the pixels p_i and p_j with weight $w(p_i, p_j)$. The utility of this graph is to define the non-local gradients ${\nabla_{p_i}}^w$ using graph gradient operators

$${\nabla_{p_i}}^w = (I(p_i) - I(p_j)) * \sqrt{(w(p_i, p_j))}; \quad p_i, p_j \in I. \tag{2.41}$$

The graph weighting function w is computed, similar to non-local means case, as the distance between the patches Ψ_{p_i} and Ψ_{p_j} about p_i, p_j respectively as

$$w(p_i, p_j) = e^{\frac{-\|\Psi_{p_i} - \Psi_{p_j}\|}{2\sigma^2}} \tag{2.42}$$

where σ is the controlling parameter. Then the graph-based regularized inpainting problem is represented as the following inverse problem [Peyré et al., 2008]

$$F^\star = min_I \quad \frac{1}{2}\|I - (1 - M)F\|^2 + \sum\nolimits_I \|\nabla I\|. \tag{2.43}$$

Instead of solving this variational problem directly, the proximal of this is modeled and solved the inpainting problem [Peyré et al., 2008]. Arias [Arias et al., 2011] in his work extended the [Peyré et al., 2008] model in two stages. The first extension is to expedite the search process for exemplar patch by adapting the PatchMatch [Barnes et al., 2009] correspondence algorithm. The other one is to incorporate four different ways to define the similarity measure [Wexler et al., 2007]. The inpainting problem is solved in two stages. The first stage is dedicated to locate the best exemplar and the second stage updates the image. Recently, [Duan et al., 2015] proposed a Non-Local Color TV method to extend this algorithm to color images where the minimization problem is solved through split-Bregman Method. These Hybrid methods solve the Object removal *inpaint category* problem under Hybrid *inpaint taxonomy* within single and large hole *inpaint class*.

The overall summary of the inpaint models and their applications in image preservation domain are presented in the Table 2.1.

Table 2.1: Summary of the Survey

Method	Prior	Suited for	Application
PDE	Smoothness	Structure	Reinvention
Exemplar	Self-Similarity	Texture	Editing
Separation	Morphological	Structure & Texture	Restoration
NLTV	Self-Similarity	Structure & Texture	Editing
Dictionary	Sparsity	Structure & Texture	Reshuffling

2.8 Overview on Spatial Interpolation Techniques

Spatial Interpolation is a mechanism used by the geographic scientists to measure the ore quality within a regulatory mine area. As it is possible to collect the desired information at finite locations within the regulatory area, it is desired to apply some interpolation methods to reliably predict the quality of ore at the remaining locations. General interpolation techniques do not pay any regard to the importance of understanding the patterns present in the data and predicts the result directly [Gentile and Meylan, 2013]. Spatial interpolation techniques, on the other hand, consider lot of characteristics of the data and model them in the course of predictions.

2.8.1 Classification of interpolation schemes

Based on some general characteristics the interpolation schemes are classified as

- point based versus area based- predicts the response either at a point or over an area
- global versus local- Develop either a single model for the entire data or multiple models over divided spatial data sets

- exact versus approximate- Interpolated values either matches with the measured value at the design sites or not
- stochastic versus Deterministic- Accommodates the randomness in the model and addresses the uncertainty in predictions or not
- inclusion of derived data and spatial relationship- Addresses the spatial variability of the data and reduces the uncertainty of predictions
- ability to quantify the error incurred in prediction

Spatial interpolation algorithms follow the Tobler's First Law of Geography [Tobler, 1970] " closer together two points are in space the more likely those points are to be similar". Nearest Neighbor algorithms and Inverse Distance Weightage (IDW) are examples of spatial interpolations.

2.8.2 Nearest Neighbor (Polygonal)

In this model, the response at an unknown location is assigned based on the value of the nearest design site or by computing the mode of the design sites available within the specified polygon around it (in the case of 2-D data). Thus this model assigns a value to the unknown location either a value available in the enclosing polygon. If the sampled data points are in a rectangular grid, then the resulting polygons will be of equal size and regularly spaced. If the measured data points are irregularly spaced, then the resulting prediction surface will be an irregular lattice of polygons. Literature is enriched with lot of polygon constructing methods that include Thiessen Polygons, Voronoi diagrams or maps, and Delauney triangulation [Ripley, 1991]. The limitation of this method is, there is a restriction on maximum and minimum values of the prediction.

2.8.3 Inverse Distance Weighting

Inverse Distance Weighting (IDW) [Brus et al., 1996] model assigns a value to the target location x_0 by taking the weighted average of the design sites $y(x_i), i = 1 : N$

available in the well defined neighborhood as

$$\hat{y}(x_0) = \sum_{i=1}^{N} w_i y(x_i) \tag{2.44}$$

where w's are weights defined as

$$w_i = \frac{1}{d(x_0, x_i)^\beta} / \sum_{j=1}^{N} \frac{1}{d(x_0, x_j)^\beta} \quad \beta \geq 0 \quad \sum_{j=1}^{N} w_j = 1. \tag{2.45}$$

In the above equation $d(x_0, x_i)$ denotes the distance between the points x_0 and x_i. The rational behind this equation is, nearer points contribute more to the prediction and the rate of influence is defined by β. For large value of β, the predictions become more similar to the closest design site and results in the interpolation surface with sharp peaks. In this sense, the β parameter controls the degree of smoothness of the interpolation surface [Brus et al., 1996]. Even value of N determines the degree of the smoothness of the interpolation. Thus IDW is a local and deterministic interpolation method. Though this method is simple to implement, it suffers from the disability to measure the uncertainty of prediction. The IDW can be made more versatile by redefining the distance in a more generalized form by introducing radial basis function (RBF). RBF, also called as kernel ϕ, is a real valued function which evaluates the response at a point x_0 depending upon the radial distance between x_0 and a set of points x_i in terms of Euclidean distance as $\phi(\|x_0 - x_i\|)$. The interpolated value $z(x_0)$ at x_0 is a linear combination of RBF evaluated at x_i's as

$$z(x_0) = \sum_{i=1}^{N} \lambda_i \phi(\|x_0 - x_i\|) \tag{2.46}$$

where the λ_i are N weights that are computed by solving the system of N equations formed by imposing the exactness of the interpolation at all x_i. The accuracy of the interpolation depends upon the compliance of the selected RBF with the underlying data which is subjective in nature. The Table 2.2 gives the most frequently used instances of RBF. The RBF based interpolation is exact and

Table 2.2: RBF $\phi(r)$ where $r = \|x_0 - x_i\|$

RBF	The Expression
Thin plate	$r^2 ln(r)$
linear	r
cubic	r^3

deterministic in nature but the selection of RBF basis for the given problem is fundamentally empirical.

2.9 Quality Metrics for Inpainting

The ill-posed nature of the inpainting problem also suffers from the lack of a proper measure to ascertain the efficacy of the inpainting algorithm and the quality of the results. The primary reason is that the inpainting algorithm reinvents the damaged area of the image without having any prior information. Most of the inpainting algorithms are evaluated only through visual inspection which verifies how pleasantly the inpainted region gels with the surrounding source region. Inpaint quality assessment efforts to extract some parameters from Φ and compare them with that of the inpainted region also turn futile in many situations. However, Peak Signal to Noise Ration (PSNR) and Structural Similarity Index Measure (SSIM) [Wang et al., 2004] are the two measures widely used by the image processing community for inpainting problems. These were defined for monochrome images and are extended to color images based on averaging.

2.9.1 Peak Signal-to-Noise Ratio (PSNR)

PSNR is the most widely used measure by the signal and image processing communities for ascertaining the quality of the reconstruction algorithm of lossy compression algorithms. Given the original monochrome image F of size $m \times n$ and the result of the inpainting algorithm I the human perceptual quality of I measured

through mean square error (MSE) as

$$MSE = \frac{1}{mn} \sum_{i=0} \sum_{j=0} (\quad F(i,j) - I(i,j) \quad)^2 \tag{2.47}$$

and the PSNR is defined on the logarithmic decibel (db) scale as

$$PSNR = 10. \lg \frac{MAX^2}{MSE} = 20. \lg(MAX) - 10. \lg(MSE) \tag{2.48}$$

Where MAX is the maximum possible value of the pixel, 255 in the case of the 8-bit monochrome image. Theoretically, if PSNR value of a result is above 25db, then the algorithm is evaluated to be good.

2.9.2 Structural Similarity Index Measure

(SSIM) Wang [Wang et al., 2004] in his work argued and demonstrated that the natural images are a perfect mix of structures and textures and show some stationary behavior in the sense that the structural features show a lot of dependence among themselves. The image quality index metric Minkowski, for example, which defines error E between the ground truth x and the inpainted image y as

$$E = \sum_i [|x_i - y_i|^p]^{\frac{1}{p}} \tag{2.49}$$

where x_i, y_i are the individual corresponding pixel values each of them is assumed to be independent in the images x and y. Wang demonstrated that such a metric which disregards the structures present in the image evaluates two visually different images derived from the same ground truth to have same Minkowski quality index. Wang [Wang et al., 2004] developed SSIM which defines the image quality by comparing the extracted three components luminance l (mean μ), contrast c (standard deviation σ) and structure s (normalized by standard deviation σ_{xy})

from both the images and computes

$$\mathrm{SSIM}(X,Y) = \frac{(2\mu_x\mu_y + K_1)(2\sigma_{xy} + K_2)}{({\mu_x}^2 + {\mu_y}^2 + K_1)({\sigma_x}^2 + {\sigma_y}^2 + K_2)}. \tag{2.50}$$

The quality of the results of image inapinting is qualified to be good if the SSIM is above 0.9 [Wang et al., 2004].

2.10 Summary

This chapter presented the complete spectrum of inpainting models and highlighted their capabilities. It is very evident from the literature that all these models directly accounts for texture synthesis, structure propagation, and both in few cases but with subtlety as to the proper selection of a model determines the quality of the results. The ill-posed nature of the problem offers a window of opportunities for the researchers to re-design certain features of existing models and to cast the problem into any capable domain. The inpainting solution attains a special status and attracts the interest of researchers if it can model the structure and texture features in a deterministic statistical area and avoids the subjectivity in selecting a sophisticated algorithm to solve the non-trivial inpainting problem instances. As per my knowledge goes, there is no attempt made so far by the inpainting community to cast the large scale inpainting problems in the statistical domain. Also, there is a natural requirement to extract middle-level information from the image to utilize it for characterizing the image and solving the problem in real time. This thesis makes an attempt to address these issues and evolves an intelligent framework to solve diversified inpainting problems.

Chapter 3

Granular Approach For Best-Exemplar

The Exemplar algorithm proposed by Morel [Masnou and Morel, 1998] was modified by Criminisi [Criminisi et al., 2004] while accounting for the self-similarity measure and arrived at the Best-Exemplar algorithm which is iterative and interior point based approach. However, Best-Exemplar method suffers from a couple of drawbacks as it essentially involves almost total exploration of the source region in each iteration as the *fill front* construction cannot reveal any clue about the target patch and will not support any constraint development on the search space. But, in many situations, a patch on the boundary, with source region point at the center and few source region pixels, can offer some pattern and dispense relevant spatial features of the target patch. This intuition led me to make contributions to this chapter. By utilizing this information, the source region divided into granules based on the histogram of pixel values, a non-parametric model. It helps in quickly locating the Exemplar element from the source region Φ. Best-Exemplar selection from Φ involves utilization of a simple similarity measure but does not consider the essential structural features present in source region pixels. This observation prompted me to propose a robust similarity measure as the second improvement.

In this chapter, features of self-similarity based Best-Exemplar are redesigned and text removal *inpainting category* problem is solved. Initial section 3.1 discusses

the aspects that spell predicaments on the performance of Best-Exemplar model. Subsequently, construction of *target front* $\partial\Omega$, the selection of target patch from it by using a priority term which is free from Gradient are discussed in section 3.2. Quick-Search algorithm- a granule based two-fold search space reduction technique which selects all candidate patches Ψ_q is developed in section 3.2.3. Then the section 3.3 presents the proposed L_0- a Pseudo Norm, based distance measure to choose the Best-Exemplar $\Psi_{\tilde{p}}$. The implementation details of the proposed model is presented in 3.5. The detailed iterative algorithm *Granule Approach for Inpainting* (GAI) is presented in 3.6. The efficacy and scalability of the proposed algorithm is demonstrated by carrying out the experiments on benchmark images, and presenting the results using the quality metrics SSIM and PSNR, in section 3.7. Finally, the summary is presented in 3.8.

3.1 Design Issues of Best-Exemplar Inpainting

The topology of the inpaint region Ω may be irregular with boundary $\partial\Omega$ and the source region: Φ is vast, see Figure 3.1. Now, the direct adoption of Exemplar approach [Efros and Leung, 1999] to fill the Ω leads to various kinds of artifacts depending on the selection of boundary pixels in different orders in different runs, see Chapter 2, Figure 2.12(c) & 2.12(d). Criminisi [Criminisi et al., 2004] has addressed the uncertainty associated with the Exemplar method through prioritizing the boundary pixels and picking up the Best-Exemplar as highlighted in Chapter 2, Section 2.5.2. The essential but unspecified design issues of Best-Exemplar approach [Criminisi et al., 2004] are grossly listed as follows:

- Construction of target front/boundary $\partial\Omega$ about Ω
- Specification of the target patch selection from the $\partial\Omega$ which is free from uncertainties.
- Implementation of secondary data measures e.g. gradient, structure tensors.

- Optimized exploration of Φ in search of the Best-Exemplar through constructing efficient data structures
- Specification of a prudent measure to compute the similarity between the Ψ_p and the set of candidate patches Ψ_q from Φ
- Ensuring the graceful inward propagation of structural elements which reach the $\partial\Omega$ without any distortion
- Realizing the local textural preservation and global structural consistency.

The subsequent sections evolve a Granule driven strategy to address major chunk of the above identified design issues that reduce the cost of the Best-Exemplar search algorithm and define a new similarity measure aimed at improving the fidelity of the inpainting. This improved Best-Exemplar is called Granular Approach for Inpainting (GAI). It includes a search space reduction technique and a pseudo-similarity based measure for picking the Best-Exemplar as discussed next.

3.2 Proposed Granule Based Search Space Reduction

The text removal from images is a very typical instance of the *inpainting category* problem and provoked me to have a re-look at certain features of Best-Exemplar model. To start with, text removal *inpainting category* problem, as shown in Figure 3.5(a), has inpaint region which is not a single connected region but of many disconnected regions, rendering the fill front construction algorithm to end up with a lot of smaller contours. The Best-Exemplar involves the computation of priority term for each and every boundary pixel. If the text to be removed is overlaid covering the larger portion of the image, then the Exemplar algorithms suffer from segmentation problem as the size of Ω is large. In addition to this, if the characters in the text are very nearby, then the patch size must be smallest

possible. On the other hand, if larger size patches are considered, then there is a high probability of texture garbage and structure distortions in the results. Also, the Best-Exemplar selection involves the exploration of the entire source region which renders the algorithm to run slow. Importantly, there is a high risk of non-coherent patches getting propagated due to the involved greedy policy in selecting the Best-Exemplar from multiple Exemplars, satisfying the similarity criterion. The proposed Granule Based approach addresses these issues as elaborated in the upcoming sections.

3.2.1 Proposed Target Front Specification

Given the input image I with source region Φ and inpaint region Ω, a mask is constructed- a binary image in which 1 represents the inpaint region location, and 0 represents a source region location. The source region pixels that are just one pixel away from Ω are isolated and maintained on the *target front* ($\partial\Omega$) data structure as shown in Figure 3.1. Such a set can be easily constructed by collecting all pixel locations from the mask with value 0 and having at least one pixel out of its eight neighbors possessing value 1. The target front contributes towards assigning the priority to all pixels in a deterministic manner and in resolving the granule selection while searching for Best-Exemplar. One observation here is, if the patch size is modeled to be 3×3 then the number of source region pixels range from 2 to 8 in all regular problem cases. And in each iteration of the algorithm, one patch gets updated while effecting at least 1 to 7 pixels of the inpaint region. Moreover, the target front is updated in every iteration. If the cardinality of Ω is K then the total number of iterations required, before stopping the inpaint algorithm, ranges from $K_0 = \frac{K}{7}$ to $K_1 = K$.

3.2.2 Target Patch Identification

Throughout this discussion, patches of size 3×3 are chosen. For every pixel p on the *target front*, a patch is constructed having p at the center. Then two

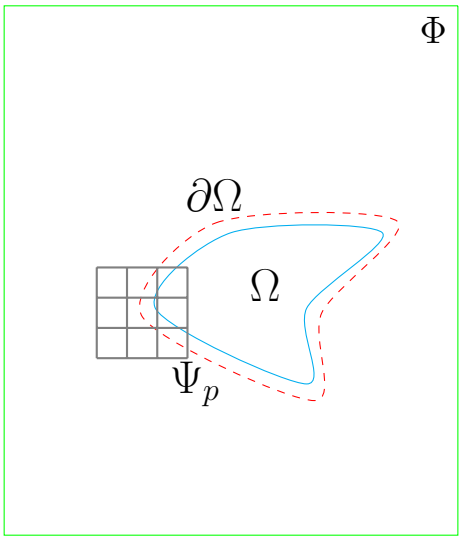


Figure 3.1: Target Front construction

quantitative terms are collected: Data, DT(p)- The difference between maximum and minimum pixel values associated with that patch and Confidence, C(p)- the number of pixels belonging to the source region respectively. The sum of these two terms is assigned as the priority Pr to pixel p as

$$Pr(p) = \mathrm{C}(p) \quad + \quad \mathrm{DT}(p). \tag{3.1}$$

The DT term ensures that patch containing structural information hitting the $\partial\Omega$ gets a higher priority and leads to propagation of the neighboring structures inwards. For certain pixels on target front, the DT term may be zero. The notable point here is, Criminisi method was silent on gradient computation, which led to various implementations that turned detrimental for ensuring the consistency. Figure 3.2 highlights the implication of forward, backward and central difference methods of the derivative term, on the overall result of the Best-Exemplar method [Buyssens et al., 2015]. In the proposed model every *target front* pixel is assigned with the priority term in each iteration. The patch $\Psi(p)$ with largest number of source region pixels is the most reliable patch and is elected as the target patch for processing.

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

Figure 3.2: (a) Bungee input image (b) Result of forward difference (c) Backward difference (d) Central Difference.

3.2.3 Granules construction using Histogram

Upon selecting the *target patch* Ψ_p from the target front, a search for Best-Exemplar is invoked over Φ. This step involves computing the distance between Ψ_p and every overlapping patch Ψ_q collected from Φ. Depending on the size of source region, the computation time may range from tens of seconds to hours [Le Meur et al., 2011]. The proposed methodology scores a point over Best-Exemplar approach in this step. The simplicity of this model lies in optimization of the Best-Exemplar exploration step by utilizing the absolute intensity values that are directly available on a non-parametric statistical tool- histogram. Before using the histogram, the fundamental definitions and principles related to it are given below.

Given an 8-bit Gray Level Image with intensity values $i \in [0 - 255]$, then a function is constructed, that maps the intensity value i onto the number of pixels with intensity value i. If this mapping is depicted in the bar chart form, it is called a histogram

$$h(i) = \text{Card}\{\ p\ /\ I_p = i\ \}$$

where h term denotes the histogram and Card refers to the cardinality of the associated set.

By definition, histogram abstracts the locations and can only compactly represent the data about dynamic range and pixel distribution over the intensity values as frequencies. For high-resolution images with a wide range of intensity values,

the histogram is very broad. So instead of working on individual intensity values, granules are constructed from them by dividing the pixel values into ranges and histogram depicts the frequencies over each range. This technique refers to bin based histogram construction. Formally, given 8-bit image I with gray values from $[0-255]$ that are divided into k bins with $0 = b_0 \leq b_1, ..., \leq b_k = 255$ then the binned histogram for I is the function

$$h(i) = \text{Card}\{\ p/b_i \leq I_p\ \leq b_{i+1}\ \}. \tag{3.2}$$

The selection of the number of bins k and the width of each bin is a subjective matter. In all the experiments conducted in this chapter, bins with equal width are created. The generation of granules in the proposed algorithm facilitates the Best-Exemplar Search operation, as discussed next.

3.2.4 Quick-Search

The pixels are grouped into a finite number of Granules, 30 for example, based on the range of values and the Histogram of them. This non-parametric method helps to reduce the search space. The dynamic range of natural images is often vast and the *target patch* center pixel value will fall into a single bin with good support. That means, if the pixel value is near to the mode of the bin, then the search space can be further reduced based on the relative frequency of the bin. Moreover, if the selected bin is the tallest tower, then we can have one more degree of freedom to select the pixels from a compact range. The granules are constructed using Histogram with 30 bins over the dynamic range $[0-255]$ and the frequencies f_n, denoting the number of pixels falling into each bin, are maintained in a linear array. In the case of natural images with high contrast, each bin accommodates a smaller number of pixels, when compared to the total number pixels present in Φ. Further, each bin is represented by the center B_c of the bin, which is the average value of the intensities of all pixels that fall into that bin. Using this arrangement, to search for the Best-Exemplar, for a patch Ψ_p, its intensity value

is mapped onto the nearest bin center. This nearness is defined as the absolute difference between the intensity of p and each bin center B_c. Once the nearest bin B_d is selected, it gives an overall number of patches that are required to process and pick the Best-Exemplar. Thus the granule selection constitutes the first level of search space reduction in the proposed model. Subsequently, the frequency f_d of the bin B_d is compared with frequency f_{min} of bin B_{min} with least frequency. Then a simple qualitative metric $Degree$ is derived as follows

$$\text{Degree} = \lceil \frac{f_{min}}{f_d} \rceil \times 30. \tag{3.3}$$

The purpose of Degree term in Equation 3.3 is, it helps to reduce the number of pixels to be processed further and localizes the search with compact support. This aspect is illustrated by taking up a numerical instance of the term $Degree$. Suppose $f_{min} = 50$ and $f_d = 100$ then $Degree$ is 3. Thus, all pixels from the selected bin with a ± 3 deviation in intensity values about the p are isolated and the patch around them are extracted into a data structure. In few instances, even pixels from two neighboring bins will be considered for processing. Thus this model accomplishes the adaptive thresholding mechanism and is termed as Quick-Search. Quick-Search extracts the granules from the source region and decides the range of the pixel values that must appear at the center of the each candidate patch. Let G be the number of granules derived from Φ, and the S be the number of source region patches, then it is trivial to observe that the number of patches associated with a granule g is determined by the distribution of the associated range of pixel values spanned by g. And it is much smaller than S in the case of all natural images. Hence, this ideology of selecting the granule g based on the clue offered by the central pixel of the Ψ_p on the target front $\partial\Phi$, and confining the search operation to that bin reduces the time complexity substantially in each iteration of the Best-Exemplar algorithm. In addition to this, the proposed model exploits the low contrast feature of the input image by further sampling the selected g based on the $Degree$ of tolerance which is determined by the frequency of g.

Algorithm 2 Quick-Search

Input: Input Image I, the associated Mask M, the number of granules k

Output: Range of Intensity values considered for Best-Exemplar search

1. `Construct the Granules by generating the histogram, given` k, `and maintain the frequency` f `of each bin with center` B_c
2. `Construct the target front` $\partial\Omega$, §3.2.1
3. `Select the` ***target patch*** Ψ_p, §5.8
4. `Select the Granule` g `in which` p `falls` §3.2.3
5. `Compute` $Degree$ `using the Equation 3.3`
6. `Determine the range of intensity values over which the Best-Exemplar search is invoked`

3.3 Proposed Pseudo-Similarity Measure

In this step, the Best-Exemplar shall be picked out from all candidate patches selected from the previous step, which is carried out by computing the distance between the Ψ_p and the candidate patches through a well-defined similarity measure. The literature is enriched with various types of distance measures that each one caters to a spatial feature.

3.3.1 L_2 Norm and its limitations

The simpler and most widely used, square of Euclidean measure (SSD) between two points can be extended to compute the similarity between two patches also. In the case of 2-D data the Euclidean distance measure is geometrically represented by a circle and for higher dimensional data this is a sphere. It is intuitive to visualize that there are finitely many points at the same distance from the given reference point that are lying on the circumference of the circle. Thus, the

Euclidean measure gives the overall dissimilarity between two higher dimensional data points with a disregard to the deviations between the individual pair of dimensions. Hence it is very much suitable to pick patches for smooth portions in the image [Xu and Sun, 2010]. But, the real world images have structural and texture areas also. Hence, a pseudo-norm similarity measure which offers the differential respect to the pairwise difference of pixels from two patches is introduced to select the Best-Exemplar.

3.3.2 Proposed L_0-Norm Based Similarity Measure

The proposed norm aligns the target patch Ψ_p with candidate patch Ψ_q and computes the absolute difference between the corresponding pairs of pixels. Further, it gives a special attention for the pixels with differentially larger values among the available pixels within the Ψ_p. The proposed similarity measure takes Dev as the user input, related to the extent of deviations allowed between the pair of pixels while determining the similarity. Then the distance evaluating logic computes the pixel-wise absolute differences into on a linear array $D[\imath \times \jmath]$ that thresholds each difference onto 0 or 1 by the logic given Equation in 3.4.

$$\text{if } ((\Psi_p)(i) - (\Psi_q)(i)) < \text{Dev} \quad \text{then} \quad D[k] = 0 \quad \text{else} \quad D[k] = 1. \tag{3.4}$$

Now the distance between the given two patches is the number of non-zero entries present in the array D. This distance measure constitutes the pseudo-norm. The proposed similarity measure assumes a robust form through the enhancement

$$\text{if } \Psi_p \geq \mu(\Psi_p) \text{ and } (\Psi_p(k) - \Psi_q(k)) < Dev \text{ then } D[k] = \text{-Dev}. \tag{3.5}$$

where $\mu(\Psi_p)$ denotes the average of the known pixel values in Ψ_p. Overall, the proposed distance measure picks up the Best-Exemplar based on the distance

measure, given in Equation 3.6:

$$\Psi_{\tilde{p}} = \underset{\Psi_q \in \Phi}{\arg\min} \quad |\Psi_p, \Psi_q|_{l_0}. \tag{3.6}$$

Thus, the proposed pseudo-norm does a favor to smooth areas, while catering to sharp changes in the pixel values within the Ψ_p that possibly represent a structure. This norm can impute the missing values of the patch by selecting the Best-Exemplar in a robust manner compared to SSD-based measure. The computational complexity of this distance measure is just n subtractions when compared to n^2 multiplications and n^2 subtractions in the case of SSD.

3.3.3 Best-Exemplar Selection from Multiple Imputations

Further, if any iteration in the algorithm could generate multiple imputations with same similarity value, then the spatial offset of each candidate patch location about Ψ_p location is considered. The underlying idea for reviewing this policy is that natural images possess Second Order Stationary property [Hastie et al., 2001]. According to this property, the neighboring pixels that are separated by a finite lag distance are highly correlated. Hence, a spatially nearest imputation from all possible multiple imputations is selected as the Best-Exemplar. This approach also favors the smooth propagation of edges present in the vicinity.

3.4 Post Processing

As mentioned before the quality of the results are highly influenced by the nature of the algorithm, contents available in the image, size and shape of the inpaint region and the initialization of parameters. As the proposed framework is a greedy model, artifacts could creep in. Thus it is essential to employ the post processing step to remove them in an efficient way through spatial domain filters.

3.4.1 Isotropic Diffusion:

As a last step in the GAI framework, post processing is employed. The artifacts generated by the algorithm are local in nature and anisotropic diffusion process can fix these distortions. Oliveira [Oliveira et al., 2001] developed two diffusion kernels to remove the small noisy areas from the image. These kernels diffuse the information equally in all directions. Hence they are called isotropic kernels. The last step in this framework involves the convolution of the inpainted image with these kernels to get the regularized results.

3.5 Implementation Details

The experiments are conducted on standard images to remove text, scratches, and objects. Special emphasis is given to text removal problem wherein the input image overlaid with text over a large portion of the image. The effective implementation of the Quick-Search is possible with the utilization of a linear data structure. Initially, the entire source region, inpaint region and *target front* pixels are classified by utilizing the mask. Then, patches, of size 3×3, are constructed about each of the source region and the *target front* pixels. These patches are organized as records on two data structures, Table 3.1 ST and Table 3.2 BT respectively. These records include the center pixel location of each patch (abscissa and ordinate) that form the key attributes of the table, its intensity value and the intensity values of the eight neighboring pixels that are collected in an anti-clockwise direction. This arrangement complements the histogram to locate the pixels based on intensity values, which is not a direct capability of the histogram. The BT table gets updated in each iteration of the algorithm. Once any target patch from *target front* is replaced with the Best-Exemplar drawn from the ST table, the BT table is updated by marking that record as an inpainted patch. This step, in turn, involves updating the corresponding fields of all records in BT that get effected with the replacement of Ψ_p with $\Psi_{\tilde{p}}$. Further, the target front must be reconstructed. Thus, the locate and replace based algorithm eventually, erodes

the BT table to end up with zero records with inpaint region pixels to mark the end of the algorithm. The ST table shown below accommodates the source region pixel locations x_a, y_a and with pixel value p_{x_a,y_a} and its 8 neighbors as $n_{a_1}, ...n_{a_8}$. Similarly, the BT table gives the same representation to the boundary patches wherein the unknown location is represented with 526- an invalid gray value. In both the tables the first two fields x_a, y_a constitute the key attributes.

Table 3.1: ST table Records

x_a	y_a	p_{x_a,y_a}	n_{a1}	n_{a2}	n_{a3}	n_{a4}	n_{a5}	n_{a6}	n_{a7}	n_{a8}
x_b	y_b	p_{x_b,y_b}	n_{b1}	n_{b2}	n_{b3}	n_{b4}	n_{b5}	n_{b6}	n_{b7}	n_{b8}
x_c	y_c	p_{x_c,y_c}	n_{c1}	n_{c2}	n_{c3}	n_{c4}	n_{c5}	n_{c6}	n_{c7}	n_{c8}

Table 3.2: BT table Records

x_j	y_j	p_{x_j,y_j}	n_{j_1}	526	526	n_{j_4}	n_{j_5}	n_{j_6}	526	526
x_k	y_k	p_{x_k,y_k}	n_{k_1}	n_{k_2}	526	n_{k_4}	526	n_{k_6}	526	n_{k_8}

3.5.1 Tuning the Model Parameters

As mentioned before Histogram is a non-parametric model. But in the proposed algorithm GAI, there are few model parameters that are supposed to be properly initialized by the user. These are the size of the patch, number of bins, the Dev and have lot of influence on the computation time and quality of the result. The experiments that are conducted in this chapter initializes these parameters in a objective manner. The number of bins has been initialized to 30. The Dev is taken from the range $[1 - 30]$.

3.5.2 Handling of Color Images

The granule construction based on histogram can present only the distribution of intensity values, not their actual spatial information. This is one of the disadvantages of histogram. But the utilization of histogram in this contribution is limited to figure out the number of pixels in the source region, whose intensity values are

with the user specified term *Dev*. Moreover it is not used to assess the similarity between two patches, to solve a classification problem or object recognition problem, for which scale, translation invariant properties are highly desirable [Lowe, 1999]. With regard to color image, this issue becomes more prominent. For color images, the histogram summarizes the joint distribution of model based color channels and the overall distribution of the colors across the entire image. Moreover, histogram is not capable of directly representing any individual feature present in the image. Given this argument, the inpainting of color images do not involve the construction of color Histogram. Rather, the GAI framework converts the given input color image of RGB model into a gray scale image using a built-in routine 'moving windowgray' of MATLAB. This transform maps the color components R, G and B associated with a pixel on to the gray scale by taking the weighted sum of the individual components as given below

$$\mathrm{g} = 0.2989 * \mathrm{R} + 0.5870 * \mathrm{G} + 0.1140 * \mathrm{B}$$

where g refers to the gray scale intensity. It is very clear to note that the color triple of all pixels in the color image are manipulated into the intensity values through the same transform. Hence it is functionally meaningful to locate the Best-Exemplar in the grayscale image and utilize the locations of target patch and Best-Exemplar, to manipulate the underlying color image internally and simultaneously. This transformation offers the second advantage with the GAI. Here the issue related to the selection of color model is not discussed as it is a topic on its own and is beyond the scope of the work. The upcoming section demonstrates the efficacy of the proposed algorithm by conducting the experiments on the standard images.

3.6 The Complete Algorithm of GAI

The detailed algorithm GAI makes use of the routine Quick-Search, which hashes the source domain twice and picks up the Best-Exemplar based on L_0 norm.

Algorithm 3 GAI

Input: Input Image I and the associated Mask M

Output: Inpainted image

1. `Decompose the` I `into` Ω `and` Φ
2. `Apply Quick-Search algorithm,` §3.2.4
3. `Select the Best-Exemplar` $\Psi_{\tilde{p}}$ `using the proposed` L_0 `norm` §3.3.2
4. `Copy the`$\Psi_{\tilde{p}}$ `into` Ψ_p `at locations` r, $\forall r \in \Phi_p \cap \Omega$
5. `Mark` Ψ_p `as inpainted and update` Ω `and` Φ
6. `Repeat steps 2 to 5 until` Ω `ends up with no pixel.`

The object removal problem can be efficiently solved by mapping this GAI in the context of the Global Optimization model. The proposed implementation directly supports the `Maximum A Posterior` formulation. Once the Best-Exemplar is selected for a target patch, instead of using it to replace the target patch, it is marked as a potential Best-Exemplar which stands for the potential pixel values of inpaint region pixels. Thus consecutive iterations of GAI can assign a set of possible intensity values to each pixel forming the possible set of labels to it. Then the resulting situation can be viewed as a scaled down instance of the Message Passing algorithm over Markov Random Field. Maximum A Posteriori solution based formulation can be implemented using the Belief Propagation algorithm. This extension will reduce the space and time complexities substantially as it has to process only 8 labels for each inpaint location instead of 255 possible labels.

3.7 Experiments and Results

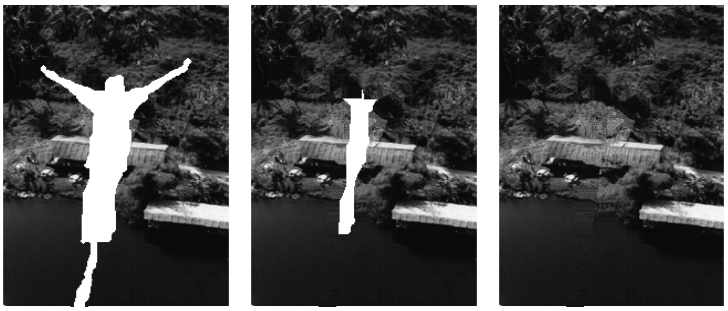
The efficacy of the proposed models is demonstrated by conducting the experiments on benchmark images from 'The Berkeley Segmentation Dataset and Bench-

mark' [Martin et al., 2001]. Results are analyzed using Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) metrics that are widely considered to be apt [Roth and Black, 2005], by inpainting community for assessing the quality. Proposed approaches demonstrate their viability and fidelity in text removal, scratch removal and object removal categories with man-made constraints.

The experiments in this thesis are carried out in a very generic manner. All the experiments were conducted using MATLAB toolbox on Intel architecture with 4 GB RAM.

Experiment: 3.1

In this experiment the GAI algorithm was launched on Bungee image and the the intermediate results are presented in Figure 3.3 on the gray version of the input image. The complexity associated with this image was, it had highly non-

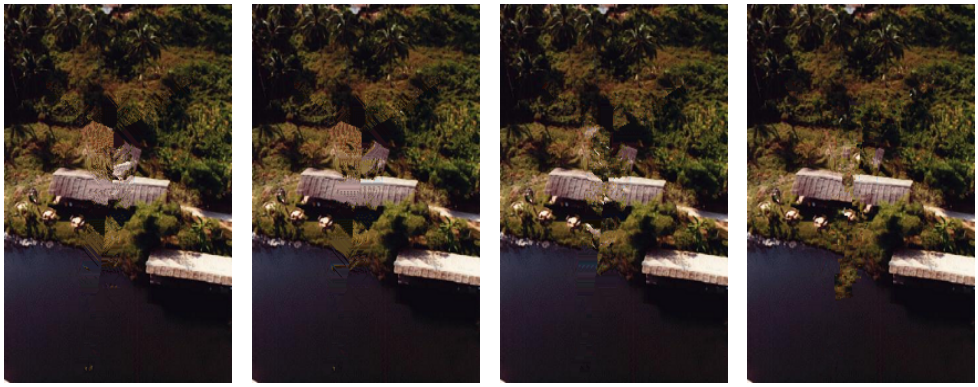


(a) (b) (c)

Figure 3.3: (a) Input image of size with object highlighted in white (b) Result of GAI after 10 iterations (c) and the final Result

stationary information and possesses structures, textures and smooth areas. The Figure 3.4 demonstrated the effect of *Deg* on the overall results. It is very clearly observable from this experiment that, large value of *Deg* led to smoothing of results. The results of this experiment were visually pleasing and the presented the capability of the GAI to preserve the structures and avoided the texture garbage to a larger extent. Quality measures were not derived for this experiment as the ground truth was not available for this image.

Experiment: 3.2 In this experiment, see Figure 3.5, text removal problem



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

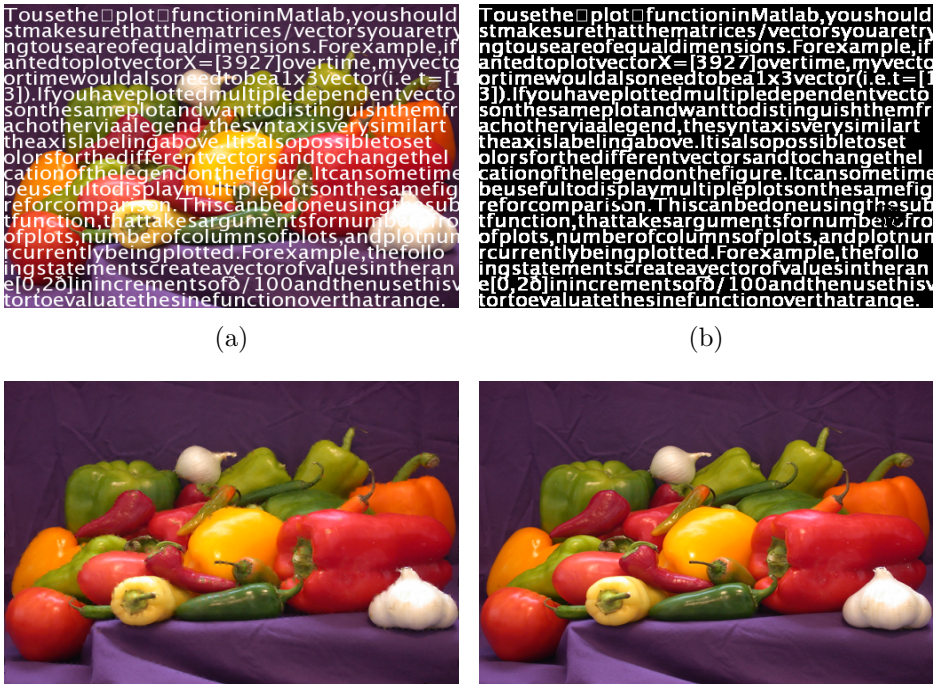
Figure 3.4: (a) The result of GAI with Degree 50 (b) Degree 40 (c) Degree 30 (d) Result of Criminisi model

was solved. Prior to invoking the algorithm text is overlaid on on 'peppers'- a benchmark image covering the 34% of the image. The size of input area was relatively larger, and in addition to the text spanned the entire image and made the problem more ill conditioned. The experiment was conducted by taking a patch of size 3×3. Figure 3.5(c) presented the result of the GAI. The same problem was given as input to the implementation of Criminisi algorithm which could not manage the entire input image and ended up with segmentation fault. The PSNR of the output was 33.38 db and SSIM was 0.947. Table 3.3 presents the reduction of search space in different experiments. The number of source region patches utilized for selecting best exemplar varies from patch to patch based on the size of the granule selected. The column 4 presents the minimum, column 3 presents the maximum number of patches considered for selecting the Best Exemplar, where as the column 2 presents the total number of patches available in the source region.

Table 3.3: The reduction of search space in GAI

Image	Patches	Max	Min
Buggy	88388	8014	1452
Fruits	73147	7700	1716
Children	174923	5923	1230
Bungee	53506	5614	1067

Experiment: 3.3 In this experiment, see Figure 3.6, a widely solved problem



(a) (b)

(c) (d)

Figure 3.5: (a) Input image of size 384 X 512 overlaid with text (b) Mask (c) Result of GAI model (d) Ground Truth

was taken which has only the structural features. It was evident that the GAI could successfully remove the text from the image while preserving the smooth and finer details. The source region has 88388 pixels out of which the GAI processed 1452 to 8014 pixels over all during all iterations in this experiment, which is a substantial reduction in search space size. Observe the Figure 3.6(c), the curvature of the wheel was preserved and at many places the long structures were reconstructed. The quantitative result of this experiment had the SSIM as 0.93.

Experiment: 3.4 In this experiment, see Figure 3.7 a benchmark image with finer textures was taken up to solve scratch removal, see Figure 3.7(a) and small object removal, see Figure 3.7(d) problems. From the results it was very clear that the inpaint region had been filled, see Figure 3.7(c) & 3.7(f), with meaningful content which was very coherent with the surroundings and the PSNR measures realized in this experiment were 28.85 and 29.37 respectively. Though these measures seemed to be inferior, the visual inspection qualified the results to be pleasing

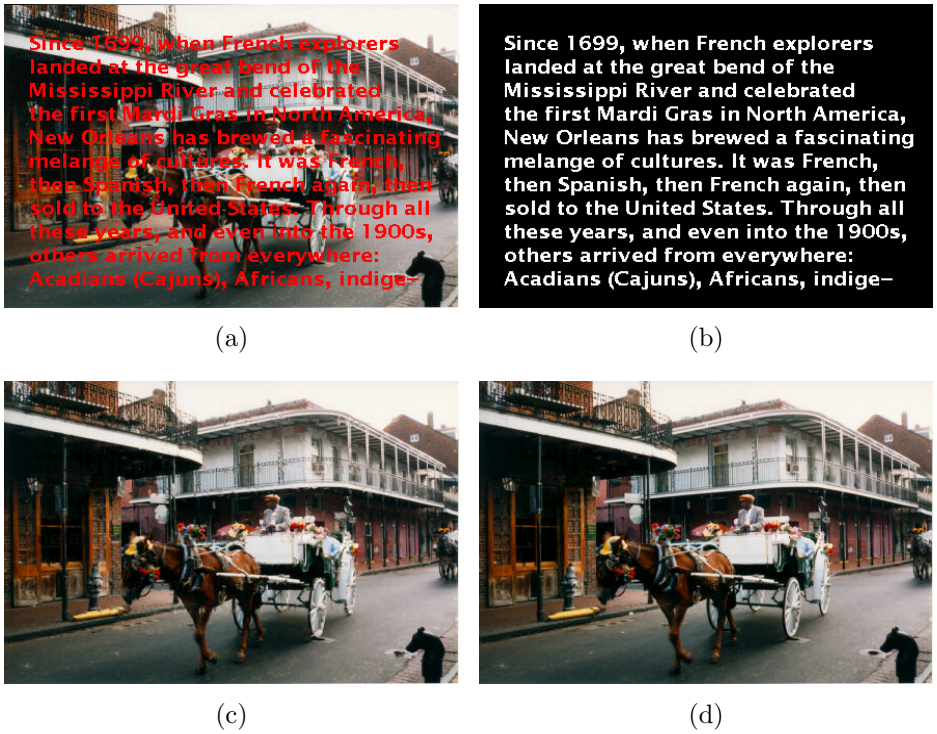


Figure 3.6: (a) Input image of size 297 X 438 with a hole (b) Mask (c) Result of GAI model (d) Ground Truth

ones. The overall capabilities of GAI are presented in Table 3.4 in terms of quality measures.

Table 3.4: The summary of GAI results in SSIM and PSNR measures

Image	SSIM	PSNR	Nature
Buggy	0.93	26.3	Structures and Smooth
Fruits	0.947	33.38	Cartoon
Texture-A	0.93	28.85	Texture
Texture-B	0.93	29.39	Texture

Experiment: 3.5 In this experiment, see Figure 3.7, the scratch removal from an ancient image was carried out. For this image the scratches to be removed were denoted in the associated mask. There were portions ranging from thin lines to relatively bigger patches to be removed. This image exhibited non-stationary behavior along the surroundings of inpiant region. The inpaint region disturbed the very sensitive information around the eyes and chins of the baby girl sitting in the middle, see Figure 3.7(c). The GAI model could reconstruct the sensitive

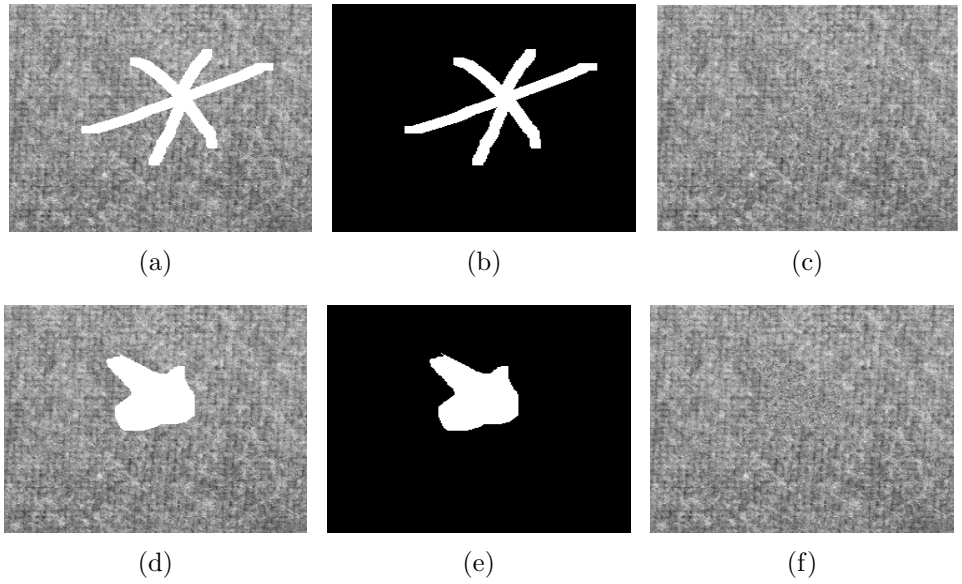


Figure 3.7: (a) Texture-A of size 243 X 243 (b) Mask (c) Result of scratch removal using GAI model (d) Texture-B (e) Mask (f) Result of image completion using GAI model

information, which preserved the original definition of the image successfully.

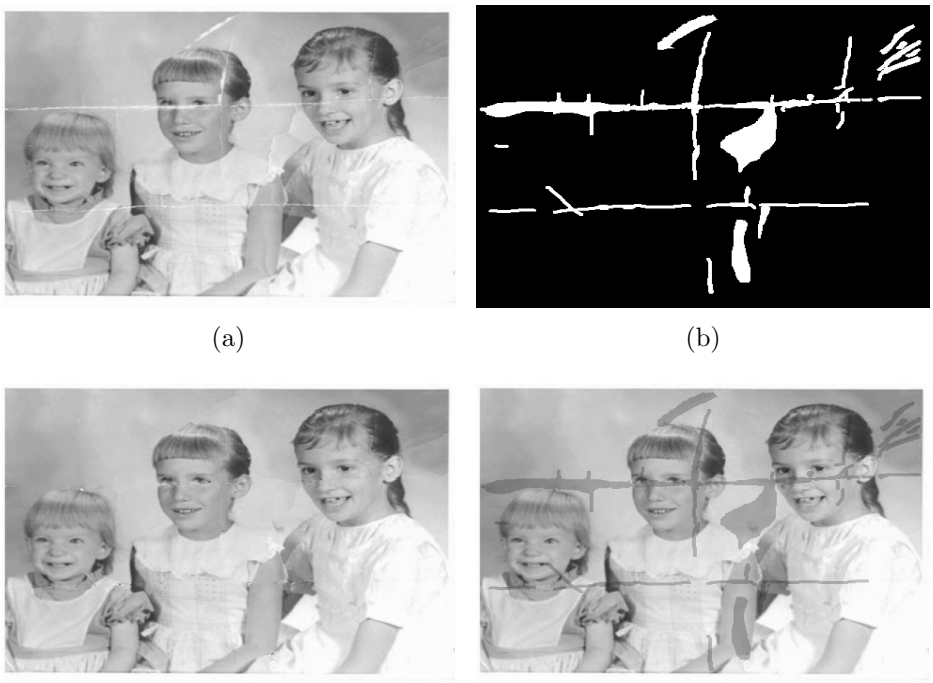


Figure 3.8: (a) Input image of size 405 X 483 with scratches (b) Mask (c) Result of GAI model (d) Input image highlighting the inpaint region

Experiment: 3.6 In this experiment Figure 3.9 an image with fine texture was processed for object removal. The input was composed of finer textures and the GAI was able to remove the object without any visible texture garbages.

(a) (b)

Figure 3.9: (a) Input image of size 362 X 263 (b) Result of GAI

3.8 Summary

This chapter proposed and implemented the GAI- framework which could solve the scratch removal and text removal *inpainting category* problems that fall into multiple but small holes *inpainting class*. The novelty of GAI framework resides in defining the target front and simplification of deriving the priority for each pixel on the target front. GAI incorporated the granule approach for reducing the search space size while selecting a Best-Exemplar which is about 10-60 folds faster than that of original Best-Exemplar method. The L_0 norm reduces the computation time of similarity between target patch and candidate patch as it does not involve square root, yet it is sound in capturing the structural features. The effectiveness of GAI is established from the results of the experiments that are quantitatively good and visually plausible. But in some cases, because of its greedy nature, GAI was able to address only local structures not the over all unity. Hence there is a need to incarcerate the structure aware feature. The framework can be extended to solve large scale object removal problems using *Kd*-trees, *vp*-trees, etc.

Chapter 4

Spatial Anisotropic Interpolation Approach

Literature projects the efforts made by researchers to inpaint the natural images with texture and structures by separating them as independently modeled components and processing them individually to solve the inpainting problem. The partial results are then subjected to a fusion algorithm so as to get the overall result. Another set of algorithms handled this issue by representing the textures and structures simultaneously in a sparse manner. But, both of these schools of thought were computationally expensive and resource hungry. Hence, a methodology for efficient inpainting encapsulating structures and textures is desirable. Geo statistics is an area which models the underlying random process of any system in terms of well-defined statistical properties and offers a bunch of inference algorithms both in parameter space and data space.

This chapter is devoted for a Geo-statistical framework to solve the inpainting problem. It starts with an overview of random fields and their characterization. Subsequently, the existing classes of inpainting algorithms are investigated in 4.2 that emulate an essential physical phenomena 'anisotropy'; a major topic of interest in image inpainting domain. Then, Design and Analysis of Computer Experiments (DACE) framework for expressing the spatial anisotropy, in a quantitative manner, through spatial correlation is discussed in 4.3. The inpainting problem is

modeled by solving the two convolved regression mechanisms within the Ordinary kriging framework in 4.3.1. The model building of Kriging based on the proposed Systematic Sampling technique is illustrated in 4.4. The text removal aspects, through the inferences from the moving average based *Spatial Anisotropic Interpolation* (SAI) model, in an algorithmic form, are elaborated in 4.7. The proposed SAI model is evaluated by conducting the experiments on benchmark images. The quality of results is expressed using PSNR and SSIM metrics and compared with other anisotropy related methods in 4.8. The last section 4.9 paves the way to subsequent sections.

4.1 Random Fields

Computer Vision often models the pixel gray value with a notion of a random variable, out of mathematical necessity. And each random variable is associated with a probability density function with a well defined mean and variance. This association offers flexibility for the imputing algorithms to carry out their task of missing value prediction with some level of variance which is expressed as the error. Further, these algorithms can be made to possess the capability to manage the error by modeling it over some probability distribution. For better handling, even the error term is also represented as a random variable [Cressie, 1993].

It is also possible to extend this association to a collection of pixels arranged in space or a rectangular form. Therein the pixel values constitute the responses, which are in turn associated with random variables; then the resulting model is called as a Random Field. On top of this arrangement, we can even impose certain structural properties and foster the random variables to interact with each other, in a controlled manner [Wackernagel, 2003]. Literature has a rich collection of structure learning models realized by sampling properties like Gibb's Sampling [Geman and Geman, 1984] in which the model parameters are learned from the data itself. Once the model building is completed, then the imputation algorithms infer the missing values. In many situations, the underlying problem will be

ill-posed. That means the model can impute multiple solutions for the given problem instance. Then, the problem turns out to be combinatorial optimization problem [Hastie et al., 2001]. In many situations, it is a standard technique to add a regularization term into the model and to fine tune the imputations. The regularization usually is expressed as the prior knowledge of some physical phenomena which reduces the artifacts in the results. For example, the prior can be about the designing the interaction patterns and imposing it onto the model. Within the context of image inpainting problem, it is quite essential to assume further that the underlying field exhibits some sort of stationary characteristics, so that the interactions and relations among pixels become well-defined [Black et al., 1998]. Particularly, the analysis of the interaction pattern is restricted to only two locations and hence it is well to express the same by the correlation among the pair of pixels. Next subsections formalize all these models.

4.1.1 Second Order Stationary and Intrinsic Random Fields

Formally speaking, the Random Field $Z(X)$ is a spatial arrangement of random variables $Z(x_i), i = 1 : n$ associated with spatial location x_i. A particular instance $z(x_i)$ of the random variable $Z(x_i)$ is called as Regionalized Variable. As mentioned before, to ensure that inference process is robust, certain spatial properties are imposed on the Random Field.

Definition 4.1- **Stationary Random Field:** Z(X) is called a stationary if the distribution associated with this field is independent of the location. Hence the parameters of the distribution such as the mean and variance remain the same over changing locations. This assumption is not a rational one to be imposed on natural Images, because any natural image does not have all pixels possessing same intensity value throughout, i.e. the variable is not persistent. Hence second order stationary property is introduced which is defined as follows.

Definition 4.2- **Second Order Stationary Field:** The random field Z(X) is second order stationary if the three following assumptions are satisfied:

1. The mean is constant: i.e. $E[Z(X)] = \mu$

2. The variance is invariant over position: i.e. $Var[Z(X)] = \sigma^2$
3. The covariance depends only on the absolute distance h between the locations of pair of sites:

$$COV(Z(x_i), Z(x_i + h)) = COV(h).$$

From the properties of covariance function mentioned above it is evident that if $h = 0$ then $COV(0) = Var[Z(X)] = \sigma^2$. This property ensures that the neighboring pixels separated by small lag h are highly correlated. With an intent to turn this stationary property into a tool with a purpose and bring it more close to real world situation, intrinsic property is introduced.

Definition 4.3- **Intrinsic Random Field:** The random field Z (X) is intrinsically stationary if the following assumptions are satisfied:

1. The mean is constant:

$$E[Z(X)] = \mu$$

2. The variance over the lag h depends only on h i.e.

$$var[Z(x_i) - Z(x_i + h)] = 2\gamma(h)$$

where 2γ is called variogram. The variogram represents the spatial continuity of the random field in a pictorial form and it offers an empirical means to assess the average dissimilarity between two points that are separated by lag h. For solving high-level problems like classification, a functional form can be adapted to the variogram. If covariance function is defined over a random field, then the intrinsic and second ordinary random fields are related through the following equation

$$\gamma(h) = C(0) - C(h).$$

Definition 4.4- **Positive Definiteness:** The covariance functions $C(h)$ is valid if and only if it is positive definite. A matrix C is positive definite if for every

non-zero vector a the expression

$$a^\top C a \geq 0$$

is true [Shewchuk, 1994]. This property ensures that C is invertible, and C^{-1} is evolved through Cholesky Decomposition [Cressie, 1993]. This offers a computational advantage over the direct inversion method.

4.1.2 Introduction to Metamodels

A metamodel [Cressie, 1993] is a notion that is capable of combining the capabilities of numerical methods with a stochastic process. It provides an abstraction of a latent physical phenomenon which is realized either through interpolation or approximation. Examples of meta-models include Response Surface, Spline, Radial Basis Functions, Kriging, Neural Networks, etc. The utility of a chosen metamodel is, it acts as a surrogate to the underlying actual experiment. The immediate advantage with these surrogate models is, it is not required to conduct the experiment physically at all design sites to get the responses, but can be predicted with a minimum Mean Square Error (MSE). Before taking up construction of the metamodel for solving inpainting problem, we have a relook at its characteristics and map them onto a random field.

4.2 Anisotropy as a Characteristic of Inpainting

Any natural image poised to have a lot of information in it. The finer and essential features one can observe from any natural image are; Though the neighboring pixels are highly correlated with each other, still there may be many slowly varying features. These features can be about showing a fractal nature, representing cartoon areas or containing textural elements. From inpainting point of view, we expect the interpolating algorithm to possess both the structure preserving and texture interpolation capabilities; perhaps all existing algorithms fail in this re-

spect. Thus the inpainting algorithm should formally incorporate these aspects into the design and perform considerably well. Researchers haven't given serious attention to exploiting the knowledge representation capabilities of the random fields and the ergodic feature readily available in digital images to address the inpainting problem using anisotropy models.

4.2.1 Anisotropy

To address the design challenges of inpainting problems, as discussed in Chapter 2 various physical phenomena are modeled and incorporated into the solution space. One of the common behavioral features observed in existing models is, they effectively tackle the anisotropy aspect and exploit it to spread the information. The quality of the inpainting result depends on the class of the problem and the scale of the anisotropy property, which governs the selection of the operational tool. In some cases, the entire image data shows a systematic behavior across the image and the notion of the trend that judiciously models such global features. There is a possibility that the components present in the image also possess small scale features and show a regular pattern. Local operators such as gradient, incorporated in the inpainting, can make use of sophisticated mathematical tools to extract these patterns present around the inpaint region and attempt to propagate that information along different directions in various amount. These tools are, in fact, borrowed from Computational Fluid Dynamics and from Variational Calculus that are useful to solve network flows and heat diffusion problems. The various forms of anisotropy- diffusion and total variation, realized in literature are discussed below.

4.2.2 Anisotropic Diffusion

Authors [Bertalmio et al., 2001] [Grady and Schwartz, 2003] solved the inpainting problem by diffusing the information available at boundary of inpaint region, in different amounts in different directions and called it as anisotropic diffusion. The

Algorithm 4 presents the numerical implementation of the anisotropic diffusion. [Bertalmio et al., 2001].

Algorithm 4 Anisotropic Diffusion

Input: Input Image I and the associated Mask M

Output: Inpainted image

1. $I_{t+1} = I_t + step * \frac{\partial I}{\partial t}$
2. $\frac{\partial I}{\partial t} = \nabla(\triangle).\nabla^{\perp} I$
3. `stop if condition` $\frac{\partial I}{\partial t} = 0$ `otherwise goto step 1`

The Anisotropic diffusion algorithm diffuses the Laplacian of data along the isophotes that are collinear with the normal at the point on the boundary. This algorithm propagates the information across the edges in the image and ends up smoothing the edges. Authors in [Grady and Schwartz, 2003] realized the anisotropic diffusion by solving the Laplace Equations with Dirichlet Boundary conditions in a combinatorial manner. They have utilized the combinatorial versions of diffusion operators and launched them on a graph. The graph construction itself addresses the anisotropy property by giving different weights to different edges by using Gaussian Weighting function. These methods could realize the inpainting only on finer gaps and small holes.

4.2.3 Anisotropic TV

Tony F Chan [Chan et al., 2002] modeled the inpainting by extending the object disocclusion solution of Masnou and Morel [Masnou and Morel, 1998]. Instead of using the Laplacian operator, due to Bertalmio, Chan applied the Gradient operator as shown Equation 4.1.

$$E(I) = \int_I |\nabla I||dxdy; \qquad I = f \in \Phi. \tag{4.1}$$

Subsequently, recasting of the Rudin Osher Fethami [Rudin et al., 1992] model for solving spatially varying denoising problem resulted in a solution to inpainting problem as

$$E(I) = \int_{\Omega} |\nabla I| dxdy + \frac{1}{2} \int_{\Phi} \lambda(x)(I - f)^2. \tag{4.2}$$

The minimization of the energy function in Equation 4.2, $E(I)$ is intractable because it involves two regularization terms defined regarding L_1 and L_2 norms respectively. Osher [Goldstein and Osher, 2009] redefined this model by introducing an auxiliary variable d, involving ∇I as

$$min_{ij} \sum_{i,j} d_{i,j} + \frac{1}{2} \sum_{\Phi} \lambda(x)(I - f)^2 \qquad \text{subject to} \qquad d = \nabla I. \tag{4.3}$$

Equation 4.3 represents a constrained problem and is solvable through Split Bregman Method [Goldstein and Osher, 2009] which is available in two flavors- isotropic and anisotropic. Authors in [Getreuer, 2012] implemented the anisotropic TV minimization problem and solved the inpainting instance. This method preserves the edges only in case of smaller and narrow inpaint regions and fails to synthesize the textures.

4.3 Proposed Anisotropic Interpolation Approach

Consider an image I of size $a \times b$ having source region Φ and inpaint region Ω. Then the source region pixel locations are represented as a set of design sites $X = [x_1, ..., x_m]^T$; $x_i \in \Re^n$ and the set of corresponding pixel values, called as responses, is represented as $Y = [y_1, ..., y_m]^T$; $y_i \in \Re^q$. With regard to image inpainting problem we have $n = 2$ and $q = 1$ (in case of gray image) or $q = 3$ (for color image). Then an attempt to solve the inpainting problem by casting it as a conventional interpolation problem fails due to various reasons. Regular interpolations can impute or estimate the missing values at inpaint locations based on the pixels values available in the input sample at the design sites and equal weightage is given to all design sites. Regression is a mechanism to predict or impute the

dependent variable from known explanatory variables. And the mainstay of the regression theory is the Linear Regression, which fits a trend surface trough the design sites using polynomials and models the underlying pattern. The order of the polynomial and the nature of the polynomial have a lot of bearing on the ability of the model during predictions. A lower order polynomial suffers from the problem of underfitting the trend whereas a higher order one leads to overfitting with high likelihood. Also, these methods don't consider the spatial location of the design sites at which the imputation is required. That means the model built by such an interpolation scheme is very sensitive to the outliers [Gentile and Meylan, 2013] [Santner et al., 2003].

4.3.1 Design and Analysis of Computer Experiments

Design and Analysis of Computer Experiments (DACE) [Lophaven and et al., 2002] being a knowledge representation metaphor, offers a 'surrogate computer model' feature for any computer experiment. This model is endowed with Kriging approximation, for data resulting from computer experiment. Here, a computer experiment corresponds to a collection of pairs. Each pair includes an input location and the associated response from different runs of a computer model. Both the input and the response from the computer model are likely to be in higher dimensions.

Having introduced to some simple definitions related to Random Fields and meta models, now kriging model is derived for missing value imputation which exploits the spatial correlation of the underlying Random Fields. This phenomenon is called spatial interpolation [Cressie, 1993]. Kriging is superior to ordinary regression in two respects. The first one is that it relates the independent variable with dependent and establishes the relation between the dependent variable and the spatial lag. The second one is its capability to assess the quality of prediction with estimation on prediction error [Stein, 1999]. These are realized through a combination of Polynomial and Gaussian Regressions. The former one represents the structural component of the random field through a global feature namely,

trend. The latter one models the error term, which denotes the deviation of individual prediction from the trend, by addressing the spatial correlation component. Here is where kriging scores a point over regular regressions. In general regressions, the prediction errors are modeled through white noise (see section 2.5.2). But Kriging assumes the error as a random field which is modeled as a stationary and correlated field. Its capability to give more accurate results is due to adapting its parameters with the changing inputs, in contrast to the regular regression which uses a single estimated parameter list [Grady and Schwartz, 2003]. That means, kriging model fitting imposes a rule that spatial correlation between neighboring points is more when compared to the farther points. Hence this interpolation is a local method. Further, the trend plays a vital role in selecting an appropriate kriging model for the interpolation problem at hand.

4.3.2 Trend Modeling

Widely used kriging models are simple (SK), ordinary (OK) and Universal (UK) [Burrough and A, 1998]. In the case of simple kriging, the trend has a constant known mean throughout the design space. In ordinary kriging, the trend is modeled as a constant but with unknown mean. Universal kriging assumes that the trend is a regression polynomial F of known order but with unknown coefficients, β. This model is also suitable to handle the non-stationary property which is further elaborated in the next chapter. Universal Kriging model can predict the response $\hat{y}$ at a design site x as

$$\hat{y}(x) = z(x) + \sum_{j=1}^{p} \beta_j F_j. \tag{4.4}$$

The first term $z(x)$, $x \in \Re^n$ on the right hand side in Equation 4.4 models the residual, a regionalized variable of random field $Z(X)$. The residual is assumed to follow the second order stationary with zero mean and finite covariance $\sigma^2 R(\theta, x_i, x_l)$, for $i, l = 1...m$. Note that here instead of covariance, the spatial correlation R with θ as its parameter is employed as the variogram construction for the anisotropic

process is difficult. The latter term denotes the trend which is modeled through a known basis function F with unknown coefficients β_j. Thus, kriging simultaneously models the given trend and the correlation structure [Stein, 1999]. The fundamental working principle in kriging is to launch a Bayesian approach to model the uncertainty about the pixel value as a stationary field in the form of Gaussian random field [Grady and Schwartz, 2003]. In polynomial regression, if the known basis is a set of polynomials, then the corresponding trend is a regression polynomial. For text removal problem the universal kriging is selected. This aspect is substantiated in the upcoming discussions.

Now, the objective is to tune the unknown regression coefficients β_j, such that the minimum residual at a new point during prediction step is possible. Thus kriging exploits the error term in the modeling process. This is realized in two stages. The first step is to ensure that residue is unbiased. Then the second step is to solve the polynomial regression to give unbiased solution while honoring the spatial correlation. This step is crucial for regularizing the ordinary least square solution.

To simplify the discussion, consider only a scalar response at a given design site in $2-D$. If the trend functions F is expressed as a full second-order multivariate polynomial

$$F(x_i, y_i) = 1 + x_i + y_i + x_i^2 + y_i^2 + x_i y_i$$

and the nature of predictor is restricted to be linear then the linear regression model can predict the value at new site x as $\hat{y}(x) = C^T Y$ with $C \in \Re^m$. Then, Equation 4.5 models the error at x as

$$\begin{aligned} \hat{y}(x) - y(x) &= C^T Y - y(x) \\ &= C^T(F\beta + Z) - f(x)^T\beta + z \\ &= C^T Z - z + (F^T C - f(x))^T \beta. \end{aligned} \tag{4.5}$$

To turn the predictor as the unbiased, Equation 4.6 gives a set of constraints to be satisfied.

$$\begin{bmatrix} c_1 & \dots & c_m \end{bmatrix} \begin{bmatrix} 1 & x_1 & y_1 & x_1^2 & y_1^2 & x_1y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & y_m & x_m^2 & y_m^2 & x_my_m \end{bmatrix} = \begin{bmatrix} 1 & x_0 & y_0 & x_0^2 & y_0^2 & x_0y_0 \end{bmatrix} \tag{4.6}$$

i.e.,

$\Sigma_{i=1}^{n} c_i = 1$ (Ordinary kriging unbiasedness condition)

$$\left.\begin{array}{l} \Sigma_{i=1}^{n} c_i x_i = x_0 \\ \Sigma_{i=1}^{n} c_i y_i = y_0 \\ \Sigma_{i=1}^{n} c_i x_i^2 = x_0^2 \\ \Sigma_{i=1}^{n} c_i y_i^2 = y_0^2 \\ \Sigma_{i=1}^{n} c_i x_i y_i = x_0 y_0. \end{array}\right\} \text{(additional UK unbiasedness conditions)}$$

Upon introducing the collocation matrix $F = \{F(x_1)...F(x_m)\}^T$, Correlation Matrix between every pair of design sites as $R = \{R(x_i, x_l)\}$ for $1 \leq i,\ l \leq m$ and Correlation Matrix between every design site s and an untried location x as $r(x) = [R(x, x_1), ...R(x, x_m)]^T$, now Equation 4.5 can be simplified to

$$MSE(\hat{y}(x)) = \sigma^2(1 + C^T(x)RC(x) - 2C^T(x)r(x)). \tag{4.7}$$

Subject to constraint in Equation 4.6.

Then the corresponding Lagrange equation of the minimization function in Equation 4.7 with constraint is expressed as

$$L(C, \lambda) = \sigma^2(1 + C^T(x)RC(x) - 2C^T(x)r(x)) - \lambda^T(x)(F^TC(x) - f(x)). \tag{4.8}$$

The minimization problem is solved by equating the gradient of Equation 4.8 with

respect to C, to zero which results in a system of equations

$$\begin{bmatrix} 0 & F^T \\ F & R \end{bmatrix} \begin{bmatrix} \hat{\lambda} \\ C \end{bmatrix} = \begin{bmatrix} f \\ r \end{bmatrix} \tag{4.9}$$

where $\hat{\lambda} = \frac{\lambda}{2\sigma^2}$.
The solution of Equation 4.9 for $C(x)$ gives

$$\begin{aligned} \hat{\lambda} &= (F^T R^{-1} F)^{-1} F^T R^{-1} r - f \\ C &= R^{-1}(r - F\hat{\lambda}). \end{aligned} \tag{4.10}$$

Then the model predicts the response of the mode at x from Equation 4.10 as

$$\hat{y}(x) = r^T R^{-1} Y - (F^T R^{-1} r - f)^T (F^T R^{-1} F)^{-1} F^T R^{-1} Y. \tag{4.11}$$

Equation 4.11 gives the unbiased estimate of $y(x)$. With this the implementation of the Gaussian error regression is complete. Now the second regression aspect, multivariate polynomial regression is taken up. The Generalized Least Square Solution (GLS) [Lophaven and et al., 2002] to the polynomial regression

$$F\beta = Y \tag{4.12}$$

gives

$$\hat{\beta} = (F^T R^{-1} F)^{-1} F^T R^{-1} Y \tag{4.13}$$

The solution given in Equation 4.13 is superior to Ridge regression [Hastie et al., 2001] which is about regularizing the Ordinary Least Square (OLS) estimation of β, for a parameter λ with the quadratic term as

$$\min_\beta (\frac{1}{2} ||Y - \beta^T X||^2 + \lambda ||\beta||^2). \tag{4.14}$$

The solution of this objective function,

$$\beta = (XX^T + \lambda I)^{-1} \lambda Y^T \tag{4.15}$$

simply inflates the covariance matrix, in parameter space by a constant amount in all directions which favors the smaller values of β by rotating the solution vector towards them. Hence Ridge regularization is biased. In contrast to this observed biased phenomena, GLS offers an unbiased solution by regularizing the OLS solution based on the spatial correlation R. This term makes the GLS unbiased. The computational advantage with GLS is, we can approximate the variance-covariance information with the Correlation and implement the kriging model. Then the overall predictor from Equations 4.14 and 4.15 is given in 4.16

$$\hat{y}(x) = F^T \hat{\beta} + r^T R^{-1} (Y - F\hat{\beta}) \tag{4.16}$$

with a variance of estimation error as expressed in Equation 4.17

$$\sigma^2(x) = \frac{1}{m} (Y - F\hat{\beta})^T R^{-1} (Y - F\hat{\beta}). \tag{4.17}$$

I, now summarize the similarities and differences between the standard linear regression model and the DACE model. The two models share a common mathematical framework consisting of regressors and errors, but the emphasis is quite different. Linear regression focuses entirely on the regressors and their coefficient estimates and makes simplistic assumptions about the errors (independence). In contrast, DACE makes simplistic assumptions about the regressors (just a constant term in the simple case) and focuses entirely on the correlation structure of the errors. Thus, regression and DACE are probably best thought as diametric opposites. Regression is about estimating regression coefficients that (together with the assumed functional form) completely describe what the function is. DACE is about estimating correlation parameters that describe how the function typically behaves.

The salient features of kriging interpolation are listed below [Gentile and Meylan, 2013]

- Predictions based on a spatial statistical analysis of the data
- Best Linear Unbiased Estimator (BLUE)
- Many forms of Kriging are available, applicable to various data configurations
- Automatically accounts for clustering
- Remains efficient under conditions of sparse data
- Can take into account variation bias toward specific directions (anisotropy)
- Able to quantify interpolation errors (Kriging variance)

4.3.3 Sampling Scheme

Thorough knowledge about the stationary properties of the data is critical for selecting an appropriate interpolation method and fine-tuning of the results produced by the experiments. The characteristics of the data either directly measurable or indirectly derived from the data play a crucial role in selecting an appropriate interpolation model and to analyze the results produced by the experiments. Such characteristics include spatial representativeness, sampling scheme, data collection mechanism and its accuracy, and establishing the existence of spatial relationships among the data items. The basic feature of spatial interpolation is to augment the pixel to pixel information with the spatial interactions among the pixels as a secondary piece of information and use them to interpolate at new locations. If the number of source region pixels is very high, then for the interpolation model, maintaining them on suitable data structures is a space complexity issue, which turns out to be an important point to be addressed more carefully in case of kriging based models. Kriging involves computation of inverse for large scale correlation matrix and learning the model parameters in an iterative manner

resulting it to be a memory-hungry surrogate. This aspect refers to as "curse of dimensionality" in the literature. Hence to manage this situation, it is essential to select a subset of representatives from the entire set of responses available, without losing the essential patterns present in the data of the underlying experiment. Robust sampling algorithms are necessary to address this problem. The scale is the ratio between the number of the informative pixels maintained on the experiment to the original number available. This measure quantitatively speaks about the area a sampled pixel represents in the original image. There are a lot of sampling schemes designed for changing the scale. Concerning upscaling, the sampling aims at increasing the support which is equivalent to decreasing the spatial resolution.

4.3.4 Latin Hypercube Sampling (LHS)

LHS is a statistical method to generate the sample from multidimensional distributions [McKay, 1992]. Consider a problem of generating a sample of size four from a 2-D image of size $M \times N$. The first step in LHS is to divide its dimensions into four equally probable intervals along row and column, resulting in $4 \times 4 = 16$ bins, see Table 4.1. Then generate the four samples such that only one sample is drawn from each row and column ensuring exactly four samples. Thus LHS involves deciding the number of sample points required and for each sample selected, remember the row and column. This sampling is not suitable for inpainting as it cannot capture the structural properties of the image.

Table 4.1: LHS Demo

	Col1	Col2	Col3	Col4
Row1	*			
Row2			*	
Row3		*		
Row4				*

4.4 Proposed Systematic Sampling scheme

The prediction variance, which is already discussed in Equation 4.17, is not just dependent upon the observed pixels themselves, rather it is dependent upon the spatial arrangement of these pixels. This fact is exploited to manage the memory related aspects effectively. In the proposed Systematic Sampling scheme the image is subjected to raster scanning along rows and columns respectively. The pixels from the source region are scanned from left to right and then from top to bottom for sampling the pixels, that are separated by a predefined lag either 1 or 2 or 3. Based on the number of design sites available, shape and size of inpaint region the lag is determined. This aspect is depicted in Figure 4.2. Upon completion of one complete scan, for lag 2 or 3, the size of design space gets diminished by 4 or 9 folds, but the number of inpaint region pixels remain the same. Effectively, the DACE model is built using sampled source region pixels alone, which requires less space. The underlying assumption that the natural images conform to the intrinsic random field property and favors the proposed sampling scheme. According to this assumption, the nearby pixels are highly correlated. So instead of working on such redundant responses the proposed model suggests to include only the subsample of source region, which contributes to the desired image feature like structure or texture, into the model fitting process. In Figure 4.1, the proposed systematic sampling is presented for lag one and two, respectively along one direction. The row 4 and row 3 show the scenario of having consecutive source region pixels and the sampling with lag one and two respectively. The row 2 and row 1 depict the presence of inpaint region pixels and their implication on Systematic Sampling with lag 1 and lag 2. Among all source region pixels based on the sampling frequency, the continuous source region pixels are collected to constitute the design space and utilized for model building. But, about inpaint region pixels, all of them are considered at the time of prediction.

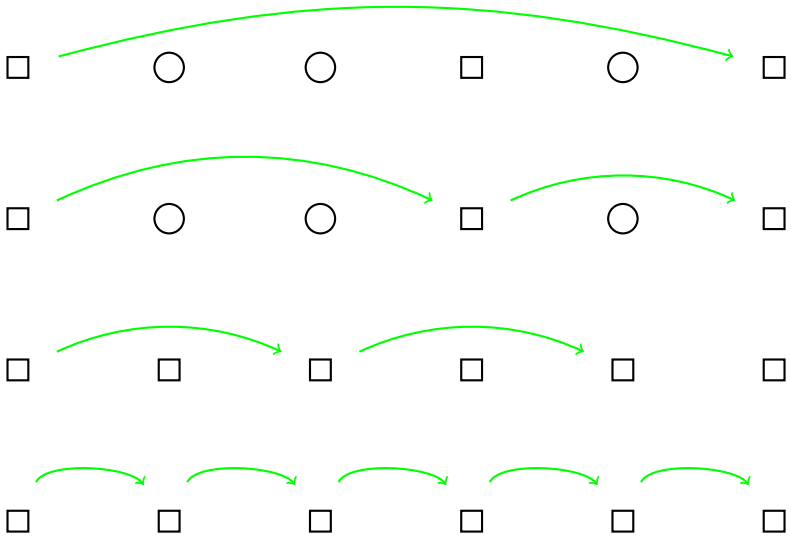


Figure 4.1: The proposed Systematic Sampling Scheme- □- represents the source region location and ∘- represents the inpaint region location

4.5 Trend as Abstraction of Global Features

Natural Images are composed of a lot of regions showing different spatial features. The image as a whole might exhibit some trend, alongside the small scale features of it also may show some local trends. Even in such mixed situations, the polynomial regression in kriging predicts the missing values accurately because of its adaptivity to the space varying trends. These aspects are, further discussed and demonstrated in section 4.8. Universal Krigig (UK) models the trend through a polynomial basis and solves the associated regression problem using GLS. The selection of the polynomial basis is a crucial issue here. Unfortunately, the nature and order of the polynomial shall be determined in a subjective manner. If a lower order polynomial is fitted, then it may result in underfit leading to high variance in the prediction [Grayson and Blschl, 2004]. On the other hand, if higher order polynomial is fit then it may represent the underlying trend in an imperfect manner which causes the overfitting [Grayson and Blschl, 2004]. If p is the order of the polynomial F and if the cross terms are involved then the number of coefficients of the polynomial basis is β_n is $n = \frac{(p+1)(p+2)}{2}$. In case of Ordinary Kriging (OK) the trend is modeled as a piecewise constant.

4.6 Quantitative Measure for Spatial Anisotropy

The aspect of spatial interpolation is intimately related to the spatial correlation among the data items at the design sites. If the correlation pattern changes with distance and direction, then the data is said to show the anisotropy behavior [Oliver et al., 1989]. This dependency is verifiable by plotting the empirical variogram, and it is possible to model it through an appropriate theoretical function and the corresponding variance-covariance matrix is constructed.

4.6.1 Correlation as a Measure for Spatial Anisotropy

In DACE environment this aspect is simplified by employing the correlation as the model for expressing the anisotropy. The involved issue in kriging is to learn the correlation model from the data. DACE uses a wide variety of theoretical correlation functions that are available in the literature. Table 4.2 gives few correlation functions in 1-D with θ as the parameter. The correlation parameter θ

Table 4.2: Variants of Correlation Models with $d_p = |x_p - y_p|$

Input	Form	
Exp	$EXP(-\theta_p \lvert d_p \rvert)$	
Gauss	$EXP(-\theta_p {d_p}^2)$	
Lin	$max\{0, 1 - \theta_p \lvert d_p \rvert\}$	
Cubic	$1 - 3{\xi_p}^2 + 2{\xi_p}^3$	$\xi_p = min\{1, \theta_p \lvert d_p \rvert\}$

is the model parameter and d_p denotes the absolute difference between the spatial locations x_p and y_p. If θ is a scalar then it then the correlation among spatially arranged pixels is dependent upon only the distance, not the direction. This kind of spatial phenomenon is called as called as isotropic correlation. On the other hand, if θ is a 2-D vector, then it insists that it also depends on direction, which induces anisotropic property. The major issue in kriging and widely discussed in the literature as 'epistemic uncertainty' [Lockwood and Anitescu, 2012] which arises from the lack of knowledge about the actual value of the parameter. Hence it is essential to learn the parameter θ, associated with the model from the data

itself. DACE model assumes that the error follows the Gaussian model with θ as an unknown parameter. The optimal choice of θ is obtained by Maximum Likelihood Estimation (MLE) principle best implemented in DACE [Ingalls et al., 2004] as the maximizer of

$$-\frac{1}{2}(m\ ln\sigma^2 + ln|R|) \tag{4.18}$$

based on Hooke and Jeeves [Hooke and Jeeves, 1961] pattern search mechanism which is free from the computation of derivatives. Typically, this parameter learning algorithm is executed on a sample of data collected from LHS as it can better handle epistemic uncertainties. In the case of n dimensional data, the anisotropy is expressed as the tensor product of correlation functions defined along each dimension. For example in case of 2-D exponential correlation the basis can be defined over the distance d as

$$R(x, y) = \prod_{p=1}^{2} \exp^{(-\theta_p|d_p|)} \tag{4.19}$$

From Equation 4.19, it is quite simple to observe that for a significantly large value of θ_p only the very neighbor by pixels are correlated, and the correlation decreases exponentially with distance for a given θ_p. Hence proper modeling of both these parameters is indispensable to develop a surrogate with high fidelity. Another important aspect is about modeling the distance parameter which is possible in many ways. In its simple form, absolute diffidence of the individual coordinate is taken as the distance. On the other hand, sophisticated methods create a manifold on the input image and the distance between two locations is computed using Dijkstra's algorithm [Tenenbaum et al., 2000] instead of just the absolute differences of ordinates and abscissa.

4.6.2 Regularization

The fidelity of the DACE is due to its capability to model the Universal Kriging through developing Gaussian regression and polynomial regression simultaneously.

The error is assumed to possess the stationary behavior. The associated regression is modeled to be unbiased. But the OLS estimates of the standard error are biased and hence to achieve the compatibility between the two regressions, it is suggested to regularize the polynomial regression about the correlation matrix [Lophaven and et al., 2002]. Thus the polynomial regression turns into a generalized and unbiased. Such a regularized solution also shows certain essential characteristics that are useful to solve large scale inpaint problems as discussed in the next chapter.

4.6.3 Proposed Moving Window Approach

Text removal problem seeks a model which considers only locality during inpainting. Hence it is sufficient to propagate the information that is available in the vicinity of the inpaint region and preserve the trend observed in the source region. In addition to that, if overlaid text spans the entire image covering the major portion of image pixels, then, due to the memory hungry feature of DACE, it is not possible to develop a single kriging model. These two aspects favor the sufficiency to develop the DACE model only on every non-overlapping sub-image of appropriate size and to fit Ordinary Kriging to represent the trend. Thus the moving window based DACE model can solve the large scale text removal problem and is named as *Spatial Anisotropic Interpolation* (SAI) approach.

4.7 The Complete Algorithm of SAI

Let I be input image of size $a \times b$ and the corresponding mask is M in which each inpaint region location is given the value 1 and source region location is assigned a value 0. Then extract all non-overlapping sub-images of size $k_1 \times k_2$ (in our experiments 30×30) from image I, and corresponding sub-masks form the mask M. Then SAI is applied on each sub-image as listed below, which involves only one index variable i so that the notations appear to be simple.

Algorithm 5 Spatial Anisotropic Interpolation (SAI)

Input: Set of all tiles I_i and M_i extracted from I and M

Output: Inpainted image

1. Repeat steps 2 to 10 for each sub-image I_i and M_i
2. Extract source region Φ_i and inpaint region Ω_i
3. Apply Systematic Sampling on source region pixels §4.4
4. Choose anisotropic correlation kernel R
5. Apply ML on sample to learn correlation parameter (θ)
6. Select a regression model φ of interest.
7. Fit DACE model for chosen sub-sample, φ, R and θ
8. Normalize the inpaint locations and put them on a linear array
9. Predict the responses at inpaint locaions by applying DACE model
10. Retain the values obtained in step 9 into inpaint region of I_i

The Algorithm SAI is applied on sub-images for which the ratio ρ between the size of inpaint region and source region is within the range $0 \leq \rho \leq 0.4$

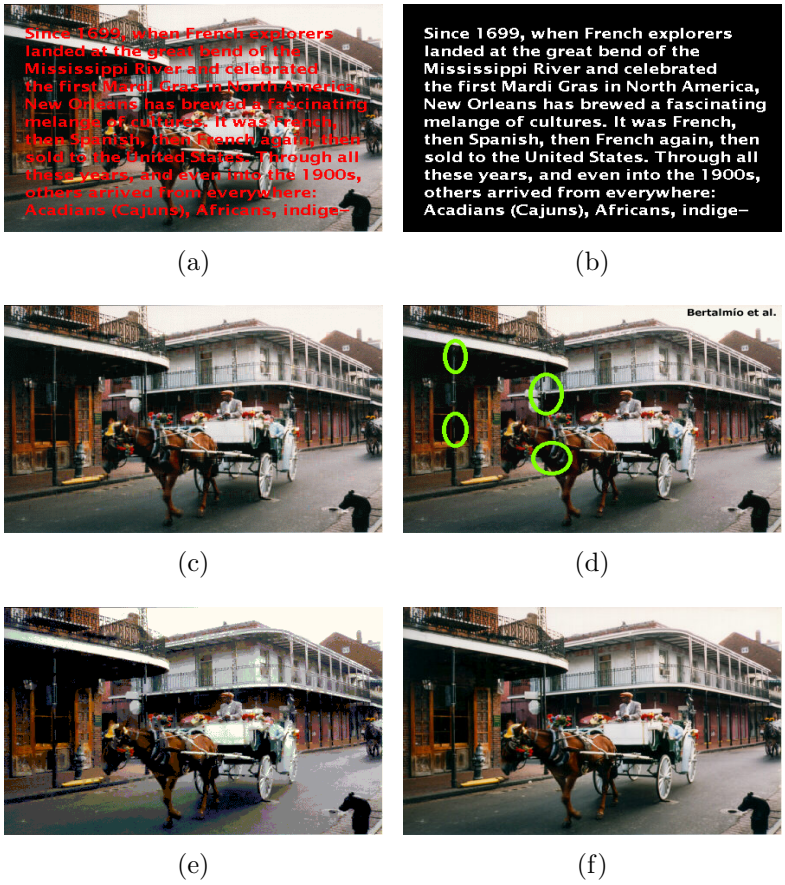
4.8 Experiments and Results

Experiments were conducted on a wide set of images with structural and texture elements on which text was imposed. Addition of text to original image in experiment 2 and experiment 3 was achieved by using tools developed by [Getreuer, 2012]. The proposed model had been exhaustively run while utilizing different correlation kernels and regression polynomials of different orders. The results presented below involved exponential correlation model and constant polynomial

regression. Upon experimenting with a bunch of images it is observed that most natural images exhibit higher correlation along the rows than the columns. This observation led us to initialize the θ appropriately and to avoid the invocation of ML based parameter learning algorithm on each tile during DACE model building. In order to compare the results of proposed SAI with standard models available in literature, experiments were conducted by utilizing tools developed by [Getreuer, 2012] and [Roth and Black, 2005]. The comparison of results accrued by SAI with the standard models available in the literature is made by utilizing the implementation of [Bertalmio et al., 2001] and [Getreuer, 2012] available in the public domain.

4.8.1 Images with Structures

Experiment: 4.1 The results presented, in Figure 4.2 were compared through visual inspection and the quality parameters are extracted from the results of SAI model. It was evident to notice that the proposed method SAI successfully preserves all linear structures, (refer to Figure 4.2(c)) like poles and the window frames whereas all other methods fail to preserve these features. For demonstration, such broken structures were highlighted in Figure 4.2(d), the result of [Bertalmio et al., 2001], within pale green ellipses. The same disability was seen in results of [Getreuer, 2012] and [Roth and Black, 2005] as shown in Figure 4.2(e) and Figure 4.2(f). The results of this experiment established that even the pole, a linear structural element in Figure 4.3(a) was totally obscured by characters 'l' and 'w' at different locations, still SAI was able to perfectly preserve the pole (Figure 4.3(d)). In addition to this, SAI preserved the portion of the wheel (compare Figure 4.3(b) with Figure 4.3(e)) with Non-linear structures. Figure 4.3(c) and Figure 4.3(f) presented the strength of SAI to retain the back of the knee and heel, despite very little information was available.

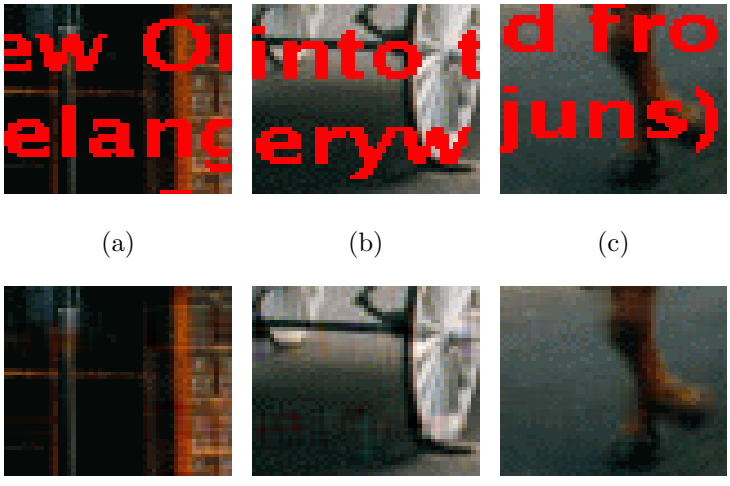


(a) (b)

(c) (d)

(e) (f)

Figure 4.2: (a) Input image of size 297 X 438 overlaid with text (b) The corresponding Mask (c) Result of the proposed SAI model (d) Result of [Bertalmio et al., 2001] method (e) Result of [Getreuer, 2012] method (f) The ground truth



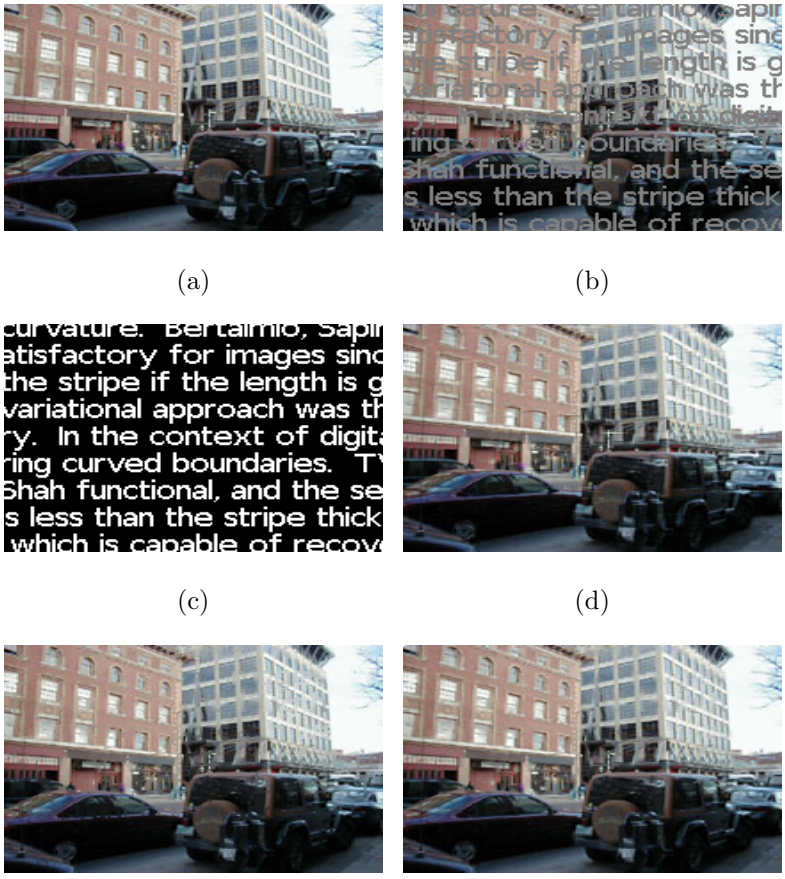
(a) (b) (c)

(d) (e) (f)

Figure 4.3: A close look at results of Experiment 4.1

In the upcoming experiments the performance of SAI against algorithm [Getreuer, 2012] was compared by deriving either Peak Signal to Noise Ratio (PSNR) or Structural Similarity Index Measure (SSIM) [Wang et al., 2008], which ever was relevant to the content in the input image. These metrics were regarded to be appropriate to inpainting problem by image processing community. Image restoration algorithm is analyzed to be effective if the resulting PSNR is between 30db and 50db. The SSIM value lies in the range [0 - 1] and is considered to be effective if its value is very close to 1.

Experiment: 4.2



(a) (b)

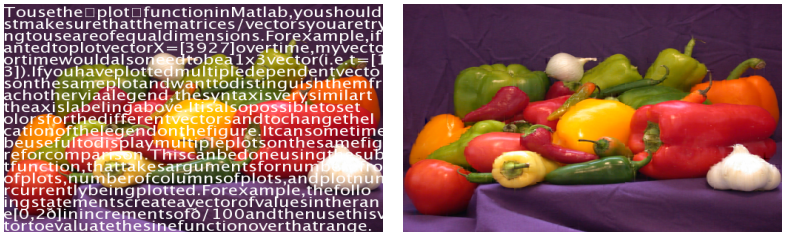
(c) (d)

(e) (f)

Figure 4.4: (a) Original image of size 213 × 163 (b) Input image (c) Associated mask (d) Result of SAI (e) Result of [Getreuer, 2012] method (f) Result of [Roth and Black, 2005] method

In this experiment, an image which had linear structures was taken up. In Figure 4.4(a) both vertical and horizontal edges were obscured by the text. A close look at the building appearing on the right hand side in Figure 4.4(d) revealed that crossovers on it were perfectly retained by SAI. In addition to that, the circular logo on the rear glass of parked vehicle was completely reconstructed as shown in Figure 4.4(d), which was not possible in Bregman method or Roth [Roth and Black, 2005] method (see Figures 4.4(e) and 4.4(f)). SSIM, a structure related metric was extracted from the results of SAI and Bregman experiments. SAI achieved a distinguishable SSIM 0.938 to that of [Getreuer, 2012] method which achieved SSIM 0.918.

Experiment: 4.3 The input image in this experiment, Figure 4.5(a) had smooth



(a) (b)

Figure 4.5: (a) Input image of size 384×512 (b) Result of SAI

areas, non-linear structures and textures. The result of SAI was presented in Figure 4.5(b) with PSNR 35.05 db.

4.8.2 Images with Textures

Experiment: 4.4 The input image in this experiment had Non-linear structures as well as textures that are spread across the image. The result of SAI in Figure 4.6(d) spoke about its strength in terms of curvature preserving capability at peaks of the mountain, compare with Figure 4.6(a). The textures were interpolated perfectly even though the inpainting was carried out on non-overlapping tiles. The results of SAI and Bregman method were compared by extracting the PSNR, a texture related metric. The PSNR realized by SAI was 38.87 db whereas [Getreuer, 2012]

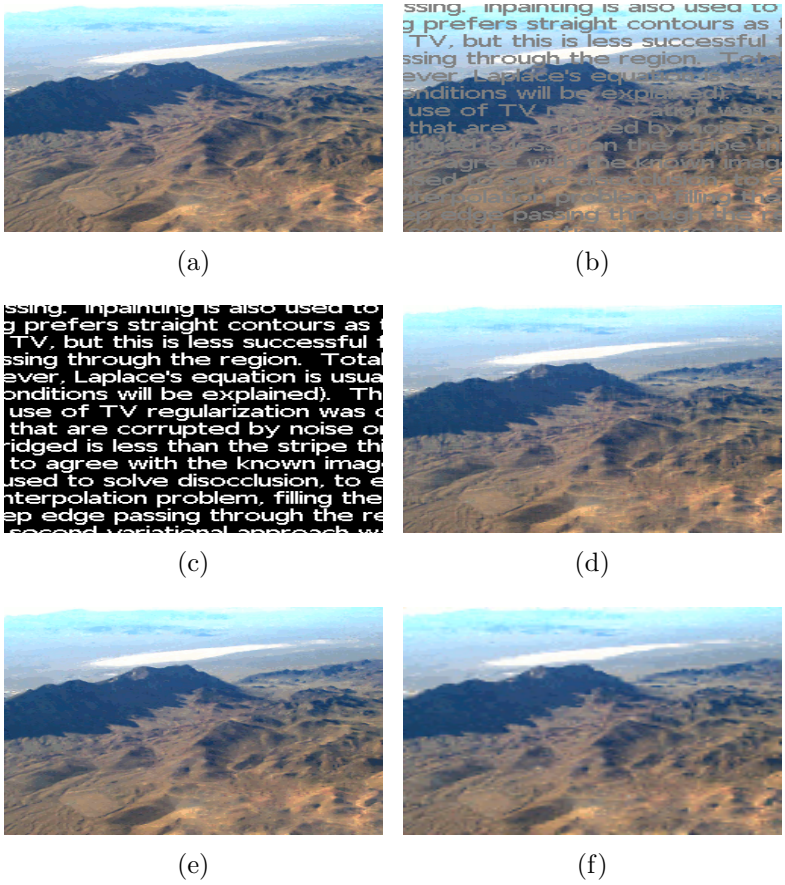
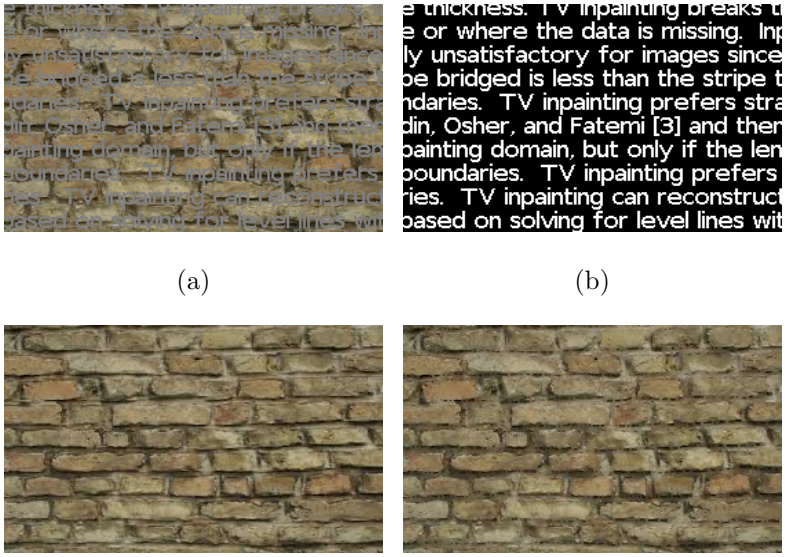


Figure 4.6: (a) Original image of size 250 × 250 (b) Input image (c) Associated mask (d) Result of SAI (e) Result of [Getreuer, 2012] method (f) Result of [Roth and Black, 2005] Method.

method was 39.0 db.

Experiment: 4.5 The input image in this experiment, see Figure 4.7(a) had linear structures and textures. The result of SAI in Figure 4.7(c) showed the textures were synthesized perfectly even though the inpainting was carried out on non-overlapping tiles. The results of SAI and Bregman method were compared by extracting the PSNR, a texture related metric. The PSNR realized by SAI was 35.23 db whereas for [Getreuer, 2012] method was 32.80 db, see Figure 4.7(d).



(a) (b)

(c) (d)

Figure 4.7: (a) Input image of size 183 × 275 (b) Associated mask (c) Result of SAI (d) Result of [Getreuer, 2012] method

Table 4.3 gives the summary of the results of the experiments conducted using SAI, in this chapter.

Table 4.3: Summary of SAI results

Input	Measures	SAI	Getreuer
Buggy	PSNR	28.23	23.72
	SSIM	0.84	0.75
Buildings	PSNR	34.13	33.8
	SSIM	0.938	0.918
Mountains	PSNR	33.48	34.3
	SSIM	0.924	0.93
Brick-wall	PSNR	33.54	27.8
	SSIM	0.93	0.94

4.9 Summary

In this chapter, the image inpainting problem was modeled as a moving window based Spatial Anisotropic Interpolation (SAI) problem and solved the text removal application. SAI incorporated a suitable Systematic Sampling model for inpainting problem which reduces the space requirements while building the model. SAI could handle both the situations related to handling the textures and structures efficiently. It was also observed that exponential correlation modeled the spatial anisotropy very well and yields the predictions with minimum error variance that was demonstrated through the quantitative measures PSNR and SSIM.

Chapter 5

Scalable SAI Using Elastic Net Regularization Approach

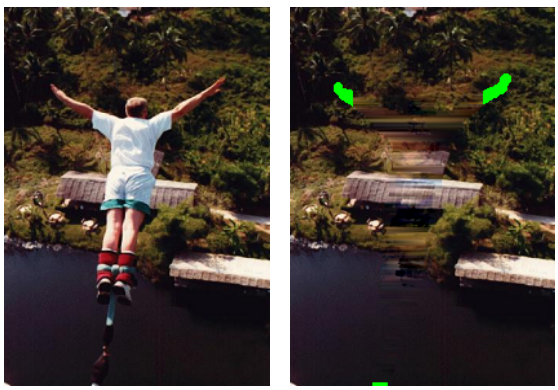
The third contribution of this thesis is to extend the proposed SAI model to *Scalable SAI* which solves the object removal *inpainting category* problem, in two folds. The first extension achieves the scalability and the other attracts an intelligent agent based trend modeling. The proposed Scalable SAI (SSAI) is purely a statistical model and avoids the necessity to model the structure and textures explicitly. This chapter starts with the demonstration of low fidelity solution of SAI model in 5.1. The Systematic Sampling in SAI is extended to 'Not Only Symmetric Hierarchical Sampling', which retains the interpolation feature of kriging and succinctly represents the spatial anisotropic correlation through two sampling kernels, as detailed in 5.2. The Hurst classifier is presented in 5.3.2 and the feature selection based on *Elastic net* regularization is demonstrated in 5.4.3. The PushBack operation of the NoSHS is presented in 5.5. The overall object removal algorithm is presented in 5.6. The quality of the results of SSAI is presented in 5.7.

5.1 Object Removal and Spatial Interpolation Approach

Let I be the image with an object which is marked as unwanted by the user with the help of a mask. Then the image completion algorithm should replace the specified object with some content, which is coherent with the neighboring areas. The complexity of this task increases with the size of the object. In many real world images, the target object may obscure some other structural feature of neighboring object making the problem more difficult. In the worst case, more than one structure or texture might surround the target object which refers to a non-stationary data. Then the object removal algorithm must exhibit an intelligent way of managing these issues and be capable of producing visually pleasing results.

5.1.1 Low Fidelity Solution of SAI

Invocation of moving window based SAI method on Figure 5.1(a), while utilizing short and fat patches, suffer from over smoothing of predictions as shown in Figure 5.1(b) and fails to preserve the structures. In addition, SAI suffers from low fidelity and segmentation problems in case of smaller patches as it turns out to be an extrapolation problem for which DACE is a weak candidate.



(a) (b)

Figure 5.1: (a) Input image of size 183×275 (b) Result of SAI with low fidelity

5.1.2 Issues on Sampling in DACE Modeling

In general, the resolution of an image determines the quality of the definition. Given a high-resolution image, with smooth content, then the image can be represented by a fewer number of pixels than it originally has. The simple interpretation here is, if the image resolution is reduced by half then the area covered by the pixel gets double within the image of the same size. In spatial statistics, the area covered by the pixel under the given radiometric feature is referred to as support. Hence, the selection of a sampling scheme is determined by the spatial resolution of the image and the support of each pixel. Sampling is realized either in a systematic manner or a random manner. In the case of Systematic sampling, as proposed in Chapter 4, the pixels that are separated by the given distance are collected along the rows and columns. Whereas in the random sampling; Latin Hypercube sampling, for example, an exclusive sampling criterion is decided first and the locations are randomly selected while meeting that standard. Sub-sampling (upscaling) reduces the resolution and hence the support increases proportionally. On the other hand, downscaling increases the resolution and reduces the support. Repeated sub-sampling leads to large support, and it causes the smoothing of the variability of data. There is one more quantitative parameter 'extent' that refers to the complete range over which the data spread is defined. If the extent is small, then data start exhibiting a trend which is expressed through mathematical notions. For example, the derivative based descriptors are well suited for capturing the changes over the small extents and constitutes for the anisotropy. In other words, the anisotropy refers to the local level trends prevailing in small components of the image. On the other hand, to model the large-scale patterns, the polynomial basis functions are appropriate. Hence, the interpolation scheme must accommodate the details of the *extent* if it does not remain fixed and the sampling plan assumes that samples are independent. The mainstay of this sampling scheme is the assumption that the underlying random field preserves the gray scale distribution (or histogram) of the image under sampling.

5.2 Proposed Not only Symmetric Hierarchical Sampling

The proposed object removal algorithm extends SAI in two respects. The first extension is related to the sampling scheme adapted in SAI, which works along the rows and columns, and generates only one sub-image from the given input image. The proposed 'Not only Symmetric Hierarchical Sampling' (NoSHS) utilizes one of the two proposed sampling kernels and produces multiple sub-images of equal size. The second improvement to SAI is about solving the object removal inpainting category problem, on each sub-image which is elaborated later. If the number of source region pixels is very high, then Kriging model, which involves computation of inverse for large scale correlation matrix turns the model into a memory hungry. This aspect refers to 'curse of dimensionality' in the literature. Hence, it is essential to select a subset of representatives from the entire set of available design sites. The proposed NoSHS, which is an extension of the sampling schemes that are widely used in the super-resolution analysis. NoSHS probes the entire image by a sampling kernel of size $l \times k$ for $l \in 2, 3$ and $m \in 2, 3$ and each pixel spanned by the probing element are arranged spatially along the rows and columns of the respectively derived sub-image indexed by l and k. For example, if an image I of size $M \times N$ is subjected to NoSHS for once, with $l = 3$ `and` $k = 2$, Algorithm $Extract\_6$ which simulates the asymmetric kernel, downsamples the image by factor 3 along the row and 2 along column simultaneously and produces 6 sub-images I_1, I_2, I_3, I_4, I_5 and I_6 each of size $\frac{M}{3} \times \frac{N}{2}$. Similarly, the symmetric version of NoSHS kernel, Algorithm $Extract\_4$, in one iteration with $l = 2$, $k = 2$, downsamples the input image by factor two along the row and column simultaneously and produces four sub-images I_1, I_2, I_3 and I_4 each of size $M/2 \times N/2$. Here, the sampling kernel acquires the symmetric shape. Given l and k, the chosen sampling kernel is invoked recursively for a finite number of times on each sub-image that is produced in the previous step, until the size of the sub-image is under the specified limit. Such a sampling scheme is unbiased, but the selection

of the sampling kernel by itself is content aware. Thus, the sampling scheme acts as a transfer function which retains the interpolation feature of the problem on each sub-image and results in no loss of data if the each dimension of the input image is divisible by l^j where $j = 1, 2, ..$ denotes the level of the recursion. If the data size is not compatible with this prerequisite, padding with zeros is one option. The total number of sub-images $\Bbbk$ extracted from I with $l = k$, after j levels of recursion, is expressed as a sum of Geometric progression $1, p, p^2, p^3...$ where $p = l^2$. Upon sampling, the kriging model is developed over the sampled data in each sub-image and the prediction of the intensities at the inpaint region is carried out. The Algorithm Extract_4 presents the sampling aspect in the form of an algorithm which extracts four sub-images from the input image. The reader may be little apprehensive about the saliency of the proposed sampling, which brings the farther pixels close to each other and subsequently may impact the fidelity of the polynomial regression model due to the new spatial relations. Thus it is essential for the inpainting model to address few fundamental questions as to how the regression basis selection is done, what must be the order of the selected basis and which sampling kernel to be utilized so as to preserve the anisotropy. These issues are addressed little later. The selection of sampling kernel is substantiated and well supported by the proposed random field model based on the fundamental assumption of second order stationarity. The anisotropy induces the asymmetric correlation along the columns and rows and the tensor product of the correlation structure assumes the elliptic shape in which case high-frequency sampling can done along the major axis, and low-resolution sampling is possible along the minor axis. Inherently, the proposed NoSHS could preserve the underlying properties of the random field. Hence, the proposed NoSHS also favors the underlying stationary property and improves the fidelity of prediction process. For the sake of visualization, in Figure 5.3 the matrix I of size 4×4 presents hypothetical image at highest level (zero level) composed of four different symbols which serves as the input to the NoSHS algorithm. The Extract_4 algorithm performs a probe on image at level zero and extracts every alternate symbol present along

Algorithm 6 Extract_6

Input: Input Image I
Output: Extracted Sub-images $I_i; i = 1 : 6$

1. $l := 1; \quad k := 1;$
2. $i_2 := 2;\ j_2 := 2;\ j_3 := 3;$
3. $for \quad i_1 := 1 \ : \quad M - 1 :$ `step by 2`
4. $i_2 := i_1 + 1;$
5. $for\ j_1 := 1 \quad : N - 2 :$ `step by 3`
6. $j_2 := j_1 + 1; \quad j_3 := j_1 + 2;$
7. $I_1(l, k) := I(i_1, j_1);$
8. $I_2(l, k) := I(i_1, j_2);$
9. $I_3(l, k) := I(i_1, j_3);$
10. $I_4(l, k) := I(i_2, j_1);$
11. $I_5(l, k) := I(i_2, j_2);$
12. $I_6(l, k) := I(i_2, j_3);$
13. $k := k + 1;$
14. `end`
15. $l := l + 1; \quad k := 1;$
16. $i_2 := i_2 \quad + \quad 2;$
17. `end`

the rows and columns simultaneously into four different sub-images $I_i, i = 1, ..., 4$ and arranges them relatively on level one each of size 2×2. In the case of large images this procedure is repeated on each sub-image at the latest level to produce four even smaller four sub-images and these are arranged in a rectangular array until the the sub-images size reduces to a manageable size. The inpainting process is applied on sub-images available at the lowest level. Figure 5.2 presents the input image- 5.2(a) and the corresponding 4 sub-images- 5.2(b-e) extracted through

Algorithm 7 Extract_4

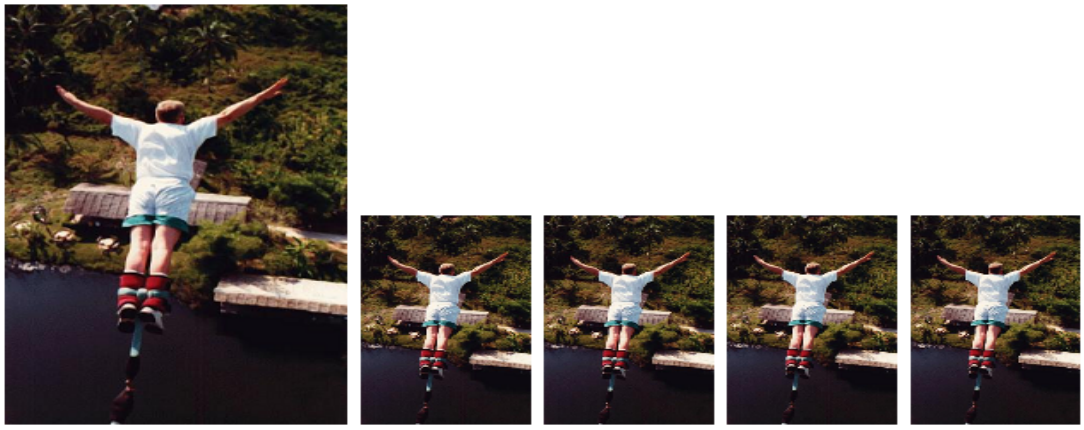
Input: Input Image I
Output: Extracted Sub-images $I_i; i = 1:4$

1. $l := 1; k := 1;$ `i2 := 2;` $j_2 := 2;$
2. $for\ \ i_1 := 1 : M-1 :$ `step by 2`
3. $for\ \ j_1 := 1 : N-1 :$ `step by 2`
4. $I_1(l,k) := I(i_1, j_1);$
5. $I_2(l,k) := I(i_1, j_2);$
6. $I_3(l,k) := I(i_2, j_1);$
7. $I_4(l,k) := I(i_2, j_2);$
8. $k := k+1;\ \ j_2 := j_2 + 2;$
9. `end`
10. $l := l+1; \quad k := 1;\ \ i_2 := i_2 + 2;\ \ j_2 := 2;$
11. `end`

NoSHS with a symmetric version.

5.3 Modeling the Trend

Let I be the input image with a large target object to be removed, and M be the corresponding binary mask with one denoting the target object spatial information and zero denoting the source region locations. These two images are subjected to k-level NoSHS resulting in $m = l^{k^2}$ sub-images I_m and sub-masks M_m of equal size at level k. Each sub-image has Ω and Φ that are smaller in size when compared to their counterparts in the original input image I. The objective of this sampling is to fit the regressions couple on each sub-image and to allow the spatial prediction of pixel values at the inpaint region locations. This aspect is different from the multi-resolution analysis, which arranges the data collected from various sources with the same spatial resolution. Here the sole reason behind the sub-sampling is



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

Figure 5.2: (a) Input image [(b)-(e)] The sub-images extracted through Hierarchical Sampling

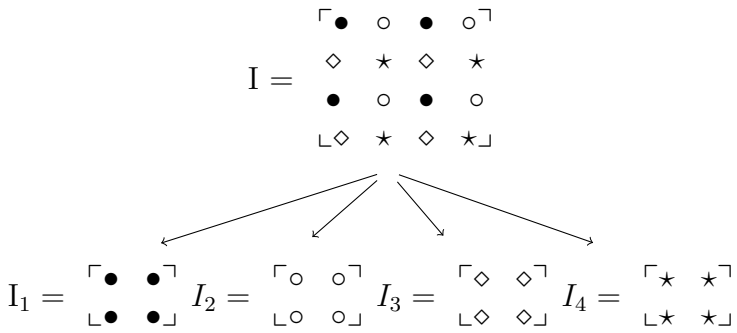


Figure 5.3: I presents a synthetic input image and I_1, I_2, I_3 and I_4 present the sub-images extracted through Symmetric kernel

to reduce the size of data model while preserving the structural correlation.

5.3.1 DACE model

Once the sub-images extracted from the input image, then the proposed model cast the object removal as a Spatial Anisotropic Interpolation problem, which is already developed in Chapter 4. The rationale behind this approach is that the interpolation mechanism involving the spatial properties satisfies certain invariant features and strongly supports the proposed sampling scheme.

The SAI fits the DACE model over the design space, as discussed in Chapter 4, in two phases. In the first step, the Gaussian regression is solved on the random

error field which is modeled around the correlation R in Equation 5.1

$$\hat{y}(x) = r^T R^{-1} Y - (F^T R^{-1} r - f)^T (F^T R^{-1} F)^{-1} F^T R^{-1} Y. \tag{5.1}$$

Then the polynomial regression in Equation 5.2

$$Y = F\beta \tag{5.2}$$

about the Correlation, R is solved through the Generalized Least Square (GLS) in the second phase as given in Equation 5.3

$$\hat{\beta} = (F^T R^{-1} F)^{-1} F^T R^{-1} Y. \tag{5.3}$$

GLS computes the regression coefficients β based on the Mahalanobis distance [Maesschalck et al., 2000]. It is a unitless measure of distance between a point p and the distribution to which the p belongs to, as the number of standard deviations. This distance is scale invariant. Also, while computing the regression coefficients, the spatial correlation is taken into account. In the proposed NoSHS mechanism, the pixels that are far apart are subjected to a transfer function which, scales down the distance between the pixels by a factor of the level l^k and the model development depends on the new spatial distance while implementing the spatial correlation function. As this transfer function is commonly applied to every the pair of pixels belonging to the intrinsic random field, the least square solution does not get affected.

5.3.2 Proposed Hurst Classifier

Selection of a suitable basis which controls the varied nature of the image content, while capturing the local details is a topic of research. The notable feature of natural images is, they often exhibit non-stationary behavior over medium scale distances. Such a characteristic cannot be captured by just increasing the order of the polynomial basis, as it results in wiggling. Hence selection of a suitable

basis which controls the non-stationary nature while capturing the local details is a challenging problem. The proposed object removal algorithm SSAI is capable of selecting a basis function from a predefined bunch of different basis functions, in an adaptive manner. From our empirical observations, the nature of the polynomial can be a constant, higher order polynomial or spline wavelet, each of them with varying capabilities while modeling the trend. The decision as to which basis must be selected from the collection of basis, is addressed by extracting the Hurst exponent H [Nicolis, 2006]. $H \in [0 \quad 1]$, which is a self-similarity based measure capable of manifesting the nature of the content. Hurst exponent describes the probability with which the autocorrelation among the pixels of the image changes asymptotically with the lag h as follows

$$E[\frac{R(h)}{S(h)}] = Ch^H \qquad as \qquad \mathrm{h} \rightarrow \quad \infty \tag{5.4}$$

where R refers to the range of values spanned by h and S is the standard deviation of the data, and C is a constant. In literature, it is possible to estimate the H value in various ways such as Box Counting, Rescaled Range analysis [Carbone, 2007], Wavelet Spectral Density, etc. Subsequently, it is possible to extend it for 2-D images also [Van De Ville et al., 2005]. For the sake of completeness and clear understanding of Hurst exponent one estimation method- R/S is discussed for 1-D data. Given $y_i; \quad i = 1 : n$ the pixel responses values at locations $x_i; \quad i = 1 : n$ with mean m then calculate the deviations d_i of each y_i from m using $d_i \quad = \quad y_i \quad - \quad m$. Subsequently, the series of partial sums p_i over d_i are calculated in a cumulative manner as follows

$$s_k = \sum_{l=1:k} d_l \qquad k = 1 : n.$$

Then the difference of $max\{s_k\}$ and $min\{s_k\}$ denotes Range, R. Then the slope of 'bestfit' line for $\log(\frac{R}{S})$ vs $\log(n)$ represents the *estimate* of Hurst exponent H. The inherent feature of this model is represented as Fractal coefficient which

expresses the self-similarity measure of Brownian motion [Van De Ville et al., 2005]. The *estimated* Hurst exponent value, characterize the underlying process that generated the data into three classes- stationary, unknown or non-stationary. The proposed Hurst classifier utilizes the value of H to categorize the images into three types as follows. If $H \in [0 \quad 0.3]$ then the image is categorized as non-stationary (chaotic component) image, else if $H \in (0.3 \quad 0.5)$ then it is a second order stationary image, otherwise the image is considered to be smooth (deterministic component). SSAI selects Spline Wavelet as the basis if the data present in image exhibits the non-stationary property. The insight in choosing the Spline wavelets is, as explained in subsequent subsections, they have the ability to provide a platform to deal with the varying characteristics at local and global levels through decomposition and reveals the structural breakups that correspond to the non-stationary feature of the underlying image. In the case of Second Order stationary images, a higher polynomial basis is deployed. Otherwise, if the content is smooth, the trend is modeled through a constant. This adaptive feature in SSAI is a novel contribution in inpainting domain.

5.3.3 Higher Order Polynomial Basis as the Trend

Though the issues related to sampling are justified through an established theory, still another involved problem of multivariate polynomial regression prevails. The assumption is that the underlying random field models the trend as a polynomial; as a means to realize the large scale changes in the image. However, the order of the basis is not addressed formally. There are two attitudes towards resolving this issue discussed in the literature. The first approach is to assume that the underlying random field possesses some unknown trend. Then a Bayesian [George and Mcculloch, 1993] [Chipman et al., 1997] model based ranking is performed on features. This approach learns the interactions between the spatial data and ranks the interactions. Then, a regression function is built by a forward selection method. In the next step, a kriging model is used to measure the prediction accuracy of the chosen regression model in an iterative fashion. This ideology is

incorporated very recently in Blind kriging [Joseph et al., 2008]. It is a 'multiple kriging' model with the conservative approach. As a first step, the data is fit using Ordinary Kriging and the correlation parameters θ are learned. Subsequently, a sequence of kriging models is fit in an iterative manner to find the polynomial basis quotients β through feature selection algorithm. Clearly, this algorithm turns out to be a computationally expensive way of fitting the model. However, Blind kriging makes use of Orthogonal Polynomials [Couckuyt et al., 2012] which are defined only for equally spaced samples. But in the case of inpainting problem, it is not rational to assume this spatial property. In large scale inpainting problems the intra pixel distance between source region pixels is not constant.

5.3.4 Handling of Non-Stationarity using Spline Wavelets

The natural images often exhibit non-stationary behavior over medium scale distances. Such a characteristic cannot be captured by just increasing the order of the polynomial basis which models the trend as it results in wiggling. Hence selection of a suitable basis which actually controls the non-stationary nature while capturing the local details is a topic of research. Splines provide a framework for dealing with interpolation and discretization problems. One can always obtain a continuous representation of a discrete signal of one or more dimensions by fitting it with an appropriate spline. The fit may be exact (interpolation) or approximate (least-squares or smoothing splines). Spline fits are usually preferable to other forms of representations (e.g. Lagrange polynomial interpolation) because they have a lesser tendency to oscillate (minimum curvature property). The multiresolution properties of splines make them the ideal candidate for constructing wavelet bases that are used for image reconstruction from projected data besides many other applications [Unser and Blu, 2000]. Polynomial splines can be expressed as linear combinations of B-spline basis functions joined smoothly at the break points (knots: $t_0, t_1, ...$), and the degree of smoothness depends on the order of spline basis. For cardinal B-splines (basis splines), these breakpoints are equally spaced. Unlike polynomials, the spline basis forms the local basis and have many useful

properties that can be applied to function approximation. The cardinal B-spline $b_m(t)$ of order m with the knot sequence $\{0, \pm1, \pm2, ...\}$ consisting of polynomials of degree $m-1$ between the knots is defined by the following convolution

$$b_m(x) = ({b_0}^1 * {b_0}^2 * \quad ... \quad * {b_0}^{m+1})(x) \quad \text{for} \quad m \geq 0$$

where b_0 is a characteristic function defined over $[0, 1)$ as $\mathrm{b}_0 = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$ and the $*$ is the convolution operator. The B-splines satisfy all the requirements of a valid scaling function of $L_2(R)$, that are the necessary and sufficient conditions

- **Two-scale relation:** $b^m(x) = \sum_{k \in Z} h(k) b^m(2x - k)$
- **Partition of Unity:** $\sum_{k \in Z} b^m(x - k) = 1$

where the filter $h(k)$ is the binomial filter, $h(k) = \frac{1}{2^m}\binom{m+1}{k}$. For example, the B-Spline basis of order 1 to 4 given in the Table 5.1 [Unser and Blu, 2000] These

Table 5.1: The table showing the basis of order from 1 to 4

	$\mathrm{b}_1(x)$	$\mathrm{b}_2(x)$	$\mathrm{b}_3(x)$	$\mathrm{b}_4(x)$
$0 \leq x < 1$	1	x	x^2	x^3
$1 \leq x < 2$	0	2 -x	$-2\mathrm{x}^2 + 6x - 3$	$-3\mathrm{x}^3 + 12x^2 - 12x + 4$
$2 \leq x < 3$	0	0	$(\mathrm{x} - 3)^2$	$3\mathrm{x}^3 - 24x^2 + 60x - 44$
$3 \leq x < 4$	0	0	0	$-\mathrm{x}^3 + 12x^2 - 48x + 64$
elsewhere	0	0	0	0

conditions ensure that B-Splines can be used to generate a MRA (Multi Resolution Analysis) of L_2 functions. The wavelet function corresponding to the scaling function $b^m(x)$can be constructed by

$$\psi(x) = \sum_{k \in Z} g(k) b^m(2x - k) \tag{5.5}$$

The filter coefficients $h(k)$ and $g(k)$ completely characterize various families of spline generated wavelets such as orthogonal, semi-orthogonal or bi-orthogonal.

The 2D spline wavelet transform is obtained by successive one-dimensional wavelet transformation along the rows and columns as the 2D spline functions are separable, i.e. $(x, y) = \psi^m(x)*\psi^m(y)$. The insight in selecting the Spline wavelets is wavelet's have ability to provide a platform to deal with the varying characteristics at local and global levels through decomposition and reveals the structural breakups that correspond to the non-stationary feature of the underlying image. The wavelet-based basis can capture the trends in the image that appear, disappear and reappear over time period. Wavelet adapts itself to capture the wide range of frequencies which correspond to textures while localizing the events that correspond to structures. With regard to 2-D data, the Spline basis is separable and can be realized through the tensor product of individual components.

5.4 Enhanced Design and Analysis of Computer Experiments

The proposed algorithm addresses the regression polynomial basis selection issue, very rationally, by employing a futuristic approach which involves kriging, just for once. Initially, for the given design space the correlation R basis is assumed which models the random field $Z(X)$ associated with the residue. Then a set of polynomial basis functions are selected to form a hybrid higher order polynomial F with complete interactions among the basis functions, in terms of cross product terms, is selected to model the global trend. For example, in case of DACE model for a polynomial basis F of order n the total number of quotients to β_k be learned is $\frac{n(n+1)}{2}$. As the object removal, problem entails with the larger number of design sites when compared to the number of candidate features of higher order polynomial the chances of overfitting is very less. This is an added advantage to the proposed model. Once the DACE model is fit for the entire design space S with responses Y, from Equation 4.13, the quotients are approximated as follows

$$\hat{\beta} = \left(F^T R^{-1} F\right)^{-1} F^T R^{-1} Y. \tag{5.6}$$

Among these quotients, some are either considerably smaller or reprehensible in contributing to the quantitative decision making of regression model leading to increasing in the variance of the prediction. In order to mitigate this problem, a strategy to retain only a few important regression coefficients and discarding the remaining is one of the possibilities. This turns the problem at hand as a regression with subset selection problem. There are plenty of ways to accomplish this aspect in the literature under the titles feature selection [Miller, 2002], pattern search [Hooke and Jeeves, 1961] etc.

5.4.1 Proposed Feature Selection Based Trend Modeling

Feature selection is a scientific procedure to determine the contributing features of the data to solve the given task. Guyon and Elissee [Guyon and Elisseeff, 2003] elaborated that there are many potential benefits of feature selection: facilitating effective data visualization, reducing data measurement requirement and training times and handle the curse of dimensionality aimed at improving the imputation performance. In machine learning and statistics, feature selection is the technique of selecting a subset of relevant features so that the redundant and irrelevant features can be eliminated. In a way, it helps improve the performance of learning models while retaining the accuracy and alleviate the effect of the curse of dimensionality. Further, it helps in enhancing generalization capability of the systems under development. I propose a feature selection based trend modeling aspect for inpainting through L_1 regularization mechanisms.

5.4.2 L1 Regularization Using LASSO

To achieve feature selection, the Least Absolute Shrinkage and Selection Operator (LASSO) developed by Tibshirani [Tibshirani, 1994] is considered. The idea behind applying this operator is to regularize the GLS solution $\hat{\beta}$ using L_1 norm of β. L_1 regularization is a feature selection method which reduces the variance of the prediction by solving the Lagrangian version of the optimization problem

with $\lambda \geq 0$ as a parameter.

$$\min_{\beta} \frac{1}{2} ||Y - \beta^T X||^2 + \lambda ||\beta||_1. \tag{5.7}$$

Equation 5.6 does not have a closed form solution as that of ridge regression. But there are efficient algorithms to solve this optimization problem. Both the terms on the right-hand side of Equation 5.6 are convex sets, and hence it has a single minimum. The nature of these regularizations is compared with the corresponding terms of the Ridge regression and depicted in Figure 5.4 [Hastie et al., 2001]. The signal processing community implemented this Lagrange equation under basis pursuit terminology.

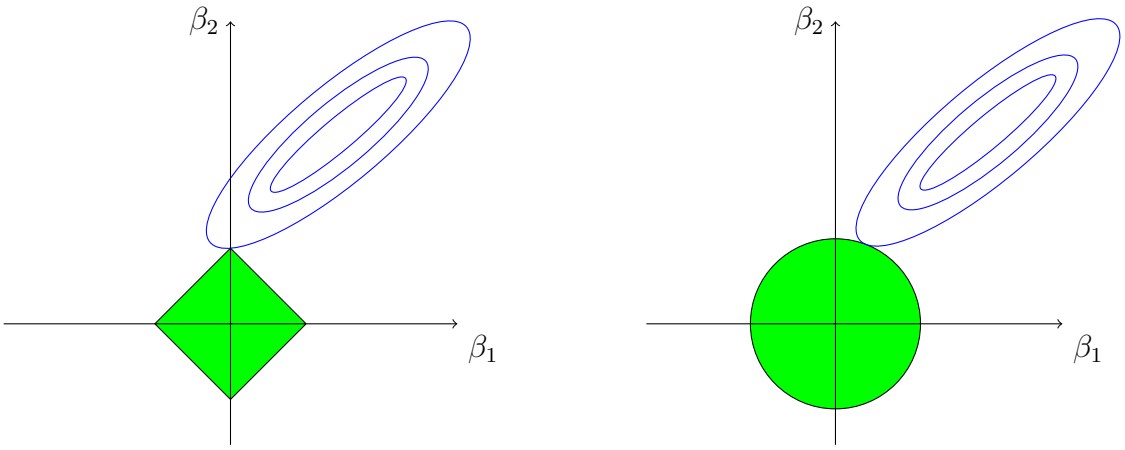


Figure 5.4: Comparison of shrinkage in LASSO with RIDGE regressions

5.4.3 Implementation Using Elastic Net

The LASSO regularization works well if the design matrix is thin and tall. That is the number of design points is much larger than the number of coefficients to be learned. But, the ailment of LASSO is if the basis functions do not form a linear independent set and if they are correlated with each other, it fails to select the desired number of features. To address these limitations Zou [Zou and Hastie, 2005] introduced the elastic net which combines the L_1 norm and L_2 norm terms in a convex manner, in order to regularize the GLS estimates of β which is defined

as follows

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}}(\|Y - X\beta\|^2 + \lambda\|\beta\|^2 + (1-\lambda)\|\beta\|_1). \tag{5.8}$$

Hence, the elastic net method involves both the LASSO and ridge regression. Here λ determines the level of sparsity. Small values of λ yields the sparse solutions. The nature of the elastic net solution compared to the LASSO and ridge in case of 2-Dim is demonstrated in the Figure 5.5. The primary goal of introducing

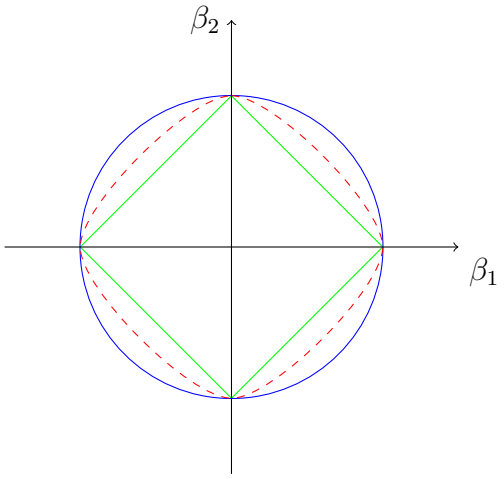


Figure 5.5: Comparison of ridge (Blue), LASSO (Green) with Elastic Net (Red Dotted) regularization schemes

elastic net regularization in the proposed model is to resolve the primary decision conflict about the order of the polynomial that represents the trend. Initially, the trend is modeled with higher order polynomial and penalized to annihilate many coefficients to zeros and selects only a few coefficients. Subsequently Efron [Efron et al., 2004] proposed vectorized version of LASSO called Least Angle Regression (LARS) to accelerate the computations. The elastic net based regularized solution of DACE model is named as Enhanced DACE (EDACE). Many implementations of LASSO [Sjöstrand, 2005] [Schmidtm, 2008] are available in public domain. The upcoming experiments based on the proposed proactive model EDACE is implemented by utilizing the [Schmidtm, 2008] model. In the next experiments section, The DACE model is run to evaluate the regression coefficients using GLS and are further penalized by elastic net regularization. The proposed EDACE

model demonstrates that the results are comparable with the best results of the hybrid models.

5.5 The PushBack Operation

Algorithm 8 PushBack_6

Input: Set of all 6 inpainted sub-images $I_i, i = 1 : 6$ that are extracted from I

Output: Reconstructed image I

1. $p := 1; q := 1;$
2. $for \quad i := 1 : m : \quad step \quad by \quad 1$
3. $for \quad j := 1 : n : \quad step \quad by \quad 1$
4. $I(p, q) := I_1(i, j));$
5. $I(p, q + 1) := I_2(i, j);$
6. $I(p, q + 2) := I_3(i, j);$
7. $I(p + 1, q) := I_4(i, j);$
8. $I(p + 1, q + 1) := I_5(i, j);$
9. $I(p + 1, q + 2) := I_6(i, j);$
10. $q := q + 3;$
11. `end`
12. $q := 1; p := p + 2;$
13. `end`

As a last step in the proposed framework, all inpainted sub-images all belonging to a particular level of the hierarchy need to be systematically combined to get the

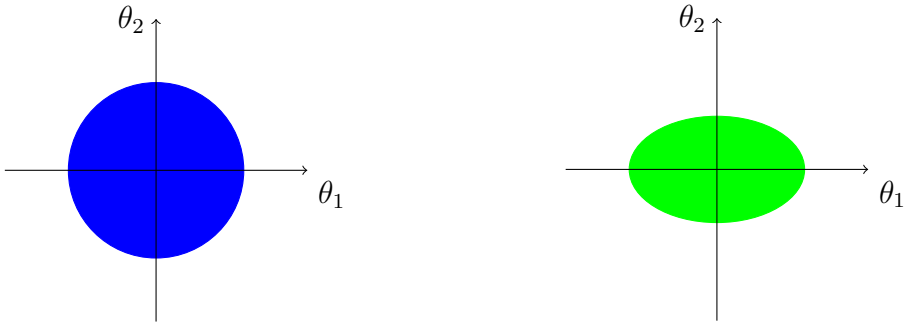
overall result. It is a simple push back of the chosen hierarchical sampling that simply rearranges the pixels from each inpainted sub-image back into their proper locations of the original input image. The algorithms PushBack_4 and PushBack_6 demonstrate the implementation of the push back mechanisms for sampling order 2 and 3 to get the original image back. If asymmetric sub-sampling of order 3 x 2 is done then reconstruction of I from all extracted sub-images $I_i, i = 1 : 6$ is achieved through the Algorithm PushBack_6. Both the algorithms PushBack_6 and PushBack_4 work upward from the lowest level of hierarchy while merging the associated sub images of the underlying image sub-plane into a sub-plane at the next higher level. At level 1 of the hierarchy there are either 4 or 6 sub-images that are merged into the overall result image.

Algorithm 9 PushBack_4

Input: Set of all 4 inpainted sub-images $I_i, i = 1 : 4$ each of size $m \times n$ that are extracted from I

Output: Reconstructed image I

1. $for \quad i := 1 : m : \quad step \quad by \quad 1$
2. $for \quad j := 1 : n : \quad step \quad by \quad 1$
3. $I(2 \times i - 1, 2 \times j - 1) := I_1(i, j);$
4. $I(2 \times i - 1, 2 \times j) := I_2(i, j);$
5. $I(2 \times i, 2 \times j - 1) := I_3(i, j);$
6. $I(2 \times i, 2 \times j) := I_4(i, j);$
7. `end`
8. `end`



(a)Isotropic correlation tensor $[\theta_1 \ \theta_2]$ – *Symmetric Sampling*

(b)Anisotropic correlation tensor $[\theta_1 \ \theta_2]$ – *Asymmetric Sampling*

Figure 5.6: The pictorial representation of Isotropic correlation and Anisotropic correlation kernels

5.6 The Complete Algorithm of SSAI

Algorithm SSAI, summarizes the steps discussed so far and presents the complete algorithm for object removal from images. If the input image possesses the isotropic correlation, (i.e. $\theta_1 = \theta_2$) then their tensor product gains circular shape (see Figure 5.6(a)) and the symmetric kernel of the proposed NoSHS, Extract_4 carries out uniform sampling and extracts four sub-images. In contrast to this, if the correlation is more prominent along the columns than the rows (i.e. $\theta_1 < \theta_2$) then their tensor product assumes elliptic shape, as shown in Figure 5.6(b). Then the asymmetric kernel of NoSHS, Extract_6 performs coarse sampling along the columns and fine-grained sampling along the rows to derive 6 sub-images. The EDACE model is developed, based on the Hurst classifier, for predicting the pixel values at inpaint locations for each sub-image. The PushBack function relative to step 4 is invoked recursively to get the overall inpainted image. Table 5.2 presents the source image and the basis selection (transparent to the user) based on the nature of the image, which is determined by the value of the Hurst exponent.

Table 5.2: The Decision Table Based on Hurst classifier

Image	H	Nature	Suitable Basis
Grid	0.89	Smooth (I)	Constant
Texture	0.14	Non-stationary (III)	Spline Wavelets
Boy	0.56	Smooth (I)	Constant
Bungee	0.36	Second-Order-Stationary (II)	Polynomial
Buggy	0.83	Smooth (I)	Constant
Trouser	0.19	Non-Stationary (III)	Spline Wavelets
Fruits	0.72	Smooth (I)	Constant
Baseball	0.37	Second-Order-Stationary (II)	Polynomial

Algorithm 10 SSAI

Input: Input image I and the associated Mask M

Output: Inpainted image

1. Derive H from I
2. Learn the parameter θ of correlation kernel R
3. Check the size of the inpaint region and determine the levels of hierarchy k
4. Apply the relevant form of NoSHS on I and M determined by the θ, §5.2, and extract $\mathbb{k}$ sub-images
5. Select the sub-images available at bottom most level of the hierarchy and put them onto array A
6. Invoke Hurst classifier to determine the trend basis F §5.3.2
7. For every sub-image on A, fit EDACE model, §5.4 and predict the responses over Ω
8. Invoke the corresponding PushBack kernel, §5.5, recursively on inpainted sub-images, relative to step 4, to produce the inpainted image.

The only requisite to apply the proposed scalable spatial anisotropic algorithm for object removal from images is the dimensions of the image shall be exponentials of either 2 or 3. Once this requisite is met then the number of levels in the hierarchical sampling is decided by the absolute size of the inpaint region. The experiments conducted in this chapter applied at most 4 levels of NoSHS. Authors in [Wang et al., 2013], confined their framework to symmetric kernels and made use of preprocessing and post processing steps. But their results revealed the low fidelity feature and over-smoothing of the inpainted region, because of the absence of anisotropy modeling element. Algorithm SSAI presents the complete algorithm for object removal from images.

5.7 Experiments and Results

The proposed SSAI was applied on a good collection of images. The intent behind these experiments was, resolving the unknown trend issue in spatial interpolation experiments through a sound feature selection based ideology. The proposed methodology was compared against the brute force based feature selection and other approaches. The results of all these experiments concluded that the proposed framework was faring on par with the exemplar based object removal algorithms. In many situations, SSAI could successfully preserve the structures and synthesize fine scale textures. The polynomial regression was modeled based on the Hurst Exponent of the source image.

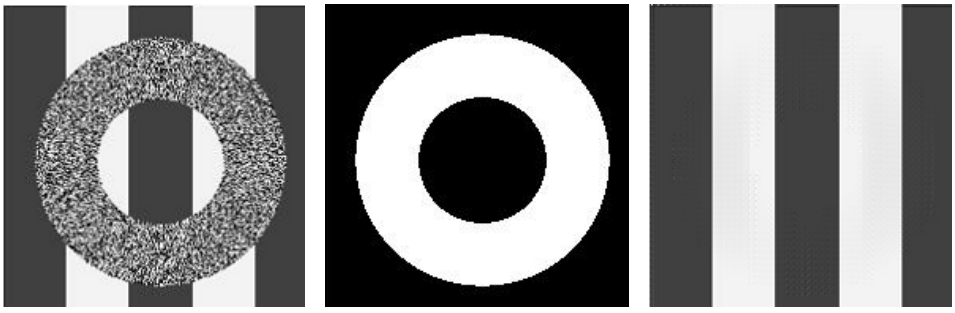
5.7.1 Structure Preserving Capability

The results presented here demonstrated the feature selection capability of SSAI. In each experiment the input image with object to be removed was marked by hand selection tool and the corresponding mask was constructed from it. All these experiments were conducted on MATALB 9 and the execution time of each experiment was tractable. The important observation out of these experiments is that if the sampling is done to a coarse level then the prediction error is minimum.

This is because, the coarse sampling facilitates to understand the mean responses rather than individual responses of design sites available in the design space of the interpolation problem. But, the potential problem of too many coarse sampling steps associated with inpainting problem is the influence of farther objects become dominant and the model fitting suffers from outliers. Perhaps, this is the one of the limitations of the proposed model. To mitigate this problem, a sampling mechanism which considers only the relevant data points can be taken up.

Experiment: 5.1

In this experiment synthetic images each with a large gap marked in white color were considered with two regions that are separated by sharp edge, see Figure 5.7(a). SSAI was able to fill the whole with a meaningful content on both sides of the edge and the linear structures are preserved as shown in Figure 5.7(c).



(a) (b) (c)

Figure 5.7: (a) Input image (b) The mask (c) The SSAI Result

In Figure 5.8(a) and 5.8(d) the confusion as to continue the horizontal structure or vertical structure was resolved by the anisotropy along the row by learning the θ parameter appropriately. This control was not possible with any other model of inpainting in the literature. In Figure 5.8(b) and 5.8(e) the staircase was preserved thoroughly and the neighboring information was propagated along the structures. The third image pair 5.8(c) and 5.8(f) demonstrate the structure preserving capability of the SSAI model while the color information on both the sides of the edges were perfectly retained.

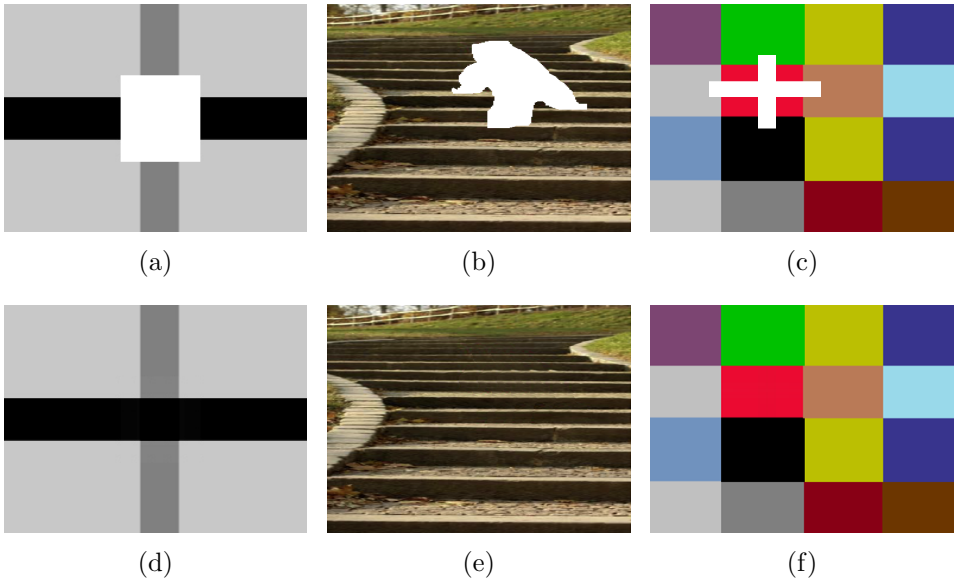


Figure 5.8: Row 1 presents the input images (a), (b), (c) with gap to be filled Row 2 presents the corresponding inpainted results (d), (e), (f) by SSAI Model

5.7.2 High fidelity of SSAI

Experiment: 5.2

Now, in this experiment, see Figure 5.9, I demonstrate the effect of the regularization term introduced in the proposed model which replaced the Euclidean distance based measure with checker board distance. Initially a fifth order polynomial was selected to represent the trend in the course of polynomial regression. This involved estimation of 21 coefficients out of which many of them are either not contributing to the decision or pushing the regression into outliers sensitive. The experiment, see Figure 5.9(a) and Figure 5.9(d), demonstrated the fact that Elastic net regularization alleviated the influence of the unnecessary coefficients of the higher order and improved the fidelity of the results, compare Figure 5.9(b) with 5.9(c) and 5.9(e) with 5.9(f).

5.7.3 Handling Textures and Structures Simultaneously

Experiment: 5.3

This experiment, see Figure 5.10(a), processed an input image with finer tex-

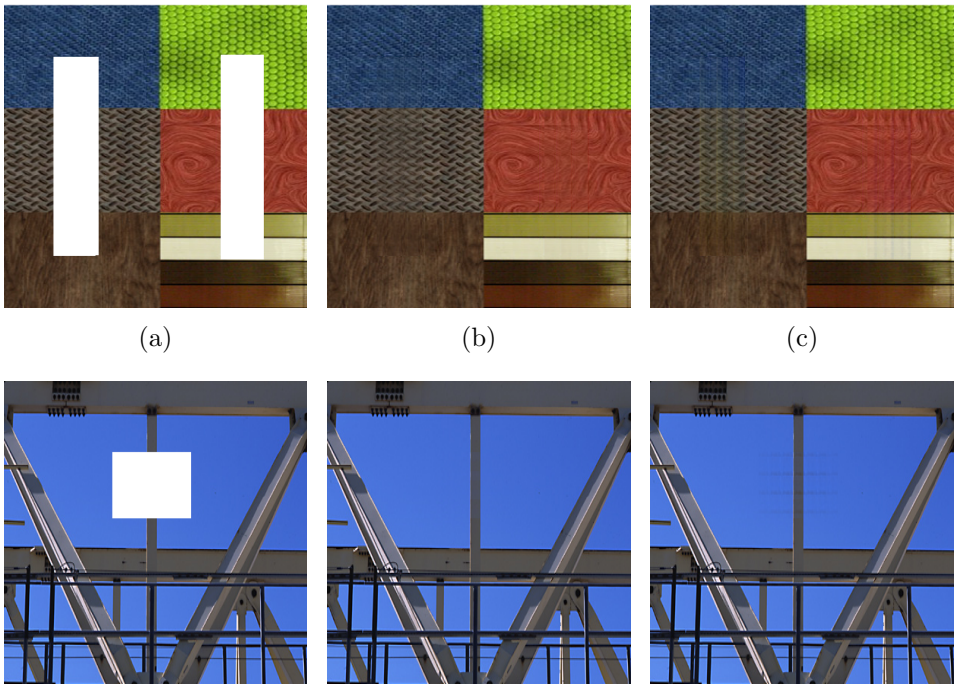


Figure 5.9: Column1 (a), (d) presents the input images, Column 2 (b), (e) presents the results of the proposed SSAI model and Column 3 (c), (f) presents the results in the absence of Elastic net regularization.

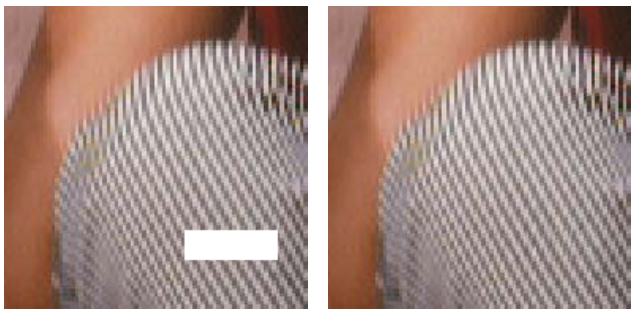
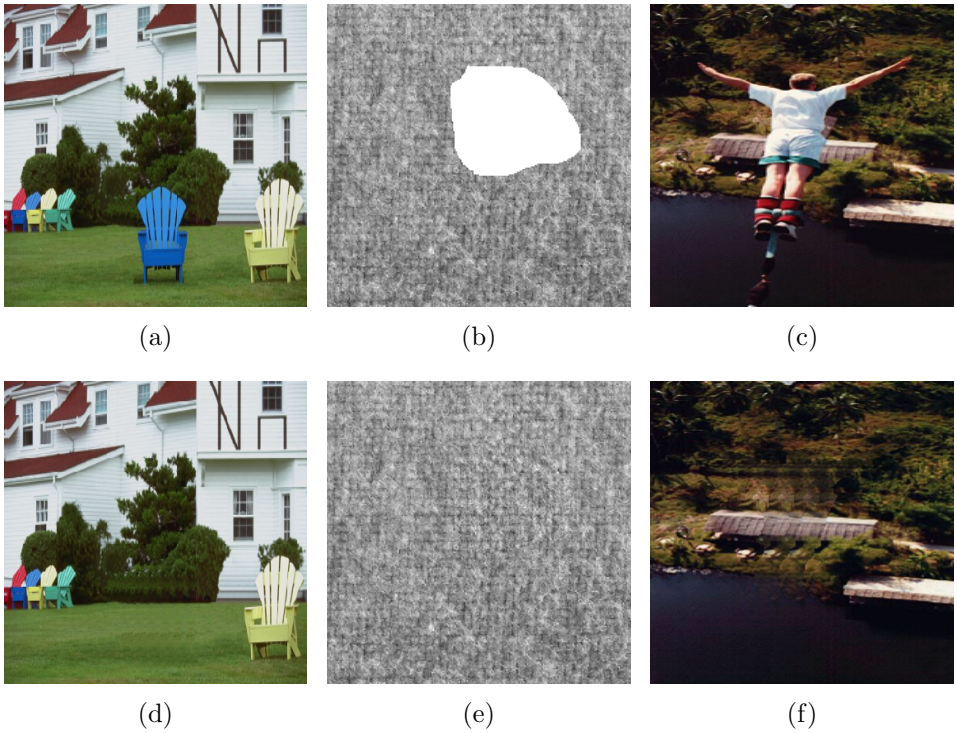


Figure 5.10: (a) Input non-stationary image with a gap in it (b) SSAI Result

tures along with well behaved non-linear structure on the trouser. Hurst classifier qualified the input image to show non-stationary nature (III) and recommended to utilize the Spline-Wavelet as the trend model. The input image had a wide gap and the task was to fill the hole through texture synthesis. SSAI was able to continue the non-linear structure and synthesized the texture associated with it

simultaneously with out modeling them separately. The result presented in Figure 5.10(b) demonstrated the capability of SSAI which was comparable with domain decomposition approach fallowed by [Starck et al., 2005].

Experiment: 5.4



(a) (b) (c)

(d) (e) (f)

Figure 5.11: Row 1 (a), (b), (c) presents the input images with non-stationary content, Row 2 (d), (e), (f) presents the corresponding results by SSAI.

In this experiment images with fine textures were taken up to evaluate the capability of SSAI algorithm. These images were collected from Berkeley image database. The texture synthesizing capability of SSAI was demonstrated in Figure 5.11(a) & 5.11(d) and Figure 5.11(b) & Figure 5.11(e). The chair was removed from the scene and the finer texture on the grass and the larger textures in the background were synthesized. Figure 5.11(c) is a standard benchmark image for measuring the goodness of any inpainting algorithm. The proposed SSAI could retain the structures on the roof of the shed which is unique to the proposed model, Figure 5.11(f) as the results of various models available in literature suffer from texture garbage and structure collapses, see Figure 2.17.

Experiment: 5.5

This experiment is aimed to demonstrate the capability of SSAI to handle even the non-stationary property of the data in the input image. In Figure 5.12 the input images, 5.12(a) & 5.12(b) are chosen such that they exhibit finer structural features and non-stationary nature. The SSAI model built in this experiment while utilizing the Spline Wavelets to handle the non-stationary data. These images are composed of textures, structures and smooth components that span the entire image. These images were processed by state of the art hybrid model, [Arias et al.,

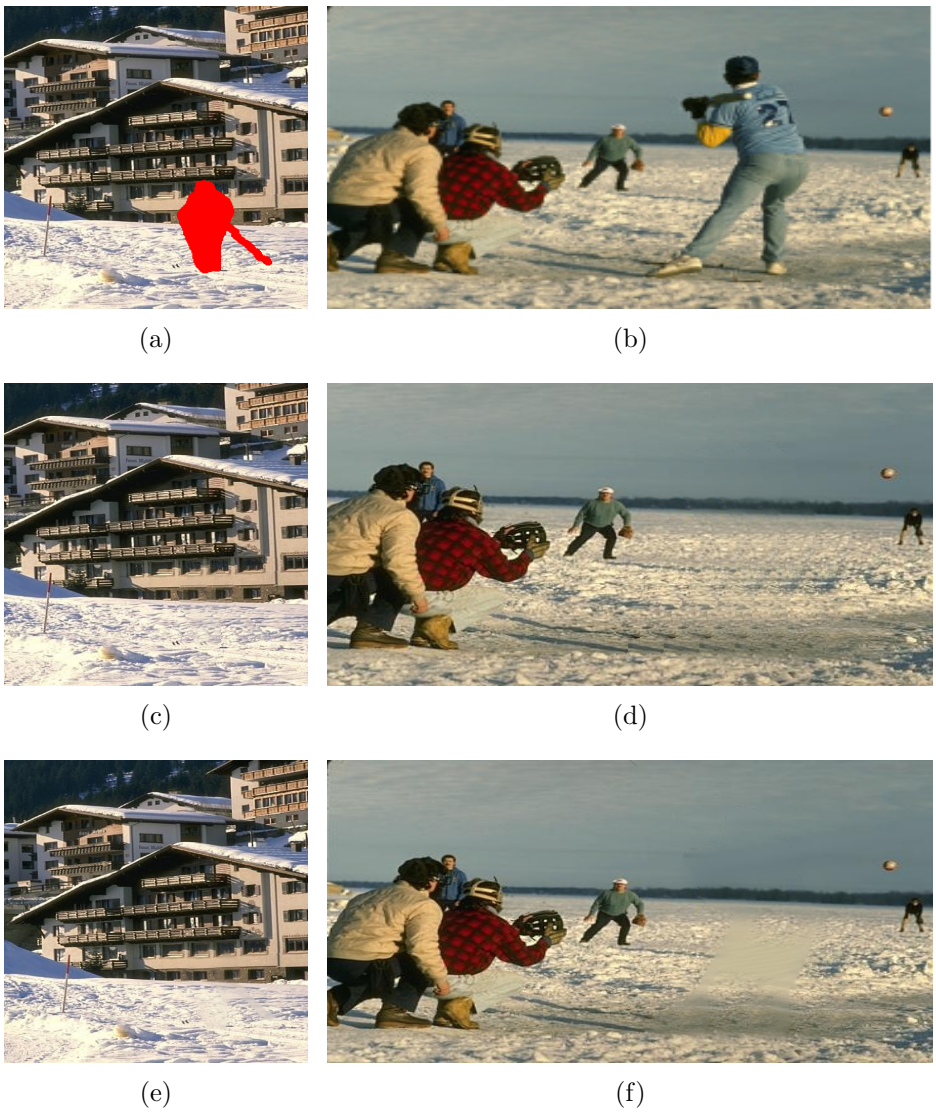
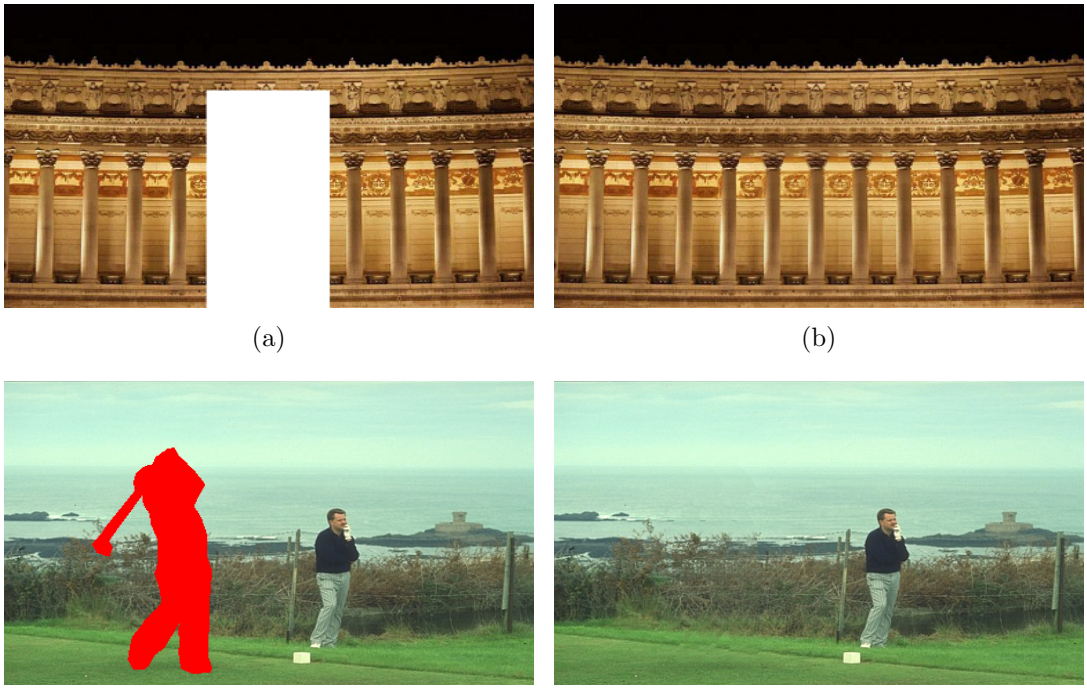


Figure 5.12: Row 1 (a), (b) presents the input images, Row 2 (c), (d) presents the results of SSAI and Row 3 (e), (f) presents the results of Aria's method.

2011], a NLTV method, which failed to synthesize the textures clearly, see Figure 5.12(e) & Figure 5.12(f). SSAI could preserve the structures and interpolated the texture perfectly, see Figure 5.12(c) and Figure 5.12(d).

Experiment: 5.6

In this experiment a series of standard images is processed for object removal, see Figure 5.13(a) & 5.13(c). The object to be removed is either highlighted or masked. The input image had multiple segments that are separated by strong edges. Within each segment finer textures were present. Figure 5.13(b) & 5.13(d) highlight that textures were synthesized and the structures were propagated without introducing any artifacts. The results of this experiment were comparable with latest Best-Exemplar method by [He and Sun, 2012]. Figure 5.14(a) presents the



(a) (b)

(c) (d)

Figure 5.13: Column 1 (a), (c) presents the Input images and Column 2 (b), (d) presents the results of SSIA model.

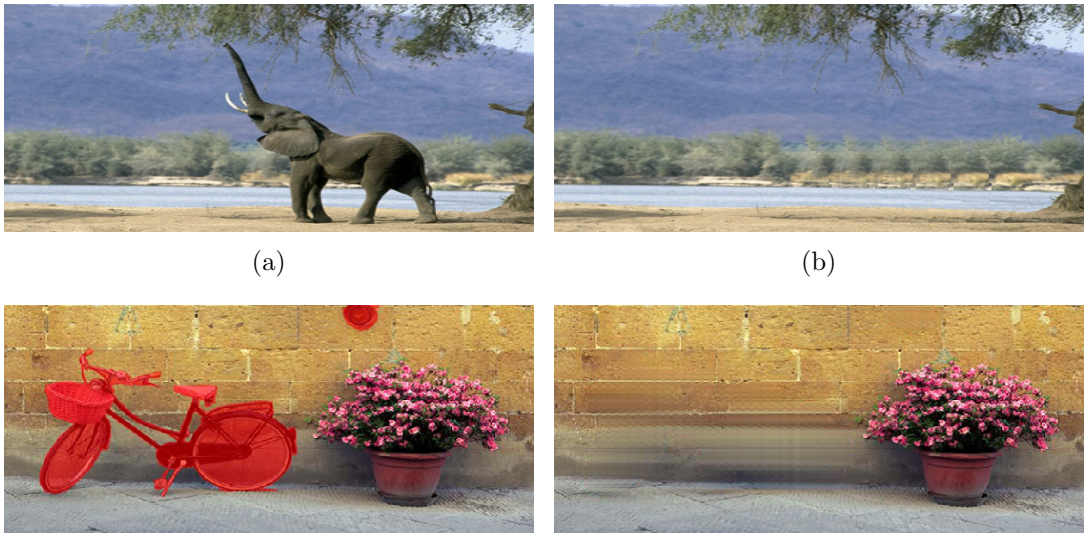
input with fine textures and structures and Figure 5,14(b) highlights the capability of SSAI in synthesizing textures and structures.

Experiment 5.7: Naturally the interpolation based inpainting expects the ob-

(a) (b)

Figure 5.14: (a) Input image with finer texture and large structures and (b) shows the result of SSAI model

ject to be removed must be free from influence of other near by objects, see Figure 5.15(a) & 5.15(b). If this influence is not controlled formally, the results of inpainting suffer from artifacts as shown in Figure 5.15(b) & 5.15(d). This experiment clearly gives a clue that only relevant design sites must be considered for model fitting. Though there are other formalisms like Nearest Neighbors (NN) well defined in literature, the proposed SSAI model manages this situation through exploiting the anisotropy property built into the model. Proper initialization of θ would reduce the influence of unnecessary object.



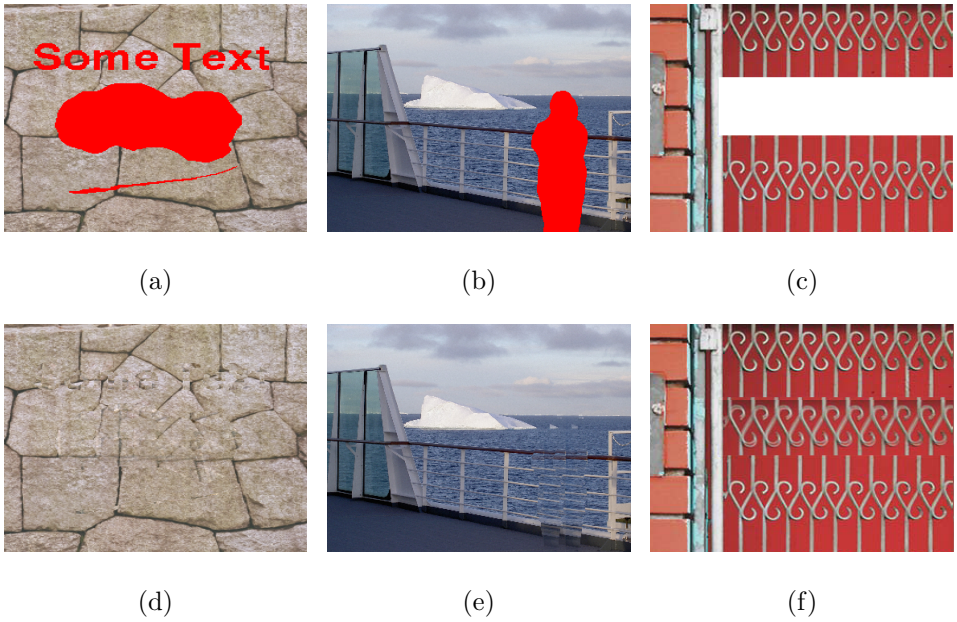
(a) (b)

(c) (d)

Figure 5.15: Column 1 (a), (c) input image and Column 2 (b), (d) presents the results of the proposed SSAI model.

Experiment: 5.8

In this experiment, see Figure 5.16, the disability of SSAI is brought out. The DACE model represents the spatial anisotropy in the form of spatial correlation. The experiments demonstrated in this thesis made use of exponential correlation kernel which considers the absolute difference between the corresponding ordinate values and the corresponding abscissa values of two design sites as the spatial distance. This distance measure could not model the radio metric values at those locations and hence failed to represent the stationary properties along non-linear structures or oblique lines present in the image, see Figure 5.16(d) and Figure 5.16(e). The anisotropy based interpolation couldn't generate the large scale structures that as demonstrated in Figure 5.16(f).



(a) (b) (c)

(d) (e) (f)

Figure 5.16: Row 1 (a), (b), (c) Input images, Row 2 (d), (e), (f) The results of the proposed SSAI model.

5.8 Summary

In this chapter the object removal problem was solved through spatial interpolation framework without involving 'locate and copy' or diffusion of information. A

widely used sampling scheme Symmetric Hierarchical Sampling by remote sensing community was tuned to reduce the complexity of the inpainting problem. Adaptive Basis selection in order to enhance the fidelity of polynomial regression is accomplished through Hurst classifier. This chapter also addressed the ill posed nature of the problem through L_0 regularization principle. The results of the proposed model established a fact that properly modeled spatial interpolation could solve large scale inpainting problems. The proposed model, on the negative side, could not propagate the non-linear structures that are oblique to the direction of assumed global trend. If the model utilizes data selection and maps the input onto manifolds, then the identified limitations may be alleviated.

Chapter 6

Inpaint Agent

By virtue of the theoretical knowledge poured into my inpainting models and empirical evidences offered by the experiments conducted in the previous three chapters, I make an attempt to propose an agent based approach. The functional behavior of this agent is, it suggests the decision rules for selection of the proposed algorithms based on the meta data extracted from the image. The entire work presented in this thesis can be encapsulated as Inpaint Agent(IA)- a Machine Learning model in 6.1 and evaluated in 6.2.

6.1 Proposed Inpaint Agent

Inpaint Agent (see Figure 6.1) takes the input image and the associated mask as inputs. Then it analyzes the morphological features of the mask and the stationary properties of input image. If IA perceives large holes in the mask, it checks the number of holes and the *inpainting class* is resolved. If the class falls into multiple and small inpaint regions then the Hurst exponent (H) is extracted from the image and the following rules are proposed. If the $H < 0.3$ then GAI is activated. Otherwise SAI model is invoked. Thus the confusion created by multiple inpainting classes is resolved deterministically. On the other hand, if the morphological feature of the mask content classifies the problem at hand as large hole problem, then the IA suggests to invoke SSAI model with Constant (SSAI-CONST) as basis

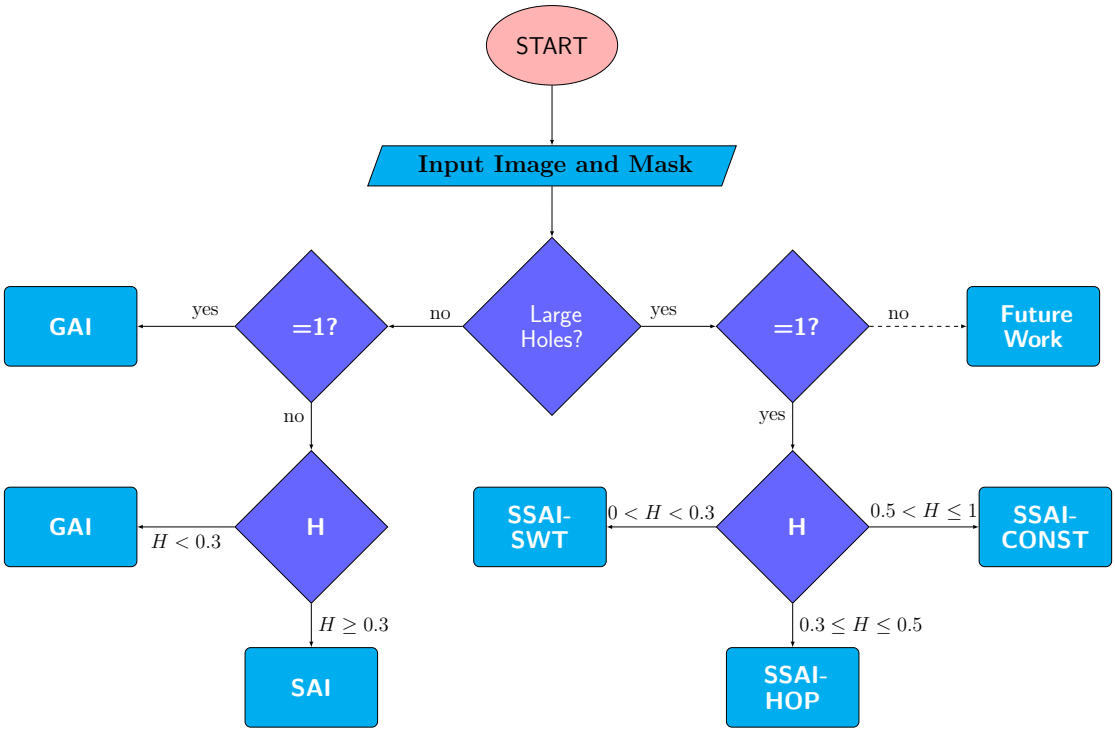


Figure 6.1: The inpaint agent framework which works based on the morphology of inpaint region and nature of the image content

function if Hurst exponent is above 0.5. Yet another rule recommends SSAI with Higher order polynomial (SSAI-HOP) if Hurst exponent is between 0.3 and 0.5. The other rule of says if if the Hurst exponent is below 0.3 then SSAI with Spline Wavelet basis (SSAI-SWT) is preferred.

6.2 Demonstration of Correctness of IA

Overall, our proposed *Inpaint Agent* model offers a complete knowledge about the selection of high fidelity inpainting models under the umbrella of *regularization* schemes. The insight gained by the analysis of mask and the nature of the input image contribute towards selecting a proper sampling plan and launching the appropriate inpainting algorithm. For elastic net regularization, the MATLAB toolbox available at [Schmidtm, 2008] is utilized. Table 6.1 validates the proposed *Inpaint Agent* model by comparing the quality metric PSNR, extracted from the results collected from the experiments conducted across chapters 3, 4

Table 6.1: Expressing the correctness of the proposed Inpaint Agent in PSNR

Input	H	GAI	SAI	SSAI	Recommended	Best
Fruits	0.72	33.38	35.05	9.68	SAI	SAI
Buggy	0.83	26.35	28.23	13.27	SAI	SAI
Trouser	0.19	22.89	32.64	39.76	SSAI	SSAI-SWT
Texture-A	0.14	29.18	28.45	27.58	GAI	GAI
Texture-B	0.14	29.39	30.01	35.05	SSAI	SSAI-SWT
Gantry	0.42	31.20	31.78	46.45	SSAI	SSAI-HOP

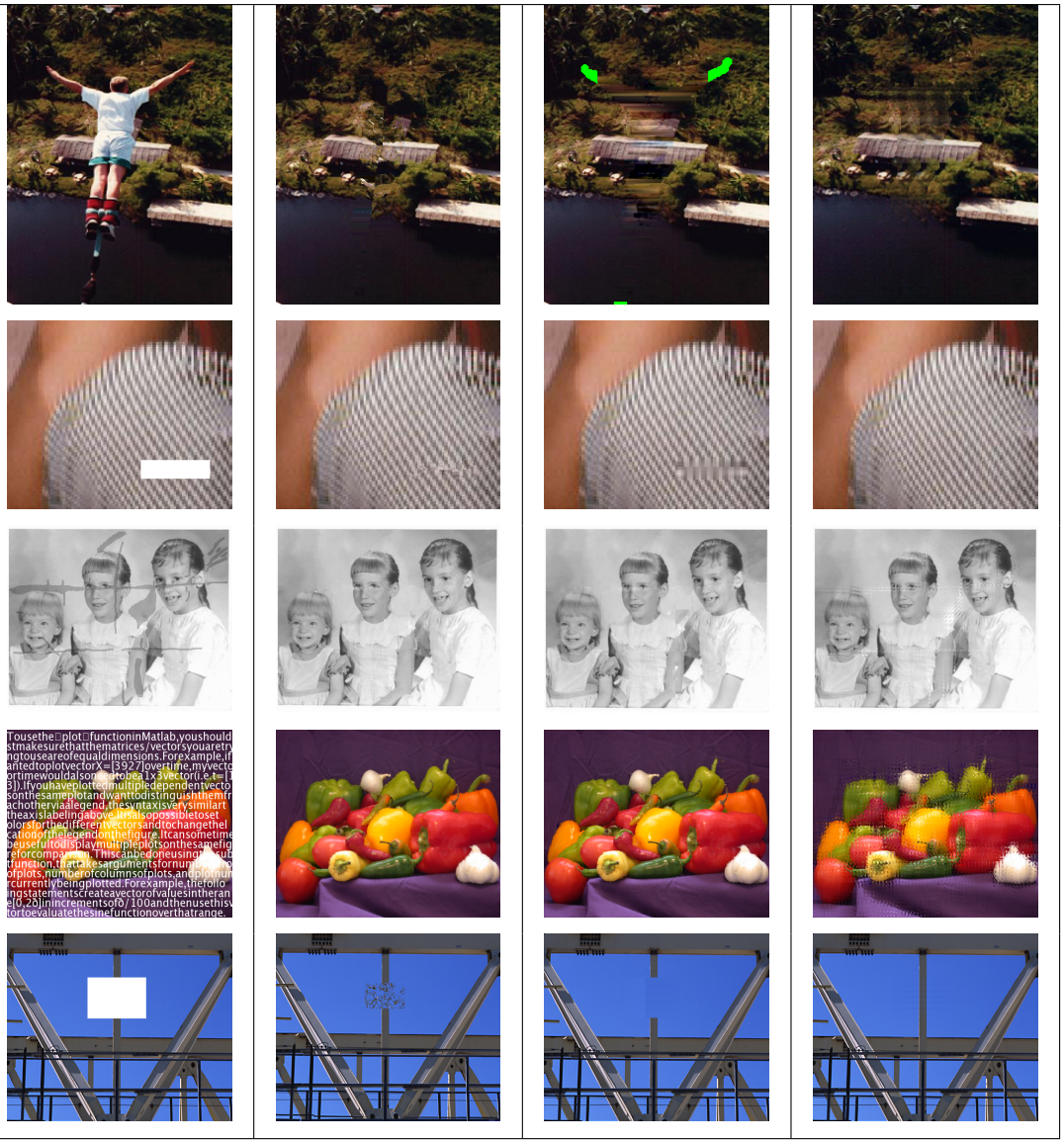
and 5. The input images for which ground truth are available and the results of IA are presented in Table 6.2. Row 5, substantiates the relevance and applicability of the proposed GAI model in the context of sophisticated models SSA and SSAI. If the input image is classified to be non-stationary and problem category is scratch removal, then the observations made from the experiments confirm that non-stationary process can be best approximated by self-similarity measures, which is inherent to GAI model. The text removal category problem is demonstrated by considering a smooth image, see row 2, which has plenty of structures which demands to model the anisotropy. The *Inpaint Agent* recommended SAI model and the experimental results substantiated the structure aware modeling of SAI. The last category of the inpainting involves solving object removal problem and the polynomial basis selection was based on Hurst exponent. The experiments were carried out on synthetic images. The results of this category almost agree with the recommendations of *Inpaint Agent* and substantiate the robustness of the DACE model over the improper selection of trend model in the case of synthetic data. The success of the underlying DACE model is attributed to the feature that the model building works in the parameter space associated with the error but not in the dataspace directly. But, the experiments on the real-world and natural images, established the fact that SSAI-SWT produces visually pleasing results when compared to other models. The summary of contributions and their applicability on different inpainting problems are presented in the Table 6.2 and Table 6.3.

Table 6.2: The nature and applicability of proposed methods

Method	Prior	Mask	Nature	Usecase
GAI	Sparsity based Self-Similarity	Sparse smaller gaps	III	Image Reinvention: Scratch Removal Text Removal
SAI	Moving window with anisotropy	Sparse smaller gaps	I, II	Image Reinvention: Scratch Removal Text Removal
SSAI	Basis Selection and Elastic Net	Large gap	I, II, III	Image Restoration: Image Completion Object Removal

Table 6.3: Comparison of results gained from different proposed algorithms

Input Image	GAI	SAI	SSAI



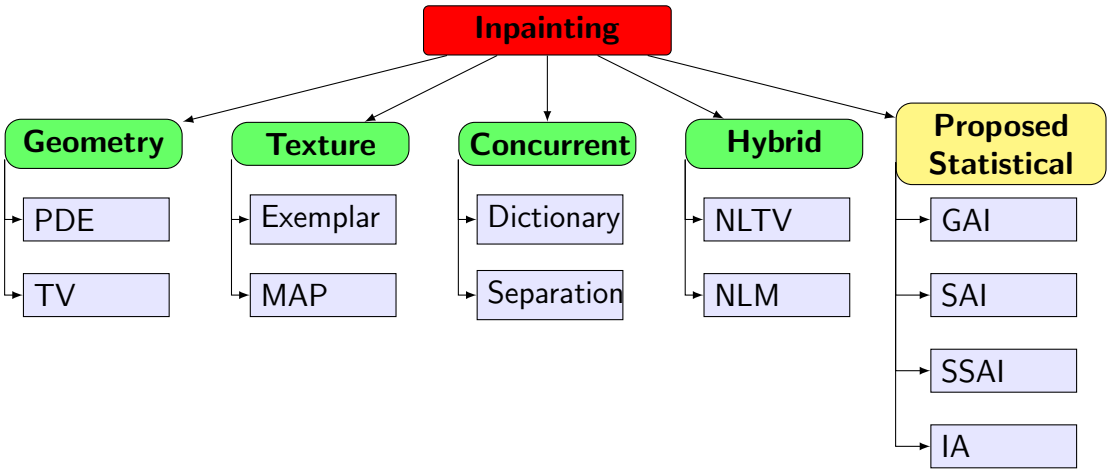


Figure 6.2: Enhanced taxonomy of image inpainting

Figure 6.2 presents the enhanced taxonomy of the image inpainting domain through the contributions made in this thesis.

6.3 Summary

This chapter proposed an intelligent learning based agent *Inpaint Agent* for solving the different categories of inpainting problem based on the morphological features of the mask. The Hurst exponent can rationally classify the images into stationary, second order stationary and non-stationary images and helps to invoke an appropriate algorithm based on the nature of the content. The proposed Inpaint Agent proved to solve the object removal problem at hand and is independent of structure and texture duo. It is quite simple to extend the SSAI to map the spatial coordinates onto high dimensional spaces to realize more structural features and to implement the models in big data realm.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

- The present study demonstrated the efficiency of the proposed regularization approaches in solving the inpainting problem. Self-similarity based regularization method discussed in Chapter 3 is different from that of the methods presented in the chapters 4 and 5. Despite the availability of latest models to address the object removal problem, inpainting solutions suffer from the fact of unease utility of them, wherein the associated mask is very sparse, inferring many disconnected components to be removed. An instance of that problem is the image entirely overlaid with text. The proposed Granular Approach based search space reduction has successfully established that the large scale image inpainting problem can be deterministically solved through L_0 - norm based distance measure. The target front construction step helped to select the target patch based on a simple variance driven priority rule and the local features. Hence the unity of the image could be preserved and the fidelity of the model increased. On the other side, it was not accommodating any spatial relationship among the pixels. This drawback insisted the algorithm to include a post-processing step, which was based on anisotropic filtering. It is evident from the literature that, the self-similarity measure based approaches subjected to a lot of enhancements to exercise high-level

knowledge in various forms for solving the object removal problems. These limitations were addressed in Chapter 4 and Chapter 5 by developing new models that elicit the relationship among the pixels and exploit them to solve the inpainting problem avoiding the post-processing step.

- For the last few decades, the image processing and computer vision domain algorithms are enriched by incorporating the mid-level information derived from the data into the model building step. This ideology has thrown new challenges to the inference algorithm designers. SAI, a moving window based DACE algorithm for inpainting was developed over random fields that satisfy some stationary properties. The model building was based on Systematic Sampling while expressing the spatial anisotropy in terms of a correlation kernel and the regularization of the least square solution. The SAI model demonstrated that the large scale text removal inpainting category problems are well addressed and the results obtained proved to be superior.

- Object removal problems are more challenging as the accuracy of the results goes down drastically with the size of the object. This characteristic compelled the researchers to come up with succinct sampling schemes. The moving window based spatial interpolation failed to solve the object removal problem due to its poor extrapolation capability. Direct extension of Systematic Sampling to Not only Symmetric Hierarchical sampling and generating multiple sub-images helped me to address the design issues effectively. This thesis addressed the trend modeling issue in inference algorithms by developing robust basis selection from a collection of basis functions through Hurst classifier. The SSAI model also utilized the features of higher order polynomial, multiresolution spline wavelet basis and a proactive feature selection approach through Elastic Net regularization. This machine learning feature enabled basis selection problem transparent to the user with increased fidelity of the results and helped to overcome the notion of blind kriging over DACE.

- Overall, this thesis addressed the inpainting problem by implementing the spatial anisotropy in three different forms i) self-similarity measure ii) interpolation and iii) sampling. The statistical approach based *Inpaint Agent* established the rules for invoking an appropriate inpainting algorithm. The experimental results confirmed the relevance and capabilities of a statistical frame work to solve varieties of inpainting problems on par with hybrid methods.

7.2 Future Work

The computer vision community would be very much interested, to have real world applications of the proposed models. Hence, the future enhancements would be (i) establishing of fuzzy membership rules for target-front pixels (ii) effective dimensionality enhancement mechanism to accommodate to implement the higher order locally varying anisotropy.

RESEARCH PUBLICATIONS

1. Raghava,M.,Arun, Agarwal., Raghavendra, Rao C. Spatial Anisotropic Interpolation Approach for Text Removal from Images, MIWAI 2013, LNCS, Vol 8271, 153-164, 2013. (ISBN Number: 978-3-642-44948-2)

2. Raghava,M.,Arun, Agarwal., Raghavendra, Rao C. A Scalable Spatial Anisotropic Interpolation Approach for Object Removal from Images using Elastic Net Regularization, MIWAI 2016, LNCS, Vol 10053, 126 - 140, Dec, 2106. (ISBN Number: 978-3-319-49396-1)

References

[Alahari et al., 2010] Alahari, K., Kohli, P., and Torr, P. H. S. (2010). Dynamic hybrid algorithms for map inference in discrete mrfs. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(10):1846–1857. 41

[Arias et al., 2011] Arias, P., Facciolo, G., Caselles, V., and Sapiro, G. (2011). A variational framework for exemplar-based image inpainting. *Int. J. Comput. Vision*, 93(3):319–347. 47, 130

[Ashikhmin, 2001] Ashikhmin, M. (2001). Synthesizing natural textures. In *Proceedings of the 2001 Symposium on Interactive 3D Graphics*, I3D '01, pages 217–226, New York, NY, USA. ACM. 36

[Aubert and Kornprobst, 2006] Aubert, G. and Kornprobst, P. (2006). *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. Springer-Verlag New York, Inc., Secaucus, NJ, USA. 4, 15, 23, 42

[Averbuch et al., 2006] Averbuch, A., Gelles, G., and Schclar, A. (2006). Fast hole-filling in images via fast comparison of incomplete patches. In *Proc. Multimedia Content Representation, Classification and Security, International Workshop, MRCS*, pages 738–744. 36

[Baatz et al., 2008] Baatz, W., Fornasier, M., Markowich, P. A., and bibiane Schnlieb, C. (2008). Inpainting of ancient austrian frescoes. In *Proc. Bridges 2008, Leeuwarden 2008*, pages 150–156. 18

[Barnes et al., 2009] Barnes, C., Shechtman, E., Finkelstein, A., and Goldman, D. B. (2009). Patchmatch: A randomized correspondence algorithm for structural image editing. *ACM Trans. Graph.*, 28(3):24:1–24:11. 47

[Benzarti and Amiri, 2013] Benzarti, F. and Amiri, H. (2013). Repairing and inpainting damaged images using diffusion tensor. *CoRR*, abs/1305.2221. 20

[Bertalmio et al., 2001] Bertalmio, M., Bertozzi, A. L., and Sapiro, G. (2001). Navier-stokes, fluid dynamics, and image and video inpainting. In *Proc. IEEE Computer Vision and Pattern Recognition (CVPR*, pages 355–362. 15, 27, 28, 42, 80, 81, 97, 98

[Black et al., 1998] Black, M. J., Sapiro, G., Marimont, D. H., and Heeger, D. (1998). Robust anisotropic diffusion. *Trans. Img. Proc.*, 7(3):421–432. 12, 16, 21, 77

[Brus et al., 1996] Brus, D., De Gruijter, J., Marsman, B., Visschers, R., Bregt, A., Breeuwsma, A., and Bouma, J. (1996). The performance of spatial interpolation methods and choropleth maps to estimate properties at points: a soil survey case study. *Environmetrics*, 7(1):1–16. 49, 50

[Buades and Coll, 2005] Buades, A. and Coll, B. (2005). A non-local algorithm for image denoising. In *In CVPR*, pages 60–65. 45

[Bugeau et al., 2010] Bugeau, A., Bertalmío, M., Caselles, V., and Sapiro, G. (2010). A comprehensive framework for image inpainting. *IEEE Transactions on Image Processing*, 19(10):2634–2645. 34

[Burrough and A, 1998] Burrough, P. and A, M. (1998). *Principles Of Geographical Information Systems*. Spatial Information Systems and Geostatistics. Oxford University Press. 84

[Buyssens et al., 2015] Buyssens, P., Daisy, M., Tschumperlé, D., and Lézoray, O. (2015). Exemplar-based inpainting: Technical review and new heuristics for

better geometric reconstructions. *IEEE Transactions on IP*, 24(6):1809–1824. 32, 37, 45, 58

[Carbone, 2007] Carbone, A. (2007). Algorithm to estimate the hurst exponent of high-dimensional fractals. *Phys. Rev. E*, 76:056703. 113

[Chambolle, 2004] Chambolle, A. (2004). An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.*, 20(1-2):89–97. 22

[Chan et al., 2002] Chan, T. F., Kang, S. H., Kang, and Shen, J. (2002). Euler's elastica and curvature based inpaintings. *SIAM J. Appl. Math*, 63:564–592. 3, 25, 26, 27, 28, 81

[Chatterjee et al., 2009] Chatterjee, P., Mukherjee, S., Chaudhuri, S., and Seetharaman, G. (2009). Application of papoulis-gerchberg method in image super-resolution and inpainting. *Comput. J.*, 52(1):80–89. 18

[Chipman et al., 1997] Chipman, H., Hamada, M., and Wu, C. F. J. (1997). A bayesian variable-selection approach for analyzing designed experiments with complex aliasing. *Technometrics*, 39(4):pp. 372–381. 114

[Couckuyt et al., 2012] Couckuyt, I., Forrester, A., Gorissen, D., Turck, F. D., and Dhaene, T. (2012). Blind kriging: Implementation and performance analysis. *Advances in Engineering Software*, 49:1 – 13. 115

[Cressie, 1993] Cressie, N. A. C. (1993). *Statistics for Spatial Data, revised edition*. John Wiley & Sons, New York. 76, 79, 83

[Criminisi et al., 2004] Criminisi, A., Perez, P., and Toyama, K. (2004). Region filling and object removal by exemplar-based image inpainting. *Trans. Img. Proc.*, 13(9):1200–1212. 30, 32, 38, 46, 54, 55

[Drori et al., 2003] Drori, I., Cohen-Or, D., and Yeshurun, H. (2003). Fragment-based image completion. *ACM Trans. Graph.*, 22(3):303–312. 32

[Duan et al., 2015] Duan, J., Pan, Z., Zhang, B., Liu, W., and Tai, X.-C. (2015). Fast algorithm for color texture image inpainting using the non-local ctv model. *Journal of Global Optimization*, 62(4):853–876. 47

[Efron et al., 2004] Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004). Least angle regression. *Annals of Statistics*, 32:407–499. 120

[Efros and Leung, 1999] Efros, A. A. and Leung, T. K. (1999). Texture synthesis by non-parametric sampling. In *Proc. ICCV '99 - Volume 2*, ICCV '99, pages 1033–, Washington, DC, USA. IEEE Computer Society. 28, 30, 55

[Elad et al., 2005] Elad, M., Starck, J.-L., Querre, P., and Donoho, D. (2005). Simultaneous cartoon and texture image inpainting using morphological component analysis (mca). *Applied and Computational Harmonic Analysis*, 19(3):340 – 358. Computational Harmonic Analysis - Part 1. 43

[Fadili and Starck, 2005] Fadili, M. and Starck, J. (2005). Em algorithm for sparse representation-based image inpainting. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 2, pages II–61–4. 44

[Geman and Geman, 1984] Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 6(6):721–741. 76

[Gentile and Meylan, 2013] Gentile, M. Courbin, F. and Meylan, G. (2013). Interpolating point spread function anisotropy. , 549:A1. 48, 83, 89

[George and Mcculloch, 1993] George, E. I. and Mcculloch, R. E. (1993). Variable Selection Via Gibbs Sampling. *Journal of the American Statistical Association*, 88(423):881–889. 114

[Getreuer, 2012] Getreuer, M. (2012). Total Variation Inpainting using Split Bregman. *Image Processing On Line*, 2:147–157. 82, 96, 97, 98, 99, 100, 101, 102

[Gilboa and Osher, 2008] Gilboa, G. and Osher, S. (2008). Nonlocal operators with applications to image processing. *Multiscale Modeling & Simulation*, 7(3):1005–1028. 47

[Goldstein and Osher, 2009] Goldstein, T. and Osher, S. (2009). The split bregman method for l1-regularized problems. *SIAM J. Img. Sci.*, 2(2):323–343. 28, 82

[Grady and Schwartz, 2003] Grady, L. and Schwartz, E. (2003). Anisotropic interpolation on graphs: The combinatorial Dirichlet problem. Technical Report CAS/CNS-TR-03-014, Department of Cognitive and Neural Systems, Boston University, Boston, MA. Submitted to IEEE Pattern Analysis and Machine Intelligence. 19, 80, 81, 84, 85

[Grayson and Blschl, 2004] Grayson, R. and Blschl, G. (2004). *Spatial patterns in catchment hydrologyobservations and modelling*. John Wiley & Sons, Ltd. Cambridge University Press. 92

[Guyon and Elisseeff, 2003] Guyon, I. and Elisseeff, A. (2003). An introduction to variable and feature selection. *J. Mach. Learn. Res.*, 3:1157–1182. 118

[Hadamard, 1923] Hadamard, J. (1923). *Lectures on Cauchy's Problem in Linear Partial Differential Equations*. Yale University Press. 12, 14

[Hastie et al., 2001] Hastie, T., Tibshirani, R., and Friedman, J. (2001). *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc., New York, NY, USA. 64, 77, 87, 119

[He and Sun, 2012] He, K. and Sun, J. (2012). Statistics of patch offsets for image completion. In *Proc. ECCV - Vol Part II*, ECCV'12, pages 16–29, Berlin, Heidelberg. Springer-Verlag. 41, 131

[Hooke and Jeeves, 1961] Hooke, R. and Jeeves, T. A. (1961). " direct search" solution of numerical and statistical problems. *J. ACM*, 8(2):212–229. 94, 118

[in Lee et al., 2006] in Lee, S., Ganapathi, V., and Koller, D. (2006). Efficient structure learning of markov networks using l1regularization. In *In NIPS*. 40

[Ingalls et al., 2004] Ingalls, R. G., Rossetti, M. D., Smith, J. S., Peters, B. A., Beers, W. C. M. V., and Kleijnen, J. P. C. (2004). Kriging interpolation in simulation: A survey. 94

[Jia and Tang, 2003] Jia, J. and Tang, C.-K. (2003). Image repairing: Robust image synthesis by adaptive nd tensor voting. In *Proc. IEEE Computer Society Conference on CVPR*, CVPR'03, pages 643–650, Washington, DC, USA. IEEE Computer Society. 20

[Joseph et al., 2008] Joseph, V. R., Hung, Y., and Sudjianto, A. (2008). Blind Kriging: A New Method for Developing Metamodels. *Journal of Mechanical Design*, 130(3):031102+. 115

[Komodakis, 2006] Komodakis, N. (2006). Image completion using global optimization. In *Proc. IEEE Computer Society CVPR '06 - Volume 1*, CVPR '06, pages 442–452, Washington, DC, USA. IEEE Computer Society. 32, 39

[Kumar et al., 2008] Kumar, N., Zhang, L., and Nayar, S. K. (2008). What is a good nearest neighbors algorithm for finding similar patches in images? In *Proc. ECCV)*. 36

[Le Meur et al., 2011] Le Meur, O., Gautier, J., and Guillemot, C. (2011). Examplar-based inpainting based on local geometry. In *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pages 3401–3404. 37, 38, 59

[Levin et al., 2003] Levin, A., Zomet, A., and Weiss, Y. (2003). Learning how to inpaint from global image statistics. In *9th IEEE International Conference on Computer Vision (ICCV 2003), 14-17 October 2003, Nice, France*, pages 305–312. 38

[Li, 2009] Li, S. Z. (2009). *Markov Random Field Modeling in Image Analysis*. Springer Publishing Company, Incorporated, 3rd edition. 39

[Lockwood and Anitescu, 2012] Lockwood, B. A. and Anitescu, M. (2012). Gradient-enhanced universal kriging for uncertainty propagation. *Nuclear Science and Engineering*, 170(2):168–195. 93

[Lophaven and et al., 2002] Lophaven, S. N. and et al. (2002). Dace – a matlab kriging toolbox – version 2.0. 83, 87, 95

[Lowe, 1999] Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *Proc. ICCV - Volume 2*, ICCV '99, pages 1150–, Washington, DC, USA. IEEE Computer Society. 67

[Maesschalck et al., 2000] Maesschalck, R. D., Jouan-Rimbaud, D., and Massart, D. (2000). The mahalanobis distance. *Chemometrics and Intelligent Laboratory Systems*, 50(1):1 – 18. 112

[Martin et al., 2001] Martin, D., Fowlkes, C., Tal, D., and Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. 8th Int'l Conf. Computer Vision*, volume 2, pages 416–423. 69

[Masnou and Morel, 1998] Masnou, S. and Morel, J. (1998). Level lines based disocclusion. In *ICIP (3)*, pages 259–263. 14, 22, 23, 26, 54, 81

[McKay, 1992] McKay, M. D. (1992). Latin hypercube sampling as a tool in uncertainty analysis of computer models. In *Proceedings of the 24th Conference on Winter Simulation*, WSC '92, pages 557–564, New York, NY, USA. ACM. 90

[Meyer, 2001] Meyer, Y. (2001). *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*. American Mathematical Society, Boston, MA, USA. 41, 42

[Miller, 2002] Miller, A. J. (2002). *Subset selection in regression*. Monographs on statistics and applied probability. Chapman & Hall/CRC, Boca Raton. 118

[Nicolis, 2006] Nicolis, Orietta ; Garutti, C. . V. B. (2006). 2-d wavelet-based spectra with application in analysis of geophysical images. 113

[Nitzberg et al., 1993a] Nitzberg, M., Mumford, D., and Shiota, T. (1993a). *Filtering, Segmentation and Depth*, volume 662 of *Lecture Notes in Computer Science*. Springer. 14

[Nitzberg et al., 1993b] Nitzberg, M., Mumford, D., and Shiota, T. (1993b). *Filtering, Segmentation and Depth*, volume 662 of *Lecture Notes in Computer Science*. Springer. 22

[Oliveira et al., 2001] Oliveira, M. M., Bowen, B., McKenna, R., and sung Chang, Y. (2001). Fast digital image inpainting. In *Proc. VIIP 2001*, pages 261–266. ACTA Press. 20, 65

[Oliver et al., 1989] Oliver, M., Webster, R., and Gerrard, J. (1989). Geostatistics in physical geography. part i: Theory. *Transactions of the Institute of British Geographers*, 14(3):pp. 259–269. 93

[Perona and Malik, 1990] Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(7):629–639. 16

[Peyré et al., 2008] Peyré, G., Bougleux, S., and Cohen, L. D. (2008). Non-local regularization of inverse problems. In Forsyth, D. A., Torr, P. H. S., and Zisserman, A., editors, *Proc. of ECCV'08*, volume 5304 of *Lecture Notes in Computer Science*, pages 57–68. Springer. 47

[Rajwade et al., 2013] Rajwade, A., Rangarajan, A., and Banerjee, A. (2013). Image denoising using the higher order singular value decomposition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(4):849–862. 45

[Ripley, 1991] Ripley, B. D. (1991). *Statistical inference for spatial processes*. Cambridge university press. 49

[Roth and Black, 2005] Roth, S. and Black, M. J. (2005). Fields of experts: A framework for learning image priors. In *In CVPR*, pages 860–867. 40, 69, 97, 99, 100, 101

[Rudin et al., 1992] Rudin, L. I., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268. 21, 22, 41, 82

[Santner et al., 2003] Santner, T. J., Williams, B. J., and Notz, W. (2003). *The design and analysis of computer experiments*. Springer series in statistics. Springer, New York. 83

[Schmidtm, 2008] Schmidtm (2008). 120, 136

[Shewchuk, 1994] Shewchuk, J. R. (1994). An introduction to the conjugate gradient method without the agonizing pain. Technical report, Pittsburgh, PA, USA. 79

[Sjöstrand, 2005] Sjöstrand, K. (2005). Matlab implementation of LASSO, LARS, the elastic net and SPCA. Version 2.0. 120

[Starck et al., 2005] Starck, J.-L., Elad, M., and Donoho, D. L. (2005). Image decomposition via the combination of sparse representations and a variational approach. *IEEE Transactions on Image Processing*, 14(10):1570–1582. 44, 129

[Stein, 1999] Stein, M. (1999). *Interpolation of Spatial Data: Some Theory for Kriging*. Springer Series in Statistics. Springer New York. 83, 85

[Tenenbaum et al., 2000] Tenenbaum, J. B., de Silva, V., and Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319. 94

[Tibshirani, 1994] Tibshirani, R. (1994). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58:267–288. 118

[Tobler, 1970] Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46:234–240. 49

[Tschumperle and Deriche, 2005] Tschumperle, D. and Deriche, R. (2005). Vector-valued image regularization with pdes: a common framework for different applications. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(4):506–517. 45

[Unser and Blu, 2000] Unser, M. and Blu, T. (2000). Fractional splines and wavelets. *SIAM Review*, 42(1):43–67. 115, 116

[Van De Ville et al., 2005] Van De Ville, D., Blu, T., and Unser, M. (2005). Isotropic polyharmonic B-Splines: Scaling functions and wavelets. *IEEE Transactions on Image Processing*, 14(11):1798–1813. 113, 114

[Wackernagel, 2003] Wackernagel, H. (2003). *Multivariate Geostatistics: An Introduction with Applications*. Springer Verlag. 76

[Wang et al., 2013] Wang, M., Yan, B., and Ngan, K. N. (2013). An efficient framework for image/video inpainting. *Signal Processing: Image Communication*, 28(7):753 – 762. 125

[Wang et al., 2008] Wang, Y., Yang, J., Yin, W., and Zhang, Y. (2008). A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. Img. Sci.*, 1(3):248–272. 22, 99

[Wang et al., 2004] Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E. (2004). Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on*, 13(4):600–612. 51, 52, 53

[Weickert, 1999] Weickert, J. (1999). Coherence-enhancing diffusion filtering. 20

[Wexler et al., 2007] Wexler, Y., Shechtman, E., and Irani, M. (2007). Space-time completion of video. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(3):463–476. 32, 47

[Wong and Orchard, 2008] Wong, A. and Orchard, J. (2008). A nonlocal-means approach to exemplar-based inpainting. In *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, pages 2600–2603. 46

[Xu and Sun, 2010] Xu, Z. and Sun, J. (2010). Image inpainting by patch propagation using patch sparsity. *Trans. Img. Proc.*, 19(5):1153–1165. 37, 38

[Yedidia et al., 2003] Yedidia, J. S., Freeman, W. T., and Weiss, Y. (2003). Exploring artificial intelligence in the new millennium. chapter Understanding Belief Propagation and Its Generalizations, pages 239–269. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. 40

[Zou and Hastie, 2005] Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B*, 67:301–320. 119