Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 2, Issue. 4, April 2013, pg.317 – 325

SURVEY ARTICLE

A Survey on Techniques and Challenges in Image Super Resolution Reconstruction

Pandya Hardeep[1], Prof. Prashant B. Swadas[2], Prof. Mahasweta Joshi[3]
[1]Research Scholar Student, BVM, V V Nagar, gtu, India
[2]Head of Computer Department, BVM, V V Nagar, gtu, India
[3]Faculty In Computer Department, BVM, V V Nagar, gtu, India

[1] hbpandya12@gmail.com; [2] pbswadas@bvmengineering.ac.in; [3] sweta.ce2013@gmail.com

Abstract— ***In this paper we study and represent various techniques of image super resolution (SR), it is also called as super resolution reconstruction. Super resolution reconstruction produces high resolution image from sequence of low resolution images. The main aim of super resolution is to improve visual quality of available low resolution image. Also existing Low Resolution (LR) imaging can be utilized with help of super resolution reconstruction. This paper focuses on reconstruction based and learning based super resolution methods. We have represented various existing super resolution techniques, advantages and disadvantages of those techniques, recent work and recent methods of super resolution reconstruction method. Finally we have presented challenge issues and future research directions for super resolution.***

Key Terms: - Low Resolution; High Resolution; Super Resolution; Reconstruction based super resolution; learning based super resolution

I. INTRODUCTION

The physical characteristics of a sensor, e.g. its size and density of detectors that form the sensor are major factors which limit the resolution of an image. These restrictions imposed by the sensor can be overwhelmed by a method known as the super-resolution restoration of images. The super resolution method is to take more samples of the scene so as to get some extra information which can be used, while merging the samples to get a high resolution image. These samples can be acquired by sub-pixel shifts, by changing scene illumination or, by changing the amount of blur [1].

Images with high resolution (HR) are desired and often required in most electronic imaging applications. HR means that pixel density within an image is high, and therefore an HR image can offer more details that are important in many applications [2]. For example, for a doctor to make a correct diagnosis, HR medical images are very helpful.

There are various ways to increase resolution of an image. The most direct method [2] to increase spatial resolution is to reduce the pixel size by sensor manufacturing techniques. As the pixel size decreases, however, the amount of light available also decreases. It generates shot noise that degrades the image quality seriously. Another approach for enhancing the spatial resolution is to increase the chip size, which leads to an increase in capacitance [2]. This approach is not considered too much effective because large capacitance makes it difficult to speed up a charge transfer rate. So, a new approach toward increasing spatial resolution is required to overcome these limitations of the sensors and optics manufacturing technology. One promising approach is to use signal processing techniques to obtain an HR image (or sequence) from observed multiple low resolution (LR) images [2]. Recently, such a resolution enhancement approach has been one of the most active research areas, and it is called super resolution (SR) (or HR) image reconstruction or simply resolution enhancement.

The major advantage of the super resolution approach is that it may cost less and the existing LR imaging systems can be still utilized. The SR image reconstruction is proved to be useful in many practical cases [2] where multiple frames of the same scene can be obtained, including medical imaging, satellite imaging, and video applications. Synthetic zooming of region of interest (ROI) is another important application in surveillance, forensic, scientific, medical, and satellite imaging [2]. This application is most suitable for magnifying objects in the scene such as the face of a criminal or the licence plate of a car. The SR technique is also useful in medical imaging such as computed tomography (CT) and magnetic resonance imaging [1], [2] (MRI). In satellite imaging applications such as remote sensing and LANDSAT, several images of the same area are usually provided, and the SR technique to improve the resolution of target can be considered. Another application is conversion from an NTSC video signal to an HDTV signal since there is a clear and present need to display a SDTV signal on the HDTV without visual artifacts.

The basic problem is to obtain an HR image from multiple LR images. The basic assumption for increasing the spatial resolution in SR techniques is the availability of multiple LR images captured from the same scene. In SR, the LR images represent different “looks” at the same scene [1], [2], [5]. In that LR images are sub-sampled as well as shifted with sub-pixel precision. If the LR images are shifted by integer units, then each image contains the same information, and thus there is no new information that can be used to reconstruct an HR image. If the LR images have different sub-pixel shifts from each other and if aliasing is present, and then each image cannot be obtained from the others [2], [3]. In this case, the new information contained in each LR image can be exploited to obtain an HR image.

To obtain different looks at the same scene, some relative scene motions must exist from frame to frame via multiple scenes or video sequences. Multiple scenes can be obtained from one camera with several captures or from multiple cameras situated in different positions [3]. These scene motions can occur due to the controlled motions in imaging systems, e.g., images acquired from orbiting satellites. The same is true of uncontrolled motions, e.g., movement of local objects or vibrating imaging systems. SR method assumes that there are small differences between the input images [2], [3]. These differences are caused by small camera movements. In an ideal situation, one can 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 from the first image can then be interleaved with pixels from the three other images, and a double resolution image is obtained [3], [5]. This setup is shown in figure 1.1.

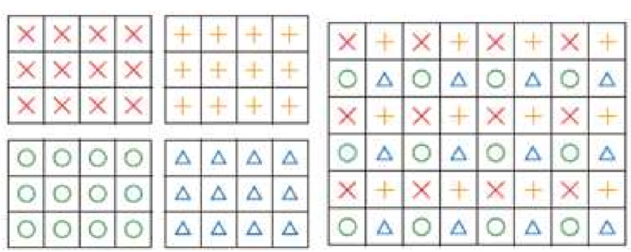


Figure-1.1:-Ideal super-resolution setup. Four images are taken with relative shifts of half a pixel in horizontal, vertical, and diagonal directions (left). Their pixels can then be interleaved to generate a double resolution image (right) [3].

Most of the super-resolution image reconstruction methods consist of three basic components [1], [2]: (i) motion estimation (ii) interpolation and (iii) blur and noise removal. Motion estimation is used to map the motion from all available low resolution frames to a common reference frame. The motion field can be modeled in terms of motion vectors or as affine transformations. The second component refers to mapping the motion-corrected pixels onto a super-resolution grid. The third component is needed to remove the sensor and optical blurring.

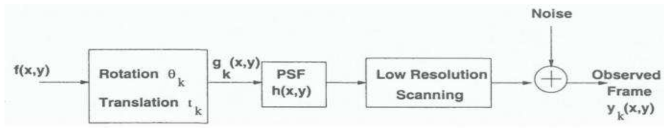


Figure 1.2: Observation model relating a high resolution image to the low resolution observed frames [1].

The observation model relating a high resolution image to the low resolution observed frames is shown in Figure 1.2. The input signal f (x,y) denotes the continuous (high resolution) image in the focal plane co-ordinate system (x,y) . Motion is modelled as a pure rotation θ_k and a translation t_k. The shifts are defined in term of low-resolution pixel spacings. Because of sampling grid changes in the geometric transformation, this step requires interpolation. Next the effects of the low resolution sensor (i.e., blur due to integration over the surface area) and the optical blur (i.e., out-of-focus blur) are modelled as the convolution of $g_k(x,y)$ with the blurring kernel h(x,y). Finally, the transformed image undergoes low-resolution k^{th} scanning followed by addition of noise yielding the low resolution observation $y_k(x,y)$ [1].

II. SR IMAGE RECONSTRUCTION ALGORITHMS

At present there are various super resolution reconstruction algorithms are available we will discuss some of them here.

A. *Non-uniform Interpolation*

The basis of non-uniform interpolation super-resolution techniques is the non-uniform sampling theory which allows for the reconstruction of functions from samples taken at non-uniformly distributed locations [2]. Early super-resolution applications used detailed camera placement to allow for accurate interpolation, because this method requires very accurate registration between images.

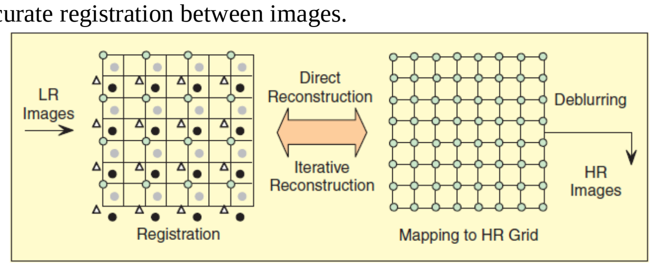


Figure 2.1: Registration-interpolation-based reconstruction [2].

A new method was developed to overcome the limitations of insufficient registration accuracy by applying multiple digital sensors with different pixel sizes [2], [4], [5]. This ensures that pixels of multiple images will not coincide regardless of camera placement. Non-uniform interpolation is a basic and intuitive method of super-resolution and has relatively low computational complexity, but it assumes that the blur and noise characteristics are identical across all low-resolution images as shown in figure 2.1 [2]. Figure 2.2 [2] shows results obtained by various interpolation methods for super resolution of image.

The advantage of this approach is that it takes relatively low computational load and makes real-time applications possible [2], [4]. However, in this approach, degradation models are limited they are only applicable when the blur and the noise characteristics are the same for all LR images [2], [5].

(a) (b)

(c) (d)

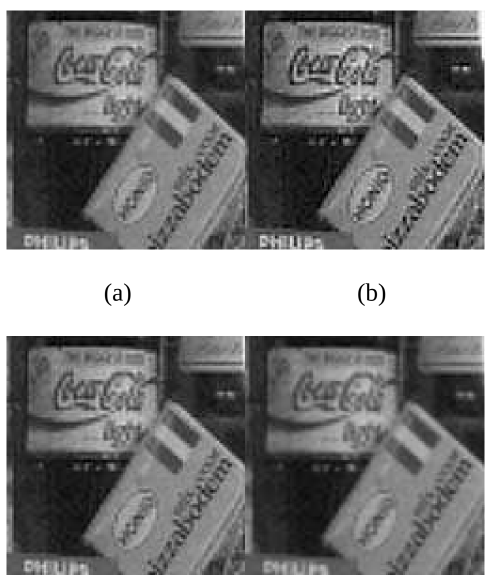
Figure 2.2: Nonuniform interpolation SR reconstruction results by (a)nearest neighbor interpolation, (b) bilinear interpolation, (c) nonuniform interpolation using four LR images, and (d) debluring part (c) [2].

B. Projection onto Convex Sets

This method is based on a linear model describing the relation of HR and LR images, a cost function is introduced and the HR image is obtained [2]. POCS algorithm has many advantages like simplicity; it can be applied to the occasion with any smooth movement, and can easily join in the prior information, so this method is widely used. But POCS algorithm is strict to the accuracy of movement estimation [2]. So in order to improve the stability and performance of the algorithm, the relaxation operator will be used to replace ordinary projector operator, at the same time it is not contributing to the resumption of the edge and details of images.

This method is based on a linear model describing the relation of HR and LR images, a cost function is introduced and the HR image is obtained. POCS algorithm has many advantages like simplicity,it can be applied to the occasion with any smooth movement, and can easily join in the prior information,so this method is widely used [2]. But POCS algorithm is strict to the accuracy of movement estimation. So in order to improve the stability and performance of the algorithm, the relaxation operator will be used to replace ordinary projector operator, at the same time it is not contributing to the resumption of the edge and details of images [2].

However, the linear model used in this method is an ill-posed problem in the sense that its transformation matrix may be singular and so a unique solution cannot be obtained. The advantage of POCS is that it is simple, and it utilizes the powerful spatial domain observation model [2]. It also allows a convenient inclusion of a priori information. These methods have the disadvantages like non-uniqueness of solution, slow convergence, and a high computational cost. Figure 2.3 shows various results of this method.



(a) (b)

(c) (d)

Figure 2.3: POCS SR results (a) by bilinear interpolation and by POCS after (b) 10 iterations, (c) 30 iterations, and (d) 50 iterations [2].

C. Frequency Domain Method

The frequency domain approach makes explicit use of the aliasing that exists in each LR image to reconstruct an HR image [3]. Tsai and Huang first derived a system equation that describes the relationship between LR images and a desired HR image by using the relative motion between LR images.

The frequency domain approach is based on the following three principles [2], [3]: (1) the shifting property of the Fourier transform, (2) the aliasing relationship between the continuous Fourier transform (CFT) of an original HR image and the discrete Fourier transform (DFT) of observed LR images, (3) and the assumption that an original HR image is band limited.

These properties make it possible to design the system equation relating the aliased DFT coefficients of the observed LR images to a sample of the CFT of an unknown input image [3]. For example, let us assume that there are two 1-D LR signals that are sampled below the Nyquist sampling rate. From the above stated principles, the aliased LR signals can be decomposed into the unaliased HR signal as shown in Figure 2.4.

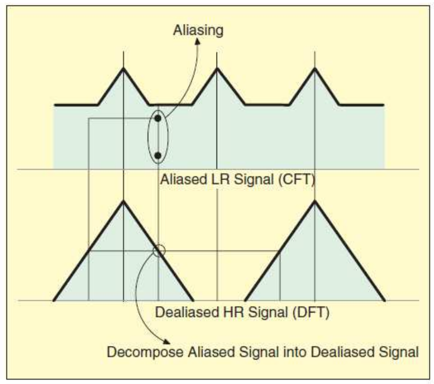


Figure 2.4: Aliasing relationship between LR image and HR image [2].

D. Sparse Representation Method

This method is based on single-image super resolution, which is based on sparse signal representation. Researchers in imaging field suggest that image patches can be well represented as a sparse linear combination of elements from an appropriately chosen over-complete dictionary [2], [7]. Learning an over-complete dictionary capable of optimally representing broad classes of image patches is a difficult problem. It is difficult to learn such a dictionary or using a generic set of basis vectors (e.g., Fourier), so for simplicity one can generate dictionaries by simply randomly sampling raw patches from training images of similar statistical nature [2]. Researchers suggest that simple prepared dictionaries are already capable of generating high-quality reconstructions, when used together with the sparse representation prior. Figure 2.5shows several training images and the patches sampled from them.

Figure 2.5: Left: three training images which are used in experiments. Right: the training patches extracted from them.

By jointly training two dictionaries for the low- and high-resolution image patches, one can enforce the similarity of sparse representations between the low-resolution and high resolution image patch pair with respect to particular dictionaries [7], [8]. So, the sparse representation of a low-resolution image patch can be applied with the high-resolution image patch dictionary to generate a high-resolution image patch [7], [8]. The learned dictionary pair is a more compact representation of the patch pairs. The effectiveness of such approach is demonstrated for both general image super-resolution (SR) and the special case of face hallucination [7]. Figure 2.6 shows some results obtained by this method

Figure 2.6: The flower and girl image magnified by a factor of 3. Left to right: input, bicubic interpolation, neighbor embedding, sparse representation, and the original image [7].

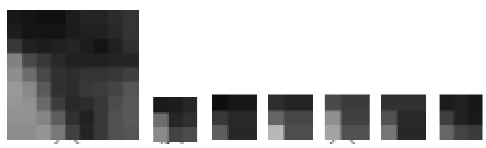
E. Super Resolution through Neighbor Embedding

This method is used for solving single-image super-resolution problems [8]. Given a low resolution image as input, objective is to recover its high-resolution counterpart using a set of training examples [8], [9], [10].

In a recent neighbor embedding method based on Semi-nonnegative Matrix Factorization (SNMF) only non-negative weighs are considered [10]. In LLE the weights are constrained to sum up to one, but no constraints are specified for their sign [10], [11]. This might explain the unstable results observed in [9], since possible negative weights can lead to having subtractive combinations of patches, which is counterintuitive.

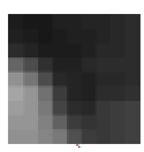
This method is based on assumption that small patches in the low- and high-resolution images form manifolds with similar local geometry in two distinct spaces. In this method each low- or high-resolution image is represented as a set of small overlapping image patches [8], [9]. Each patch is represented by a feature vector.

The feature may be contrast, correlation, entropy, variance, sum of average, sum of variance, homogeneity, variance of difference, sum of entropy, difference of entropy, change of luminance[8], [9] etc. Figure 2.7 shows one example of such a patch generation.



(a) (b) (c)

(d)

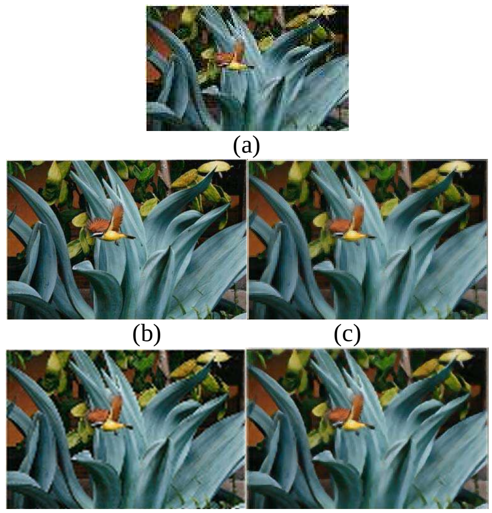


(e)

Figure 2.7:Neighbor embedding procedure applied to a low-resolution patch for 3X magnification: (a) true high-resolution patch; (b) input low-resolution patch down sampled from (a); (c) five nearest neighbor low-resolution patches from the training images; (d) high-resolution patches from the training images corresponding to the low-resolution patches in (c); (e) target high-resolution patch constructed from (d). [9]

This method is also called as learning based method for super resolution; this method has been inspired by recent manifold learning methods, particularly locally linear embedding (LLE) [8], [9]. In that method small image patches in the low and high resolution images form manifolds with similar local geometry in two distinct feature spaces. As in LLE, local geometry is characterized by how a feature vector corresponding to a patch can be reconstructed by its neighbours in the feature space [9].

Also using the training image pairs to estimate the high-resolution embedding, some researchers have enforced local compatibility and smoothness constraints between patches in the target high-resolution image through overlapping. Experiments show that this method is very flexible and gives good empirical results [8], [9], [10]**.**Figure 2.8 shows some results of this method with available high resolution training image and input low resolution image.



(a)

(b) (c)

(d) (e)

Figure 2.8: 3X magnification of the bird image: (a) input low-resolution image; (b) true high-resolution image; (c) median filtering; (d) cubic spline interpolation; (e) Neighbor Embedding method [9].

III. COMPARISON OF VARIOUS SUPER RESOLUTION TECHNIQUES

Comparisons of super-resolution techniques have been mainly concerned with what assumptions are made in the modelling of the super-resolution problem. The blurring process to be known or those regions of interest among multiple frames are related through global parametric transformations, these are the assumptions one has to make. Other models take into account arbitrary sampling lattices, physical dimension of sensor, a non-zero aperture time, focus blurring, and more advanced additive noise models. To simplify a model many times these assumptions are chosen and are usually used in a specific method [2], [12].

In addition, methods that do not make these assumptions have not demonstrated objectively that removing these assumptions gains better super-resolution reconstruction performance. Signal-to-noise ratio (SNR), Peak signal-to noise ratio (PSNR), Root mean square error (RMSE), mean absolute error (MAE), and mean square error (MSE) have all been used as objective measures of super-resolution accuracy; however, the outstanding method of presenting results is clearly subjective to visual quality [12], [13].

IV. CHALLENGE ISSUES FOR SUPER RESOLUTION

Image registration

Image registration is critical for the success of multi-frame SR reconstruction, where spatial samplings of the HR image are fused. The image registration is a basic image processing problem that is well known as ill-posed. The problem is more difficult in the SR setting, where the observations are low-resolution images with heavy aliasing artifacts [12]. The performance of the standard image registration algorithms decreases as the resolution of the observations goes down, resulting in more registration errors. Degradations caused by these registration errors are visually more annoying than the blurring effect resulting from interpolation of a single image [12].

Computation Efficiency

Another difficulty limiting practical application of SR reconstruction is its intensive computation due to large number of unknown samples, which require expensive matrix manipulations [12]. Real applications always demand efficiency of the SR reconstruction to be of practical utility.

Robustness Prospects

Traditional SR techniques are vulnerable to the presence of deviation due to motion errors, inaccurate blur models, noise, moving objects, motion blur, moving scene etc. Robustness of SR is of interest because the image degradation model parameters cannot be estimated perfectly, and sensitivity to deviations may result in visual degradations, which are unacceptable in many applications, e.g., video standard conversion [12], [13].

V. FUTURE RESEARCH DIRECTIONS

Degradation Models

Accurate degradation/observation models promise improved SR reconstructions. Several SR application areas may benefit from improved degradation models. For improved reconstruction of compressed video, degradation models for lossy compression schemes are most promising one to use [12].

Motion Estimation

SR enhancement of random scenes containing global, multiple independent and individual motion, occlusions, transparency etc. is a main focus of SR research. Obtaining this is critically dependent on robust, model based, sub-pixel accuracy motion estimation and segmentation techniques is a crucial research problem [13]. Motion is typically estimated from the observed under-sampled data.

Restoration Algorithms

MAP and POCS based algorithms are very successful. Hybrid MAP/POCS restoration techniques will combine the mathematical stiffness and uniqueness of solution of MAP estimation with the convenient a priori constraints of POCS [2], [13]. Simultaneous motion estimation and restoration gains improved reconstructions since motion estimation and reconstruction are correlated. Separate motion estimation and restoration, as is commonly done, is sub-optimal as a result of this interdependence. Simultaneous multi-frame SR restoration is expected to achieve higher performance since additional spatio-temporal constraints on the SR image ensemble may be included. In SR reconstruction this technique has limited application.

VI. CONCLUSION

The researches on super-resolution image reconstruction mainly consider the situation that degraded model is linear. Noise is not much concerned and systemic analysing method and filter designing method have not been formed yet [12], [13]. As different methods of super-resolution have been developed using models with unequal assumptions of the existing problem, and because the results provided have been primarily based on subjective measurements, it is difficult to find an unbiased comparison on what super-resolution methods are more appropriate for a given task. There must be considerations like if more than one input images are present then use multi frame super resolution approach and if one or more high resolution training images are available then use single image super resolution approach. If we do not want registration step then we can use single frame image resolution. Also if high resolution training is not available but different low resolution images are available for same scene than one must have to use multi frame super resolution. This does not provide a clear method of comparing different implementations suitability for a desired application, so one have to implement super resolution method based on problem model which can be generalized to all SR reconstruction problems.

REFERENCES

[1] SubhasisChaudhuri (Indian Institute of Technology), " Super Resolution Imaging" Kluwer Academic Publishers, pp.1-44, 2002.

[2] Sung Cheol Park, Min KyuPark,and Moon Gi Kang, "Super-Resolution Image Reconstruction: A

Technical Overview". IEEE Signal Processing Magazine May 2003.
[3] Patrick Vandewalle, Sabine Susstrunk and Martin Vetterli," A Frequency Domain approach to registration of aliased images with application to super-resolution" EURASIP journal on applied signal processing 2006.
[4] Gilman, A., Bailey, D.G., Marsland, S.R," Interpolation models for image super resolution", 4th IEEE International Symposium on Electronic Design, Test and Applications, DELTA 2008.
[5] NarasimhaKaulgud and Uday B. Desai, "Image Zooming: Use of Wavelets", The International Series in Engineering and Computer Science, Springer US Publishers,, ©Kluwer Academic Publishers.
[6] R. Y. Tsai and T. S. Huang. Multiframe image restoration and registration. In R. Y. Tsai and T. S. Huang, editors, Advances in Computer Vision and Image Processing, volume 1, pages 317–339. JAI Press Inc., 1984.
[7] Jianchao Yang, John Wright, Thomas S. Huang, Yi Ma, "Image Super-Resolution as Sparse Representation of Raw Image Patches", IEEE conference on computer vision and pattern recognition, 2008.
[8] William T. Freeman, Thouis R. Jones, and Egon C. Pasztor, "Example-Based Super- Resolution" Image-Based Modeling, Rendering, and Lighting,IEEE March/April 2002.
[9] Mei GONG, Kun HE, Jiliu ZHOU, Jian ZHANG, "Single Color Image Super-resolution through Neighbor Embedding", Journal of computational information systems 7:1 (2011) 49-56.
[10] Marco Bevilacqua, AlineRoumy, Christine Guillemot, Marie-Line AlberiMorely, "Neighbor embedding based single-image super-resolution using semi-nonnegative matrix factorization", IEEE international conference on Acoustics, speech and signal processing, 2012.
[11] Hong Chang, Dit-Yan Yeung, YiminXiong, "Super-Resolution Through Neighbor Embedding", IEEE conference on computer vision and pattern recognition, 2004.
[12] Vishal R. Jaiswal, Girish P. Potdar, Tushar A. Rane,"Recent Developments in Super Resolution", International Journal of Computer Science & Engineering Technology (IJCSET), Vol. 2 No. 4, 2011.
[13] KathiravanSrinivasan, J. Kanakaraj, "A Study on Super-Resolution Image Reconstruction Techniques", Computer Engineering and Intelligent Systems, iiste, Vol 2, No.4, 2011

