**Literature Survey**

The process of removal of fences in images is mainly divided into two phases: 1) Image de-fencing and 2) image inpainting. First, the fence in the image is detected and corresponding fence pixels are covered with a mask which we take as an input to the next phase. Then, an inpainting method is used to remove the fence to get a non-occluded, de-fenced image.

**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, several different texture representations were introduced from time to time in order to accommodate a variety of textures.

In early work, Liu et al. detected fence structures based on texture regularity by considering the fence as the deformed lattice. The fence is segregated using frequency domain processing prior to the wavelet transformation, according to the fence detection technique proposed by Hettiarachchi et al. Then, it is segmented using support vector machine classification. This detection algorithm works well for semi-regular fence structures whether the fence occlusion is in foreground or background. Yamashita et al. and Li et al. used multi-focus images to detect the fence. These algorithms require two images with and without a flashlight. Parallel de-fencing algorithm is presented in Khalid et al. The algorithm presented in Yang et al. fence is detected by color-based classifier from learned samples obtained from super pixel classification. The lack of fence types, along with simple research objects, caused the poor robustness of these methods, especially when dealing with fences of different shapes and colors.

In recent years, the research performed on fence detection in image processing has built further on the foundations already laid. Y. Wang et al. proposed a detection algorithm of fence completeness based on vertical distribution of horizon edge, but the method does not provide a method for fence extraction. Doubek et al. presented a method for image retrieval using repetitive patterns as the only feature, but the method ignored the lattice edge. The method described by M Varalakshmamma et al. aims to detect fences having different orientation, shapes, color and texture. The technique 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. However, undesirable results are produced if there is high similarity between the fence and background colour. The method described by Lou et al. has presented a method for fence detection based on image binary morphology. The method proved to be fast and efficient, which can achieve the effect of real-time conversion even when running on a poorly configured device.

The technique described by Mu et al. involves automatic detection and removal of occlusions from video clips. This method takes advantage of the fact that consecutive frames aligned frame by frame have information of the pixels in the de-fenced video. 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. The method described by Joanna et al. removes fences from video frames by using either Gabor Filter approach or a trained SVM Classifier for fence detection. By employing machine learning based techniques, it is able to work with dynamic objects as well, however, the camera should be fronto-parallel to the image. The method described by Xue et al. combines the visual information across the image sequence to produce a clean image of the desired background scene with the visual obstruction removed. It is able to remove both reflections and occlusions; however, only occlusions whose planes are parallel to the moving path of the camera can be removed efficiently. S. Jonna et al. also proposed a method that used convolutional neural networks to detect fence pixels, and proposed a semi-automated de-fencing algorithm using a video of the dynamic scene, but the applications of supervised learning algorithms are limited by its data set on which it is trained, in this case only includes some fixed-mode mesh fences.

Most of the existing segmentation techniques use tri-map as input to segment the single object from the color images. The segmentation techniques discussed so far are effective in segmenting coherent objects. But they are not as effective in segmenting the distributed objects that may cover the entire image region. In these cases, initialization of the tri-map is a very time consuming, tedious and error prone task.

**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 a 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.

Texture synthesis based algorithms are one of the earliest methods of image inpainting. They introduce the notion of patching and these algorithms are used to replace the damaged pixel using the similar neighbourhood to complete the missing region. But texture based methods did not provide satisfactory results in terms of visual appearance. Many of the ideas introduced in texture synthesis led to the development of exemplar based inpainting.

The method proposed by Criminisi et al. was one of the first to suggest the use of example patch information to fill missing patches in an image. 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. Huang et al. uses an image segmentation algorithm to segment the image based on topography before applying the Criminisi’s algorithm. The results showed that the method in Huang et al. achieved slightly higher PSNR than Criminisi’s Algorithm. Nishihara et al. uses a patch shifting scheme for cases where Criminisi’s method might fail. The method described by Nishihara et al. not only produces a higher PSNR than Criminisi’s but also has a better visual appearance. Xiao et al. understands the depth information to guide appropriate scale transformation before applying an exemplar based image inpainting technique to complete the image. The method described in Xiao et al. produced higher PSNR than all existing methods. Fan et al. uses pixel inhomogeneity factor to drive the priority function of Criminisi’s algorithm. The proposed method gave comparable results to other improved Criminisi’s algorithms in terms of visual appearance but gave significant improvements on complexity and execution time.

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. Yoshida et al. compares the results of traditional Exemplar based inpainting with Deep Convolutional Generative Adversarial Net (DCGAN) based inpainting. It is observed that using DCGAN 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. When comparing the results of Exemplar Based Inpainting techniques (EBI) with DCGANS, Image Complexity, Image Realism and Image Test Time Cost, were scored on a scale of 0-10. It was observed that EBI scored a 6.0 for complexity whereas DCGAN scored a 7.6, EBI scored a 9.3 for Image Realism whereas DCGAN scored 8.2 and EBI scored a 9.3 for Image Test Time Cost Performance whereas DCGAN scored a 7.2. Wu et al. uses a two-stage deep learning framework to inpaint parts of the human body that is obstructed by occlusions. When trained using the ATR dataset, the method was able to produce an L1 error of 9.3675, PSNR of 27.0852, SSIM of 0.9736 and an FID of 10.6247. But supervised learning based methods are again limited by the training of their dataset.

Exemplar-Based Inpainting is suitable for applications where users require quick responsiveness, for example, mobile applications that automatically “Photoshop” pictures taken. On the other hand, despite its requirement of huge commitment during training process and long runtime during testing, we think that DCGAN based approach is suitable for a different set of scenarios - scenarios that can tolerate time consuming completion process but require high quality and accuracy, such as scene reconstruction for criminal investigation, or human body imaging (tumor detection) for medical purposes, etc.