IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Volume-5, Issue-3 E-ISSN: 2347-2693

A Review on Patch Based Image Restoration or Inpainting

K. Singh[1*], J. Shaveta[2]

[1*] GGS College of Modern Technology, Punjab Technical University, Kharar, India
[2] GGS College of Modern Technology, Punjab Technical University, Kharar, India

*Corresponding Author: *firstauthor_karanveer.kurali@gmail.com, Tel.: +91-7087602074*

Available online at: www.ijcseonline.org

Received:15/Feb/2017 Revised: 22/Feb/2017 Accepted: 18/Mar/2017 Published: 31/Mar/2017

Abstract- Blocking artefacts occurs almost in every compression technology including the most renowned JPEG compression. To minimize the blocking artefact problem, several researches have been done. But adaptively lacks in those algorithms which leads to complex calculation and distortion in the image. In this paper, we have proposed adaptive neighbourhood selection in a way that balances the exactness of approximation. The proposed method is iterative and spontaneously adapts to the degree of underlying smoothness. Our proposed method also restores distorted cracked images along with compressed blocking artefacts.

Keywords : **JPEG, Artefacts , Image, DCT**

I. Introduction

Image in-painting, or image completion, is an image processing task of filling in the missing region in an image in a visually plausible way. Applications include image restoration (e.g., scratch or text removal), image coding and transmission, photo-editing (object removal), virtual restoration of digitized paintings (crack removal), etc. In literature, two categories of image in-painting approaches can be distinguished: diffusion- and patch-based.

Diffusion-based methods fill in the missing region by smoothly propagating image content from the boundary to the interior of the missing region. The problem of propagating linear structures, e.g., object lines and boundaries that are interrupted by the hole, is then often formulated in terms of solving partial differential equations. Although these approaches yield good results when in-painting long thin regions, they experience difficulties in replicating texture, which is largely due to their local nature. Patch-based 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. While these approaches share some ideas with patch-based texture synthesis they focus additionally on structure propagation either by defining the filling order using human intervention or decomposing the image into structure and texture components. Compared to diffusion-based methods, patch-based methods typically produce better results, especially when inpainting larger holes.

II. CONTEXT-AWARE APPROACH FOR INPAINTING

We introduce a general context-aware approach, which can be used with any inpainting algorithm. 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 in-painting 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-over lapping blocks, like in Fig. We denote by Bl an image block centered 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.

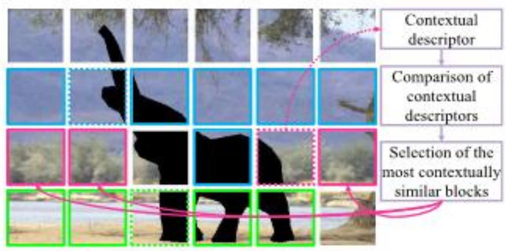

Figure. 1. Context-aware approach for inpainting

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. Let us define a measure of contextual dissimilarity H(l,m) as :

$$\bar{H}^{(l,m)} = d(\mathbf{c}^{(l)}, \mathbf{c}^{(m)}), \quad \textbf{(1)}$$

Where d(c(l),c(m)) is some distance measure between contextual descriptors c(l) and c(m). The more similar context of the blocks Bl and Bm, the lower ¯H(l,m). Let(l) denote the set of positions of the blocks that are contextually similar to Bl. In general, we can write

$$\Sigma^{(l)} = \left\{ m \,|\, \bar{H}^{(l,m)} \le \tau \wedge m \in \lambda \right\},$$

(2)

Where τ is some block similarity threshold. The constrained source region (l) is then a union of known parts of blocks indexed in (l):

$$\Phi^{(l)} = \cup_{m \in \Sigma^{(l)}} (B_m \cap \Phi).$$

(3)

Some examples of block matching for fixed-size blocks are shown in Figs. 1 and 2(a). Current blocks are denoted by a dashed-line border, and their contextually similar blocks by a solid-line border of matching color.

In practice, some blocks may be dominated by missing pixels (e.g., two central blocks in the third row of Fig. 1). We consider a block with less than half known pixels, as unreliable and we do not rely on its contextual descriptor, but we rather determine its constrained source region based on the neighboring blocks. The proposed context-aware approach is summarized as pseudo-code in Algorithm. It applies also to the adaptive blocks, just that the set λ is determined by the adaptive block division scheme.

B. Division into Blocks of Adaptive Sizes

In most natural images some image areas call for finer division than the others (see the example in Fig. 2). Moreover, the optimal size of blocks can differ from one image to another. We introduce a simple top-down splitting procedure that automatically divides the image into blocks of adaptive sizes depending on the “homogeneity” of their texture. We need to favor that the splits in horizontal and vertical directions alternate through levels in order to prevent splitting along one direction only. Therefore, we assign each block a directional flag δ(l) ∈{h,v}. This flag determines the direction, horizontal (h) or vertical (v), along which the evaluation of the block’s homogeneity will have the priority.

$$\bar{H}_d^{(l)} = \bar{H}^{(l_{1d}, l_{2d})}, \quad d = h, v. \quad (4)$$

Let Bl1d and Bl2d denote two sub-blocks of Bl along direction d (see Fig. 3). We measure the inhomogeneity of the block Bl along direction d as the contextual dissimilarity from Eq. (1):

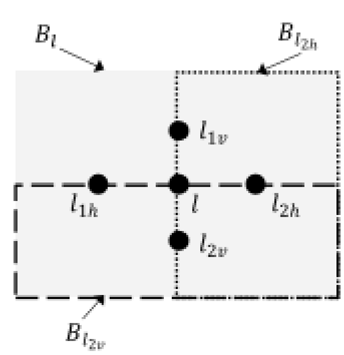

Figure. 2. A division of a block into sub-blocks.

We initialize the block splitting procedure by dividing the image coarsely into four (approximately) equal blocks and assigning them directional flags based on their longer dimension. This figure also shows an example of block matching result for adaptive-size blocks obtained by Algorithm.

C. Discussion on Context Representation

We assign contextual descriptors to image blocks of fixed or adaptive size. An alternative could be to divide the image into regions using some image segmentation technique or user input, as explored. Then, for the missing region within a given segment (region), well-matching candidate patches can be found within that segment. A difficulty with this approach is that the segment boundaries coincide with image structures whose correct propagation inside the missing region is crucial for the quality of the inpainting result. The search for its candidate patches should thus be confined to the union of these two segments (all green and orange areas), which excludes the well-matching patches that can be found on the opposite side of the missing region, because they belong to a different segment. Solution to this problem is to first connect the curves representing the segment boundaries, which leads to better inpainting results.

D. Choice of Contextual Descriptors

We considered a general formulation of contextual descriptors as some characterization of spatial content and textures within blocks. There are many ways to extract texture features, e.g., computing co-occurrence matrices using local binary patterns estimating parameters of MRF models multichannel filtering etc.

For our problem, multi-channel filtering is well suited, both in terms of performance and relatively simple implementation. Let Gn denote one filter from the bank of linear spatial filters at various orientations and scales, where n = 1,...,N f and N f is the total number of filters. After convolving the image I with such bank of filters, each pixel p is assigned an N f -dimensional vector of filter outputs, F(p) =(F1(p),...,FN f (p)), where Fn(p) = (I ∗ Gn)(p). This vector characterizes the image patch centred at that pixel.

Often, dimensionality reduction is applied to the resulting vector. For example, in filter outputs were averaged within square non-overlapping blocks to obtain coarse description of textures called a gist, which was employed for various computer vision tasks and also in our previous inpainting work. We observed, however, that an alternative approach, using the so-called text on histograms similar to those from yields similar or even slightly better results in our setting, while requiring less parameters.

$$c_n^{(l)} = \frac{1}{|B_l \cap \Phi|} \sum_{p \in (B_l \cap \Phi)} \xi[T(p) = n], \quad n = 1, \ldots, K, \tag{5}$$

Where |·| denotes the cardinality of the set and ξ is the indicator function (returning one if its argument is true and zero otherwise). The contextual dissimilarity ¯H(l,m) (Eq. (1)) can now be expressed as any distance between the histograms c(l) and c(m). Here, we employ the common χ2-test:

$$\bar{H}^{(l,m)} = \chi^2(\mathbf{c}^{(l)}, \mathbf{c}^{(m)}) = \frac{1}{2} \sum_{n=1}^{K} \frac{(c_n^{(l)} - c_n^{(m)})}{c_n^{(l)} + c_n^{(m)}} \tag{6}$$

Text on histograms as described above will be used in all the results in this paper. Note, however, that our general framework as described in Sections II-A and II-B can be used with other contextual descriptors as well. Reduction of Ringing Artefacts in Text and Graphics Regions

The overview of our method is shown in Fig.5We first consider grey-level images. The results will be extended to colour images at the end of this subsection. For each textual region, a grey value histogram is first built. Three pieces of information are derived from the histogram, namely, the gray value of the background, a threshold that separates the text and the background, and a Signal-to-Noise Ratio (SNR) level for the region.

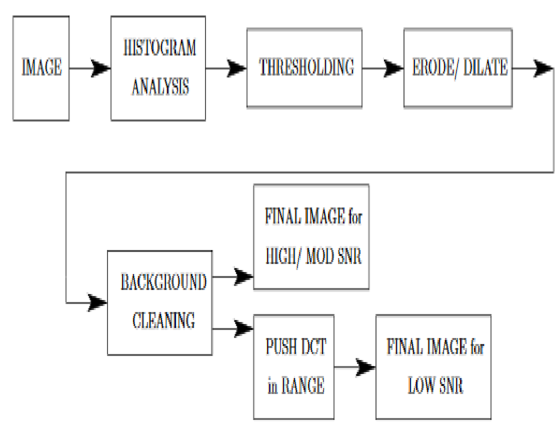

Figure 3.Overview of the method

Since in most text regions, the background pixels are dominant in number, it is easy to determine the background colour of the image by either choosing the most frequent grey level or the weighted average of several frequent grey levels as the background colour of the image region. From the histogram, we also determine a threshold value that can be used as a metric to assign each pixel as a member of the text or background.

III. Reduction of Blocking Artefacts in Pictorial Regions

In order to eliminate the blocking artefacts from a pictorial region, we first tag the pixels in the region as edge/non-edge. This can be accomplished by any standard edge detection algorithms. In addition, we also identify the pixels that lie on the 8×8 tile boundary used in JPEG compression. If a priori knowledge about the tiling boundary is not available, it can be determined by a Maximum A Posteriori (MAP) like estimator. For each non-edge pixel on the tiling boundary, a sigma filter is applied to smooth out the blocking artefacts. The sigma filter is an edge preserving smoothing filter. Its output is an average over the pixels within a small window. In calculation of the average, the pixels whose absolute intensity differences with the current pixel exceed a threshold value are excluded.

IV. Literature Review

Basak Oztan and Amal Mali et al (2015) in this paper a segmentation-based post-processing method to remove compression artefacts from JPEG compressed document images. JPEG compressed images typically exhibit ringing and blocking artefacts, which can be objectionable to the viewer above certain compression levels. The ringing is more dominant around textual regions while the blocking is more visible in natural images.

Tijana Ružic and Aleksandra Pižurica, et al (2015) This approach can be employed to improve the speed and performance of virtually any (patch-based) inpainting method. In this paper, we introduced a novel MRF-based inpainting method that uses context-aware approach to reduce the number of possible labels per MRF node and choose them in such a way that they better fit the surrounding context. Context is represented within blocks of fixed or adaptive sizes using contextual descriptors in the form of normalized texton histograms. Additionally, to divide the image into blocks of adaptive size, a novel top-down splitting procedure was introduced.

Joseph Salmon at el (2013) in this paper, in recent years, over complete dictionaries combined with sparse learning techniques became extremely popular in computer vision. While their usefulness is undeniable, the improvement they provide in specific tasks of computer vision is still poorly understood. The aim of the present work is to demonstrate that for the task of image de-noising; nearly state-of-the-art results can be achieved using orthogonal dictionaries only, provided that they are learned directly from the noisy image.

Emmanuel 2013 In this thesis, we explore alternative approaches to non-locality, with the goals of i) developing universal approaches that can handle local and non-local constraints and ii) leveraging the qualities of both non-locality and sparsity. For the first point, we will see that embedding the patches of an image into a graph-based framework can yield a simple algorithm that can switch from local to non- local di usion, which we will apply to the problem of large area image inpainting.

Sing Bing Kang at el(2014) We first show that, for natural images, PatchGPs can be effectively approximated by minimum hop paths (MHPs) that generally correspond to Euclidean line paths connecting two patch nodes. To construct the de-noising kernel, we further discretize the MHP search directions and use only patches along the search directions. Along each MHP, we apply a weight propagation scheme to robustly and efficiently compute the path distance. To handle noise at multiple scales, we conduct wavelet image decomposition and apply PatchGP scheme at each scale. Comprehensive experiments show that our approach achieves comparable quality as the state-of-the-art methods such as NLM and BM3D but is a few orders of magnitude faster.

V. Problem Formulation

In our base model, similar patches in an image are clustered and a low rank matrix is obtained. Then they have used an algorithm named singular value thresh holding (SVT) algorithm solves the low-rank approximation. Then geodesic distance is used only to weigh the patches in the matrix to find out the difference between different blocking artefacts. Limitation about the method is that it is not adaptive to the changes that are happening in the neighbouring edges or say block of pixels. So this procedure reduces the approximation of accuracy. Except that, base model doesn’t provide mechanism to reduce cracks in images which is also a noise for the image. In our approach, we will try to tackle both of these problems.

Algorithm

1. We describe a hybrid algorithm (adaptive neighboring estimator Algorithm) with a minimal number of standardized parameters.

2. These parameters based on the previous construction of the adaptive window and the corresponding estimator.

3. The key ingredient of the procedure is an increasing sequence of nested square windows, centered at xi, of size |Δi,n| = (2n + 1) × (2n + 1) pixels with n = 1, . . . , N.

4 At the initialization, we naturally choose |Δi,0| = 1 and set the fixed size of √ p × √ p patches and the parameter λα involved in the image patch comparison.

5 In addition, the estimation procedure relies on the preliminary estimation of the noise variance σb 2 robustly estimated from input data.

6 We manually set the number N (typically N = 4) of iterations to bound the numerical complexity which is of the order p × N × |Δ·,N | × |Ω| if an image contains |Ω| pixels.

7 As expected, increasing N allows for additional variance reduction in homogeneous regions.

8 The image become perfect

References

[1] T. Brox, O. Kleinschmidt, and D. Cremers, "*Efficient nonlocal means for denoising of textural patterns*", IEEE Trans. on Imag. Proc., Vol. 17(7), pp. 1083–1092, 2008

[2] J. Grazzini and P. Soille, "*Edge-preserving smoothing using a similarity measure in adaptive geodesic neighbourhoods*", Pattern Recogn., Vol. 42(10), pp. 2306–2316, 2009.

[3] L. I. Rudin, S. Osher, and E. Fatemi, "Non-linear total variation based noise removal algorithms", Physica D: Nonlinear Phenomena, Vol. 60, pp. 259 – 268, 1992.

[4] L. Zhang, W. Dong, D. Zhang, and G. Shi, "*Two-stage image denoising by principal component analysis with local pixel grouping*" Pattern Recogn., Vol. 43(4), pp. 1531–1549, 2010.

[5] Y. Wang, M. Orchard, V. Vaishampayan, and A. Reibman,"*Multiple description coding using pairwise correlating transforms,*" IEEE Transactions on Image Processing ,vol. 10, pp. 351–366, 2001.

[6] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "*Image quality assessment: from error visibility to structural similarity*", IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600–612, 2004.

[7] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," CVPR, vol. 2, pp. 60–65, 2005.

[8] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian,"*Image denoising by sparse 3-d transform-domain collaborative filtering*", IEEE Trans. on Image Processing, vol. 16, no. 8, pp. 2080–2095, Aug. 2007

[9] J. G. Apostolopoulos and N. S. Jayant, "*Post processing for Very Low Bit-Rate Video Compression*", IEEE Transactions on Image Processing. Vol. 8, NO. 8, pp. 1125-1129, (Aug. 2012).

[10] C. Wang, P. Xue, W. Lin, W. Zhang and S. Yu, "*Fast Edge-Preserved Postprocessing for Compressed Images*", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 16, NO. 9, pp. 1142-1147, (Sep. 2006).

[11]. D. G. Sheppard, A. Bilgin, M. S. Nadar, B. R. Hunt and M. W. Marcellin, "*A Vector Quantizer for Image Restoration*", IEEE Transactions on Image Processing, Vol. 7, NO. 1, pp. 119-124, (Jan. 1998).

[12]. R. Nakagaki and A. K. Katsaggelos, "*A VQ-Based Blind Image Restoration Algorithm*", IEEE Transaction on Image Processing. Vol. 12, NO. 9, pp. 1044-1053, (Sep. 2003).

[13]. Y. Liaw, W. Lo and J. Z. Lai, "*Image Restoration of Compressed Image using Classified Vector Quantization*.", Pattern Recognition. Vol. 35, pp. 329-340, 2002.

[14]. W. T. Freeman, E. Pasztor, O. Caemichael, "*Learning Low-level Vision*", International Journal of Computer Vision, Vol. 48, pp. 25-47, 2011.

[15]. J. Sun, N. N Zheng, H. Tao and H. Y. Shum, "*Image Hallucination with Primitive Sketch Priors*", IEEE Computer Society Conference on Computer Vision and Pattern Recognition. (2009).

[16]. L. Ma, Y. Zhang, Y. Lu, F. Wu and D. Zhao, "*Three-Tiered Network Model for Image Hallucination*", Accepted by International Conference on Image Processing, (2008).

[17]. S. Roweis and L. Saul, "*Nonlinear Dimensionality Reduction by Locally Linear Embeddings*", Science. Vol. 290, NO. 5500, pp. 2323-2326, (Dec. 2000).

[18] S. Schulte, V. D. Witte and E. E. Kerre, "*A Fuzzy Noise Reduction Method for Color Images*", IEEE Transactions on Image Processing, Vol. 16, No.5, pp.1425–1436, 2007.

[19]. L. I. Rudin, S. Osher, and E. Fatemi, "*Nonlinear total variation based noise removal algorithm*", Physica D: Nonlinear Phenomena, Vol. 60, pp. 259 – 268, 1992.

