Survey on image defencing and inpainting

Aditi Awasthi, Deepthi Bhat, Medhini Oak, Dr. Kayarvizhy N

Department of Computer Science and Engineering
BMS College of Engineering
Bengaluru, Karnataka, India

***Abstract* – Photographers are often hindered in their attempts at capturing pictures by unsuitable imaging conditions. One such condition is the presence of grills, fences or enclosures between the camera and the scene of interest. Due to growing security concerns, the presence of such barriers cannot be avoided. This gives rise to the need for a tool which can remove the occlusion from the clicked image and replace it with content that blends with the background.**

Such a problem is quite challenging and numerous techniques have been developed to tackle a few subproblems contained within it. The purpose of this paper is to discuss and evaluate these algorithms. After analyzing them and identifying their limitations, we conclude with several promising directions for future research.

Index Terms – inpainting, defencing, quasi-periodic

I. INTRODUCTION

Due to the inclusion of image capturing devices into portable devices like smartphones and digital cameras, the number of pictures generated has increased greatly in recent times. Amateur photographers capture numerous pictures when they find an opportune moment, but they are often hampered by unwanted obstructions like fences, grills, enclosures or reflective surfaces. While these obstructions are most often used as a layer of security, they also have drawbacks when it comes to photography since fences and enclosures cannot be removed from the frame by changing the angle of the camera or the plane of focus. The presence of these obstructions ruins the aesthetic experience of the picture. These structures distract the viewers from the actual focus of the picture, thereby ruining its visual appeal.

This problem can be tackled by using image processing to remove the fence from the photograph after it has been clicked. Even though there exist editing mechanisms available for photographers to remove obstructions from the picture, the removal of fence-like structures without losing some parts of the original picture is particularly tedious as they generally cover the entire picture. Instead of undergoing the harrowing process of manually editing such occlusions, a solution can be offered using image processing algorithms, where the fence is automatically detected and filled without or with minimal manual assistance. By using accurate de-fencing and image inpainting techniques, the occlusions can be effectively removed and filled with content that fits the original image in a natural way, thereby removing the need to compromise with distorted photographs or performing manual editing on them.

II. IMAGE DE-FENCING

There is a long history of research concerning the identification of regular and near-regular patterns in images. A fence can be classified as a texture in an image. The perception of texture has numerous dimensions. If there is a repetition of a texture element at almost regular or quasi-periodic intervals, such textures can be classified as quasi-periodic or ordered and the smallest repetitive element is called a texton or a texel. In contrast if no such repetitive element can be identified, those textures can be classified as random. Thus, a number of different texture representations were introduced from time to time in order to accommodate a variety of textures. Some of those methods are surveyed in this paper.

A. *Fence-like Quasi-periodic Texture Detection in Images*

The method described in [1] aims to automatically detect fences or fence-like objects present in the foreground of an image. An application of this is removal of occluding cages or wire mesh in pictures of animals in enclosures. The process is achieved in three stages: Frequency domain filtering, Multi-resolution Processing and SVM Classification. In frequency domain filtering, the image is considered as a two-dimensional discrete time signal. The quasi periodic signal here is the fence which can be filtered using a band-pass filter in the frequency domain. For Multi-resolution Processing, wavelet transformation is used. It uses a coarser to finer strategy where the fence masks at different levels of wavelet pyramid are combined. The detected fence mask after multi-resolution processing classifies a large number of pixels not picked up on the fence texture. Hence, some samples from the fence mask were picked and the features of those sample pixels were used to train an SVM classifier so that the fence texture can be segmented. The fence detection is complete after this stage. Exemplar based image inpainting technique is then used to fill the fence region in this approach. The proposed method works well for fence texture with different shapes, sizes, colours and orientations. Fence texture detection was successful not only for images having fence in the foreground but also for images having fence in the background.

B. A deep learning approach for video de-fencing

The aim of the proposed method described in [2] is to formulate a sparsity-based optimization framework to fill-in fence pixels in a video by using a convolutional neural network. The neural network consists of five layers, two convolutional and two pooling layers and finally a fully connected output layer. The output maps from the fourth layer are concatenated to form a vector while training and this is given to the next layer as input. The final output layer contains two neurons which corresponds to each class. These are fully connected with the previous layer by weights. A linear sigmoid function is used to modulate the responses of the output layer. This output produces a resultant score for each class. Like in traditional video defencing, it is assumed that the areas of the frame that are covered by one frame might be revealed in the subsequent frames. This motion is estimated using the optical flow and wrapping matrix. Split Bregman Iterative framework is used to obtain the best estimate of the de-fenced image. The proposed algorithm is able to inpaint video-based frames with superior accuracy when compared to traditional video inpainting technique.

C. Video De-fencing

The technique described in [3] involves automatic detection and removal of occlusions from video clips. This method takes advantage of the fact that consecutive frames aligned frame by frame has information of the pixels in the de-fenced video. Therefore, the fence-free video can be obtained by substituting the fence pixels to the pixel information in the frames with an unobstructed view. The algorithm can be divided into two parts: estimation of the term "Probability of Fence" (PoF) and pixel selection. The goal of PoF estimation is to find the confidence term of every fence pixel of every frame. This is done by calculating visual parallax by analyzing and inferring the visual flow and image appearance. After finding the affected pixels, pixel restoration is performed. This is done with an improved image alignment algorithm. This algorithm computes the optical flow by making use of a "robust temporal median filter" (R-TMF) which is able to give the correct pixel information even when the fence pixels dominate the pixel collection. This part of the algorithm determines its complexity since it is the most time consuming. The method proposed gave promising results on a dataset of actual world consumer videos with static scenes. However, this method was unable to properly restore the video when the depth of the fence was the same as the background or with moving objects in the background. Methods described in [4], [5], [6] try to improve on this algorithm.

D. Detection and Restoration of Image from Multi-Color Fence Occlusions

The method described in [7] aims to detect fences having different orientation, shapes, color and texture. The image is first converted from RGB to YCbCr. Histograms are computed for each channel of g(x, y) image. These histograms are analysed to select threshold values based on which the segmentation of the fence is done. The segmented mask is amended by using morphological operations like elimination of false positives and insertion of false negatives. The fence mask obtained after an amendment is used for masking the original image. A hybrid inpainting algorithm is used to restore the occluded area. This algorithm is able to detect the fence in the image which may be of different orientation, texture, shape, multi-color and occluded. It is reliable with an average true positive rate of more than 95% and true negative rate of more than 97% of all the figures presented.

III. Image Inpainting

Inpainting is a process of restorative conservation where missing parts of an image are reconstructed, with the ultimate goal of presenting the image without any undesirable patches. The resultant image should appear as if taken directly using camera. The closer it appears to a realistic image, the better the performance of the inpainting algorithm employed. This technique is used to remove the detected fence from the defencing stage.

A. Exemplar based image inpainting (Criminisi's Method)

The method proposed in [8] was one of the first methods to suggest the use of example patch information to fill missing patches in an image. It separates the image into target region; the region being filled in, and source region; the region with known pixel information. The target region is divided into patches of size 9 by 9 pixels. But the user should change this to a size slightly bigger than the biggest texture element or a texel. The algorithm consists of two parts: deciding the filling order of the patches and filling the patches with appropriate colour and texture. To decide the filling order of the patches, a priority term is assigned to each patch. This priority is given by a "confidence term" and by a "data term". The patch P with the highest priority is selected to be filled in. To fill this selected patch with highest priority, a patch from the source region Q is selected that is lost similar to the given patch. This is done by calculating sum of squared distances between Q and the known pixels in P. After selecting appropriate Q, the pixel information from Q is copied onto P. The confidence terms of the patches are updated and this process is repeated until all the patches are filled. The proposed method outperformed every other existing method at that time and was able to remove large objects from images. Many improvements on the original algorithm have been proposed since such as using an image segmentation algorithm to segment the image based on topography before applying the algorithm used in [9], using a patch shifting scheme for cases where Criminisi's method might fail as used in [10], understanding the depth information to guide appropriate scale transformation as used in [11], using pixel inhomogeneity factor to drive the priority function as used in [12] and many more.

B. Deep Convolutional Generative Adversarial Net (DCGAN) based inpainting

Content aware inpainting algorithms use the neighbouring pixel information to fill in target pixels. However, in some cases there is a need for the algorithm to have some intuitive knowledge as to what to fill as described in [13]. Supervised learning is used to solve that problem. DCGAN is essentially a convolutional neural network that is used to generate new content. It understands the semantics of the entire image and suggests new relevant content to complete the image. DCGANs consists of two parts, the discriminator and the generator. The discriminator comprises of four convolution two dimensional layers activated by Rectified Linear Unit and lastly a fully connected layer. The generator comprises of a linear layer along with a reshape transformation so that the input is in appropriate dimensions followed by four convolution two dimensional transpose layers activated by Rectified Linear Units. The generator creates new content to fill into the target region whereas the discriminator distinguishes between what is real and fake. It is observed that using DCGA*N* produces superior inpainting although it takes longer time to run. However, it does not apply to every scenario since its applicability is limited by the training dataset. [14] uses Convolutional neural networks to inpaint parts of the human body that is obstructed by occlusions.

IV. CONCLUSION

In this survey, we have explored the various methods which can be used to remove occluding elements from images using advanced image processing techniques involving complex mathematics. It is remarkable to see the impact of pure math-based computer applications in such varied real-world problems.

ACKNOWLEDGMENT

The work reported in this paper is supported by the college through the TECHNICAL EDUCATION QUALITY IMPROVEMENT PROGRAMME [TEQIP-III] of the MHRD, Government of India.

REFERENCES

[1] Hettiarachchi, Randima & Peters, James & Bruce, Niel. (2014). Fence-like Quasi-periodic Texture Detection in Images. Theory and Applications of Mathematics & Computer Science. 4.

[2] S. Jonna, K. K. Nakka and R. R. Sahay, "My camera can see through fences: A deep learning approach for image de-fencing," 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), Kuala Lumpur, 2015, pp. 261-265.

[3] Y. Mu, W. Liu, and S. Yan, "Video de-fencing," IEEE Trans. Circts. Sys. Vid. Tech., vol. 24, no. 7, pp. 1111–1121, 2014.

[4] Jonna, Sankaraganesh & Satapathy, Sukla & Sahay, Rajiv. (2016). Stereo image de-fencing using smartphones.

[5] Jonna, Sankaraganesh & Nakka, Krishna & Sahay, Rajiv. (2016). Towards an Automated Image De-fencing Algorithm Using Sparsity.

[6] Xue, Tianfan & Rubinstein, Michael & Liu, Ce & Freeman, William. (2015). A Computational Approach for Obstruction-Free Photography. ACM Transactions on Graphics. 34. 79:1-79:11. 10.1145/2766940.

[7] LUO, Meng-xiao & XU, Wei-sheng & YU, You-ling. (2019). Image De-fencing Based on Binary Morphology. DEStech Transactions on Computer Science and Engineering. 10.12783/dtcse/icaic2019/29429.

[8] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," IEEE Trans. Image Process., vol. 13, no. 9, pp.1200–1212, 2004..

[9] Huang Ying, Li Kai a*, Yang Ming "An Improved Image Inpainting Algorithm based on Image Segmentation" International Congress of Information and Communication Technology 2017 a Institute of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing, China 400065.

[10] S. Tae-o-sot and A. Nishihara, "Exemplar-based image inpainting with patch shifting scheme," 2011 17th International Conference on Digital Signal Processing (DSP), Corfu, 2011, pp. 1-5.

[11] Xiao M, Li G, Xie L, Peng L, Chen Q (2018) Exemplar-based image completion using image depth information. PLOS ONE 13(9): e0200404.

[12] Q. Fan, H. Liu, Z. Fu and X. Li, "Exemplar-based image inpainting based on pixel inhomogeneity factor," 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Kuala Lumpur, 2017, pp. 1164-1168.

[13] Chenduo Huang & Koki Yoshida, (2017). Evaluations of Image Completion Algorithms: Exemplar-Based Inpainting vs.Deep Convolutional GAN. Stanford University.

[14] Xian Wu, Rui-Long Li, Fang-Lue Zhang, Jian-Cheng Liu, Jue Wang, Ariel Shamir, Shi-Min Hu,(2018) "Deep Portrait Image Completion and Extrapolation".