PATCH BASED IMAGE INPAINTING USING TEXTON HISTOGRAMS AS CONTEXTUAL DESCRIPTORS

Abhijeet Chaudhari | Rupesh Dheple | Evans Kurapaty | Aniket Mulmule

[1](Computer Engineering, Savitribai Phule Pune University, Nashik, India, abhijeet.chaudhari03@gmail.com)
[2](Computer Engineering, Savitribai Phule Pune University, Nashik, India, rupeshdheple007@gmail.com)
[3](Computer Engineering, Savitribai Phule Pune University, Nashik, India, evans93_nsk@yahoo.co.in)
[4](Computer Engineering, Savitribai Phule Pune University, Nashik, India, 93aniket@gmail.com)

Abstract—*Image Inpainting is an image processing task of filling in the missing region in an image in a visually plausible way. The current methods of image inpainting consider only a segment of the image to obtain relevant data and fill the target patch. The proposed method uses patch based image inpainting, also known as textural inpainting. Textural inpainting methods fill in the missing region patch-by-patch by searching for well-matching replacement patches (i.e., candidate patches) in the undamaged part of the image and copying them to corresponding locations. Texton histograms, which give similarities between any two patches are referred to for selection of candidate patches. In the process, a huge number of candidate patches will be selected. The other drawback of existing inference methods is their inefficiency in selecting the right patch when the number of candidate patches is huge. In order to overcome this drawback, a novel optimization approach suitable for the inpainting problem with large number of patches has been proposed. This optimized approach is based on Markov Random Field Modelling.*

Keywords—Inpainting, patch-based, texture features, context-aware.

---

1. INTRODUCTION

Now a day's image has become an important media of communication in the world. There are numerous goals and applications of the image inpainting technique from the restoration of damaged images and paintings to the removal or replacement of the selected objects in the image. Digital remastering of images has become an increasingly popular field given its wide range of applications. Digital photography has increased exponentially with the presence of smartphones. At times, knowingly or unknowingly, the digital pictures clicked have unwanted objects present in it. Image inpainting methods provide a way to remove these objects from the picture without causing it much damage. The unwanted is first selected as a patch in the image, this patch is then deleted from the image and then data from the pixels surrounding the patch is used to fill the patch. This is the general modus operandi of image inpainting methods. The existing image inpainting methods are also very slow and computationally expensive to execute.

The proposed method provides greater accuracy in filling the target patch and uses as little computation as possible. Nowadays, various researches are performed in the field of image inpainting due to the varied and important applications of an automatic means of image inpainting. The main applications include undesired object removal such as removing unavoidable objects like birds or aeroplane that appear during filming, to censor an obscene gesture or action that is not deemed appropriate for the target audience, but for which it would be infeasible or expensive to re-shoot the scene. Another main application is the restoration of the images that are damaged by scratches or dust spots or frames that may be corrupted or missing. When the image is transmitted over unreliable networks, there is a possibility of losing significant portions of image frames.

2. LITERATURE SURVEY

In the inpainting based on partial differential equation, the goal was to maintain the angle of arrival. The basic idea was the smooth propagation of information from the surrounding areas in the isophotes direction. The drawback of this method was that the CPU time required for inpainting depends on the size of selected region. Therefore, was a time consuming process as it took nearly 8 to 10 minutes for inpainting. In inpainting by total variance and curvature-driven diffusion methods proposed by Chan and Shen, they used Euler-Lagrange equation and inside the inpainting domain the model simply employs anisotropic diffusion based on the contrast of the isophotes. This was used for inpainting small regions and was also good in removing noise but it did not connect broken lines or edges. The CDD model extended the TV algorithm by taking also geometric information of isophotes while defining strength of the diffusion process, allowing large area inpainting. The major drawback of the TV inpainting model was that it was unable to restore well a single object when its disconnected remaining parts were separated far apart by the inpainting domain. A.Telea proposed a fast marching algorithm that can be looked as PDE based approach without the computational overheads. It is fast and simple to implement than other PDE based methods, this method produced very similar result comparable to other PDE methods. The main limitation of this method is the blurring produced when inpainting regions thicker than 10—15 pixels, especially visible when sharp isophotes intersect the region‘s boundary almost tangentially.

Oliveira(Convolution Operator) proposed a fast image inpainting method based on convolution operator. This algorithm produces the result in few seconds with both the diffusion kernels after more than 100 iterations and faster than any image inpainting algorithms but produce blurring.

3. Proposed system

In this paper, a context-aware Global MRF-based inpainting method is proposed. The main idea is to employ contextual (textural) descriptors to guide and improve the inpainting process. Two most important contributions are (i) a novel context-aware patch selection strategy and (ii) an efficient inference method for global MRF-based inpainting. Our framework is general also in the sense that it allows the use of arbitrary contextual descriptors, e.g., those used in image retrieval, scene classification, etc. In our practical method, we choose to use normalized texton histograms computed from Gabor filter responses as contextual descriptors. Similar texton histograms (computed from responses of different filters) were previously used for image segmentation texture classification and image retrieval. The second main contribution of this paper is specific to MRF-based inpainting. We introduce a novel optimization approach, which builds upon our recent inference method to make it suitable for MRF-based inpainting with huge number of labels. Compared to the related method, this approach is faster and consumes less memory, allowing processing of larger images. Comparative results with other related inpainting methods demonstrate potentials of the proposed method for scratch or text removal and object removal.

The main idea is to guide the search for patches to the areas of interest based on contextual features. Fig. 1 illustrates this concept: contextual descriptors are assigned to image blocks, which can be of fixed size (like in Fig. 1) or adaptive. For the missing region within a given block, well-matching candidate patches will be found in the contextually similar blocks. The benefit is twofold: the search for well-matching patches is accelerated and the inpainting result is improved.

A. Context-Aware Patch Selection

Let the input image I be defined on a lattice S. Pixel positions on this lattice are represented by a single index p ∈ S, assuming raster scan ordering. Let ⊂ S denote the region to be filled (target region), and ⊂ S denote the known part of the image (source region), where ∪ = S. Suppose we divide the image into M × N square non-overlappingblocks, like in Fig. 1 (an extension to adaptive blocks is described in Section II-B). We denote by Bl an image block centred at the position l. The central positions of all the blocks form a set λ, which is determined, together with the block sizes, by the particular block division scheme. The idea of our context-aware approach is to constrain the source region for target patches from a block Bl to a region (l) ⊂ with the context well matching that of Bl. We assign to each block Bl a contextual descriptor c(l), which, in general, is some feature vector that characterizes spatial content and textures within the block.

1: **for all** B_l such that $l \in \lambda$ and $B_l \cap \Omega \neq \emptyset$ **do**
2: set $\Phi^{(l)} = \emptyset$
3: **if** B_l is reliable **then**
4: compute $\bar{H}^{(l,m)}, \forall m \in \lambda$ (Eq. (1))
5: define new source region $\Phi^{(l)}$ (Eqs. (2) and (3))
6: **else**
7: **for all** neighbouring blocks B_n **do**
8: repeat steps 2-5
9: add $\Phi^{(n)}$ to $\Phi^{(l)}$
10: **end for**
11: **end if**
12: **end for**

Figure 1 : Context Aware Patch Selection Algorithm

4. Context-aware MRF –Based Inpainitng

Now we employ the proposed context-aware approach within a novel context-aware MRF-based inpainting algorithm. After constraining the search for candidate patches to the regions of well matching context, the number of labels is still too big and most of the existing inference methods will be inefficient. We propose a novel optimization approach suitable for global inpainting problem with large number of labels.

A. Notations and Definitions

Let the patches be square image blocks of size W × W, where W = 2w + 1.2 We will treat the patches from the source region, which are the candidate patches for the target region, as labels of an MRF. By assuming an MRF model for image inpainting, as proposed in [8], spatial consistency among the candidate patches as well as their agreement with the undamaged image parts is imposed. Let a discrete lattice L consist of points, which are w pixels apart in horizontal or vertical direction on the image lattice S. Let G = (ν,ε) denote an MRF with the set of nodes ν and the set of edges ε. The MRF is imposed over the target region , meaning that ν consists of all lattice L points whose W × W neighbourhood intersects and edges ε consist of all firstorder neighbours *i, j* on the lattice L, where *i, j* ∈ ν denote MRF nodes.

B. Efficient Energy Optimization

We now propose a computationally and memory efficient inference method for the patch-based inpainting problem. The major differences with respect to p-BP [8] are the following. Firstly, instead of having a fixed label position set , we consider a context-aware label position set i ⊂ for each node i. Therefore, by applying contextaware label selection, we limit the number of labels. Secondly, we introduce new formulations of priority scheduling and label pruning, resulting in faster and more memory efficient computation. Finally, we employ a different message passing inference algorithm to obtain the final inpainting result.

5. EXPERIMENTS AND RESULTS

We evaluate the proposed method in applications of scratch and text removal, and image editing (object removal). The reference methods for comparison are chosen from all three categories: “greedy”, multiple candidate and global. For all the analysed methods we show the best inpainting result, by optimizing the patch size (where possible). Furthermore, for our method, if not stated otherwise, we use N f = 18 filters (over 3 scales and 6 orientations), K = 16 textons, block similarity threshold τ = 0.15, L = 10 chosen labels and 10 iterations of the inference algorithm. Threshold for priority, τR, is computed as a median value of SSDs between each pair of patches in the source region, as suggested in p-BP, just that in our case this source region is constrained and it differs from one block to another. For all the results, we used the division into blocks of adaptive sizes obtained by the top-down splitting procedure. This procedure was conducted until the block size reached 1/4 of the image size for images in Section IV-A and 1/8 of the image size for images in Section IV-B, because the former images contain a close-up of the object (see Fig. 6), thus finer division would not be beneficial.

A. Experiments and comparisons for Scratch and Text Removal

For the task of scratch and text removal, we use the dataset of four images from [10] (top row of Fig. 6), where the ground truth is available. The reference methods include the “greedy” approach from [7],4 commercial software Content Aware Fill of Adobe PhotoShop, based on [25] and [29], multiple candidate sparsity-based method (MCS) [10],5 and the global p-BP method [8],6 which is mostly related to ours. Peak signal-to-noise ratio (PSNR) values indicated in Fig. 6 are computed only in the missing region, with pixel values in the range [0,1]. We varied the parameter w that determines the patch size from 2 to 6 and chose the one with the highest PSNR for each method (shown in the brackets). Only Content Aware Fill does not require explicit specification of the patch size.

6. CONCLUSION

Image inpainting is a widely used image editing tool. It greatly enhances the quality and appeal of the image in which it is used. Thus it is fitting to improve the efficieny of this method by making it faster and designing it provide much more accurate results. In the exisiting implementations of image inpainting, patches are selected as blocks of a grid. This leads to more number of patches to compare thereby increasing computational time. More number of comparative patches also requires greater memory to execute and thus such implementations do not work well on lower end computers. In the proposed system, patches are selected on a context basis. Meaning all the pixels in the image which have similar properties form a patch. Thus there are fewer patches to compare while taking into account all the existing properties of the image. The system also makes use of Markov Random Field modeling which tests existing patches for probable target patches by comparing their texton histograms. Texton histograms are used as contextual descriptors for each patch. By making use of the proposed methods, the resultant implementation of image inpainting can execute well with lesser memory and gives accurate results due to detailed comparisons of the contextual descriptors.

Thus the proposed system has lesser hardware requirements and comparatively accurate results. This implementation can even be run on mobile computers with the appropriate modifications making it a portable software.

References

[1] Kaiming He, Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition.

[2] J Puzicha, Histogram Clustering for Unsupervised Image Segmentation.

[3] P Arbelaez, Contour Detection and Hierarchical Image Segmentation.

[4] N Komodakis, Image Completion Using Efficient Belief Propagation Via Priority Scheduling and Dynamic Pruning.

[5] A Antonio, Context Aware Patch Based Image Inpainting Using Markov Random Field.