Assignment 4: Seamless Editing

Joe Smith - EE Affiliate

I. INTRODUCTION

This assignment discusses and eventually implements the techniques of seamless cloning as detailed in the research paper [1]. First, a summary of the paper is included with the the key ideas and applications listed. Next, a region in grayscale image is filled using the algorithm set out. Subsequently, the introduction of a guidance field is observed with seamless cloning shown for both importing gradients and mixing multiple gradients. This application is extended to colour images. Finally, a full face swap application is implemented and potential improvements to the framework highlighted.

II. BACKGROUND

The paper looks to implement seamless editing by discussing a guided interpolation framework and showing results for various applications.
The initial work in the paper sets out the problem rigorously using mathematical notation. The first step is to describe the filling of some region Ω in the image by interpolating from the known intensity values at the boundary. The aim is to create a smooth gradient of intensity from one boundary to the other hence minimising the sum of the absolute gradient at all points. The boundary condition is that the pixels at the boundary must take on the known values $f^*|_{\partial\Omega}$

$$\min_f \iint_\Omega |\nabla f|^2 \text{ with } f|_{\partial\Omega} = f^*|_{\partial\Omega} \tag{1}$$

The gradient will be more true under the guidance of some known vector field v. Here the minimisation problem will be adapted.

$$\min_f \iint_\Omega |\nabla f - v|^2 \text{ with } f|_{\partial\Omega} = f^*|_{\partial\Omega} \tag{2}$$

In the seamless cloning section, the effect of the guidance field is shown. This field is simply the change in intensity of pixels (gradient) over the source image. The source image takes the target pixels intensities at the boundary and the other pixel intensities are interpolated by following the original gradient. The paper looks at an extension of this, creating a composite guidance field by *mixing* the gradient of the target with that of the source. At each point in the region Ω, the dominant gradient magnitude is used.
The final section in the paper details numerous other applications of this framework. One aspect is to look at modification of the guidance field (the gradient of image intensities). Retaining only dominant gradient components, i.e. those in which an edge exists between two points, results in texture flattening. Dynamic range of selections may be modified by adjusting the logarithmic guidance field, this builds on the algorithm proposed in Fattal 2002 [2]. Logarithms of luminance are approximately equivalent to perceived brightness where large gradients show drastic luminance changes. Applying the change and solving the Poisson equation for regions of the images allow illumination correction of specific foreground objects as is shown in the paper. Further, local colour correction is implemented in the case of isolating a colour object from a monochrome background. The Poisson equation is solved for some loose region Ω around the object. The monochrome destination receives colour in the region using the gradient data from the source luminance channel. The final application of the framework discussed is that of seamless tiling. A tessellating region can be adjusted by enforcing the boundary conditions at the edges of each tile.
It is evident that this paper has been written for an academic with a background in the field. Both the mathematical formalism and the use of technical jargon such as 'stroke' and 'guidance field' introduce a level of abstraction from a real implementation which makes the algorithms and techniques hard to understand. A supplementary paper referencing some source code or some pseudo code exploring the core ideas would be more useful. Due to the low resolution of images in the paper, it is difficult to assess the quality of the framework. It would be better practise if the paper referenced high resolution web versions of the images. Artefacts are evident in seamless cloning when the gradient is not mixed: this is because the edge around some object still has its initial texture. In mixing gradients, these artefacts are less apparent.

III. GRAYSCALE REGION FILLING WITHOUT GUIDANCE FIELD

The discrete solution of the Poisson equation is given in the paper. It can be seen this is just the differentiation with respect to f_p of the equation that precedes it.

$$|N_p| f_p - \sum_{q \in N_p \cap \Omega} f_q = \sum_{q \in N_p \cap \partial\Omega} f^*_q + \sum_{q \in N_p} v_{pq} \tag{3}$$

The most plausible method to solve this series of simultaneous equations is to adopt a matrix form.

$$Ax = b \tag{4}$$

The column vector x represents the image. Each pixel has an index x_i where $i = M \times N$ (rows by columns). The matrix A calculates the gradient terms. Each row is populated as the simultaneous equation requires with the number of neighbours in the diagonal column and -1 as the coefficient at each neighbour. Else, for a pixel not in the region, the intensity is left untouched with a 1 value at the diagonal of the corresponding row. The image is processed as a 1D column vector. It is known that neighbours below and above will be one cell away and those to the left and right will be distanced by the number of rows M in the image.

$$A = \sum a_{ij} \tag{5}$$

Next to each pixel x_i, there is a neighbour x_q where

$$q = \{i+1, i-1, i+M, i-M\} \tag{6}$$

$\forall x_i$

$$a_{ij} = \begin{cases} |N_p| & \text{if } j = i \quad \text{and } x_i \in \Omega \\ 1 & \text{if } j = i \quad \text{and } x_i \notin \Omega \\ -1 & \text{if } j = q \quad \text{and } x_i \in \Omega \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

As this matrix will be sparsely populated, the algorithm can be implemented speedily by pre-populating a matrix using the *sparse()* command with the row and column coordinates of non-zero elements and the values to take at each point. Additionally, the known boundary conditions will be stored in a column vector b. Each pixel on the boundary of the region Ω located at i will be stored at b_i. The result of this algorithm is to create a continuous interpolation over the region. Over smooth regions, this works well as is seen Fig 1. However, it is clear in the larger object that blurring of the bottom of the building has occurred making the concealment more obvious. In high frequency regions such as the waves of the sea, object concealment is much more obvious using this technique. In Fig 2, a blurry patch appears where the boat once was. This may be improved by adding some guidance information using a gradient field.

IV. SEAMLESS CLONING

A. Importing Gradients

Seamless cloning is achieved using the technique described in the paper. A region in a source image is selected with *roipoly()*. This source image is grown to the size of the target image using black pixel padding and a corresponding region is selected in the target. Using the intensity values around the boundary of the target region, the source image can be inserted by growing across this region through solving the Poisson equation under the presence of a guidance field. This guidance field is the record of the source gradient, i.e. the change in intensity across this region. The results are shown in Fig 3. On close inspection, at the region boundary, some artefacts are seen where the waves of the ocean fail to reach the body of the octopus. This may be improved by using the mixed gradient technique.

B. Mixing Gradients

This technique has been summarised in the background section. The absolute gradient between a pixel and each of its neighbour are compared for the source S and target image T and the dominant term used in the guidance field.

$$v = \sum_i \sum_{q \in N_p} v_{iq} \tag{8}$$

$$v_{iq} = \begin{cases} |T(i) - T(q)| & \text{if } |T(i) - T(q)| > |S(i) - S(q)| \\ |S(i) - S(q)| & \text{if } |T(i) - T(q)| \leq |S(i) - S(q)| \\ 0 & \text{if } T(i) \notin \Omega \end{cases} \tag{9}$$

The result is shown in Fig 4. A much looser selection can be made and the smooth source sea replaced with the edges in the target waves. However, the wave pattern can be seen in regions of smooth octopus skin giving the octopus less definition than the previous method.

V. SEAMLESS CLONING ON COLOUR IMAGES

It is simple to extend the cloning technique to colour images. The same algorithm is performed on a per channel basis. The results of this can be seen in Fig 5. It is observed that the effect of intensity changes to match the target is more apparent with colour images. The octopus has turned bright red. The quality of the implementation can be assessed using the *tic* and *toc* commands. The program reports the full processing, including each colour level, takes 17.043292 seconds.

VI. MULTIPLE STROKES: IMPLEMENTING A FACE SWAP ALGORITHM

The framework can be made more extensible by allowing multiple selections to be made. A popular *Facebook* application takes two faces from a photograph and swaps them. It should be feasible to adapt the Poisson cloning code to do this. The initial hurdle is to allow accurate selection of the two faces. This was implemented using the *ginput()* command to select a pixel location for the second mask and a control sequence with two dialogue boxes to allow the second face to be repositioned or the initial mask to be redrawn entirely. Next the faces had to be repositioned at their new locations. This was achieved using the *circshift()* command to find the location of the first `true` value in each 1D column matrix representation of the binary mask and shift the respective masked image data to this location. Initially it was considered that each selection would have to be solved individually. However, the two masks and guidance fields could be combined in the same matrix for much faster execution. The individual masks are maintained to check which face to populate the guidance field from and the boundary pixels found from the edge of the combined mask. It was noted that a bug was apparent in that a border of black pixels appeared in the guidance field. This was fixed using knowledge of image morphology acquired in the second assignment. The mask was eroded by a 4 pixel square *strel*. This allowed the mask to be slightly smaller and no black pixels on the edge were included in the guidance field as the gradient at all points in the image could now be computed. The results of this seamless composite are shown in Fig 6.

VII. FRAMEWORK IMPROVEMENTS

The positioning of source images should be more intuitive. A placement GUI could be implemented similar to the *roipoly()* mask select. Implementing this algorithm in the C language could allow more control over this. It should be possible to scale these source images. Different resolution images need to be matched to achieve a realistic composite. This stretch and placement tool could be integrated. Additionally, it would be beneficial if multiple source images could be added to the target simultaneously to create a more expressive result. This would require the development of a full application. The framework specified in the paper is more of a back-bone structure. For consumer use, a more complete package would need development.

VIII. CONCLUSIONS

In conclusion, all key sections of the paper [1] have been implemented. This has included region filling without a guidance field, using importing a gradient as a guidance field and mixing two gradients to make a guidance field. This has been extended to an RGB colour image. The paper has been discussed fully prior to this and a good understanding as to the operation of each algorithm shown. As an extension, a full face swapping application has been developed using multiple region filling under a gradient guidance field and good results shown. Further improvements to the framework have been suggested.

[1] P. Pérez, M. Gangnet, and A. Blake, “Poisson image editing,” *ACM SIGGRAPH 2003 Papers on - SIGGRAPH ’03*, p. 313, 2003.

[2] R. Fattal, D. Lischinski, and M. Werman, “Gradient Domain High Dynamic Range Compression,” *ACM Transactions on Graphics (TOG)*, 2002.

FIG. 1: Here shows region filling over a smooth sky region. The top image shows the flag concealed and the bottom shows concealment of the larger big ben tower

FIG. 2: The sea has edges so the attempt concealment is more obvious

FIG. 3: Seamless cloning has been achieved by importing the gradient of the octopus to a region

FIG. 4: Mixing gradients allows a looser region selection but introduces transparency in the octopus

FIG. 5: Importing the gradient on all three colour channels

FIG. 6: Here, two famous characters have their faces swapped