A literature review of different techniques in the field of image manipulation and restoration

Cathy Irwin, Department of Computer Science, Rhodes University

May 2004

Abstract

Image manipulation and restoration techniques have in the past tended to be effective for either the reproduction of texture or structure. Recent developments in this field have seen the merging of such techniques to create algorithms that are effective for a far greater variety of image completion scenarios. My research project investigates the fragment-based image completion algorithm developed by Drori, Cohen-Or and Yeshuran [2003] which is influenced by many texture synthesis, image inpainting, image manipulation and completion field algorithms, some of which will be outlined in this review. I conclude that the fragment-based image completion method is a representative example of the direction of future work in this field.

1. Introduction

The manipulation and modification of images in an undetectable manner has in the past required skilful and time-consuming artistic endeavour. Applications for such a process include the removal of scratches or defects from photographs or films, the restoration and reconstruction of damaged images, and the removal of objects from images. Filling in the missing regions in a realistic manner is an artistic challenge. Computer technology has made this process significantly easier with many widely available image-editing packages such as PhotoShop or the GIMP. These provide tools such as erasing, blurring and clone brushing. These manual techniques still require skilful application however and bring their own set of problems with them which will be discussed later (Section 2.3). Recent developments in the field of image manipulation and restoration have focussed on automating the image completion process. My research project focusses in particular on the filling in of large, textured regions once objects have been removed. The particular technique this project will attempt to implement is called fragment-based image completion and is based on the work of Drori, Cohen-Or and Yeshuran [2003]. The broad influences on their work which I will be discussing in this review are those of texture synthesis [Heeger & Bergen, 1995; Igehy & Pereira, 1997; Wei & Levoy, 2000; Efros & Freeman, 2001; Brooks & Dodgson, 2002]; image inpainting [Bertalmio, Sapiro, Caselles & Ballester, 2000]; the recent combination of these two methods [Bertalmio, Vese, Sapiro & Osher , 2004]; related types of image manipulation [Hertzmann, Jacobs, Oliver, Curless & Salesin, 2001; Oh, Chen, Dorsey & Durand, 2001; Welsh, Ashikhmin & Mueller, 2002]; and some of the more purely mathematical models of predicting completion fields [Hirani & Totsuka, 1996; Williams & Jacobs,1997; Masnou & Morel,1998; Sharon, Brandt & Basri, 2000]. I will then discuss the features of the fragment-based image completion technique and how it relates to these other methods.

2. Image Manipulation techniques

2.1 Texture Synthesis

A broad definition of texture synthesis is to use an example texture to generate a new texture of potentially unlimited size that is perceived to be the same texture as the original. New textures must be able to tile seamlessly. The usefulness of such a technique in image completion methods is obvious where large missing regions need to be filled in to match the surrounding areas. Within this broad category there are several approaches. These include a technique of incrementally building up new texture based on similar neighbourhoods in the sample texture [Wei & Levoy, 2000]; using a given mask to replace regions with synthesized texture [Igehy &Pereira, 1997]; producing new texture by stitching together blocks of example texture o; producing new textures by means of pyramid-based texture analysis [Heeger & Bergen, 1995]; and creating fundamentally new textures by means of applying editing operations to the original texture [Brooks & Dodgson, 2002]. Some of these methods are more suited to reproducing stochastic (random) textures [Heerger & Bergen, 1995; Igehy & Pereira, 1997] while others work better for more structured, deterministic textures [Efros & Freeman, 2001].

The work done by Heerger and Bergen [1995] using pyramid-based texture matching models is an important influence on many later algorithms [Igehy & Pereira, 1997; Brooks & Dodgeson, 2002]. Their method involves analysing a digitized image to compute various texture parameters. These parameters are then used in the synthesizing phase to generate the new texture so that it resembles the original. An image pyramid is generated which is a set of versions of the original texture image of varying sizes that correspond to different frequency bands [Heerger & Bergen, 1995]. Synthesis is performed by matching the histograms (distributions) of the image pyramid to modify a noise image until it has a similar distribution to the example texture. Igehy and Pereira [1997] extend this algorithm by adding a composition step which uses a mask to turn the noise image into a combination of the original texture and its own synthesized texture. The result is a smoother transition between the original texture and the newly generated texture. This reduces boundary problems where the synthesized texture needs to tile with the original and also avoids obvious repetition artefacts. As with Heerger and Bergens original algorithm, this technique is best suited to areas with stochastic textures and fails in areas of structured texture.

Markov Random Fields (MRF) is another commonly used texture model upon which many other texture synthesis algorithms are based [Wei & Levoy, 2000]. This uses a method of probability sampling to generate new texture and is effective for a wide variety of texture types. However the sampling process is computationally very expensive, so Wei and Levoy [2000] have adapted the process to use deterministic searching instead to greatly speed up the process and still achieve comparable results. The inputs consist of a sample texture and a random noise image of any desired size. Similar to the two previously mentioned algorithms, this one modifies the noise image until it matches the sample texture. This works particularly well on stochastic textures, but not where depth, 3D structure or lighting is important.

An entirely different approach is that employed by Efros and Freeman [2001]. Instead of modifying a noise image to create texture, they do what they call image quilting which is effectively taking small samples of the original texture and stitching them together in such a way as to generate unlimited amounts of new texture. This is particularly effective for generating semi-structured textures, but works well on stochastic textures too. The basic algorithm consists of firstly patching together randomly chosen blocks of the original texture, then introducing some overlap according to some measure of how that block agrees with its neighbours. Finally, looking at the error in the overlap region, a new boundary for the block is determined by calculating a minimum cost path through that error surface.

Closely related to texture synthesis is texture editing. Brooks and Dodgson [2002] have developed a technique of assessing the similarity of pixels within a texture to then apply global operations which affect the colour and brightness of all similar pixels. This warps the original texture to create a fundamentally new one. Response times are improved by using pyramids (similar to Heerger & Bergen, 1995) and neighbourhood matching (similar to Wei & Levoy, 2000) to take a multi-scaled approach.

2.2 Inpainting

Image inpainting is based on the way professional art restorators retouch paintings. Significant work on this technique has been done by Bertalmio et al.[2000]. The only user input required is the selection of the regions to be inpainted. Their algorithm then aims to smoothly propagate the information on the boundary areas of the selection inwards to fill the gap. This is achieved by iteratively prolonging the isophote lines, which are lines with equal grey values in the image, inwards from the boundary while taking into account their angle of arrival and possible curvature. This reproduces the structure of the region but does not take into account the texture of the region at all. Therefore this method is best suited to filling in scratches and small regions of disocclusion (where an unwanted object obscures an area). Application of this algorithm on large, textured areas produces flat, unrealistic results, something which the fragment-based image completion algorithm [Drori et al. 2003] attempts to address using samples from neighbourhood regions to reproduce similar texture inside the gap.

The simultaneous filling in of texture and structure into missing regions has recently been explored by Bertalmio et al. [2004]. They implement a way to combine the techniques of texture synthesis with image inpainting which they claim outperforms any results obtained by applying only one of the reconstruction techniques. The image is first decomposed into the sum of two functions, one capturing the structure and one the texture. The first of these is then reconstructed using image inpainting and the other is reconstructed using texture synthesis. The original image is then reconstructed by adding these two modified images. Theoretically, any automated texture synthesis or inpainting technique could be used in this method. This algorithm is intended for a far greater variety of image reconstruction scenarios, making it more versatile than any of the previously mentioned methods.

2.3 Other types of image manipulation

Closely related to the topic of texture synthesis and transfer are other techniques such as image analogies [Hertzmann, Jacobs, Oliver, Curless & Salesin, 2001;

transferring colour to greyscale images [Welsh, Ashikhmin & Mueller, 2002]; and certain other image based modelling and editing techniqes [Oh, Chen, Dorsey & Durand, 2001]. Common concepts utilized by these techniques and the other ones mentioned previously are their use of pyramids and the methods of getting neighbourhood samples in order to manipulate target pixels.

Transferring colour to grayscale images [Welsh et al. 2002] is an extension of texture transfer since a training colour is used similarly to the way in which a training texture would be used to generate new texture. By matching luminance and texture information in the source colour image and the target grayscale image, corresponding chromatic information is then mapped to the target image. Finding texture matches is handled similarly to the way Efros and Freeman [2001] have described. The method for finding neighbourhood samples is particularly relevant to the fragment-based image completion algorithm [Drori et al. 2003] which also needs to search for appropriately coloured and textured sample from which to generate filler information.

Generating image analogies is another variation of texture manipulation. This involves the application of any one of a variety of complex image filters to modify the style of one image to match that of an example image such as an oil painting, watercolour, embossing or blur to name a few. Hertzmann et al. [2001] propose such a technique based on the recent developments in texture synthesis [Heeger & Bergen, 1995; Wei & Levoy, 2000; Efros & Freeman, 2001]. Firstly, multiscale pyramids of the source and target images are constructed. Then, from the coarsest resolution to the finest, neighbourhood statistics of each pixel are compared to find the best match between the unfiltered source and target image. The features in the corresponding place in the filtered source are then applied to the target, generating the desired effect. This method is similar to the way the fragment-based image completion algorithm [Drori et al. 2003] searches for appropriate source pixels.

Oh et al. [2001] deal particularly with overcoming the limitations of common clone brushing tools used to fill in gaps in an image. Clone brushing is when a piece of source region is copied directly to the destination region. This can however cause texture foreshortening and can produce extraneous artefacts unless the intensity, orientation and depth of the source and target regions match exactly. Oh et al. [p.433, 2001] have developed a way to use the 3D information in an image to overcome these problems by means of a distortion free clone brush and a texture-illuminance decoupling filter.

2.4 Mathematical exploration into completion fields

Other relevant investigations into image completion include those by Sharon, Brandt and Basri [2000] and Williams and Jacobs [1997], who focus on shape and curve completions across gaps; Hirani & Totsuka [1996] who explicitly relate shape completion to image restoration; and Masnou and Morel [1998] who use a novel lines level structure to resolve disocclusion. The latter works well for images which have prominent discontinuities.

Williams and Jacobs [1997] exploit the randomness of stochastic textures by assuming that the probability distribution of possible boundary completion shapes can be modelled by a random walk in a lattice across the image plane. The probability that such a walk will join two boundary fragments is computed as the product of two vector field convolutions. The smoothest resulting curve, called the curve of least energy, is used to complete the image. Similarly, Sharon et al. [2000] detect the curve of least energy but use multiscale procedures to speed up the numerical analysis process.

In image processing terms, the frequency domain captures global features and large textures while the spatial domain captures local continuity and structure. As the previous examples highlight, many techniques are only effective on one of these domains (ie. they reproduce either texture or structure, not both). Hirani and Totsuka [1996] have developed an algorithm which works in both of these domains in order to restore images in a natural way. Based on the theory of projections onto convex sets (POCS) they also use sample regions from the source image to restore the noisy pixels. This method claims to be effective for not only stochastic regions [Williams & Jacobs, 1997], but for areas containing randomly placed prominent lines.

3. Fragment Based Image Completion

The fragment-based image completion algorithm by Drori et al. [2003], which is the focus of investigation in my research project, uses the visible parts of an image as a training set to infer the unknown parts when a portion of the image is removed. Firstly, an inverse matte that contains the entire extracted region is defined by the user. This inverse matte defines a confidence level for each pixel with those pixels that are closer to the known regions having higher confidence values. All the confidence values increase during the completion process. An approximation of the low confidence areas is generated using a simple smoothing process known as fast approximation. This rough region is then augmented with familiar details taken from areas of higher confidence. At each step a target fragment, which consists of a circular neighbourhood around the pixel, is completed by adding more detail from an appropriate source fragment which has higher confidence. The source fragments are selected from the immediate vicinity of the unknown region. As the process continues, the average confidence of the pixels converges to one, completing the image in a manner which takes both texture and some sense of structure into account. Unlike the work of Oh et al. [2001] however, this algorithm is limited by a lack of knowledge about the underlying 3D structure of the image which can lead to inaccurate results for complex structures. The example based image synthesis methods of Hertzmann et al. [2001] and Efros & Freeman [2001] also use training sets within the source image. Here these sets determine the likelihood of a particular context appearing in the missing region. Self-similarities in the image are found through a combination of transformations (translation, scale, rotation and reflection) related to the texture self-similarity algorithm of Brooks and Dodgson [2002].

4. Conclusion

Automated image restoration techniques mean that no particular training or skills are needed to perform complex image manipulations. This has significant advantages for the ordinary computer user. Future work in this field will focus on speedup and improving the efficiency of algorithms as well as further investigation into the combination of techniques to work more effectively for a broader variety of situations. In particular, there is much emphasis on developing algorithms that effectively complete both stochastic and structured regions. The fragment-based image completion algorithm [Drori et al. 2003] is a good representation of this direction of research. There is also considerable potential for the extension of techniques, such as those discussed here, to film and video restoration and manipulation.

Bibliography

[1] Bertalmio, M., Sapiro, G., Caselles, V. & Ballester, C. Image Inpainting in Proceedings of ACM SIGGRAPH (2000): 417-424. ACM Press.

[2] Bertalmio, M., Vese, L., Sapiro, G. & Osher, S. Simultaneous structure and texture image inpainting in IEEE Conference on Computer Vision and Pattern Recognition (2004), to appear.

[3] Brooks, S. & Dodgson, N. Self-similarity based texture editing ACM Transactions on Graphics, 21, 3 (2002): 653-656

[4] Drori, I., Cohen-Or, D & Yeshurun, H. Fragment-Based Image Completion in Proceedings of ACM SIGGRAPH (2003): 303-312. ACM Press.

[5] Efros, A & Freeman, W. Image Quilting for Texture Synthesis and Transfer in Proceedings of ACM SIGGRAPH (2001): 341-346. ACM Press.

[6] Heeger, D. J. & Bergen, J. R. Pyramid-based texture analysis and synthesis in proceedings of ACM SIGGRAPH (1995): 229-238. ACM Press.

[7] Hertzmann, D. J., Jacobs, C. E., Oliver, N. Curless, B. & Salesin, D. H. Image Analogies in proceedings of ACM SIGGRAPH (2001): 327-340. ACM Press.

[8] Hirani, A.N.& Totsuka, T. Combining frequency and spacial domain information for fast interactive image noise removal in proceedings of ACM SIGGRAPH (1996): 269-276. ACM Press.

[9] Igehy, H. & Pereira, L. Image replacement through texture synthesis In IEEE international conference on Image processing, 3(1997): 186-189

[10] Masnou, S. & Morel, J.M. Level- lines based disocclusion in proceedings of 5th IEEE International Conference on Image Processing (1998): 259-263

[11] Oh, B. M., Chen, M., Dorsey, J. & Durand, F. Image based modelling and photo editing in proceedings of ACM SIGGRAPH (2001): 433-442. ACM Press.

[12] Sharon, E., Brandt, A. & Basri, R. Completion Energies and Scale. IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 10. (2000): 1117 -1131.

[13] Wei, L. Y. & Levoy, M. Fast texture synthesis using tree structured vector quantization in proceedings of ACM SIGGRAPH (2000): 479-488. ACM Press.

[14] Welsh, T., Ashikhmin, M. & Mueller, K. Transferring color to greyscale images ACM Transactions on Graphics, 21, 3 (2002): 277-280.

[15] Williams, L. & Jacobs, D. W. Stochastic Completion Fields: A neural model of illusory contour shape and salience Neural Computation 9, 4. (1997): 837-858.