SURVEY ON DIFFERENT TECHNIQUES IN IMAGE INPAINTING

Gurhans Pal
M.Tech Student
Department of Electronics & Communication Engineering,
E-Max Group of Institutions, Kalpi - Naraingarh Road, Gola, Tehsil - Mullana, Dist. Ambala, Haryana, INDIA
E-mail: Guruhanspalsaroha@gmail.com

Er. Vikas Sharma
Asstt. Professor
Department of Electronics & Communication Engineering,
E-Max Group of Institutions, Kalpi - Naraingarh Road, Gola, Tehsil - Mullana, Dist. Ambala, Haryana, INDIA
E-mail:Vikas741@gmail.com

Abstract: In image processing and computer vision we try to do the restoration with digital images, this is called image completion, image disocclusion or (digital) image inpainting. The goal of image inpainting is filling the target region in the image with visual plausible information. Filling this region with information that could have been in the image. There are several applications for image inpainting, one of these is the restoration of old images and movies by removing cracks from these images and movies. The removal of objects from images, like time stamps or a person. Another application can be image inpainting as the pre-process of other computer vision or image processing applications, an example of this is the use of image inpainting in image-based material editing. It can also be used to fill missing parts in image communication, for example image compression, lost packets retrieval and zooming. In this paper we are surveying with comparative look on different techniques on image inpainting.

Keywords: Image inpainting

1.0 Introduction: The remarkable demand for universal distribution and consumption of image and video content over various networks has pressured the digital consumer electronics industry to launch an almost infinite variety of electronic devices capable of acquiring, processing, editing and storing these very attractive types of content. The combination of this trend with the increasing demand for higher qualities and larger screen resolutions has posed another challenging problem to the image coding research community: to significantly increase the current image compression efficiency for the various, relevant target qualities. In this context, better exploiting the Human Visual System behaviours and characteristics is largely recognized as an appealing way to target further improvements in terms of image compression efficiency. With this purpose in mind, inpainting-based image coding solutions have recently emerged as a novel coding paradigm to further exploit the image visual redundancy, thus allowing increasing the compression efficiency while assuring the required perceived image quality.

Image Inpainting: This field consists of three main topics: the restoration of images by texture synthesis, restoration of pictures by propagating linear structures and the restoration of films. This master thesis primarily focuses on the second technique: the restoration of images by propagating linear structures. The main challenges in image inpainting are the continuation of structure, finding the correct information and the speed of inpainting. An example of the continuation challenge is an image with a fence; when a user wants to inpaint an unknown part of a fence, the fence should continue.

Figure 1: Example of image inpainting from [1], the original image is shown on the left and the inpainted image on the right.
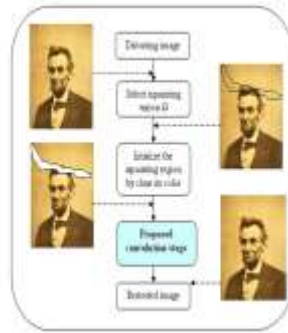
Figure 2: block diagram of image inpainting [1]

The introductory block diagram of image inpainting is shown in figure 2. In inpainting algorithms we have to find the perfect information to fill in a gap. That information can come from multiple places, adjacent pixels, other parts of the image or other images, it can be challenging to find the best match for each situation. Current research of image inpainting techniques consists of some different approaches. Some approaches are based on pixel by pixel updating by computing partial differential equations. Others are based on copying parts of the input image to the unknown region or even information from a dataset of images.

2.0 Major Challenges in Image Inpainting

There are three major challenges found in 2D image inpainting.

2.1 Domain Complexity:

For 2D images, the missing inpainting domains can be very arbitrary depending on particular applications at hand. For text removal, inpainting domains consist of various fonts of 26 characters (for English), 10 numerics, and punctuation signs. For disocclusion in computer vision, inpainting domains are determined by objects on the foregrounds, which in real life can have any shapes and sizes. For fine art retouching in museums, missing domains often correspond to the cracks on ancient paintings, which could occur quite randomly due to bad weather conditions or natural aging of pigments. For super resolution or zooming, on the other hand, missing information is scattered across the entire image domains, and the inpainting regions look more like porous media, i.e., scattered blank holes surrounded by available image features.

The available image information is therefore often given on complicated sets instead of finite discrete ones. These complex sets could contain 2D sub domains, 1D structures such as regular curve segments, and automatic or isolated points. An ideal inpainting scheme should be able to simultaneously benefit from all these different types of information, in order to interpolate or reconstruct the original images as faithfully as possible.

2.2 Image complexity: The complexity of images as functions causes further challenges for image inpainting. At coarse scales, as in Chan & Shen (2002), images can be well approximated by Sobolev smooth functions in multiresolution wavelet analysis. But in small scales, images carry fully localized details spatially organized into coherent structures. For human and computer vision, it is often these singular features of images, e.g., edge jumps, corners, T-junctions, and features of fractal dimensions that convey crucial visual information. An ideal inpainting scheme has to respect these geometric or inhomogeneous features. As mentioned in literature Chan & Shen (2005), the main challenge from modeling and computational point of view is that there seems to be no single Banach or Hilbert function space that can conveniently manage all these fine properties.

2.3 Pattern complexity: On the other hand, a realistic inpainting scheme cannot be merely built upon the fine properties of images as functions. It must also respect visually meaningful patterns (e.g., mirror symmetry).

Imagine, for example, a human face image with the entire area of the left eye and the eyebrow missing. Without any prior knowledge, it depicts a human face, a general inpainting scheme reasonably does the following: first, inspect the available skin tone information in the surrounding vicinity, and then on the missing area fill in the average tone with minor variation to smoothly match the available information. However, this may result in a horrible face image.

In this imaginary scenario, even a layman can come up with a much available or its mirror image with respect to the nose-mouth central axis. However, such trivial human inpainting activity already involves some crucial ingredients of intelligence, i.e., (a) the onsite recognition that the image (despite its incompleteness) depicts a human face and (b) the off-site prior knowledge that a typical human face has approximate mirror symmetry. Thus, the inpainting problem inherits same challenges such as pattern recognition and artificial intelligence. As mentioned in Chan & Shen (2002), Figure 1.4 and Figure 1.5 display two examples that highlight such challenges.
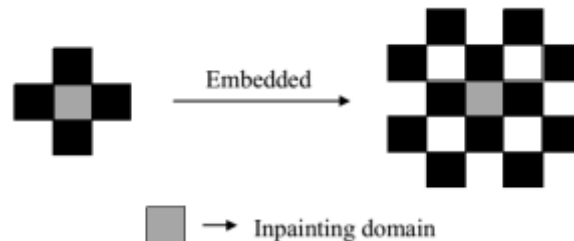
Figure 3 The effect of local and global pattern recognition

on image inpainting. The black color seems to be a reasonable inpainting solution in the left panel, whereas for the right chessboard pattern, the white color becomes more plausible [14]
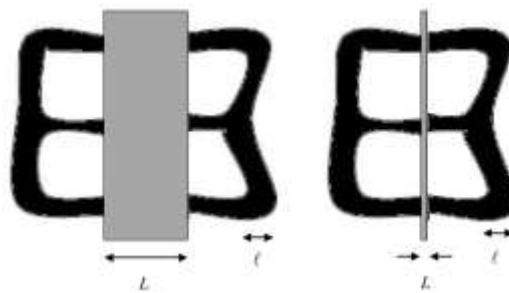
Figure 4 The effect of aspect ratios on image inpainting [14]**.**

3.0 General Guidelines for Image Inpainting

Vision research has been generally classified into different levels, low, middle, and high, though there exists no clear-cut boundaries. Tasks with low complexities such as denoising and image enhancement, or most problems in image processing, belong to lower level vision. On the other hand, inference-oriented and learning adapted tasks, such as pattern recognition, classification, and organization, are typically higher level vision problems. Most contemporary works have been primarily focused on lower level or middle level ones, because higher level ones are built upon them and in practice are still more challenging in terms of both modeling and computation. The general guidelines which can be very helpful in making novel contributions in modeling and computation of lower level inpainting algorithms are,

(a) Locality: Inpainting is local. The missing information is inpainted based only on image information available in the vicinities of missing areas.

(b) Functionality: Inpainting is functional. The inpainting models or algorithms depend only on properties of images as functions and not on any higher level pattern recognition inputs.
(c) Automation: Inpainting must be as automatic as possible. The lesser human input demands, more powerful models for real applications.

(d) Genericity: Inpainting should be able to deal with as many generic images as possible. It means that as long as information lost is local, the most generic incomplete images can be successfully inpainted to certain satisfactory precision.

(e) Stability: Inpainting should be stable, meaning that it must have built-in mechanisms to resist any minor degradation of the available image information such as noise and blur, which are very common in applications.

4.0 Clustering Inpainting Tools

As for the majority of signal processing problems, the various ways to address the inpainting problem can be organized, clustered and classified depending on the technical approach, concepts and tools used. Based on the literature review made for the purpose of understanding and structuring the problem at hand, this means image and video inpainting, some clustering dimensions emerged as more relevant. While there is no single good clustering approach, having some appropriate organization for inpainting solutions helps in understanding their relationships, notably similarities and differences between available and emerging solutions. In this context, the main dimensions proposed to cluster and classify the technologies and solutions for image and video inpainting, regardless of the inpainting perspective adopted, are depicted in Figure 5.
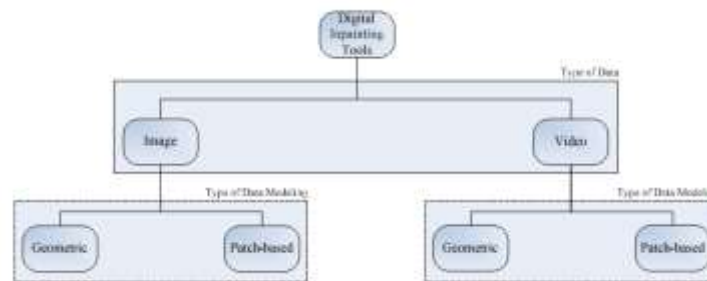

Figure 5 – Proposed clustering for digital inpainting tools [26].

As shown in Figure 5, the dimensions adopted to organize and classify digital inpainting tools are:

4.1 Type of Data : This has been considered to be the first dimension for clustering as it clearly distinguishes two related, yet different worlds: image and video. The type of data implicitly defines the nature of the redundancy to be exploited, notably spatial redundancy for image and spatial and temporal redundancies for video.

4.2 Type of Data Modeling : This dimension regards the type of data modeling used for the data to be inpainted. As such, it allows identifying two different, yet complementary approaches that are vastly shared by the inpainting research community: geometric modeling and patch-based modeling. In this context, modeling implies a set of models and underlying assumptions which will allow digitally addressing the inpainting problem. As shown in Figure 1.6, these types of data modeling are the same for image and video data, reflecting the fact that they are largely agnostic to the type of data. However, for video, they might not be applicable exactly in the same way or with the same computational cost as for images, because video sequences bring out even more challenging requirements to cope with when performing inpainting. In the following, the various inpainting tools clustering branches and leafs will be briefly discussed.

5.0 Image Inpainting Tools

In the proposed clustering, digital inpainting tools are divided into two main categories, depending on the type of data being considered, notably images and video sequences. The main differences between these two types of data are the world they live in and the amount of information to be processed. Images live in a 2D world which is spatially constrained by two coordinates, whereas video sequences live in a 3D world in the sense that they are governed by both the spatial image coordinates and time, which is considered to be the third dimension.

6.0 Geometric Modeling

Regarding image data, digital inpainting tools can be further classified using the second dimension adopted, i.e. the type of data modeling. The types of data modeling considered in image inpainting tools can be either geometric or patch-based, as shown in Figure 4. Geometric modeling is based on the propagation of structural image properties, e.g. edges, from the source to the target area to inpaint the missing information in the image. Moreover, this type of data modeling works at pixel level and is good at restoring small defects and thin structures. From the available set

of structural image properties, edges are the most vastly used features in inpainting tools as the human visual system strongly relies on them to identify and understand the objects' attributes and their mutual associations. In this context, geometric models can be seen, in some degree, as edge-continuing models, where edge information is first extracted and then diffused inwards the hole to be inpainted. As for structure propagation, partial differential equations (PDE) are vastly used by the inpainting research community since they allow interpolating structural image properties in a smooth fashion, achieving good results in terms of the restored image perceived visual quality and coding efficiency, despite being computationally demanding.

7.0 Patch-based Modeling

Patch-based modeling is considered to be the most relevant alternative to geometric modeling among the digital inpainting tools available in the literature. This is justified by the fact that the manipulation of patch-based models is straightforward, in contrast, for instance, with parametric models for which it is hard to find appropriate mathematical methods to perform inpainting. Patch-based modeling is supported by texture synthesis procedures which are responsible for searching the most texture-compatible source fragments matching the textural information of the target pixels' vicinity. Unlike geometric modeling, patch-based modeling does not work at pixel level, but rather at a more global level and, therefore, it is best suited for filling-in large regions. The underlying assumption is that patch-based models consider the patch as the fundamental element in the image instead of the pixels, in order to perform searching and template matching in an efficient and adequate fashion. Regardless of the modeling adopted, both geometric and patch-based models standing alone allow achieving good results. However, to fully take advantage from the best of both worlds in terms of the restored image perceived visual quality and coding efficiency, hybrid models may be considered by the inpainting research community.

8.0 Video Inpainting Tools

As aforementioned, although image inpainting poses challenging difficulties, video inpainting discloses even more demanding problems. First of all, the amount of data in video sequences is much greater than for images and, second, temporal consistency is a must, due to the HVS's sensitivity to motion. Moreover, when considering videos, both spatial and temporal redundancies must be exploited; this directly impacts on the tools to be used in pure video inpainting and in inpainting-based video compression. As shown in Figure 4, video inpainting tools can also be further classified by the second dimension adopted, notably using the same categories as for image inpainting tools. However, this does not mean that these types of modeling are applicable exactly in the same way, in the same proportion or even with the same computational cost, independently of the type of data addressed. If that was the case, video inpainting tools would strictly consider the video as a set of uncorrelated images and, in consequence, they would simply be a naïve extension of image inpainting tools, discarding the temporal correlation between frames. What happens instead is that the type of data modeling adopted by video inpainting tools needs to be adapted to cope with temporal consistency (avoiding intense flickering and a noticeable poor sense of motion), hence allowing to effectively exploit both the spatial and temporal redundancies.

9.0 Literature Survey

Image inpainting is an iterative process which ends when any target condition is matched or maximum number of iterations is completed. Bertalmio et al.[2] introduce the term image inpainting to computer science. In the algorithm the region that has to be inpainted will be filled-in by information of the region surrounding the gap. The curves of equal intensity (isophotes) arriving at the boundary are propagated inwards. Because this is the first and one of the most important papers in image inpainting.

Criminisi et al.[1] propose an algorithm inspired by the algorithm by Bertalmio et al.[2] and texture synthesis [3]. In contrast to the paper of Bertalmio et al. the unknown region is inpainted patch by patch. The sequence of which patch should be inpainted is based on the isophote information and the amount of known information surrounding the patch. This popular method increased the result of image inpainting at lot. Because of that this method is further explained in chapter 4.

In the algorithm of Oliveira et al.[4], the unknown region of the image is convolved with a Gaussian kernel. To prevent edges to be blurred the user manually specifies barriers for the diffusion. This algorithm is much faster than the other algorithms but it can only be used with very small unknown regions and is less accurate. An example of this algorithm is shown in figure 6.
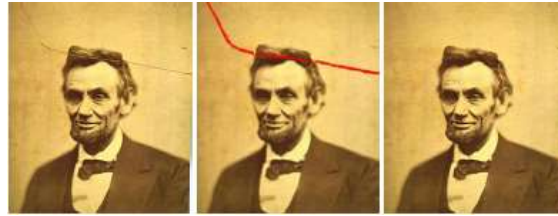
Figure 6: Olivera method [4], input, mask and result, note the diffusion barriers near the boundaries of the hair of Abraham Lincoln.

The algorithm of Perez et al.[5] showed how gradient domain reconstruction can be used in image editing applications. The actual pixel values for the unknown pixel values are computed by solving a Poisson equation that locally matches the gradients while obeying the _xed Dirichlet (exact matching) conditions at the seam boundary. Poisson Image Editing can be used best for seamless inserting and local illumination changes but can also be used for image inpainting. The author of [6] proposed a new inpainting algorithm based on propagating an image smoothness estimator along the image gradient. Similar to the algorithm of Bertalmio et al.[2]. The image smoothness is estimated as the weighted average of the known image neighborhood of the pixel to inpaint. The fast marching method (FMM) is used to create a distance function to the initial boundary. The pixels of the unknown region are inpainted in the order of the distance to the boundary, proceeding from the smallest to the largest. This method is fast but creates blurry effects with larger unknown regions. An example of this algorithm is shown in Figure 2.3.

Figure 7: Telea method [6], input plus mask and result.

Texture synthesis is the process of algorithmically constructing a large digital image from a small digital sample image by taking advantage of its structural content [3][7]. Texture synthesis can be also used to all in unknown regions in images, the algorithm of Criminisi et al.[1] is partially based on these algorithms. An example of the algorithm by Efros et al.[3] is shown in Figure 8.

Figure 8: Texture Syntheses by Efros et al.[3], input plus mask and result.

10.0 Guided Method

Sun et al.[8] introduces a new direction in image inpainting. In the authors algorithm the user is also able to specify support lines. These support lines specify where important lines in an image should be continued. This algorithm first alls in the unknown information around the support lines by dynamic programming or belief propagation, depending on the structure of the support lines. After that the rest of the image is filled in by texture propagation. The result of this algorithm is good but we can not compare these with others because this algorithm requires extra input. An example of this algorithm is shown in Figure 9.

Figure 9: Sun method [8], input, mask, first stage and result

11.0 Multiple Images

Hays et al.[9] developed a total different way of image completion, instead of searching in the input image, the algorithm searches throughout a whole database of images to and information to all the missing region. The new area is pasted in using Poisson blending [5]. The authors state that their results look better then the algorithm of Criminisi et al.[1] but the algorithm needs a database of two million images for only three scenes. This algorithm gives visual pleasing results but is very slow and is unpractical because it needs the large dataset. An example of this algorithm is shown in Figure 10.

Figure 10: Hays method [9], input, mask, first stage and result

12.0 Video Inpainting

Image completing methods can also be used for videos, because a video is a set of multiple images. Naively inpainting each frame will not result in the best result important information about the current frame can be found in the adjacent frames. Patwardhan et al.[10] presented an algorithm for video inpainting of a scene taken from a static camera. This method is an extension of the paper of Criminisi et al.[1]. It extends the idea from a single image to a set of images but also taken in to account the adjacent frames. In this algorithm the frame is separated into a background and a foreground in which the foreground is first inpainted. After this step the background is inpainted, each patch is copied to every frame to get a consistent background. This technique has some nice results but it has some restrictions, it requires a fixed camera position and a stationary background with some moving foreground. Wexler et al.[11] have proposed a method for space-time completion of large damaged areas in a video sequence. They pose the problem of video completion as a global optimization problem with a well-defined objective function.

In the algorithm every local patch should be found in the remaining part of the video and globally all these patches must be consistent with each other spatially and temporally. An example of this algorithm is shown in Figure 2.7.
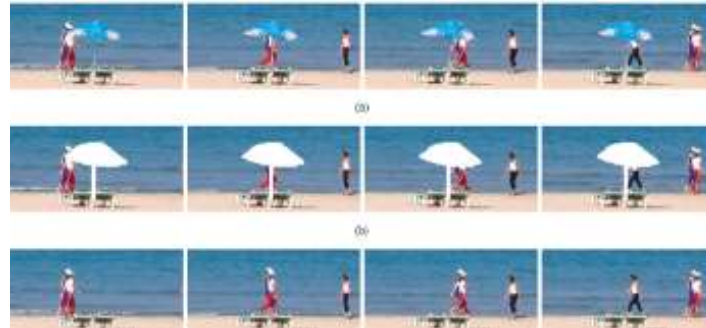

Figure 2.7: Wexler method [11], input, mask and result sequence.

13.0 Conclusion

In this survey paper we have discussed various techniques implemented by researchers for image inpainting.

14.0 References

[1].M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, \Image inpainting," in Proceedings of the 27th annual conference on Computer graphics and interactive techniques, SIGGRAPH '00, (New York, NY, USA), pp. 417-424, ACM Press/Addison-Wesley Publishing Co., 2000.

[2].A. Criminisi, P. P_erez, and K. Toyama, \Region _lling and object removal by exemplar-based image inpainting," IEEE Transactions on Image Processing, vol. 13, pp. 1200-1212, 2004.

[3].A. A. Efros and T. K. Leung, \Texture synthesis by non-parametric sampling," in Proceedings of the International Conference on Computer Vision-Volume 2 - Volume 2, ICCV '99, (Washington, DC, USA), pp. 1033, IEEE Computer Society, 1999.

[4].M. M. Oliveira, B. Bowen, R. Mckenna, and Y. sung Chang, \Fast digital image inpainting," in Proceedings of the International Conference on Visualization, Imaging and Image Processing (VIIP 2001), pp. 261-266, ACTA Press, 2001.

[5].P. P_erez, M. Gangnet, and A. Blake, \Poisson image editing," in ACM SIGGRAPH 2003 Papers, SIGGRAPH '03, (New York, NY, USA), pp. 313{318, ACM, 2003.

[6].A. Telea, \An image inpainting technique based on the fast marching method," Journal of Graphics, GPU, and Game Tools, vol. 9, no. 1, pp. 23-34, 2004.

[7].L.-Y. Wei and M. Levoy, \Fast texture synthesis using tree-structured vector quantization," in Proceedings of the 27th annual conference on Computer graphics and interactive techniques, SIGGRAPH '00, (New 61 York, NY, USA), pp. 479-488, ACM Press/Addison-Wesley Publishing Co., 2000.

[8].J. Sun, L. Yuan, J. Jia, and H.-Y. Shum, \Image completion with structure propagation," ACM Trans. Graph., vol. 24, pp. 861-868, July 2005.

[9].J. Hays and A. A. Efros, \Scene completion using millions of photographs," ACM Transactions on Graphics (SIGGRAPH 2007), vol. 26, no. 3, 2007.

[10].K. A. Patwardhan, G. Sapiro, and M. Bertalmo, \Video inpainting of occluding and occluded objects," in International Conference on Image Processing, pp. 69-72, 2005.

[11].Y. Wexler, E. Shechtman, and M. Irani, \Space-time video completion," Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, vol. 1, pp. 120-127, 2004.

[12].P. Perona and J. Malik, \Scale-space and edge detection using anisotropic di_usion," IEEE Trans. Pattern Anal. Mach. Intell., vol. 12, pp. 629-639, July 1990.

[13].T.-H. Kwok, H. Sheung, and C. C. L. Wang, \Fast query for exemplarbased image completion," IEEE Transactions on Image Processing., vol. 19, pp. 3106-3115, December 2010.

[14].Q. Chen, Y. Zhang, and Y. Liu, \Image inpainting with improved exemplar-based approach," in Proceedings of the 2007 international conference on Multimedia content analysis and mining, MCAM'07, (Berlin, Heidelberg), pp. 242-251, Springer-Verlag, 2007.

[15].Anupam, P. Goyal, and S. Diwakar, \Fast and enhanced algorithm for exemplar based image inpainting," Image and Video Technology, Pacific-Rim Symposium on Image and Video Technology, vol. 0, pp. 325-330, 2010.

[16].T. Ružic, A. Pižurica, "Context aware image inpainting with application to virtual restoration of old paintings," IEICE International Conference Information and Communication Technology Forum ICTF, 2013, 29-31 May, Sarajevo, Bosnia

[17].Raluca Vreja and Remus Brad," Image Inpainting Methods Evaluation and Improvement", The Scientific World Journal, Volume 2014.

[18].Jun Zhou, Antonio Robles-Kelly, "Image Inpainting Based on Local Optimisation" 2010 International Conference on Pattern Recognition.

[19].Sarab M. Hameed , Nasreen J. Kadhim, Mahmood A. Othman," Image Inpainting Based On Particle Swarm Optimization" Iraqi Journal of Science, Vol.50, No.2, 2009.

[20].Zhaoxia Wang, Quan Wang, CS Chang, Ming bai, Zhen Sun, Ting Yang," Image Inpainting Method based on Evolutionary Algorithm", International Journal of Digital Content Technology and its Applications. Volume 5, Number 4, April 2011.

[21].K. Sangeetha, Dr. P. Sengottuvelan," A Novel Exemplar based Image Inpainting Algorithm for Natural Scene Image Completion with Improved Patch Prioritizing", International Journal of Computer Applications, Volume 36– No.4, December 2011

[22].Seema Kumari Singh , Prof J.V Shinde," Optimum Patch Selection Using GA in Exemplar Based Image In-painting" (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (2) , 2015.

[23].B. Fergani, H.Bensuici, "A discrete firefly algorithm for geometric image inpainting". International Conference on Advanced Technologies For Signal & Image Processing Atsip 2014, At Sousse-Tunisia. Chongwu Tang, Xi Hu, Li Chen, Guangtao Zhai, Xiaokang Yang," Sample-Based Image Completion Using Structure Synthesis", IEEE 2013

[24].Shivali Tyagi, Sachin Singh," Image Inpainting By Optimized Exemplar Region Filling Algorithm" International Journal of Soft Computing and Engineering (IJSCE), Volume-2, Issue-6, January 2013

[25].Nuno Miguel Ventura Couto, "Inpainting-based Image Coding: A Patch-driven Approach", Dissertation submitted for Master in Electrical and Computer Engineering in 2010

[26].Sander van de Ven," Image Inpainting", master Thesis in Utrecht University,2012