Image Completion using Criminisi Algorithm

Aryadarsh S
Department of Electronics & Communication Engineering
Marian College of Engineering, Trivandrum, India

Dhanya Mathew
Assistant Professor
Marian College of Engineering, Trivandrum, India

Neevan R
Assistant Professor
IHRD Engineering College Attingal, Trivandrum, India

Abstract

Image completion means, Restoring the lost part of an image by using available information from the remaining parts of the image itself, it is otherwise called as Image inpainting. There are a number of applications for inpainting, they vary from the restoration of damaged paintings and photographs to the removal/replacement of selected objects. In this paper, we present an algorithm that enhances and extends a previously proposed algorithm and it provides faster inpainting. Using our approach, one can use this to inpaint large regions (e.g. to remove an object etc.) as well as it is used to recover small portions (e.g. restore a photograph by removing cracks etc.). The inpainting method is based on the exemplar based approach. The basic idea behind this approach is to find exemplars (i.e. patches) from the image and replace the lost data with it. This technique can be used for the restoration of damaged photographs or damaged film. In contrast with previous approaches, the technique here introduced has an advantage. That is, it does not require the user to specify where the novel information comes from. This is automatically done (and in a fast way), thereby it allows the system to simultaneously fill-in numerous regions which contain completely different structures and surrounding backgrounds. This paper looks forward to improve the algorithm so that the computational complexity is further improved while retaining the quality of inpainting. Here the inpainting algorithm presented here is not meant to be used for inpainting images, but for videos also. We are also investing methods to improve this algorithm to make it more robust so that it can be used with videos in this paper itself.

Keywords- Image Inpainting, isophotes, priority term, source region, target region, Video Inpainting

I. INTRODUCTION

Images are the visual representation of the various things that we see and experience in our lives. These images when profoundly affected by degradations excluding noise, it would be advantageous if those images could be completed in a visually plausible manner. This is what can be termed as Image Completion.

Images as well as videos, which are mere sequences of images, are very essential in today's world for communication and transmission. At this juncture, we would like to state that any disintegration produced in images degrade the quality and texture that it may become unusable. In order to retain such images that have suffered from any kind of disturbance, our method can be implemented and thus, the images can be completed and revived

Image completion or Image Inpainting is an important and challenging computer vision task. The goal of an image completion algorithm is to reconstruct the missing regions, those are contained in an image in such a way that, it should be visually plausible to an observer. In most cases, the missing region in the image (the target region) is filled in by using the information from the rest of the image (the source region). Image completion is an important part of many computer vision applications such as scratch removal, object removal and reconstruction of damaged architectural parts in an image. Moreover, image completion can be applied in the field of photo editing and image restoration

Image completion is a highly researched area in computer vision [Criminisi 2004, He 2012, Hsin 2010, Hung 2008, Komodakis 2007, Li 2011]. Although due to the complexity of the images, results leave a lot to be desired [1]. Ideally, any algorithm that is designed to solve the image completion problem should follow the given characteristics [2]:

- it should be able to successfully complete complex natural images,
- it should be able to handle incomplete images with (possibly) large missing area
- all these should take place in a fully automatic manner, i.e., without intervention from the user

The restoration can be usually done by using two approaches, image inpainting and texture synthesis. The meaning of the first approach is restoring of missing and damaged parts of images in a way that the observer who doesn't know the original image, will not be able to detect the difference between the original and the restored image. It is called inpainting because; it is the process of painting or filling the holes or cracks in an artwork.

The second approach is filling unknown area on the image by using the texture information from the surroundings or from input texture sample. Texture synthesis techniques could be employed for the restoration of digitized photographs especially if a damaged area needs to be filled with some pattern or structure.

However, texture synthesis usually fails, if the area to be reconstructed contains an additional colour or intensity gradient. We would like any image completion algorithm to be able to handle the related problem of texture synthesis as well. It is exactly due to all of the above requirements that image completion is, in general, a very challenging problem. Nevertheless, it is very useful in many areas, e.g., it can be used for computer graphics applications, image editing applications, film postproduction, image restoration, etc.

In our project, we introduce a novel algorithm for digital inpainting technique, which is used for still images that attempts to enhance the basic techniques used by professionals for restorations. After the user selects the regions to be restored, the algorithm automatically fills-in these regions with information surrounding the target area. The fill-in process is done in such a way that, the isophote lines arriving at the region's boundaries are completed inside. When compared with previous approaches, the technique here introduced does not require the user to specify where the new information comes from. This is automatically done (and in a fast way), thereby allowing to simultaneously fill the numerous regions which contain completely different structures and surrounding backgrounds. In addition, no limitations are imposed on the topology of the region to be inpainted

II. BASIC METHODOLOGIES

Roughly speaking, there have been three main techniques so far, for dealing with the image completion problem:

- Statistical-based methods,
- PDE-based methods,
- Exemplar-based methods.

A. Statistical-Based Methods

These methods are mainly used for texture synthesis. Typically, what these methods do is that, given an input texture, they try to describe the process by extracting some statistics by the use of compact parametric statistical models. Then, in order to synthesize a new texture, these methods start with an output image containing pure noise, and keep perturbing that image, until its statistics match with the estimated statistics of the input texture. Besides the synthesis of still images, parametric statistical models have also been proposed for the case of image sequences. However, the main drawback of all methods, those are based on parametric statistical models is that, as already mentioned, they can be applied only to the problem of texture synthesis, and not to the general problem of image completion. But even in the restricted case of texture synthesis, they can synthesize only textures which are highly stochastic and usually fail to do so for textures which contain structures as well. Nevertheless, in cases, where parametric models are applicable, they allow greater flexibility with respect to the modification of texture properties. Furthermore, these methods can be very useful for the process which is reverse to texture synthesis, i.e., the analysis of textures.

Fig. 1: Image inpainting methods, when applied to large or textured missing regions, very often over smooth the image and introduce blurring artefacts.

B. PDE-Based Methods

These methods, on the other hand, try to fill the missing region of an image through a diffusion process, by smoothly propagating information, from the boundary towards the interior portions of the missing region. According to these techniques, the diffusion process can be done by solving a partial differential equation (PDE), which is typically non-linear and of high order. This class of methods was first introduced by Bertalmio et al. in [4], in which case the authors were trying to fill a hole in an image by propagating image Laplacian functions in the isophote direction. Their algorithm was trying to mimic the behaviour of professional rest orators in image restoration. Furthermore, recently, Bertalmio et al. [5] have proposed the method to decompose an image into two components. The first component is representing structure and is filled by using the PDE based method, while the second component represents texture in the image and is filled by use of a texture synthesis method. Finally, Chan and Shen [6] have used an elastic based vibrational model for filling the missing part of an image.

However, the main disadvantage of almost all the PDE based methods is that, they are mostly suitable for the image inpainting situations. This method is usually used, where the missing part of the image consists of thin, elongated regions. In order to use PDE method, the content of the missing region should be smooth and non-textured. For this reason, when these methods are applied to images where the missing regions are large and textured, they usually over smooth the image and

introduce blurring artefacts. On the contrary, we would like our method to be able to handle images that contain possibly large missing parts. We would also like our method to be able to fill arbitrarily complex natural images, i.e., images which contain textures, structures or even a combination of both.

C. Exemplar-Based Methods

Finally, the last class of methods is the so-called exemplar-based techniques, which actually have been the most successful technique among inpainting techniques. These methods try to fill the unknown region simply by copying content from the observed part of the image. Recently, however, there have been a few authors who have tried to extend these methods to image completion also. But, in this method, there is a major drawback related to the greedy way of filling the image, which can often lead to visual inconsistencies. However, these two are the main handicaps of related existing techniques. First of all, the confidence map is computed based on heuristics and ad hoc principles, which may not be applied in the general case, and as the second step, once an observed patch has been assigned to a missing block of pixels, that block cannot change its assigned patch after that. This last fact reveals the greediness of these techniques, which may again lead to visual inconsistencies.

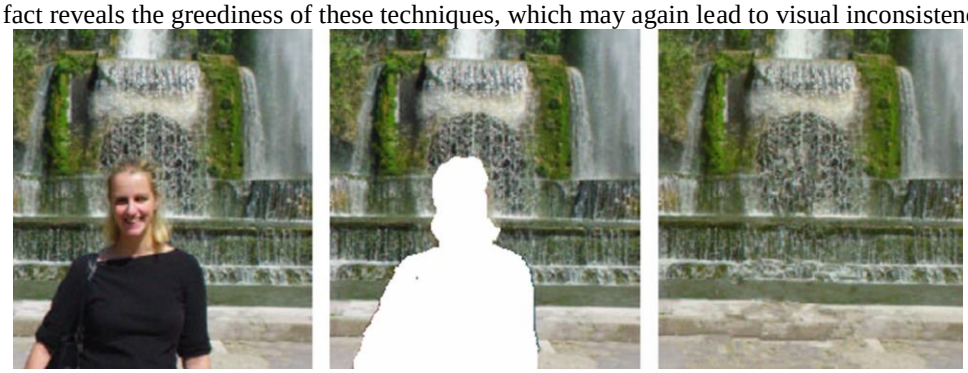

Fig. 2: an example for region filling using exemplar method of inpainting

III. LITERATURE REVIEW

Bertalmio et al (2000) [3] first presented the notion of digital image inpainting and third order Partial Differential Equations (PDE) are used to propagate the known image information into the missing regions along the direction of isophote. Many algorithms (C. Ballester et al (2001), M. Bertalmio et al (2001), M. Bertalmio et al (2000), T. Chan and J. Shen (2001), S. Masnou et al (1998)) address the region filling issue for the task of image inpainting where speckles, scratches, and overlaid text are removed from a damaged image. These image inpainting techniques are used to fill the holes in images by propagating linear structures into the target region by using diffusion technique. In this paper a third-order PDE have been introduced to perform geometric based inpainting on images. The image inpainting algorithm of Bertalmio et al (2000) is not contrast invariant, as it was pointed out by Chan et al (2001). In this paper (Marcelo Bertalmio (2006)), the inpainting problem was reformulated as a particular case of image interpolation in which the author intended to propagate level lines in the algorithm.

A. Disadvantages

- Inpainting method is slower for large areas to be filled
- Large object removal is difficult.

Chan et al (2001) [7] present the Total Variation (TV) inpainting model in Levin et al (2003), based on the Euler Lagrange equation, uses anisotropic diffusion based on the contrast of the isophote. This model is designed for inpainting small regions; it does a good job at removing noise, but couldn't repair large regions also Images may contain textures with arbitrary spatial discontinuities, but the sampling theorem considers the spatial frequency content that can be automatically restored. Thus, in the case of missing or damaged areas, one can only expect for the production of a plausible rather than an exact reconstruction. Therefore, in order for an inpainting model to be reasonably successful for a large class of images the regions to be inpainted must be locally small. As the regions become smaller, simpler models can be used to approximate the results produced by more sophisticated ones locally. Another important observation used in the design of our algorithm is that, the human visual system can tolerate some amount of blurring in areas which are not associated with high contrast edges.

B. Limitations are

- Applicable only for small scratches
- Much iteration is required

Criminisi et al (2004) [1] designed an exemplar based inpainting algorithm by propagating the known image patches (i.e., exemplars) into the missing patches. To handle the missing region with composite textures and structures, patch priority is defined to encourage the filling order of patches on the structure. In the past, this problem has been addressed by two classes of algorithms

- “texture synthesis” algorithms for generating large image regions from sample textures, and

– "Structure reconstruction" techniques for filling in small image gaps.
– "texture synthesis" algorithms for generating large image regions from sample textures, and
– "Structure reconstruction" techniques for filling in small image gaps.

The former has been demonstrated for "textures" - repeating two-dimensional patterns with some stochasticity; the latter focus on linear "structures" which can be defined as one- dimensional patterns, such as lines and object contours. This paper presents a new and efficient algorithm that combines the advantages of these two approaches. We first note that exemplar-based texture synthesis includes the essential process required to replicate both texture and structure; the success of structure propagation, however, is highly dependent on the order in which the filling process proceeds. We propose a best-first algorithm in which the confidence in the synthesized pixel values is propagated in such a way that, it is similar to the propagation of information in inpainting. The actual colour values are computed using exemplar-based synthesis. In this paper the simultaneous propagation of texture and structure information is achieved by using a single, efficient algorithm. Computational efficiency can be achieved by a block based sampling process.

IV. DIGITAL VIDEO COMPLETION

Video inpainting started as a natural extension of image inpainting algorithms and it has garnered a great deal of attention due to its applications in video error concealment in video transmission, multimedia editing and visualization, video stabilization and new applications such as video modification for privacy protection. A straightforward development of image inpainting algorithms to video inpainting is to treat the underlying video data as a set of distinct images and apply image inpainting algorithms to each frame individually. This mode of operation does not take full advantage of the high temporal correlation that occurs in video sequences and hence the qualities of video inpainting across the frames are usually unsatisfactory. For example, one of the earliest efforts in extending the Partial Differential Equation (PDE) based image inpainting to video was performed by Bertalmio et.al [4]. The focus of this method is to fill in the hole spatially by extending the edges and filling the hole with smoothed colour information by a decision process using Navier-Stokes equation. It does not take into effect the temporal information available in the video and treats the video as individual images. Due to extensive smoothing, it does not reproduce the texture information and suffers from severe blurring artefacts. Consequently, this method is effective only in restoring small scratches or spots occurring in archival footage. Cheung et.al introduced a space-time patch model based on probabilistic learning with applications to inpainting.

A priority based exemplar approach proposed for image inpainting by Criminisiet.al in was modified by Patwardhan et al. to video inpainting in. This method is capable of inpainting videos under a set of constrained camera motion. Initially, the input video is separated into background layer and foreground object layer utilizing the optical flow. Hole regions identified in the foreground layer are first inpainted by a priority-based exemplar process before proceeding to complete the damaged regions in the background layer. In this patch based exemplar method, damaged patches around the boundary of the hole are filled by a priority based mechanism. The appropriate candidates for filling the damaged areas are selected by minimizing a 5- dimensional distance metric based on the pixel colour values and the optical flow vectors while being effective in completing regions with sparsely.

A video completion scheme based on motion layer estimation followed by motion compensation and texture completion has been proposed. After removing a particular motion layer, motion compensation method is used to complete moving objects and non-parametric texture synthesis is used to complete the static background regions. The inpainted layers are then warped into every video frame for the completion of the holes. Video completion by motion field transfer - transfer of spatial-temporal patches of motion field instead of direct colour sampling has been introduced recently. This technique is extremely sensitive to noise as they involve local motion estimation by a derivative-based process. It has difficulty in, inpainting large motion as their motion estimation techniques focus solely on measuring small local movement. As the scheme transfers only motion information, it suffers from blurring artefact due to the use of a re-sampling process to estimate colour information also. A video completion algorithm for perspective camera under constrained motion has been proposed recently. The foreground and background layers are separated and objects in foreground volume are rectified to compensate for perspective projection. The pixels in the foreground are completed by modelling it as a graph labelling problem and a dynamic programming is used to solve it.

V. METHODOLOGY

In our paper, we adopt notation similar to that used in the inpainting literature. The region to be filled, i.e., the target region is represented as Ω, and its contour is denoted as $\delta\Omega$. The contour evolves inward as the algorithm progresses, and so we refer to it as the "fill front". The source region, Φ, which remains fixed throughout the algorithm, and the algorithm provides samples used in the filling process. We now focus on a single iteration of the algorithm to show how structure and texture are simultaneously handled by exemplar-based synthesis. Suppose that the square template Ψp, centred at the point p is to be filled. The best-match sample from the source region comes from the patch $\Psi q'$ which is most similar to those parts that are the actual missing parts in Ψp. In the example in fig. 3 b, we see that if Ψp lies on the continuation of an image boundary or edge, the most likely best matches will be along the same edge.

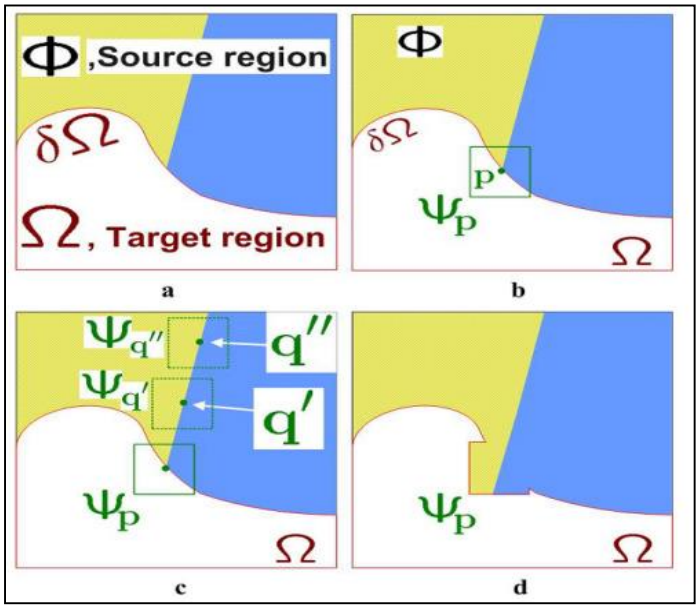

Fig. 2: Structure propagation by exemplar-based texture synthesis. All that is required to propagate the isophote inwards is a simple transfer of the pattern from the best match source patch (fig. 3.d).

From the image we can notice that isophote orientation is automatically preserved. In the figure, despite the fact that the original edge is not perpendicular to the target contour δΩ, the propagated structure has maintained the same orientation as in the source region. We focus on a patch-based filling approach this improves execution speed. Furthermore, we note that, patch based filling enhances the accuracy of the propagated structures.

A. Filling Order: A critical factor

This section demonstrates that the quality of the output image synthesis is highly influenced by the order in which the filling process proceeds. Furthermore, we list a number of desired properties of the "ideal" filling algorithm. A comparison between the standard concentric layer filling (onion-peel) and the desired filling behaviour is illustrated in Figure 3. Figures 3 b, c, d show the progressive filling of a concave target region via an anti-clockwise onion-peel strategy.

As it can be observed, this arrangement of the filled patches produces the horizontal boundary between the background images regions are to be unexpectedly reconstructed as a curve. A better filling algorithm would be one that gives higher priority to those regions of the target area which lie on the continuation of image structures, as shown in figs. 3 b', c', d'. Together with the property of correct propagation of linear structures, the Onion Peel Desiderata latter algorithm would also be more robust towards the varying shape of the target. In the figure 3 (a) is a diagram showing an image and a selected target region (given in white). The remainder of the image is the source i.e. b, c, d which explains the different stages in the filling of the target region in the concentric way of filling. (d) The onion-peel approach produces artefacts in the synthesized horizontal structure b'; c', d'. By filling the target region using an edge-driven filling order achieves the desired artefact-free reconstruction. (d) The Final edge-driven reconstruction is given here, where the boundary between the two background image regions has been reconstructed correctly.

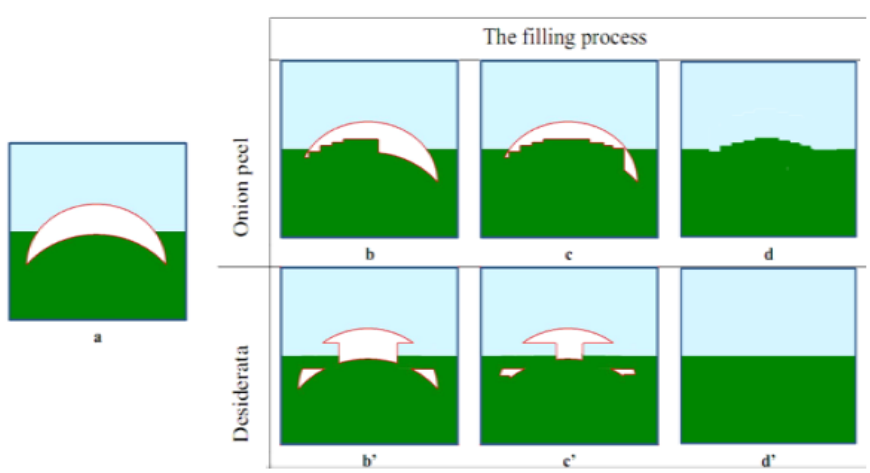

Fig. 3: The importance of the Filling order when dealing with concave target regions

This paper is an extension to earlier inpainting algorithms with a focus on improving the computational complexity of the methods along with some other improvements such as speed and accuracy. **Ω** represents the target region of an original image, i.e. the region to be inpainted. Represents the source region, i.e. the region from which information is available to reconstruct the image. Generally, $\Phi = I - \Omega$ Also, we use $\delta\Omega$ to represent the boundary of the target region, i.e. the fill front. From here that we find some patch that is to be filled.

B. Algorithm

Generally, an exemplar based inpainting algorithm includes

1) Initialize the target region. This is generally performed separately from the process of inpainting and requires the use of an additional image processing tool. This is performed by marking the target region by using some colour. Without any loss of generality, let us assume that the colour that the target region will be marked in is green (i.e. R = 0, G = 255, B = 0).
2) Find the boundary of the target region.
3) Select a patch from the region to be inpainted. The patch size should be a bit larger than the largest distinguishable texture element in the image. We have used a given patch size of 9 x 9 which can be changed with the knowledge of the largest texture element in the image. We represent the patch by Ψp.
4) Find a patch from the image which best matches the selected patch, Ψp. The process of matching can be done by using a suitable error metric. We use the Mean Squared Error to find the best matching patch.
5) Update the image information according to the patch found in the previous step.

As mentioned earlier, the result depends considerably on the third step wherein a patch is selected to be inpainted. The result that we obtain would always depends on the order of selection and thus there have been approaches that try to define this selection order so that the result is improved.

In Criminisi algorithm, the priority function used for selecting the best patch from the target region is defined in a product form:

$$P(p) = C(p) \times D(p) \quad (1)$$

Where C(p) represents the confidence term for the patch and D(p) the data term for the patch.
For each point p in the boundary $\partial\Omega$, find the following:

$$C(p) = \{ \sum q \epsilon (\Psi p \cap \Phi) \; C(q)\} / |\Psi p| \quad (2)$$

$$D(p) = |\nabla I_{\perp}^{p} . n_p| / y \quad (3)$$

Where $|\Psi p|$ is the area of the patch Ψp, which we have to fill, and γ is the normalization factor (equal to 255 for a normal gray level image), n_p is a unit vector which is orthogonal to the front $\delta\Omega$ at the point p represents the perpendicular isophote at point p. The value of n_p is found by finding the gradient for the source region. The source region represents a matrix with all ones on the points that are not present in the target region and zeros otherwise (i.e. for the points in **Ω**).

Isophote can be determined using the gradient of the image. Cheng et al. discovered that the confidence term that was defined in Criminisi algorithm decreases exponentially and thus the product form definition of the priority term needs to be replaced. Thus they modified the confidence term with the regularized confidence term. Using this confidence term the value of the confidence term is regularized to [ω, 1]. In this way the new compensated priority function will be able to resist the "Dropping effect".

Also, the addition of weights can be given to different components in the definition of priority term so that a balance between confidence and data term can be maintained.
Thus the modified priority term can now be represented as:

$$P(p) = \alpha \times Rc(p) + \beta \times D(p), \; 0 \le \alpha, \beta \le 1 \quad (4)$$

Where α and β are respectively the component weights for the confidence and data terms. Also $\alpha + \beta = 1$ and Rc (p) is the regularized confidence term:

$$Rc(p) = (1-\omega) \times C(p) + \omega, \; 0 \le \omega \le 1 \quad (5)$$

The solution to the problem that we propose consists of the computation of variance of the patches with same mean squared error. This variance that we use is the variance of the pixel values of the patch with respect to the mean of the pixels, that is the difference in intensity from the mean value, from the same patch that correspond to the pixels belonging to source region from the patch to be inpainted.

VI. CONCLUSION

For doing this project we have implemented the basic algorithm presented by A. Criminisi et. al. and then worked from thereon to improve it based on my observations and research. This presents an algorithm that can remove objects from the image in a way that it seems reasonable to the human eye. This approach employs an exemplar based inpainting in relation with a priority term that defines the filling order in the image.

In this algorithm, pixels maintain a confidence value and are chosen based on their priority that is computed using confidence and data term. The confidence term defines how much sure we are about the validity and the accuracy of that pixel whereas data term is focused towards maintaining the linear structures in the image. This approach is capable of the propagation of both linear structures and 2 dimensional textures into the target region.

We can use this technique to fill small scratches in the image/photos as well as to remove larger objects from them. In larger images also it is computationally efficient and works well. For very large image for which we are facing memory allocation problems it fails. We are also looking forward to improve the algorithm so that the computational complexity is improved while retaining the quality of inpainting and if possible, we would also like to enhance the inpainting algorithm. Also the inpainting algorithm presented here is not capable enough to be used for inpainting videos, i.e. removal of some scratches or some objects from videos. We are also exploring towards this area to make it more robust so that it can be used with videos.

ACKNOWLEDGEMENT

I would like to express my special gratitude and thanks for giving me attention and time

REFERENCES

[1] A.Criminisi, P.Perez and K.Toyama, Region filling and object removal by Exemplar-based image inpainting, Microsoft research, Cambridge (UK) and Redmond (US)
[2] Neevan R, Anil Kumar A, Image completion: An Exemplar Based Method, Dept. of computer science, College of Engineering Karunagappally
[3] Vaidehi Chhajer, Deepan Gupta, Supervisor: Dr. C. V. Jawahar, A Non-Local MRF Model for Heritage Architectural Image Completion, International Institute of Information Technology – Hyderabad.
[4] Nikos Komodakis, Georgios Tziritas, Image Completion Using Efficient Belief Propagation via Priority Scheduling and Dynamic Pruning, University of Crete, Computer Science Department, P.O. Box 2208, Heraklion, Greece.
[5] M. Bertalmio, G. Saprio, V. Caselles, and C. Ballester, (2000) “Image Inpainting,” Proceedings of the 27th annual conference on Computer graphics and interactive technique, 417424.
[6] M. Bertalm´io, L. A. Vese, G. Sapiro, and S. Osher, “Simultaneous Structure and texture image inpainting.” in CVPR (2), 2003, pp. 707–712.
[7] T. Chan and J. Shen, “Non-texture inpainting by curvature-driven Diffusions,” J. Visual Comm. Image Rep., vol. 12(4), pp. 436–449,2001.