Single Image Super Resolution using Gaussian Process and Back-Projection Algorithm

MASTERS RESEARCH PROJECT REPORT

Submitted by

DANY VARGHESE (JYANCCS002)

in partial fulfillment for the award of the degree of

MASTER OF TECHNOLOGY (M.TECH)

in

COMPUTER SCIENCE AND ENGINEERING

of

UNIVERSITY OF CALICUT

Under the guidance of

Mr. Viju Shankar

MAY 2015

Department of Computer Science and Engineering

JYOTHI ENGINEERING COLLEGE, CHERUTHURUTHY

THRISSUR 679 531

Single Image Super Resolution using Gaussian Process and Back-Projection Algorithm

MASTERS RESEARCH PROJECT REPORT

Submitted by

DANY VARGHESE (JYANCCS002)

in partial fulfillment for the award of the degree of

MASTER OF TECHNOLOGY (M.TECH)

in

COMPUTER SCIENCE AND ENGINEERING

of

UNIVERSITY OF CALICUT

Under the guidance of

Mr. Viju Shankar

MAY 2015

Department of Computer Science and Engineering

JYOTHI ENGINEERING COLLEGE, CHERUTHURUTHY

THRISSUR 679 531

Department of Computer Science and Engineering

JYOTHI ENGINEERING COLLEGE, CHERUTHURUTHY

THRISSUR 679 531

MAY 2015

BONAFIDE CERTIFICATE

This is to certify that the Masters Research Project Report entitled **Single Image Super Resolution using Gaussian Process and Back-Projection Algorithm** submitted by **DANY VARGHESE (JYANCCS002)** in partial fulfillment of the requirements for the award of **Master of Technology** degree in **Computer Science and Engineering** of **University of Calicut** is the bonafide work carried out by them under our supervision and guidance.

Mr.Viju Shankar	**Mr. Aneesh Chandran**	**Mr.Muralee Krishnan C**
Seminar Guide	Project Coordinator	Professor & Head of The Dept.
Assistant Professor	Assistant Professor	Dept. of CSE
Dept. of CSE	Dept. of CSE	

Department of Computer Science and Engineering

JYOTHI ENGINEERING COLLEGE, CHERUTHURUTHY

THRISSUR 679 531

MAY 2015

DECLARATION

I hereby declare that the work presented in this Masters Research Project titled **"Single Image Super Resolution using Gaussian Process and Back-Projection Algorithm"** is based on the original work done by under the guidance of **Mr. Viju Shankar**, Assistant Professor, Department of Computer Science & Engineering and has not been included in any other thesis submitted previously for the award of any degree.

Place : **Dany Varghese**

Date :

To

My Parents

Varghese K J & Elsy Varghese

ACKNOWLEDGEMENT

I take this opportunity to express my heartfelt gratitude to all respected personalities who had guided, inspired and helped me in the successful completion of this Masters Research Project. First and foremost, I express my thanks to **The Lord Almighty** for guiding me in this endeavour and making it a success.

I take immense pleasure in thanking the **Management** of Jyothi Engineering College and **Dr. k k Babu**, principal, Jyothi Engineering College for having permitted me to carry out this Masters Research Project. My sincere thanks to **Prof. Muralee Krishnan C**, Head of the Department of Computer Science and Engineering for permitting me to make use of the facilities available in the department to carry out the seminar successfully.

I express my sincere gratitude to **Mr. Aneesh Chandran** and **Ms. Jisna V A**, Project Co-ordinators for their invaluable supervision and timely suggestions. I am very happy to express my deepest gratitude to my mentor **Mr. Viju Shankar**, Assistant Professor, Department of Computer Science and Engineering, Jyothi Engineering College for his able guidance and continuous encouragement.

Last but not least I extend my gratefulness to all teaching and non teaching staffs who directly or indirectly involved in the successful completion of this thesis work and to all my friends who have patiently extended all sorts of help for accomplishing this undertaking.

ABSTRACT

In most digital imaging applications, high resolution images or videos are usually desired for later image processing and analysis. The desire for high image resolution stems from two principal application areas: improvement of pictorial information for human interpretation; and helping representation for au- tomatic machine perception. Image resolution describes the details contained in an image, the higher the resolution, the more image details. The resolution of a digital image can be classified in many different ways: pixel resolution,spatial resolution, spectral resolution, temporal resolution, and radiometric resolution. In this context, we are mainly interested in spatial resolution.

Super-resolution (SR) are techniques that construct high-resolution (HR) images from several observed low-resolution (LR) images, thereby increasing the high frequency components and removing the degradations caused by the imaging process of the low resolution camera. The basic idea behind SR is to combine the non-redundant information contained in multiple low-resolution frames to generate a high-resolution image. A closely related technique with SR is the single image interpolation approach, which can be also used to increase the image size. However, since there is no additional information provided, the quality of the single image interpolation is very much limited due to the ill-posed nature of the problem, and the lost frequency components cannot be recovered. In the SR setting, however, multiple low-resolution observations are available for reconstruction, making the problem better constrained. The non-redundant information contained in the these LR images is typically introduced by sub-pixel shifts between them. These sub-pixel shifts may occur due to uncontrolled motions between the imaging system and scene, e.g., movements of objects, or due to controlled motions, e.g., the satellite imaging system orbits the earth with predefined speed and path.

CONTENTS

List of Figures

List of Abbreviations

- **A/D** : Analog to Digital
- **CCD** : Charged-coupled Device
- **CMOS** : Complementary metal oxide semiconductor
- **EDI** : Edge Directed Interpolation
- **GP** : Gaussian Process
- **GPR** : Gaussian Process Regression
- **HR** : High Resolution
- **IBP** : Iterative Back Projection
- **ICBI** : Iterative Curvature-Based Interpolation
- **iEDI** : improved Edge Directed Interpolation
- **ISO** : International Organization for Standardization
- **LR** : Low Resolution
- **LW/PH** : Line Widths per Picture Height
- **MAP** : Maximum a Posteriori
- **ML** : Maximum Likelihood
- **MLE** : Maximum Likelihood Estimate
- **MRF** : Markov Random Field
- **MRI** : Magnetic Resonance Imaging
- **MTF** : Modulation Transfer Function
- **NEDI** : New Edge Directed Interpolation
- **NMF** : Non-negative Matrix factorization
- **PCA** : Principal Component Analysis
- **PET** : Positron Emission Tomography
- **POCS** : Projection onto convex sets
- **PSF** : Point Spread Function
- **PSNR** : Peak Signal to Noise Ratio
- **SFR** : Spatial Frequency Response
- **SR** : Super Resolution

CHAPTER 1

Introduction

Over the past ten years, digital cameras have gone through a fast evolution towards extremely compact models, containing sensors with a steadily increasing number of pixels. From about 0.3 mega-pixels (million pixels) in 1993, the number of pixels on the CCD or CMOS sensor in a digital camera has increased to 39 mega pixels in some of the latest professional models. This pixel count has become the major selling argument for the different camera manufacturers [4].

The number of pixels in a digital image is also often referred to as the resolution of an image. The ever-increasing demand for more pixels, or higher resolution, in combination with the availability of more and more computational power, has generated a large interest in super-resolution imaging. The goal in super-resolution imaging is to take multiple 'low' resolution images of the same scene, and combine them to generate a 'higher' resolution image. In this way, a photographer could for example take a series of four images using a four mega-pixel camera, and combine them to obtain an image as if it would be taken with a sixteen mega-pixel camera. And who would not be interested in such a feature?

In practice, such a combination of information from multiple images is not trivial. There are two main problems that need to be solved in a super resolution algorithm. First, all the input images need to be correctly aligned with each other on a common grid[39]. Next, an accurate, sharp image has to be reconstructed from the gathered information. If one of these two steps is not well done, the resulting image is not good, and no gain in resolution is obtained. In this thesis, we mainly address the first problem, more specifically the alignment of aliased input images. An image is aliased if there are not enough sampling points (pixels) to represent the high frequencies in the scene. This typically results in artificial patterns or jagged edges in

the image. If the images are not too severely aliased, we will show that it is possible to use the aliasing-free part of the images to align them one by one to a reference image. If there is severe aliasing, the different images need to be aligned jointly. In that case, the alignment is a highly non-linear problem. Multiple solution methods for such an alignment are proposed. The presented image alignment methods can be applied to different application domains, such as consumer digital cameras, satellite imaging etc.

This thesis addresses the problem of image super-resolution, which is the process by which one or more low-resolution images are processed together to create an image (or set of images) with a higher spatial resolution.

Single image super-resolution refers to the particular case where single images of the scene are available. In general, changes in these low-resolution images caused by camera or scene motion, camera zoom, focus and blur mean that we can recover extra data to allow us to reconstruct an output image at a resolution above the limits of the original camera or other imaging device [1] .

Problems motivating super-resolution arise in a number of imaging fields, such as satellite surveillance pictures and remote monitoring , where the size of the CCD array used for imaging may introduce physical limitations on the resolution of the image data. Medical diagnosis may be made more easily or more accurately if data from a number of scans can be combined into a single more detailed image. A clear, high-quality image of a region of interest in a video sequence may be useful for facial recognition algorithms, car number plate identification, or for producing a "print-quality" picture for the press.

1.1 Motivation & Scope

The motivation behind SR is quite clear: there are many situations where the resolution that can give a sensor is limited because of physical or economical constraints[34] . SR can

improve the resolution in many cases where other techniques are not feasible. The ample range of applications that we exposed in the previous section exemplifies this. The specific objectives that we pursue in this dissertation are explained below. In the super-resolution literature, the representation of the system as a discrete linear system has been the mainstream in most recent publications. The trend is to solve the problem as a whole, performing jointly as most tasks as possible: registration, interpolation, de-blurring and restoration [8] . Although supposedly we could then achieve a global optimum that would be unattainable otherwise, this leads to a very complex problem with too many degrees of freedom.

1.2 Application

Super-resolution techniques can be applied in various domains. As described above, in consumer imaging, one could imagine a digital camera that takes a burst of pictures instead of a single picture[2] . From these images, which have typically small relative shifts due to the shaking of the user's hands, a high resolution image can be reconstructed. Similarly, in satellite or spatial imaging, a set of satellite images could be combined to display fine details that are not distinguishable in any of the input images [34]. Super-resolution methods can also be used to create high resolution still pictures or video from video sequences. In surveillance cameras, additional details can be revealed by combining multiple video frames to create a single high resolution image. The same techniques can also be applied to improve the resolution of existing (low resolution) video content for use in high definition television sets.

A similar approach for one-dimensional signals is used in high rate analog to- digital (A/D) converters. If the rate at which the analog signal has to be sampled becomes too high, it is physically very difficult to build such converters. Instead of a single converter at a high rate, multiple converters at a lower rate are then used in parallel [16]. Each of the low rate converters has a small relative offset, such that the high rate signal can be reconstructed by combining the different low rate signals. In the ideal case, with for example two low rate converters, the

samples of the second converter are taken exactly in the middle between the samples of the first converter. The two signals can then be interleaved to obtain a signal at twice the rate. However, the precise synchronization of such converters is very difficult [10]. This is exactly the same problem as the alignment of images in super-resolution imaging, but now for one-dimensional signals. The methods described in the following chapters can therefore also be applied to such problems.

Applications for the techniques of super-resolution restoration from image sequences appear to be growing rapidly as the theory gains exposure. Continued research and the availability of fast computational machinery have made these methods increasingly attractive in applications requiring the highest restoration performance. If we consider that in many applications the increase in sensor resolution comes at a high cost and that increase could be carried out using SR techniques, we can easily see that in many cases is preferable to employ SR instead of increasing physically the sensor resolution. Of course in some cases this is difficult: complicated registration of the images or resolution already limited by the diffraction limit may prevent the use of SR techniques. But as SR theory starts to stabilize, applications start to soar.

SR restoration techniques have already been applied to problems in:

- Black and white photography [33]
- Satellite imaging [25]
- Astronomical imaging [20]
- Video enhancement and restoration [40]
- Video standards conversion [27]
- Confocal Microscopy
- Applied jointly with mosaicing
- Aperture displacement cameras

- Diffraction tomography
- Restoration of MPEG-coded video streams
- Magnetic Resonance Imaging (MRI)
- Positron Emission Tomography (PET)
- Applied to photographic camera color images
- Forward looking infrared cameras

1.3 Challenges

Super-resolution algorithms face a number of challenges in parallel with their main super-resolution task. In addition to being able to compute values for all the superresolution image pixels intensities given the low-resolution image pixel intensities, a super-resolution system must also be able to handle:

- Image registration : small image displacements are crucial for beating the sampling limit of the original camera, but the exact mappings between these images are unknown. To achieve an accurate super-resolution result, they need to be found as accurately as possible.
- Lighting variation : when the images are aligned geometrically, there may still be significant photometric variation, because of different lighting levels or camera exposure settings when the images were captured.
- Blur identification -:before the light from a scene reaches the film or camera CCD array, it passes through the camera optics. The blurs introduced in this stage are modelled by a point-spread function. Separating a blur kernel from an image is an extensively-studied and challenging problem known as Blind Image Deconvolution. This can be even more challenging if the blur varies spatially across the image.

CHAPTER 2

Literature Survey

2.1 Introduction

Image super-resolution is a well-studied problem. A comprehensive review was carried out by Borman and Stevenson [7] in 1998, though in the intervening period a wide variety of super-resolution work has contributed to many branches of the super-resolution problem.

There are several popular approaches to super-resolution overall, and a number of different fundamental assumptions to be made, which can in turn lead to very different super-resolution algorithms. Model assumptions such as the type of motion relating the low-resolution input images, or the type of noise that might exist on their pixel values, commonly vary significantly, with many approaches being tuned to particular assumptions about the input scenes and configurations that the authors are interested in. There is also the question of whether the goal is to produce the very best high-resolution image possible, or a passable high-resolution image as quickly as possible.

The formulation of the problem usually falls into one of two main categories: either Maximum Likelihood (ML) methods, i.e. those which seek a super-resolution image that maximizes the probability of the observed low-resolution input images under a given model, or Maximum a Posteriori (MAP) methods, which make explicit use of prior information, though these are also commonly couched as regularized least-squares problems.

The goal of super-resolution (SR) is to estimate a high resolution (HR) image from one or a set of low-resolution (LR) images. This inverse problem is inherently ill-posed since many HR images can produce the same LR image. SR methods can be broadly categorized into three

classes as in [23] interpolation-based methods, learning-based methods and reconstruction-based methods. Interpolation based methods (e.g., [23, 18, 41]) are fast but the results are lack of fine details. Reconstruction based methods (e.g., [6, 22, 26, 28, 19] apply various smoothness priors (e.g., [3, 31, 17]) and impose the constraint that when properly down sampled, the HR image should reproduce the original LR image. Alternatively, in learning-based methods (e.g., [8, 42, 14, 29]), detailed textures are hallucinated by searching through a training set of LR/HR image or patch pairs. However, note that the training images need to be carefully selected, otherwise unexpected details can be found [12].

We begin with the historically earliest methods, which tend to represent the simplest forms of the super-resolution ideas. From these grow a wide variety of different approaches, which we then survey, turning our attention back to more detailed model considerations, in particular the ways in which registration and blur estimation are handled in the literature. Finally, we cover a few super-resolution approaches which fit less well into any of the categories or methods of solution listed so far.

2.2 Early super-resolution methods

Super-Resolution was first proposed in a seminal paper by Tsai and Huang (1984)[37], who suggested a frequency domain approach to super-resolution. This initial approach assumes a translational movement among the LR images and is based on the following principles:

- The shifting property of the Fourier transform.
- The aliasing relationship between the continuous Fourier transform of the ideal HR image and the discrete Fourier transform of the observed LR images.
- The assumption that the ideal HR image is band-limited.

The field of super-resolution began growing in earnest in the late eighties and early nineties. In the signal processing literature, approaches were suggested for the processing

of images in the frequency domain in order to recover lost high-frequency information, generally following on from the work of Tsai and Huang [37] and later Kim et al. [24]. Roughly concurrent with this in the Computer Vision arena was the work of Peleg, Keren, Irani and colleagues [22], which favoured the spatial domain exclusively, and proposed novel methods for super-resolution reconstruction.

2.2.1 Simple super-resolution schemes

Enhancement by frame-fusion is possible when we have several images of the same surface and a dense correspondence (from either geometric or general registration) between the images. It attempts to reverse all of the degradations at once, recovering a de-blurred, high spatial-resolution surface texture.

The simplest methods attempt to re-sample all the observed image data into a single coordinate frame. The re-sampled data is merged by averaging or median filtering to obtain a low-noise, high-density image. A standard, single-image de-blurring step completes the process.

Early work was done by Gross [38], based on generalized sampling theory. They assume that the input images undergo relative shifts which are known precisely a priori. The low-resolution images are interpolated and merged onto a finer grid, before deblurring by convolution with a kernel derived from the inverse of the blur operator. A similar technique has been proposed by Rudin et al. [30]. They use a hierarchical block-matching algorithm to obtain a dense optic flow field. One view is chosen as a reference frame and its resolution is increased to the desired level using a simple interpolation kernel. The other images are warped and merged into the reference frame according to the optic low field. Finally, a standard single-image deblurring algorithm is applied to obtain the super resolution result.

2.2.2 Methods using a generative model

More recent methods use a generative model of the camera transfer function which determines how a real surface is transformed, filtered and sampled to form an image; and also an accurate set of registrations between the input images. They proceed by finding a high resolution image which, when transformed according to the registration parameters and degraded according to the camera transfer function, produces a set of simulated images which are as similar as possible to the actual observed low-resolution images.

Under the assumption of Gaussian distributed image noise, these methods produce the maximum likelihood estimate (MLE) of the super-resolved surface intensities. It is also common to impose a prior model of the surface intensities (such as a spatial-smoothness constraint), in which case the intensity estimates obtained are the maximum a posterior estimates (MAP).

Algorithms proposed vary in their registration methods, their use of prior texture models, and the numerical methods used to converge to the required estimate. Irani and Peleg [22] consider images obtained from a scanner which have undergone rotation and translation (a Euclidean transformation). They consider optical blur, and obtain the kernel of the point spread function (PSF) by imaging a small dot. Their cost function to be minimized is the sum of squared differences in intensity values between the simulated low-resolution images and the actual ones (see equation 2.1 in which $\hat{g_n}$ is the nth simulated image and g_n the associated actual low-resolution image), and the super-resolved texture (typically double the resolution of the input images) is found by a simple iterative update scheme similar to steepest descent. Mann and Picard extended Irani and Peleg's algorithm to fully projective image registration, obtained using a coarseto-fine texture correlation strategy.

$$\text{cost}=\sum_n \sum_{(x,y)} (\hat{g_n}(x,y) - g_n(x,y))^2 \tag{2.1}$$

Elad & Feuer also make use of a generative model within a POCS based approach to

super-resolution. They assume that accurate registration is known a priori, and fail to present any results based on real images. Tom & Katsagellos [36] propose a frequency domain method based on a generative model, in which the registration (limited to translation) and super-resolution image are estimated simultaneously. Again, no real image results are presented. Delleart et al. [13] propose an on-line method for the continuous super-resolution update of a patch of texture as it is tracked. The affine tracking algorithm uses the current estimate of the patch as a template, and the super-resolution update is performed using a very simplified Kalman-style update scheme. Shekarforoush [15] also propose several sequential update schemes for super-resolution estimation using a generative model. Additionally, they propose a novel method for estimating the blur function based on analysis of the cross-power spectrum of the images. Their method is applied to real image sequences, with unconvincing results.

2.3 Frequency domain methods

The basic frequency-domain super-resolution problem of Tsai and Huang [37] or Kim et al.[24] looks at the horizontal and vertical sampling periods in the digital image, and relates the continuous Fourier transform of the original scene to the discrete Fourier transforms of the observed low-resolution images. Both these methods rely on the motion being composed purely of horizontal and vertical displacements, but the main motivating problem of processing satellite imagery is amenable to this restriction.

The ability to cope with noise on the input images is added in by Kim et al. in [17], and Tekalp et al. [10] generalize the technique to cope with both noise and a blur on the inputs due to the imaging process.

Tom and Katsaggelos [36] etake a two-phase super-resolution approach, where the first step is to register, deblur and de-noise the low-resolution images, and the second step is to interpolate them together onto a high-resolution image grid. While much of the problem is

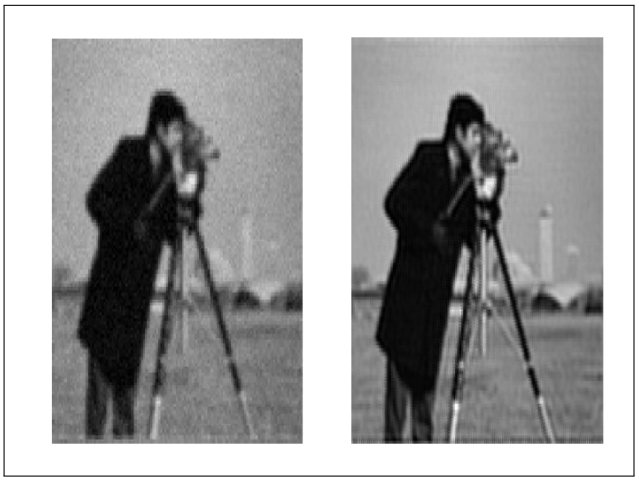


Figure 2.1: Example of Frequency-domain super-resolution
taken from Tom and Katsaggelos [36]. Left: One of four synthesized low-resolution images. Right: Super-resolved image. Several artifacts are visible, particularly along image edges, but the overall image is improved.

specified here in the spatial domain, the solution is still posed as a Frequency domain problem.

Lastly, wavelet models have been applied to the problem, taking a similar overall approach as Fourier-domain methods. Nguyen and Milanfar [15] propose an efficient agorithm based on representing the low-resolution images using wavelet coefficients, and relating these coefficients to those of the desired super-resolution image. Bose et al. [32] also propose a method based on second-generation wavelets, leading to a fast algorithm. However, while this shows good results even in the presence of high levels of input noise, the outputs still display wavelet-like high frequency artifacts.

2.4 spatial domain methods

Spatial domain methods enjoy better handling of noise, and a more natural treatment of the image point-spread blur in cases where it cannot be approximated by a single convolution operation on the high-resolution image (e.g. when the zoom or shear in the low-resolution image registrations varies across the inputs). They use a model of how each offset low-resolution pixel is derived from the high-resolution image in order to solve for the high-resolution pixel

values directly.

The potential for using sub-pixel motion to improve image resolution is highlighted by Peleg et al. [21]. They point out that a blurred high-resolution image can be split into 16 low-resolution images at a zoom factor of 4 by taking every 4th pixel in the horizontal and vertical directions at each of the 4×4 different offsets. If all 16 low-resolution images are available, the problem reduces to one of regular image deblurring, but if only a subset of the low-resolution frames are present, there is still a clear potential for recovering high-frequency information, which they illustrate on a synthetic image sequence.

A method for registering a pair of images is proposed in Keren et al. a38. The registration deals with 2D shifts and with rotations within the plane of the image, so is more general than that of . However, their subsequent resampling and interpolation scheme do little to improve the high-frequency information content of the super-resolution image.

Progress was made by Irani et al. [21], who used the same registration algorithm, but proposed a more sophisticated method for super-resolution image recovery based on back-projection similar to that used in Computer Aided Tomography. They also propose a method for estimating the point-spread function kernel responsible for the image blur, by scanning a small white dot on a black background with the same scanner.

To initialize the super-resolution algorithm, Irani et al. take an initial guess at the high resolution image, then apply the forward model of the image measurement process to work out what the observed images would be if the starting guess was correct. The difference between these simulated low-resolution images and the real input images is then projected back into the high-resolution space using a back projection kernel so that corrections can be made to the estimate, and the process repeated. It is also worth observing that at this point, the algorithms proposed for spatial-domain super-resolution all constitute Maximum Likelihood methods.

Later work by Irani et al. [22] builds on their early super-resolution work, though the focus shifts to object tracking and image registration for tracked objects, which allows the super-

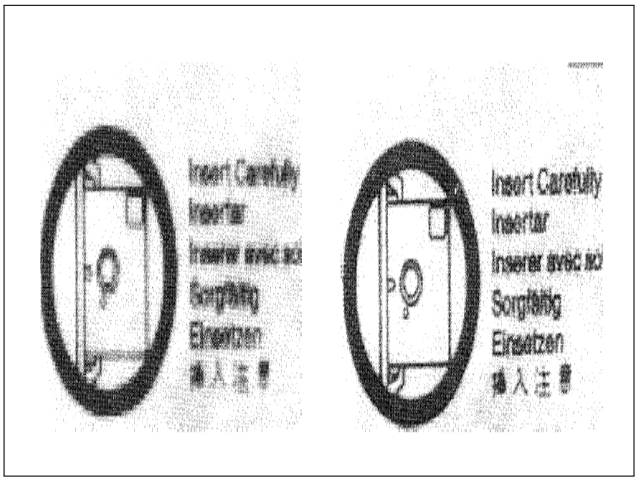


Figure 2.2: Example of early spatial-domain super-resolution
taken from Irani et al. [59]. Left: One of three low-resolution images (70 × 100 pixels) obtained using a scanner. Right: Super-resolved image. The original images were captured using a scanner.

resolution algorithm to treat some rigid moving objects independently in the same sequence. Another similar method by Zomet et al. [1] proposes the use of medians within their super-resolution algorithm to make the process robust to large outliers caused by parallax or moving specularities.

A good example of limiting the imaging model in order to achieve a fast superresolution algorithm comes from Elad and Hel-Or . The motion is restricted to shifts of integer numbers of high-resolution pixels (though these still represent sub-pixel shifts in the low-resolution pixels), and each image must have the same noise model, zoom factor, and spatially invariant point-spread function, the latter of which must also be realisable by a block-circulant matrix. These conditions allow the blur to be treated after the interpolation onto a common high-resolution frame ,this intuition is exactly the same as in [13, 20], but the work of [33] formalizes it and explores more efficient methods of solution in more depth.

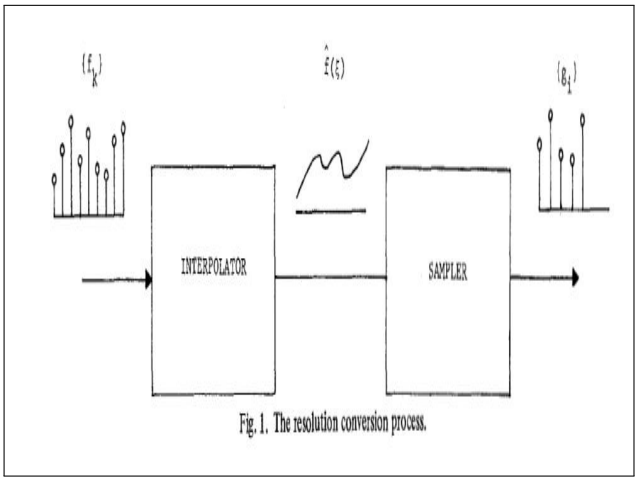


Figure 2.3: The resolution conversion process

2.5 interpolation-based methods

In relation to the many applications of interpolation in signal processing (see [11), the need for sampling rate adaption constantly arises in image processing. Examples of such applications are image resolution conversion and image change of scale. Rigorously speaking, the process of decreasing the data rate is called decimation and increasing the data samples is termed interpolation.

Conceptually, the resolution conversion process can be regarded as a two-step operation. First, the discrete data is reconstructed (or interpolated) into a continuous curve, then it is sampled at a different sampling rate. This is shown in Figure 2.5. Nevertheless, the above steps are only a mental picture for illustrating the underlying principle. In real digital processing, the procedure of reconstruction by interpolation and sampling at a different rate can be done in one operation. (There is never a continuous curve existing inside a digital processor.)

Image interpolation is a process that estimates a set of unknown pixels from a set of known pixels in an image. It has been widely adopted in a variety of applications, such as resolution enhancement, image demosaicing and unwrapping omni-images. The kinds of distortion and levels of degradation imposed on the interpolated image depend on the interpolation algorithm,

as well as the prior knowledge of the original image. Two of the most common types of degradation are the zigzag errors also known as the jaggies, and the blurring effects.As a result, high quality interpolated images are obtained when the pixel values are interpolated according to the edges of the original images.

The paper [13] proposes an edge-directed interpolation algorithm for natural images. The basic idea is to first estimate local covariance coefficients from a low-resolution image and then use these covariance estimates to adapt the interpolation at a higher resolution based on the geometric duality between the low-resolution covariance and the high-resolution covariance. The edge-directed property of covariance-based adaptation attributes to its capability of tuning the interpolation coefficients to match an arbitrarily oriented step edge. A hybrid approach of switching between bilinear interpolation and covariance-based adaptive interpolation is proposed to reduce the overall computational complexity. Two important applications of the new interpolation algorithm are studied: resolution enhancement of grayscale images and reconstruction of color images from CCD samples. Simulation results demonstrate that our new interpolation algorithm substantially improves the subjective quality of the interpolated images over conventional linear interpolation.

They propose a novel non iterative orientation adaptive interpolation scheme for natural-image sources. Our motivation comes from the fundamental property of an ideal step edge (known as geometric regularity [13]), i.e., that the image intensity field evolves more slowly along the edge orientation than across the edge orientation. Geometric regularity has important effects on the visual quality of a natural image such as the sharpness of edges and the freedom from artifacts. Since edges are presumably very important features in natural images, exploiting the geometric regularity of edges becomes paramount in many image processing tasks. In the scenario of image interpolation, an orientation-adaptive interpolation scheme exploits this geometric regularity.

Wing-Shan introduced a modification of the new edge-directed interpolation method that eliminates the prediction error accumulation problem by adopting a modified training win-

dow structure, and extending the covariance matching into multiple directions to suppress the covariance mismatch problem. Simulation results show that the proposed method achieves remarkable subjective performance in preserving the edge smoothness and sharpness among other methods in the literature. It also demonstrates consistent objective performance among a variety of images.

The kinds of distortion and levels of degradation imposed on the interpolated image depend on the interpolation algorithm, as well as the prior knowledge of the original image. Two of the most common types of degradation are the zigzag errors (also known as the jaggies), and the blurring effects.As a result, high quality interpolated images are obtained when the pixel values are interpolated according to the edges of the original images. A number of edge-directed interpolation (EDI) methods have been presented in the literature. Some of them match the local geometrical properties of the image with predefined templates in an attempt to obtain an accurate model and thus estimate the unknown pixel values. However, these algorithms suffer from the inherent problem with the use of edge maps or other image feature maps, where the edges and other image features are difficult if not impossible to be accurately located. The poor edge estimation limits the visual quality of the interpolated images. Other EDI methods make use of the isophote-based methods to direct the edge interpolation to conform the pixel intensity contours.These algorithms are highly efficient in interpolating sharp edges (with significant intensity changes across edges).

However, the interpolation performance is degraded with blurred edges, which are commonly observed in natural images. To cater this problem, edge enhancement or sharpening techniques are proposed.9 However, the use of an edge map is indispensable and noise amplification is aroused with the application of postprocessing techniques. Besides using edge maps, some EDI methods direct the interpolation by further locating the edge orientation with the use of a gradient operator. These methods are effective in eliminating the blurring and staircase problems by detecting the edge orientation adaptively. However, they suffer from the inherent problem of using an edge map, and the gradient operator is not fully adaptive to the image

structure.

2.6 Example based methods

Previous super-resolution approaches rely on aggregating multiple frames that contain complementary spatial information. Generic image priors are usually deployed to regularize the solution properly. The regularization becomes es- pecially crucial when insufficient number of measurements is supplied, as in the extreme case, only one single low-resolution frame is observed. In such cases, generic image priors do not suffice as an effective regularization for SR [2]. A recently emerging methodology for regularizing the ill-posed super resolution reconstruction is to use examples, in order to break the super resolution limit caused by inadequate measurements. Different from previous approaches where the prior is in a parametric form regularizing on the whole image, the example-based methods develop the prior by sampling from other images, similar to [24],[30] in a local way.

One family of example-based approaches is to use the examples directly, with the representative work proposed by Freeman . Such approaches usually work by maintaining two sets of training patches, sampled from the high resolution images, and sampled from the low resolution images correspondingly. Each patch pair (xi; yi) is connected by the obser vation model yi = DHxi + v. This high and low resolution co-occurrence model is then applied to the target image for predicting the high resolution image in a patch-based fashion, with a Markov Random Field (MRF) model as shown in Figure 2.6 The observation model parameters have to be known as a prior, and the training sets are tightly coupled with the image targeted. Patch size should also be chosen properly. If the patch size is very small, the co-occurrence prior is too weak to make the prediction meaningful. On the other hand, if the patch size is too large, one may need a huge training set to and proximity patches for the current observations.

A naive way to do super-resolution with such a coupled training sets is, for each low

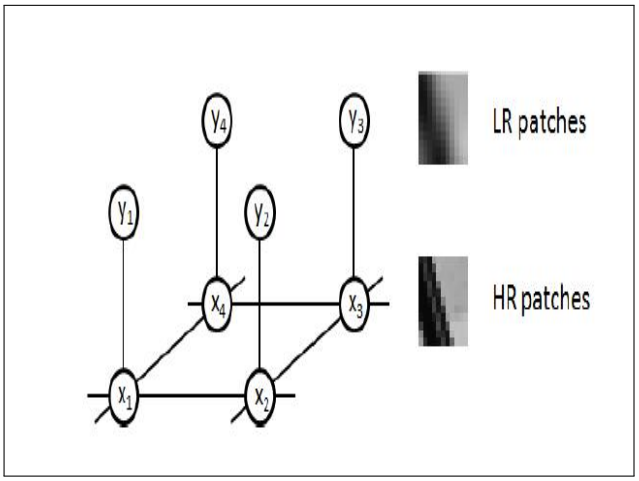


Figure 2.4: The MRF model for single frame super-resolution.

resolution patch in the low-resolution image, and its nearest neigh- bor and then put the corresponding from HR Image. Unfortunately, this simple approach will produce dis- turbing artifacts due to noise and the ill-posed nature of super-resolution [25]. Relaxing the nearest neighbor search to k-nearest neighbors can ensure that the proximity patch we desire will be included. Freeman et al. [31] proposed a belief propagation [108] algorithm based on the above MRF model to se- lect the best high resolution patch found by k-nearest neighbors that has best compatibility with adjacent patches. Sun et al.[35] extended this idea using the sketch prior to enhance only the edges in the image, aiming to speed up the algorithm. The IBP [39] algorithm is then applied as a post processing step to ensure data consistence on the whole image. further followed this line of work and proposed a statistical model that can handle unknown PSF.

One criticism with the aforementioned methods with direct examples is that operating on local patches cannot guarantee global optimality of the es- timation. Another kind of example-based approach seeks to perform MAP estimation with local priors on the image space sampled from examples. The pioneering work by Baker and Kanade [5] formulated an explicit regulariza- tion which demands proximity between the spatial derivatives of the unknown image to those of the found examples. The examples are formed by a pyramid derivative set of features, instead of raw data directly. Similar method is ap- plied to text super-resolution in [35]. Elad [11] presented a global MAP estimation where the example-based regularization is

given by a binary weighted average instead of the nearest neighbor, bypassing outliers due to noise. This work is further extended and elaborated in [21], where the binary weighting scheme is relaxed. Another noteworthy approach for example-based approach is by Protter , generalized from the nonlocal means denois- ing algorithm [8]. Instead of sampling examples from other training images, the algorithm explores self-similarities within the image (or sequence) and ex- tract the example patches from the target image (or sequence) itself. A recent work by Glasner et al. further explored self-similarities in images for SR by combining the classical algorithm based on subpixel displacements and the example-based method based on patch pairs extracted from the target image.

The use of examples can be much more effective when dealing with nar- row families of images, such as text and face images. A group of algorithms have emerged targeting face super-resolution in recent years due to its im- portance in surveillance scenarios. Face super-resolution is usually referred as face hallucination, following the early work by Baker and Kanade [6]. Capel et al. [9] proposed an algorithm where PCA [45] subspace models are used to learn parts of the faces. Liu et al. [8] proposed a two-step approach toward super-resolution of faces, where the arrst step uses the eigenface [10] to generate a medium resolution face, followed by the nonparametric patch- based approach [31] in the second step.

Example-based regularization is effective in our SR problem when insuf- Âŕcient observations are available. There are still a number of questions we need to answer regarding this kind of approaches. First, how to choose the optimal patch size given the target image. Perhaps a multi-resolution treat- ment is needed. Second, how to choose the database. Different images have different statistics, and thereby need different databases. An efficient method for dictionary adaptation to the current target image may suggest a way out. Third, how to use the example-based prior more efficiently. The computation issue could be a difficulty for practical applications. Readers are suggested to refer to [25] for more detailed analysis on example-based regularization for inverse problems.

2.7 Single-image methods

Single-image super-resolution methods cannot hope to improve the resolution by overcoming the Nyquist limit, so any extra detail included in the high-resolution image by such algorithms must come from elsewhere, and prior information about the usual structure of a high-resolution image is consequently an important source.

The single-image super-resolution method proposed by Freeman is considered state-of-the-art. It learns the relationship between low- and high-resolution image patches from training data, and uses a Markov Random Field (MRF) to model these patches in an image, and to propagate information between patches using Belief Propagation, so as to minimize discontinuities at high-resolution patch boundaries.

Tappen et al. also use Belief Propagation to produce impressive single-image super-resolution results, this time exploiting a high-resolution image prior based on natural image statistics to improve the image edge quality in the high-resolution outputs. The example in Figure 2.9 shows results from [19] for both this method and the original Freeman example-based super-resolution method.

Sun et al. [35] perform what they term Image Hallucination on individual lowresolution images to obtain high-resolution output images in which plausible highresolution details have been invented based on a "primal sketch" prior constructed from several unrelated high-resolution training images. The results display plausible levels of high-frequency information in the image edges at a zoom factor of three in each spatial direction, though in some cases appear to suffer from an edge oversharpening phenomenon similar to that described above.

Jiji et al. work with wavelet coefficients as an image representation, rather than using the pixel values in various frequency bands to estimage highfrequency image components as in Freeman et al.'s work. Wavelets allow for more localized frequency analysis than global image filtering or fourier transforms, though regularization is required to keep the outputs visually

smooth and free from wavelet artifacts. Such single-frame wavelet methods are also improved upon by Temizal and Vlachos [11], who use local corellations in wavelet coefficients to improve their performance.

It is important to note that while none of these single-image methods needs to perform registration/motion estimation between multiple inputs, these methods still highlight the great importance of having a good model and of exploiting prior knowledge about working in the image domain.

CHAPTER 3

Preliminaries

First of all, we need to define what we understand by the term 'resolution'. If we take a single image, and multiply its size by four by repeating each pixel value four times, do we gain resolution? On the other hand, let us apply a blurring filter to an image. The resulting image still has the same size, but does it have the same resolution?

The above examples show that there is more to resolution than just counting the number of pixels that are present in the image. It is related to the ability to distinguish details in the image, in other words, to its resolving power. The International Organization for Standardization (ISO) has described a precise method to measure the resolution of a digital camera [5]. The visual resolution can be measured as the highest frequency pattern of black and white lines where the individual black and white lines can still be visually distinguished in the image. It is expressed in line widths per picture height (LW/PH). The standard also describes a method to compute the spatial frequency response (SFR) of a digital camera. The spatial frequency response is the digital imaging equivalent of the modulation transfer function (MTF) used in analog imaging systems. It describes the variation between the maximum and minimum values that is visible as a function of the spatial frequency (the number of black and white lines per millimeter). It can be measured using an image of a slanted black and white edge, and is expressed in relative spatial frequencies (relative to the sampling frequency), line widths per picture height, or cycles per millimeter on the image sensor.

3.1 Image Resolution

The resolution of a digital image can be classified in many different ways. It may refer to spatial, pixel, temporal, spectral or radiometric resolution. In the following work, it is dealt mainly with spatial resolution.

A digital image is made up of small picture elements called pixels. Spatial resolution is given by pixel density in the image and it is measured in pixels per unit area. Therefore, spatial resolution depends on the number of resolvable pixels per unit length. The clarity of the image is directly affected by its spatial resolution. The precise method for measuring the resolution of a digital camera is defined by The International Organization for Standardization (ISO) [4,5]. In this method, the ISO resolution chart is sensed and then the resolution is measured as the highest frequency of black and white lines where it is still possible to distinguish the individual black and white lines. Final value is commonly expressed in lines per inch (lpi) or pixels per inch (ppi) or also in line widths per picture height (LW/PH). The standard also defines how to measure the frequency response of a digital imaging system (SFR) which is the digital equivalent of the modulation transfer function (MTF) used for analog devices.

The effort to attain the very high resolution coincides with technical limitations. Charged coupled device (CCD) or complementary metal-oxide-semiconductor (CMOS) sensors are widely used to capture two-dimensional image signals. Spatial resolution of the image is determined mainly by the number of sensor elements per unit area. Therefore, straightforward solution to increase spatial resolution is to increase the sensor density by reducing the size of each sensor element (pixel size). However, as the pixel size decreases, the amount of light impact on each sensor element also decreases and more shot noise is generated [6]. In the literature [7], the limitation of the pixel size reduction without obtaining the shot noise is presented.

Another way to enhance the spatial resolution could be an enlargement of the chip size. This way seems unsuitable, because it leads to an increase in capacitance and a slower charge

transfer rate [9]. The image details (high frequency content) are also limited by the optics (lens blurs, aberration effects, aperture diffractions etc.). High quality optics and image sensors are very expensive. Super-resolution overcomes these limitations of optics and sensors by developing digital image processing techniques. The hardware cost is traded off with computational cost.

3.2 Super-resolution imaging

If we want to increase the resolution of an image using super-resolution techniques, we essentially want to be able to distinguish more details in the final image. By adding images of the same scene, we try to add information to the reproduction. Typically this information is high frequency content of the scene.

There are different ways to add such high frequency information to an image. If we know that the image is of a certain type (faces, text, drawings, etc.), we can use that knowledge to add frequency content. Such an approach is called a model-based approach. For example, if we know that the images represent printed text, we can try to recognize characters, and replace them by sharp, high quality characters. The knowledge of the image model allows us to compute high frequency information. In this thesis, we will investigate abstract approaches to super-resolution. They use other information than a precise image model, and are therefore applicable to more general types of images. More specifically, we will compute the high frequency information from the aliasing that is present in the images.

Super-resolution techniques use a number of low resolution input images to generate a high resolution image. This assumes that there are some (small) differences between the input images. Most often, these differences are caused by small camera movements. In an ideal situation, we could assume that of four images taken, the second to fourth image have a horizontal, vertical, and diagonal shift of half a pixel compared to the first image. The pixels

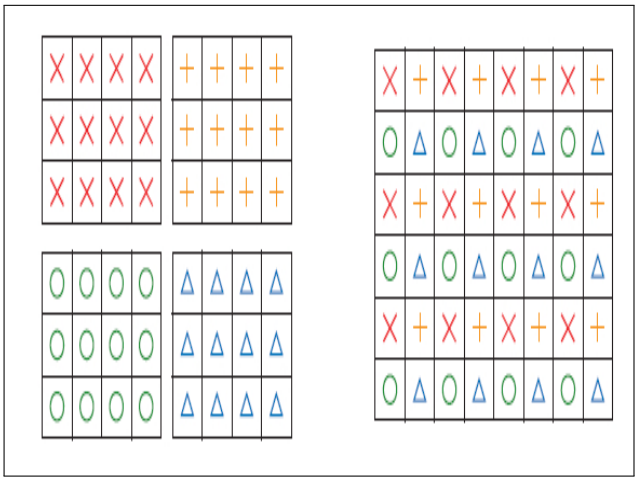


Figure 3.1: Ideal super-resolution setup
Four images are taken with relative shifts of half a pixel in horizontal, vertical, and diagonal directions (left). Their pixels can then be interleaved to generate a double resolution image (right).

from the first image can then be interleaved with pixels from the three other images, and a double resolution image (in both dimensions) is obtained (see Figure 3.2).

In general, however, the shifts between the images are not exactly half a pixel, and can take any arbitrary value. Moreover, in most applications the motion parameters are unknown, and need to be computed first. In the next chapters, we will present methods to compute these motion parameters.

3.3 How is super-resolution possible?

Reconstruction-based super-resolution is possible because each low-resolution image we have contains pixels that represent subtly different functions of the original scene, due to sub-pixel registration differences or blur differences. We can model these differences, and then treat super-resolution as an inverse problem where we need to reverse the blur and decimation.

Each low-resolution pixel can be treated as the integral of the high-resolution image over a particular blur function, assuming the pixel locations in the highresolution frame are known, along with the point-spread function that describes how the blur behaves. Since pixels are

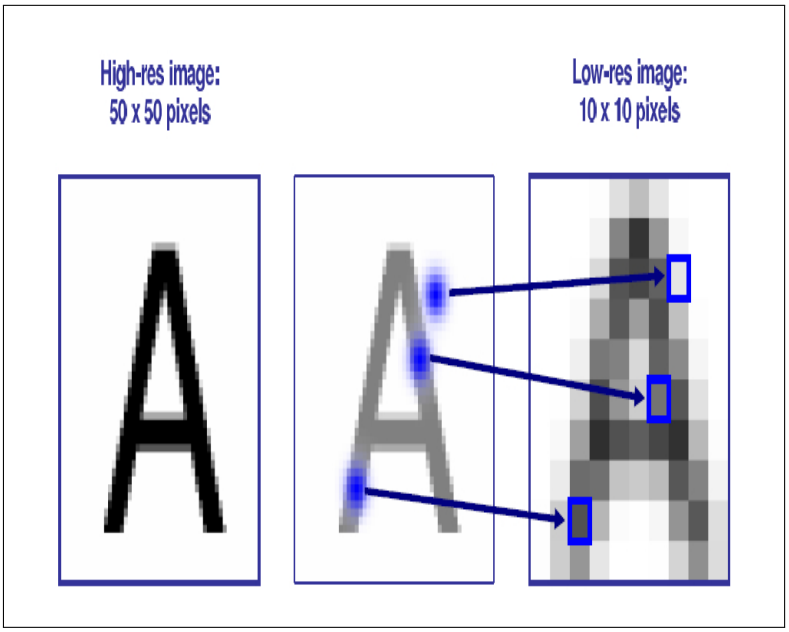


Figure 3.2: The creation of low-resolution image pixels.
The low-resolution image on the right is created from the high-resolution image on the left one pixel at a time. The locations of each low-resolution pixel are mapped with sub-pixel accuracy into the high-resolution image frame to decide where the blur kernel (plotted as a blue Gaussian in the middle view) should be centred for each calculation.

discrete, this integral in the high-resolution frame is modelled as a weighted sum of high-resolution pixel values, with the point- spreaad function (PSF) kernel providing the weights. This image generation process is shown in Figure 3.3.

Each low-resolution pixel provides us with a new constraint on the set of highresolution pixel values. Given a set of low-resolution images with different sub-pixel registrations with respect to the high-resolution frame, or with different blurs, the set of constraints will be non-redundant. Each additional image like this will contribute something more to the estimate of the high-resolution image.

In addition to the model of Figure 3.3, however, real sensors also have associated noise in their measurements, and real images can vary in illumination as well as in their relative registrations. These factors must also be accounted for in a superresolution model, so the full picture of how a scene or high-resolution image generates a low-resolution image set looks more like that of Figure 3.3.

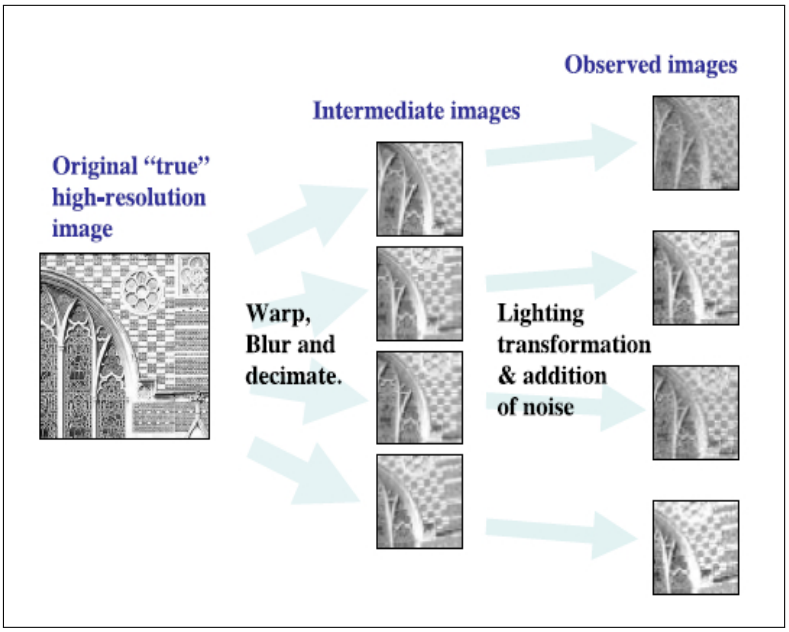


Figure 3.3: One high-resolution image generates a set of low-resolution images. Because the images are related by sub-pixel registrations, each observation gives us more additional constraints on the values of the high-resolution image pixel intensities.

Super-resolution (SR) represents a class of digital image processing techniques that enhance the resolution of an imaging system. Information from a set of low resolution images (LR) is combined to create one or more high resolution images (HR). The high frequency content is increased and the degradations caused by the image acquisition are reduced. The LR images have to be slightly different, so they contain different information about the same scene. More precisely, SR reconstruction is possible only if there are sub pixel shifts between LR images, so that every LR image contains new information [1].

Sub pixel shifts can be obtained by small camera shifts or from consecutive frames of a video where the objects of interest are moving. Multiple cameras in different positions can be used. The basic principle of SR is shown in Figure 3.3. The camera captures a few LR images. Each of them is decimated and aliased observation of the real scene. During SR reconstruction, LR images are aligned with sub pixel accuracy and then their pixels are combined into an HR image grid using various non-uniform interpolation techniques.

By a special type image, model-based approach can be applied. High frequency content

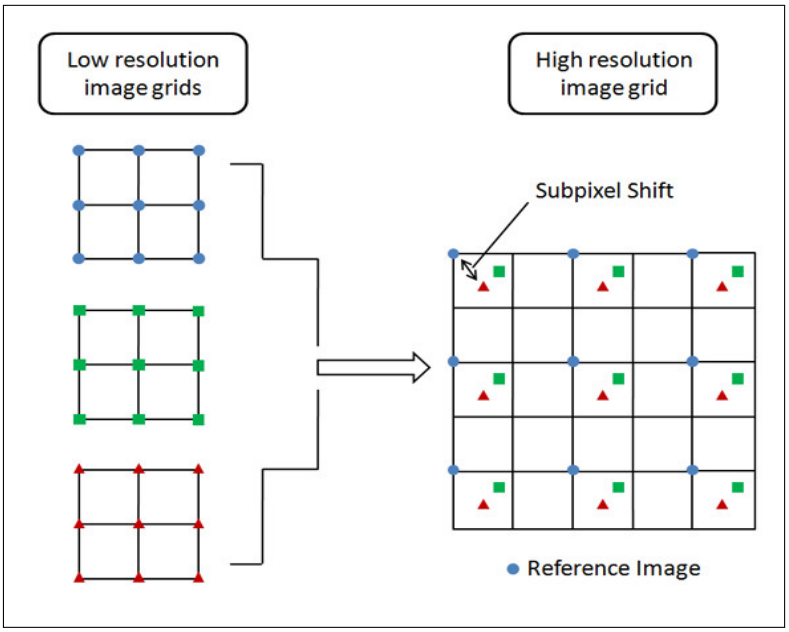


Figure 3.4: Basic principle of super-resolution reconstruction

is calculated on base of the knowledge of the model. For example, if the algorithm recognizes text, the letters can be replaced by sharper ones. In this work, attention is devoted to other methods. These methods use other information than knowledge of an image model and they can be applied to general images. Interpolation techniques based on a single image are sometimes considered as closely related to SR.

These techniques indeed lead to a bigger picture size, but they don't provide any additional information. In contrast to SR, the high frequency content can't be recovered. Therefore, image interpolation methods are not considered as SR techniques [8].

CHAPTER 4

Problem Definition

4.1 Introduction

Given that theories and observations are the two pillars of science, scientific research operates at two levels: a theoretical level and an empirical level. The theoretical level is concerned with developing abstract concepts about a natural or social phenomenon and relationships between those concepts (i.e., build "theories"), while the empirical level is concerned with testing the theoretical concepts and relationships to see how well they reflect our observations of reality, with the goal of ultimately building better theories. Over time, a theory becomes more and more refined (i.e., fits the observed reality better), and the science gains maturity. Scientific research involves continually moving back and forth between theory and observations. Both theory and observations are essential components of scientific research. For instance, relying solely on observations for making inferences and ignoring theory is not considered valid scientific research.

A construct is an abstract concept that is specifically chosen (or "created") to explain a given phenomenon. A construct may be a simple concept, such as a person's weight.

A term frequently associated with, and sometimes used interchangeably with, a construct is a variable. Etymologically speaking, a variable is a quantity that can vary (e.g., from low to high, negative to positive, etc.), in contrast to constants that do not vary (i.e., remain constant). However, in scientific research, a variable is a measurable representation of an abstract construct. As abstract entities, constructs are not directly measurable, and hence, we look for proxy measures called variables.

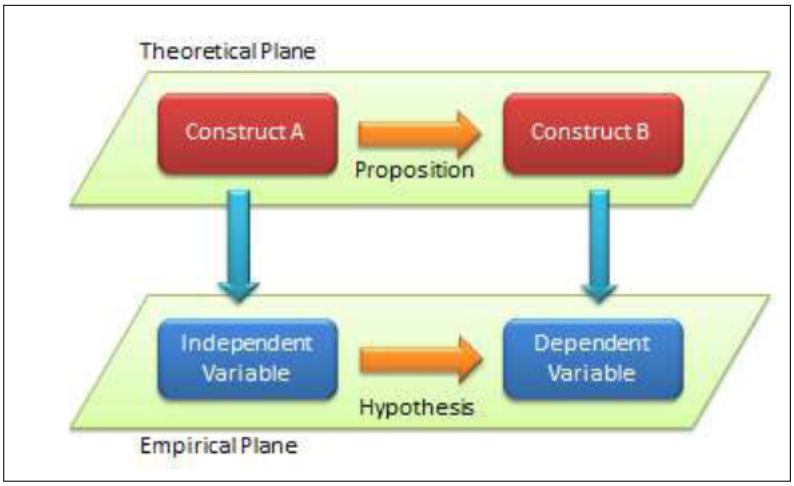


Figure 4.1: The theoretical and empirical planes of research

Depending on their intended use, variables may be classified as independent, dependent, moderating, mediating, or control variables. Variables that explain other variables are called independent variables, those that are explained by other variables are dependent variables, those that are explained by independent variables while also explaining dependent variables are mediating variables (or intermediate variables), and those that influence the relationship between independent and dependent variables are called moderating variables. As an example, if we state that higher intelligence causes improved learning among students, then intelligence is an independent variable and learning is a dependent variable. There may be other extraneous variables that are not pertinent to explaining a given dependent variable, but may have some impact on the dependent variable. These variables must be controlled for in a scientific study, and are therefore called control variables.

In this chapter we are discussing about the research topic in the research methadolgy perspective.

4.2 Research problem

My research problem is to create the High Resolution image from single low resolution image using mathematical model.

We have already discussed about the resolution of images in the previous chapters and studied different methods in the literarure survey. So we are trying to create an efficient approach for SR methods. Since most of the SR research is going on the muti-frame images , we selected single image methods.

Most of the methods using interpolation techniques but their complexity of hardares are too high, so by using some mathematical model we can create SR image effectively.

4.3 Theoretical & Empirical Plane

The main aim of our research problem is to increase the spatial resolution. The basic theory behind the approach is to increase either the value of pixel intencity of create more pixel in the neighbour hood of one pixel. So ultimate aim to predict some values between 2 near values(Pixel).

4.3.1 Constructs

As we told earlier, the ultimate representation of image is in terms of some pixel values or data points. So we have 2 constructs. One is the pixels in the input image and the second one is the predicted pixel values of high reolution. So the proposition between this two constrcts can be considers as prediction of more pixels.

4.3.2 Variables

If we move from the theretical level to higher level, the image of LR is the independend variable and the HR image is the dependend variable.The relation between these variables generating mutilple data points.

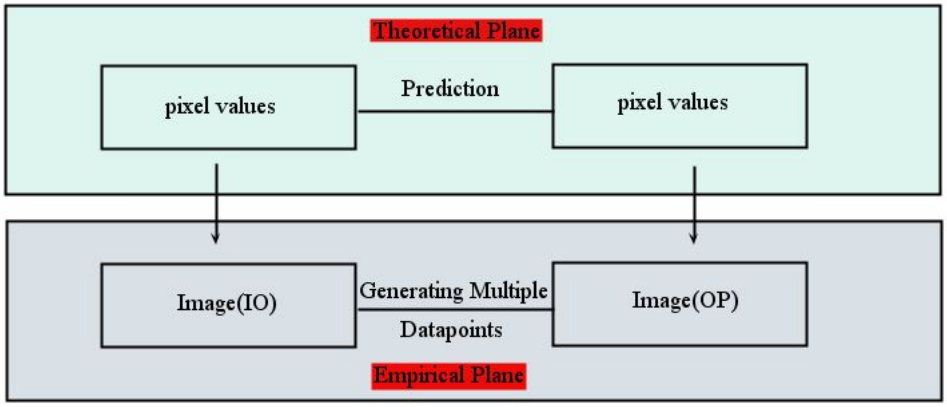


Figure 4.2: Planes Of RM

4.4 Research Design

Here we are telling how we are organizing our entire thesis work. We have already expliained about the research problem as well as the constructs and variables.

The basic theory behind the proposition is the mean value between toe points. Simple mean gives the new value in between the 2 values(Average). We started with the mean, the statical tool. We also improved the different statistical tool.

There are several factor which depend controls the process. The colour components are the main one.

4.4.1 Data Collection

We have collected the data from previous works.

4.5 Hypothesis

If use the mathematical models then we can create the HR image more effectively.

CHAPTER 5

Gaussian Process

A Gaussian process is a generalization of the Gaussian probability distribution. Whereas a probability distribution describes random variables which are scalars or vectors (for multivariate distributions), a stochastic process governs the properties of functions. Leaving mathematical sophistication aside, one can loosely think of a function as a very long vector, each entry in the vector specifying the function value f(x) at a particular input x.

Gaussian Process (GP) defines a distribution over function f, where f is a mapping from the input space χ to κ, such that for any finite subset of χ, its marginal distribution $P(f(x_1), f(x_2), ..., f(x_n))$ is a multivariate normal distribution, where x an input vector. In this section we give a brief review of GPR for implementation purpose; further details can be found in [17].

Gaussian Process is parameterized by its mean function m(x) and covariance function $k(x_i;x_j)$ such that

$$\boxed{\text{f} \mid X \sim N(m(x);K(X;X))} \tag{5.1}$$

or equivalently,

$$\boxed{\text{f(x)} \sim GP(m(x);k(x_i;x_j)) :} \tag{5.2}$$

where rows of the design matrix X are input vectors, f is a vector of function values and K(X;X) denotes the n by n matrix of covariances such that$K_{ij} = k(x_i;x_j)$.

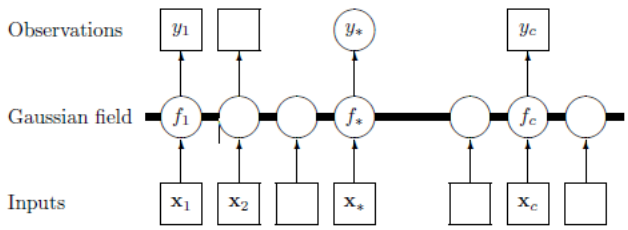


Figure 5.1: Graphical model (chain graph) for a GP for regression.
Squares represent observed variables and circles represent unknowns. The thick horizontal bar represents a set of fully connected nodes. Note that an observation yi is conditionally independent of all other nodes given the corresponding latent variable, fi. Because of the marginalization property of GPs addition of further inputs, x, latent variables, f, and unobserved targets, y*, does not change the distribution of any other variables.

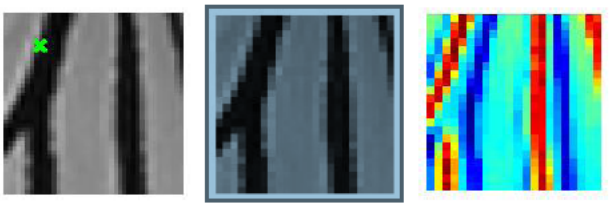
5.1 Covariance Functions

The covariance function plays the central role in GPR as it encodes our assumptions about the underlying process by defining the similarity between functions.

Covariance function is the crucial ingredient in a Gaussian process predictor, as it encodes our assumptions about the function which we wish to learn. From a slightly different viewpoint it is clear that in supervised learning the notion of similarity between data points is crucial; it is a basic similarity assumption that points with inputs x which are close are likely to have similar target values y, and thus training points that are near to a test point should be informative about the prediction at that point. Under the Gaussian process view it is the covariance function that defines nearness or similarity.

Figure 5.1(c) shows the covariance matrix K for one edge pixel x (the crossed point) in Figure 5.1(a) predicted by training points X in the LR training patch, where K_{ij} = k(x;Xij) and i, j correspond to pixel indices in Figure refs1(b). Observe that high responses (red regions) from the training patch are largely concentrated on edges, which justifies our choice of the covariance function. In addition, it is noted that high-responsive regions do not only include points on the same edge as the test point, but also other similar edges within the patch. Thus, the process can

capture both local and global similarity within a patch. In general, pixels embedded in a similar structure to that of the target pixel in terms of the neighborhood tend to have higher weights during prediction.



(a) Test point (b) Training patch (c) Covariance matrix

Figure 5.2: Covariance Function.
(b) is a patch from the LR input image and (a) is its corresponding patch from the upsampled image (bicubic interpolation). The crossed point in (a) is a test point x to be estimated from the traning set X in (b). (c) shows its covariance matrix, where each pixel's intensity value is proportional to the covariance between x and the training target at the same location in (b), evaluated by the covariance function, Equation 10

5.2 BACK-PROJECTION

Back-projection [1] method can minimize the reconstruction error efficiently by an iterative algorithm. In each step, the current reconstruction error is back-projected to adjust the image intensity. Although this method can improve the image quality greatly, it suffers from some unsatisfying artefacts, such as the ringing effect and the chessboard effect. The underlining reason is the usage of isotropic back-projection kernel. In fact, due to the under-determined nature of the SR task, there exist a lot of minimizers for the reconstruction error. It is very likely that the isotropic back-projection kernel leads to unsatisfactory results, since the edge information is totally ignored throughout the update procedure.

n this thesis, the bilateral back-projection method is pro- posed to solve the problems associated with the original one , when it is applied to single image SR. We first show that, for any given positive integer scaling factor, the original back-projection algorithm can minimize the reconstruction error efficiently under certain conditions. Then, the idea of bilateral filtering

is employed to guide the back-projection process.

CHAPTER 6

System Design

This chapter gives the over all idea about the super resolution technique we are implememted in this thesis.

6.1 Proposed System

In our regression-based framework, patches from the HR image are predicted pixelwise by corresponding patches from the LR image . GPR provides a way for soft-clustering of the pixels based on the local structures they are embedded . Given the intuitive interpretation of hyperparameters of the covariance function in our case, we can optimize their values through MAP estimation.

The steps we are following to create the HR image are

1. Upsample the image using bicubic
2. Apply the back Projection Algorithm
3. Calculate Neighburhood
4. Blur and downsanple
5. Mix the details with Upsampled image
6. Deblur the image

6.1.1 Single Image SR

The Figure 5 in the previous chapter shows a chain graph representation of GPR for image SR in our setting, where each 3×3 patch from the input image forms a predictor-target training pair. The observed y is the intensity of the pixel at the center of a 3×3 patch and x is an eight-dimensional vector of its surrounding neighbors. In order to adapt to different regions of the image, it is partitioned into fixed-sized and overlapped patches (e.g., 30×30), and the algorithm is run on each of them separately. The patch-based results are then combined to give the whole HR image.

We predict the HR image using a two-stage coarse-tofine approach, which is consistent with the image formation process. As shown in [2, 18], the imaging process can be modeled as a degradation process of the continuous spatial domain depending on the camera's Point Spread Function (PSF). After discretization, this process can be further expressed as

$$\boxed{\text{L} = (\text{f} * H) \downarrow^{d} \cap H = f * H and L = \cap H \downarrow^{d}} \tag{6.1}$$

where L and H denote the LR and HR image respectively, $\tilde{H}$ denotes the blurred HR image, f is the blur kernel and $\downarrow^{d}$ denotes the downsampling operator with a scaling factor d. As a reasonable inversion of the degradation process, in the first stage we intend to upsample the LR image without loss of details and obtain $\cap H$. In the second stage we use L and its blurred version $\cap L$ (downsampled from $\cap H$) to simulate the blurring process, thus refine the estimates to recover H. The above ideas can be implemented by Algorithm 1 in Figure **??**.

Figure 6.2 gives a real example of the resolving process. Figure 6.2(a), 6.2(b) and 6.2(c) show the first stage, where both the training targets sampled in Figure 6.2(a) and their neighbors in Figure 6.2(b) come from the LR input patch. Neighbors for predicting pixels in the HR image are obtained by upsampling Figure 6.2(a) using bicubic interpolation. Figure 6.2(d), 6.2(e) and 6.2(f) show the second stage, where the training targets 2(d) are the same as those in the first

Algorithm 1 Algorithm for Implementation

SRGPR(L)

$\tilde{H} \leftarrow Upsample(L)$

$\tilde{L} \leftarrow \tilde{H} \downarrow (BlurandDownsample)$

$H \leftarrow Deblur(\tilde{H}, L, \tilde{H})$ return H

Function : Upsample(L)

Bicubic Interpolation :$H_b \leftarrow L \uparrow$

Partition L into n overlapped patches $p_1, p_2, ...p_n$

for $p_L = p_1, p_2, ...p_n$: Sample pixels in p_L to obtain the target vector y Put the eight neighbours of each element of y in X_{NL} as a row vector for training Train a GPR model M using $\{y, X_{NL}\}$ Put the eight neighbours of each pixel of H_b in X_{NH_b} as a row vector for prediction

$p_{\tilde{H}} \leftarrow M(X_{NH_b})$

end

return $\tilde{H}$ constructed from $p_{\tilde{H}}$

Function : $Deblur(\tilde{H}, L, \tilde{H})$

Partition $\tilde{L}$ into n overlapped patches $p_1, p_2, ...p_n$

for $p_{\tilde{L}} = p_1, p_2, ...p_n$: Obtain the same target vector y in p_L Put the eight neighbors in $p_{\tilde{L}}$ of each element of y in $X_{N\tilde{L}}$ as a row vector for training Train a GPR model M using $\{y, X_{N\tilde{L}}\}$

Put the eight neighbours of each pixel of $\tilde{H}$ in $X_{N\tilde{H}}$ X_{NH_b} as a row vector for prediction

$p_H \leftarrow M(X_{N\tilde{L}})$

end

return H constructed from p_H

stage, while the neighbors come from the blurred patch (Figure 6.2(e)). Figure 6.2(f), the deblurred result shows sharper edges than Figure 6.2(c) from stage one.

6.2 System Design

The algorthm is already explained in the previous section. The flow of methods in given in the following diagram.

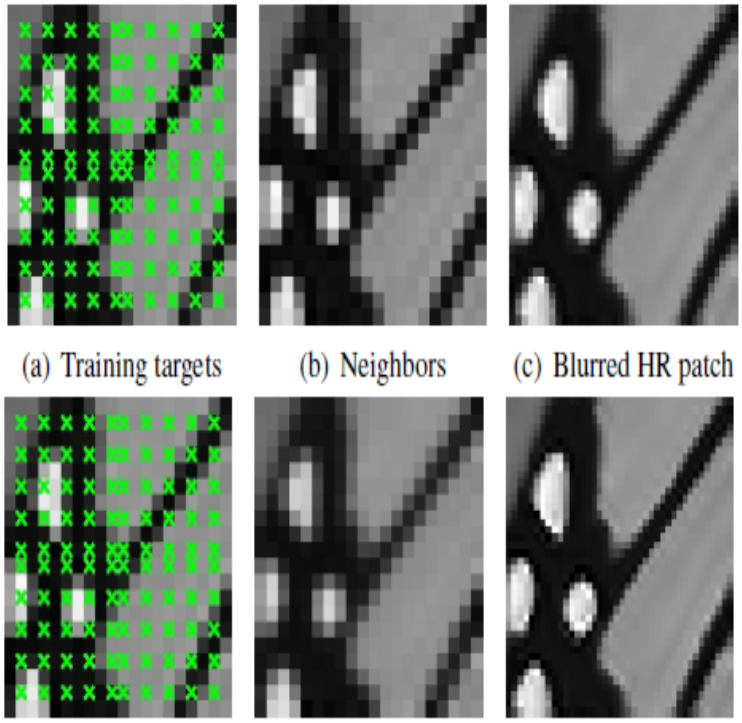


Figure 6.1: Single image SR framework

(a) and (d) are input LR patches, in which the crossed points are target pixels for training produced by simple grid sampling. (b) is the same patch as (a), where neighbors are taken as the intensity values of the crossed points' surrounding pixels. (e) is obtained by blurring and downsampling (c), the output of stage one. Neighbors are taken in the same way as in stage one from (e). (f) shows the deblurred result of the second stage.

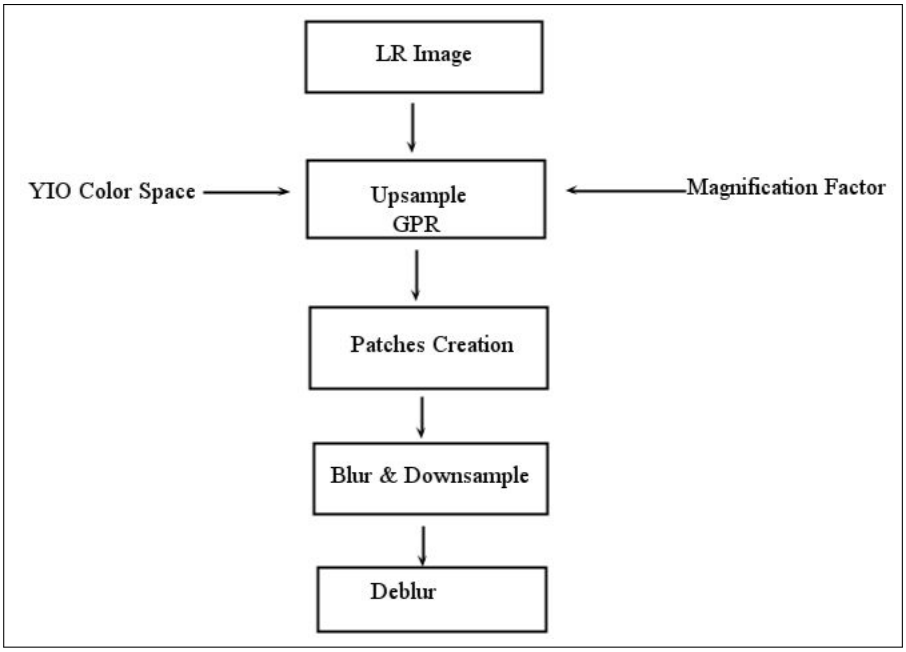


Figure 6.2: Data Flow Diagram of the System

6.3 System Specification

For the simulation of the thesis work we used following tolls

- Matlab : Version 2012a
- C programming Language

CHAPTER 7

Simulation & Testing

We test the method on a variety of images with a magnification factor from 2 to 10. In most of our experiments we began by using the original image as LR input and magnify it with a scale factor of 2. For further magnification, we used the previous output image as the input LR image and solved its HR image.

In the second stage of our method, we made use of the original LR image for deblurring, which also serves as a constraint that when downsampled, the HR image should reproduce the LR image. We set the patch size to 20×20 for the original LR image and increase it according to the magnification factor in later magnifications. For each patch, approximately 2 or 3 area is overlapped with neighboring patches.

After running our algorithm on each patch, we combined them by linear smoothing in the overlapping areas. When processing color images, we worked in the YIQ color space and only applied our SR method to the illuminance (Y) channel since humans are more sensitive to illuminance changes. The other two chrominance channels (I and Q) are simply upsampled using bicubic interpolation. The three channels were then combined to form our final HR image.

7.1 Results

In this section we are including the output images of results we obtained during the simulation.

Figure 7.1: Butterfly Image Input

Figure 7.2: Butterfly Image Output

Figure 7.3: Clock Image Input

Figure 7.4: Clock Image Output

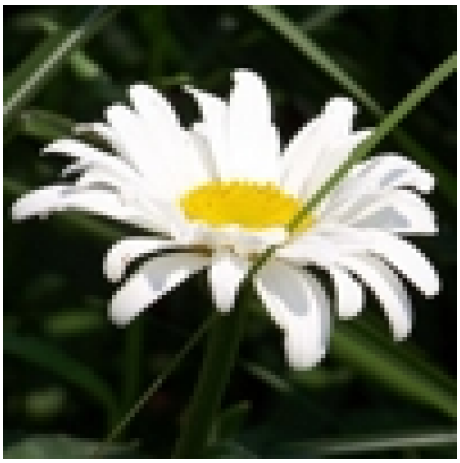


Figure 7.5: Daisy Image Input

Figure 7.6: Daisy Image Output

Figure 7.7: Tower Image Input

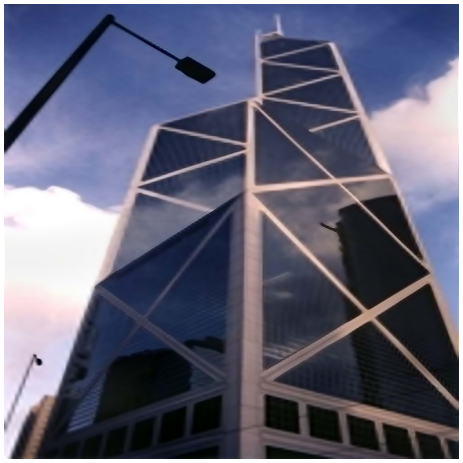


Figure 7.8: Tower Image Output

7.2 Testing

The simplest and most widely used full-reference quality metric is the mean squared error (MSE), computed by averaging the squared intensity differences of distorted and reference image pixels, along with the related quantity of peak signal-to-noise ratio (PSNR). These are appealing because they are simple to calculate, have clear physical meanings, and are mathematically convenient in the context of optimization. But they are not very well matched to perceived visual quality.

There is another tool SSIM(Strcutual Similarity Index) is a method for measuring the similarity between the images. The ssim index is a full reference metric; in other words, the meassuring of image qulaity based on an initial uncompressed or distortion free image as referrnce. SSIM is designed to improve on traditional method like PSNR and MSE.

The testing results of various images are given in the Table 7.1

Image	RMS	MSSIM
Daisy	17.02	0.85
Tower	9.68	0.79
Clock	11.02	0.86
Butterfly	10.01	0.89

Table 7.1: RMS errors and MSSIM scores of different Images

CHAPTER 8

Conclusion & Future Work

In this paper we present a novel method for single image super-resolution. Instead of imposing sophisticated edge priors as in [6, 19], we use a simple but effective covariance function to capture the local structure of each pixel through the Gaussian process regression. In addition, by using a non-parametric Bayesian model, the prediction process is able to adapt to different edges in the local image patch instead of setting a general parameter learnt from a large number of images. The experimental results show that our method can produce sharp edges with minimal artifacts. We propose a two-stage framework, where the first stage aims to recover an HR image based on the structural information extracted from the LR image and the second stage intends to refine the obtained HR image by learning from a training set constructed by the LR image itself. In comparison to the general training set as used in [7, 24], our results capture most of the details presented in the LR image and the visual quality is comparable to some example-based methods. Though presented in the context of super-resolution, our framework of self-directed learning can be extended to other image reconstruction scenarios such as denoising and deblurring by modifying the covariance function and the training pairs. Finally, the performance may be further improved through using more complex covariance function to include extracted features and to express explicitly the local structure of a pixel.

8.1 Future Work

Another diffculty limiting practical application of SR reconstruction is its intensive computation due to large number of unknowns, which require ex- pensive matrix manipulations. Real applications always demand efficiency of the SR reconstruction to be of practical utility,

e.g., in the surveillance video scenarios, it is desired for the SR reconstruction to be real time. Efficiency is also desirable for SR systems with users in the loop for tuning parameters. Many SR algorithms targeting efficiency fall into the previously discussed interpolation-restoration approach.

However, such algorithms require precise image registration, which is com- putation intensive in the first place. Moreover, these algorithm can only handle simple motion models effciently up to now, far from application in real com- plex video scenarios. For videos with arbitrary motions, [12] suggests promis- ing directions for seeking efficient algorithms. It is also interesting to see how parallel computing, e.g., GPU, and hardware implementations affect the fu ture applications of SR techniques.

REFERENCES

[1] A. Rav-Acha A. Zomet and S. Peleg. âĂIJrobust super-res olution,âĂİ. *in Proceedings International Conference on Computer Vision a nd Pattern Recognition (CVPR)*, , Dec. 2001.

[2] A. Almansa. Echantillonnage, interpolation et d ÌĄetec tion. applications. Master's thesis, en imagerie satellitaire.Ecole Norm ale Sup ÌĄerieure de Cachan,, 2002.

[3] H. A. Aly and E. Dubois. Image up-sampling using total- variation regularization with a new observation model. In *IEEE Trans. on IP*, volume 14(10), pages 1647–1659,, 2005.

[4] Vishal Vishwanath Anagire. Reengineering of super-resolution. *IJCST*, 2012.

[5] S. Baker and T. Kanade. âĂIJlimits on super-resolution and h ow to break them,âĂİ. *EEE Transactions on Pattern Analysis and Machine Intelli- gence*, , vol. 24, no. 9, pp. 1167-1183, Sept. 2002.

[6] S. Baker and T. Kanade. Limits on super-resolution and how to break them. *IEEE Trans. on PAMI*, 2002 (1167-1183).

[7] S. Borman and R. Stevenson. âĂIJspatial resolution enhance ment of low- resolution image sequences - a comprehensive review with di rections for future research,âĂİ 8. *University of Notre Dame*, 1998.

[8] H. Shum C. Liu and W. Freeman. Face hallucination: The- ory and practice. *IEE Proceedings. IP*, pages 75(1):115–134,, 2007.

[9] D. Capel. Image mosaicing and super-resolution. *, ser. Distinguished Dis- sertations.*, Springer, 2004.

[10] V. Divi and G. Wornell. âĂIJsignal recovery in time-interl eaved analog- to-digital converters,âĂİ , , pp. 593-596. In *in Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2004.

[11] M. Elad and A. Feuer. âĂIJrestoration of a single superreso lution image from several blurred, noisy, and undersampled measured ima ges,âĂİ. *IEEE Transactions on Image Processing*, , vol. 6, no. 12, pp. 1646-1658, Decem- ber 1997.

[12] R. Fattal. Upsampling via imposed edges statistics. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2007)*, 26(3):95:1–95:8,, 2007.

[13] Charles E. Thorpe Frank Dellaert and Sebastian Thrun. Super-resolved tracking of planar surface patches. Master's thesis, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 1998.

[14] D. Y. Yeung H. Chang and Y. Xiong. Super-resolution through neighbor embedding. *CVPR*, 1:275–282,, 2004.

[15] J. Zerubia H. Shekarforoush, M. Berthod and M. Werman. , âĂIJsub-pixel bayesian estimation of albedo and height,âĂİ. *International Journal of Computer Vision*, 19:289–330, 1996.

[16] T. C. Hofner. Boost your sampling rate with time-inter leaved data con- verters. *Sensors Magazine*, February 2001.

[17] H. S. Hou and H. C. Andrews. Cubic splines for image inter- polation and digital filtering. *IEEE Trans. Acoustics, Speech & Signal Proc. ASSP-26:*, pages 508–517., 2010.

[18] H. S. Hou and H. C. Andrews. Cubic splines for image inter- polation and digital filtering. In *IEEE Trans. Acoustics, Speech & Signal Proc.*, ASSP-26:508-517. 449.

[19] K. W. Hung and W. C. Siu. New motion compensation model via frequency classification for fast video super-resolution. *ICIP*, 2009.

[20] B. R. Hunt. Quantifying the super-resolution capabilities of the clean image processing algorithm. In *n Proceedings of SPIE I, vol. 300. SPIE,*, 2004.

[21] M. Irani and S. Peleg. âĂIJimproving resolution by image re gistration,âĂİ. *CVGIP: Graphical Models and Image Processing*, , vol. 53, no. 3, pp. 231- 239, May 1991.

[22] M. Irani and S. Peleg. Motion analysis for image enhance- ment: Resolution, occlusion and transparency. j. *ournal of Visual Communication and Image Representation*, 1993.

[23] Z. Xu J. Sun and H. Y. Shum. Image super-resolution using gradient profile prior. *CVPR*, pages 449–455, 2008.

[24] S. P. Kim and W. Y. Su. Recursive high-resolution reconstruction of blurred multiframe images. *IEEE Transactions on Image Processing*, 2(4):534–539,, 1993.

[25] M. T. Merino and J. NÃžÃśez. Super-resolution of remotely sensed images with variablepixel linear reconstruction. In *IEEE Transactions on Geoscience and Remote Sensing, vol. 45, no. 5, pp. 1446-1457,*, 2007.

[26] B. S. Morse and D. Schwartzwald. Image magnification us- ing level-set reconstruction. in. *CVPR*, 1:pages 333 –340, 2001.

[27] Sezan M. I. Patti, A. J. and A. M. Tekalp. High resolution standards conversion of low resolution video. In *In International Conference on Acoustics, Speech, and Signal Processing, vol. 4.*, 1995.

[28] J. Jia Q. Shan, Z. Li and C. K. Tang. Fast image/video upsampling. *ACM Transactions on Graphics (SIGGRAPH ASIA)*, pages 449 , 451, 2008.

[29] X. Tang Q. Wang and H. Y. Shum. Patch based blind image super-resolution. *CVPR*, 2013.

[30] Osher S. Rudin, L. I. and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D,*, 60:259–268, 1992.

[31] W. Xu-Y. Wu S. Dai, M. Han and Y. Gong. Soft edge smoothness prior for alpha channel super resolution. *CVPR*, 2007.

[32] N. K. Bose S. P. Kim and H. M. Valenzuela. , âĂIJrecursive re construction of high resolution image from noisy undersampled multifram es,âĂİ. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, , vol. 38, no. 6, pp. 1013-1027, June 1990.

[33] R. Schultz and R. Stevenson. Extraction of high-resolution frames from video sequences. In *EEE Transactions on Image Processing,*, 1996. vol. 5, no. 6, pp. 996-1011,.

[34] ALFONSO SANCHEZ-BEATO SUAREZ. Super - resolution: Multi—frame registration and interpolation using optimal projections on functional spaces. Master's thesis, Departamento de InformÃątica y AutomÃątica Escuela TÃl'cnica Superior de IngenierÃ■a InformÃątica, 2008.

[35] Z. Xu Sun and H. Shum. âĂIJimage super-resolution using gradient profile prior,âĂİ. *in IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, , 2008, pp. 1-8.

[36] B. C. Tom and A. K. Katsaggelos. âĂIJresolution enhanceme nt of monochrome and color video using motion compensation,âĂİ. In *EEE Trans- actions on Image Processing*, volume 10, pages 278–287, 2001.

[37] R. Y. Tsai and T. S. Huang. Multiframe image restoration and registration. *Advances in Computer Vision and Image Processing,*, 1(2):317–339,, 1984.

[38] H. Ur and D. Gross. Improved resolution from subpixel shifted pictures. *VGIP: Graphical Models and Image Processing,*, 54(2):181–186,, 1992.

[39] Patrick VANDEWALLE. super-resolution from unregistered aliased images. Master's thesis, Master in Electrical Engineering, Katholieke Universiteit Leuven, Belgique de nationalitÃl' belge, 2006.

[40] pp. 232-232 vol. 2006, no. 1. Video-to-video dynamic super-resolution for grayscale and color sequences. *EURASIP Journal on Applied Signal Processing,*, 2006.

[41] C. W. Kok W. S. Tam and W. C. Siu. A modified edge di- rected interpolation for images. *Journal of Electronic Imag- ing , 19(1), 013011:1-20,*, page 449, 2010.

[42] T. R. Jones W. T. Freeman and E. C. Pasztor. Example- based super-resolution. i 22(2) : 56-65,. *EEE Comput. Graph. Appl.*, March 2002.