Computer Science Depatrment of The University of Auckland
CITR at Tamaki Campus (http://www.tcs.auckland.ac.nz)

CITR-TR-30 October 1998

Image Stitching - Comparisons and New Techniques

Chia-Yen Chen [1]

Abstract

In this work, we are mainly dealing with the stitching of panoramic images. However, the methods described and evaluated here can also be used for different applications in image mosaicing.

[1] The University of Auckland, Computer Science Department, CITR,
Tamaki Campus (Building 731), Glen Innes, Auckland, New Zealand

1 Introduction

Recently a specialised form of image mosaicing known as *image stitching*, has become increasingly common [Shu97], especially in the making of panoramic images. Stitched images are used in applications such as interactive panoramic viewing of images, architectural walk-through, multi-node movies and other applications associated with modelling the 3D environment using images acquired from the real world.

Panoramic image stitching is the process performed to generate one panoramic image from a series of smaller, overlapping images. Fig. 1 shows the main steps in producing a panoramic image.

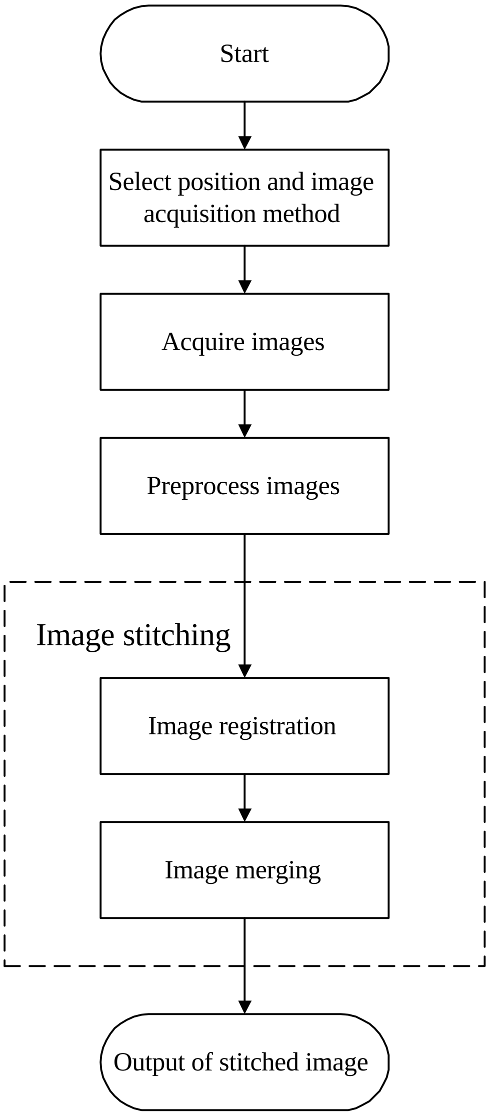


Figure 1: *The flowchart of producing a panoramic image.*

In Fig. 1, the first step in the generation of a panoramic image is to select the position and acquisition of images. In this step, a decision needs to be made on the type of resultant panoramic images. According to the required panoramic images, different image acquisition methods are used to acquire the series of images.

After the images have been acquired, some processing might need to be applied to the images before they can be stitched. For example, the images might need to be projected onto a surface, which can be a mathematical surface such as a cylindrical, spherical, or planar surface. Distortions caused by the camera lenses also need to be corrected before the images are processed further.

In this work, the process of image stitching has been divided into two steps, image registration and image merging. During image registration, portions of adjacent images are compared to find the translations which align the images. Once the overlapping images have been registered, they need to be merged together to form a single panoramic image. The process of image merging is performed to make the transition between adjacent images visually undetectable.

A panoramic image is generated after the images have been stitched. By generating panoramic images with image stitching, the images can be acquired using a relatively inexpensive camera and the angle of view covered by the panoramic image can be determined by the user. The stitched image can also be of higher resolutions than a panoramic image acquired by a panoramic camera.

Our first objective is to provide a detailed understanding of the generation of panoramic images and the steps involved in implementing an image stitcher. To achieve this goal, the three main procedures required in the generation of panoramic images, *i.e.*, image acquisition, image registration and image merging, are discussed. It is also our objective to evaluate existing image stitching methods and methods given in this work such that a quantitative indication of the performances of the various methods can be provided. It is then possible to select methods which are suitable for panoramic image stitching based on the evaluations.

In this work, we are mainly dealing with the stitching of panoramic images. However, the methods described here can also be adapted and applied for different applications of image mosaicing.

2 Image acquisition

Different image acquisition methods can be used to acquire input images that produce different types of panoramic images, depending on the type of panoramic images required and the availability of equipment.

Three set-up's to acquire images for panoramic image generation are described and discussed in this section. In the first set-up, the camera is set upon a tripod and the images are obtained by rotating the camera. The second set-up places the camera on a sliding plate and the images are obtained by shifting the camera on the sliding plate. The third set-up is where the camera is held in a person's hands and the person takes the images by turning around on the same spot, or walking in a direction perpendicular to the camera's view direction.

In all three set-up's, a still image camera has been used to take the images. The camera co-ordinate system is shown in Fig. 2, where the Z-axis points towards the object of interest and the Y-axis passes through the optical axis of the camera.

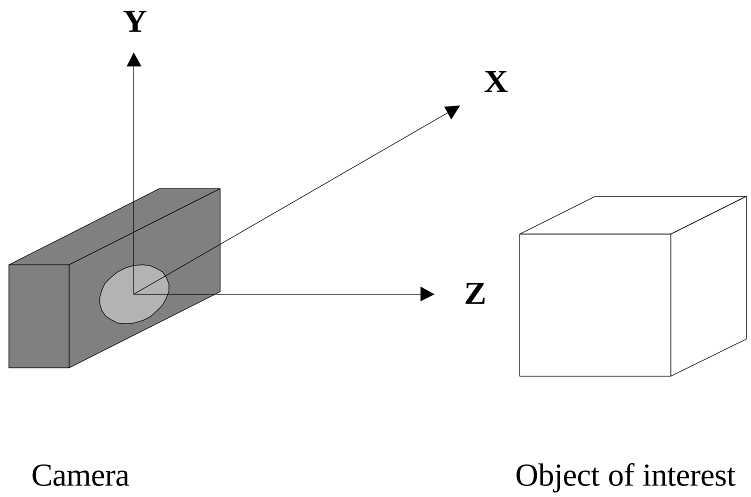


Figure 2: *Camera co-ordinate system.*

The camera's angles of view in the horizontal and vertical directions decide each image's coverage in the horizontal and vertical directions. The angles of view are defined in Fig. 3, where the angles £c and £r respectively represent the camera's angles of view in the horizontal and the vertical directions.

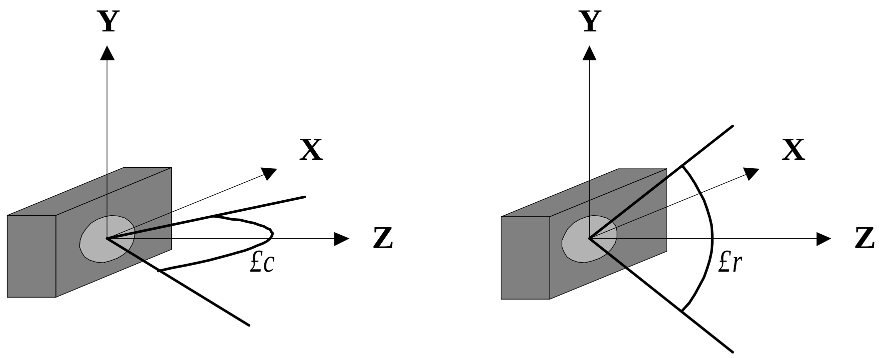


Figure 3: *The horizontal and vertical angles of view.*

2.1 Acquisition by camera rotations

In this acquisition method, the tripod is set levelly at a chosen position and stays in the same position throughout the acquisition of the images. After securing the camera on the tripod, the camera is focused on the objects of interest and rotated with respect to the vertical axis in one chosen direction. One image is taken with each rotation of the camera until the desired range has been covered. Fig. 4 shows the set-up of the tripod and the camera for image acquisition by camera rotations. In an ideal set-up, the Y-axis should pass the optical centre of the camera and there should not be any camera rotations except for the rotation about the Y-axis between successive images.

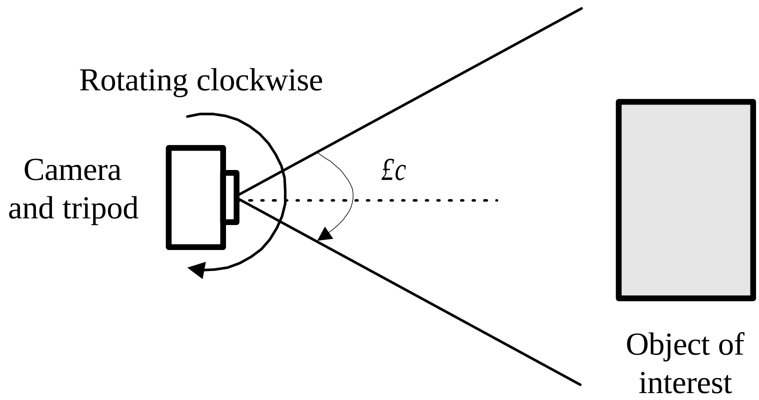


Figure 4: *Camera and tripod for acquisition by camera rotations.*

Each image in the series acquired for panoramic image stitching partially overlaps the previous and the following images. The size of the overlapping region is an important factor in image stitching. As S. E. Chen suggested in 1995, it is desirable to have 50% of the image overlap with the previous image and the other 50% of the image overlap with the following image [Che95]. A larger overlapping region allows adjacent images to be merged more easily in the image merging step of image stitching. Fig. 5 illustrates the geometry of two overlapping images as viewed from above. In Fig. 5, L represents the width of the acquired image and l represents the width of overlapping region between adjacent images. The ratio of l to L is dependent upon the angle of rotation between successive images, £\, and the horizontal angle of view of the camera, £c.

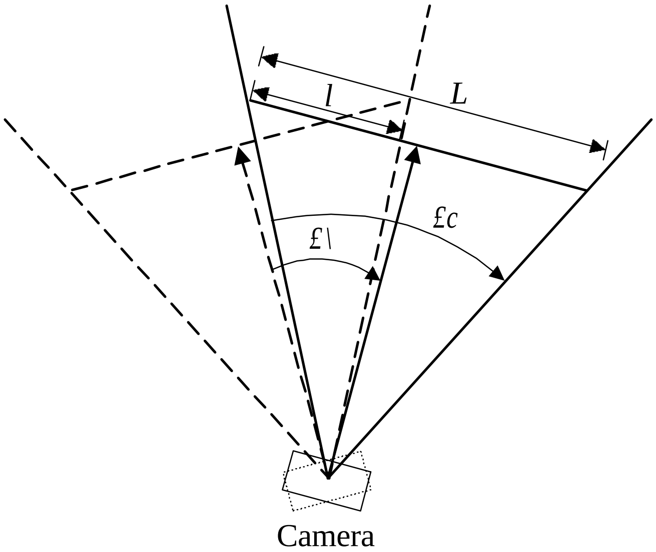


Figure 5: *Geometry of overlapping images.*

The ratio of the width of the overlapping region to the width of the image can be estimated by Eqn. 1,

$$
\frac{l}{L} = \frac{\frac{L}{2}\left[1-\frac{\tan(\alpha-\theta/2)}{\tan(\theta/2)}\right]}{L}
= \frac{1}{2}\left[1-\frac{\tan(\alpha-\theta/2)}{\tan(\theta/2)}\right] . \tag{1}
$$

However, the actual size of the overlapping region might differ from the calculated if the camera is tilted, *i.e.*, there is camera rotation other than the required rotation about the Y-axis. Rotations other than in the specified direction cause problems in the image stitching and affect the quality of the resultant panoramic image. Therefore, in the acquisition of images by camera rotation, it is undesirable to have rotations in directions other than about the specified axis. Another factor which needs to be taken into consideration is the fact that the objects in the real world are projected onto a 2D plane. Hence, by rotating the camera during image acquisition, the distances between points might not be preserved.

Since the camera is rotated between successive images, the orientation of each imaging plane is different in this acquisition method. Therefore, images acquired by this method need to be projected onto the same surface, such as the surface of a cylinder or a sphere, before image registration can be performed.

One advantage with acquiring images by rotation is that the camera can stay in one place during the acquisition of the series of images. Rotation of the camera does not require a large amount of measurement and can be performed easily. However, the acquired images are not on a same surface and need to be projected onto a same surface before being passed on to the image stitcher. As with most projections, the quality of the images degrades after the projection. This problem is more significant when the images are obtained using cameras with shorter focal length, or wider field of view. Nevertheless, due to the simplicity of the set-up and the operations, this method is preferred in acquiring images for the generation of panoramic images.

2.2 Acquisition by camera translations

In this acquisition method, the camera is shifted in a direction parallel to the imaging plane. On a smaller scale, the camera can be mounted onto a sliding plate to achieve the translations. The camera and the sliding plate are placed directly in front of the objects of interest and an image is taken with each translation of the camera until the series of images cover the desired range. Fig. 6 shows the set-up of this method, where the camera is aligned with the sliding plate so that the imaging plane is parallel to the orientation of the sliding plate.

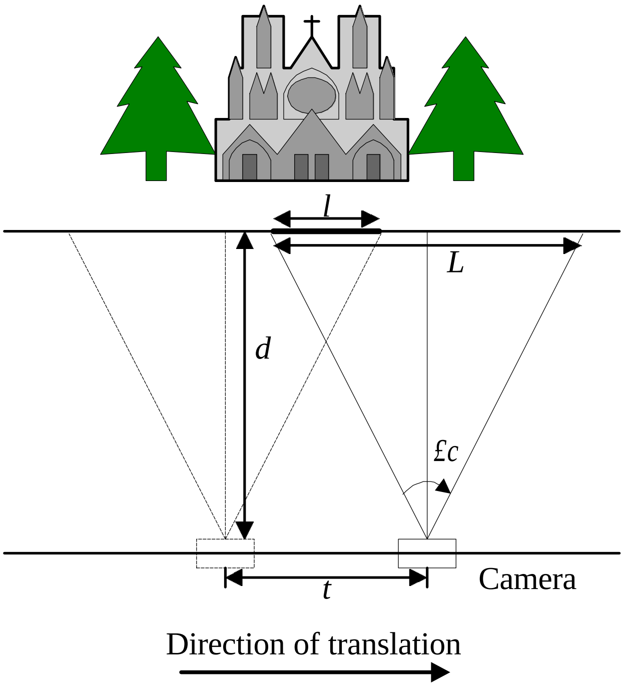


Figure 6: *Geometry for image acquisition by camera translations.*

Given the camera translation, *t*, the distance between camera and object of interest, *d*, and the camera's horizontal angle of view, £c, the ratio of the overlapping region to the whole image can be estimated by Eqn. 2,

$$\frac{l}{L} = 1 - \frac{t}{2d\tan(\theta/2)} \quad . \tag{2}$$

Nevertheless, the actual size of the overlapping region between successive images is determined by the accuracy in setting up the camera. In acquiring images by translation, it is important to ensure that the image planes are parallel to the direction of camera translations. Otherwise, the size of the objects in the images varies as the camera is shifted, causing problems in image stitching.

One disadvantage of this method is that the required translation, *t*, increases as the distance between the camera and the object of interest, *d*, increases, if the acquired images are to have the same sized overlapping region. Therefore, acquiring images when the object of interest is far away from the camera is more difficult due to the magnitude of required translations. Furthermore, since the acquired images are all on the same plane, the panoramic images obtained from images acquired by translation does not provide the feeling of looking into a 3D environment as in the case for the panoramic images obtained from images acquired by rotation.

2.3 Acquisition by a hand held camera

This acquisition method is comparatively easy to perform, the user simply holds the camera and takes images of his/her surroundings by either rotating on the same spot or moving in a predefined direction roughly parallel to the image plane. However, images acquired by this method can be difficult to stitch, due to the amount of unwanted camera rotation or translation during the acquisition of images.

In the case of the user turning with the camera to obtain the images, the user acts as the tripod of the camera in Section 2.1. However, rather than rotating about, or approximately about the vertical axis of the camera, there are inaccuracies in the alignment of the rotating axis with the vertical axis. It is also difficult to control the angles rotated between successive images. Therefore, the sizes of the overlapping regions between adjacent images have a greater variation than for images acquired by a camera mounted on a tripod.

When the user holds up the camera and moves in one direction to acquire the images, the user acts as the sliding plate in acquisition by translation in Section 2.2. However, in this situation, it is even more difficult to control the distance shifted between each image and keep each image on the same imaging plane. Therefore, apart from the difference in the size of the overlapping regions, the image planes of the acquired images have different orientations and cause problems in image stitching.

It is often desirable to have a larger overlapping regions between adjacent images to reduce the effects of the above mentioned problems in acquiring images free-handed. Larger overlapping regions implies that the camera rotations, or translations between successive images are smaller, thus reducing the amount of inconsistencies between images.

Nevertheless, acquiring images by a hand held camera is very easy to manage and can be performed in many locations where it might be difficult to set up equipment such as the tripod or the sliding plate. If care is taken during the acquisition of the images, it is possible to produce panoramic images of similar quality to those generated with images acquired by mounted cameras.

2.4 General problems in image acquisition

One of the most commonly faced problem in image acquisition is the intensity shift between adjacent images. In an ideal case, the same region or object should have the same intensity values in adjacent images. However, due to the variation in the lighting intensity, or the angle between the camera and the light source, the intensity values for the same region or object are different in adjacent images. Other causes for the intensity shift between images include the contrast adjustment done during the development of photographs, as well as during the scanning of the photographs, both of which can be avoided if a digital camera is used to acquire the images in the first place.

Another problem also associated with the lighting condition is the appearance of high-lights on reflective regions, such as on regions of glasses or shining metal. The occurrence of high-light reduces the contrast in the affected region and causes the region to be "blotted out" in the acquired image.

During the time needed to adjust the equipment to the next position after the acquisition of each image, objects within the scene might have moved from their previous positions. Therefore, considerations should be taken when movable objects are to be included in a series of images. Since it will be quite difficult to correctly register the images, once an object in the images has moved from its firstly perceived position and orientation.

The images might also suffer from lens distortions depending on the lens used to acquire the images. Pincushion and barrel distortion are two of the common distortions caused by the lens [Bra95, Ros82]. These two types of distortions can be corrected by using the same lens to take an image of a grid. By using the known parameters of the original grid, it is possible to find a transformation which maps the distorted grid in the acquired image to the original grid. The transformation can then be applied to each of the images taken with the same lens to correct the distortion.

3 Image registration

To form a larger image with a set of overlapping images, it is necessary to find the translations to align the images. The process of image registration aims to find the translations to align two or more overlapping images such that the projection from the view point through any position in the aligned images into the 3D world is unique.

An image registration method usually consists of four main components [Bro92]. They are the *feature set*, *similarity measure*, *search set* and *search strategy*. These four components respectively define what is used to compare the images, how to evaluate the similarity between the images, the range of possible transformations between the images and how to decide the next transformation for evaluation based on the current similarity measure. By varying the contents of these four components, different registration methods with different behaviours can be constructed.

The *feature set* refers to the set of features to be used in the comparison of the images. In our work, we use the term "feature" to mean the characteristics defined by the colour intensity values of the image. The set of features includes the intensity values, contours, textures and so on. A feature set must be selected for each image registration method. The chosen features are extracted from the images and compared in the registration of the images.

The *similarity measure* is a function which returns a scalar value that provides an indication of the similarities between two features. By similarity between the features, we mean the similarity in orientation, size and colours of the features. The intensity values of the selected feature from the images are used to calculate the similarity measures. The values of the similarity measures are used to select transformations for aligning the images.

The *search set* is a set of possible transformations for aligning the images. It contains transformations such as horizontal or vertical translation, rotations, or other more complex transformations obtained by a combination of translations and rotations. The transformations contained in the search set are evaluated by the similarity measures to decide the best transformation required to align the given images.

The *search strategy* is the algorithm that decides how to select the next transformations from the search set.

3.1 Basic method

Over the years, a number of image registration methods have been proposed. These methods usually involve pattern matching to find the transformations required to align the images [Bar72, Bro92, Che97, Gon87, Gam88, Kas90, Rit96, Ros82, Shi87, Son93]. In this section, selected components from previously proposed methods are used to construct a simple image registration method.

The images used in this section are from the Tamaki campus of Auckland university and they are acquired with a rotating camera set on a tripod. The camera has a focal length of 35 mm and is set in the landscape position throughout the acquisition of the images. Since the images are obtained by camera rotation, they have been projected onto a cylindrical surface before image registration. The series of images are registered in a clockwise order, *i.e.*, from left to right, using the components defined above. Fig. 7 shows an example of a pair of adjacent images.

Since the input images are all coloured images, the averaged intensities of the red, green and blue components of the image are used in the feature set. By using the averaged intensities of the three channels, image registration only needs to be performed on one set of intensities, yet each of the three colour channels contributes to the image registration process.

The similarity measure of the image registration method is defined next. The sum of the absolute differences of the averaged intensities has been selected to be used as the similarity measure in this image registration method.

Next, the type of transformations in the search set need to be defined. According to the image acquisition methods used, most of the transformations between overlapping images are translations. It has been assumed that there is negligible rotations and translations in unspecified directions during the acquisition of the images. Therefore, the search set has been defined so that it contains only the translations in the horizontal and vertical directions.

The search strategy we have decided to use in this image registration method is an exhaustive search strategy which calculates the similarity measure for each transformation in the search set. In this way, the optimal similarity measure is guaranteed to be the globally optimal measure.

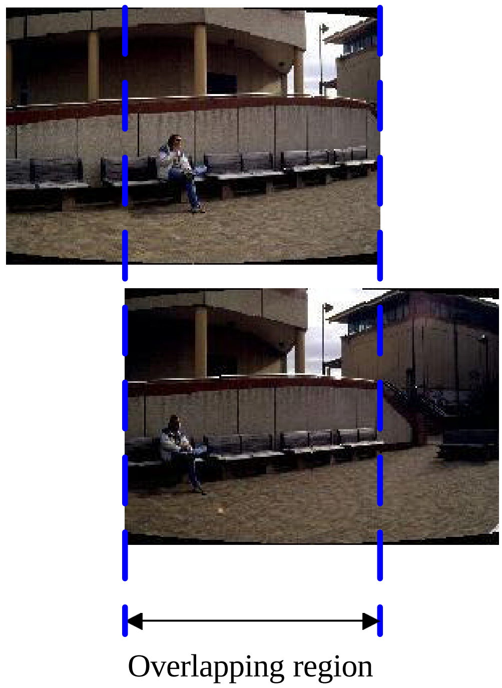


Figure 7: *Example of input images.*

In the acquisition of this series of images, the angle of rotation between successive images, £\, is 15. The horizontal angle of view of the camera, £c, is 54.4. Given these two values, the ratio of the overlapping area of this series of images is estimated to be 70% of the whole image, according to Eqn. 1. From the estimated ratio of overlapping region, a window on the right hand side of the left image is defined to be 50% of the width and height of the input images, to ensure that the window is within the overlapping region of the input images.

Another point which needs to be mentioned is that the centre of the image tends to contain more information in general, *i.e.*, the centre regions are less likely to be of uniform intensity values. Hence by placing the window towards the centre of the image, the similarity measures calculated from the contents in the windows provides a more reliable indication to the actual similarity of the windows. Registering the images on the left hand side first also means that the overlapping region between I_k and I_{k+1} is located towards the right hand side of I_k and the left hand side of I_{k+1}. Therefore, in this work, W_k has been defined on the right hand side of I_k and centred vertically in the middle of the image. Keeping the components and window properties in mind, we now describe the image registration method in more detail.

Let I_k be the image obtained from averaging the intensity of the red, green and blue channels of the kth image in the sequence of input images, where k can be from 1 to the total number of images in the series of images. Let W_k be the $m \times n$ window defined on I_k, with the top left hand corner at position (a,b) of I_k, as shown in Fig. 8.

The image on the right hand side of the kth image, *i.e.*, image I_{k+1}, is transformed by a selected transformation from the search set. A window, W_{k+1}, of the same size and shape as W_k, is defined on I_k. The position of W_{k+1} on I_{k+1} is obtained by applying the inverse of the selected transformation to (a,b).

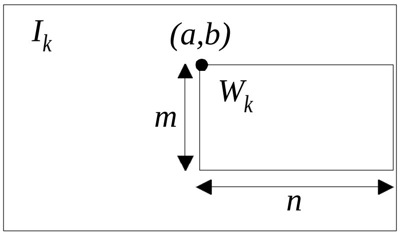


Figure 8: *W_k is the m by n window at position (a,b) in I_k.*

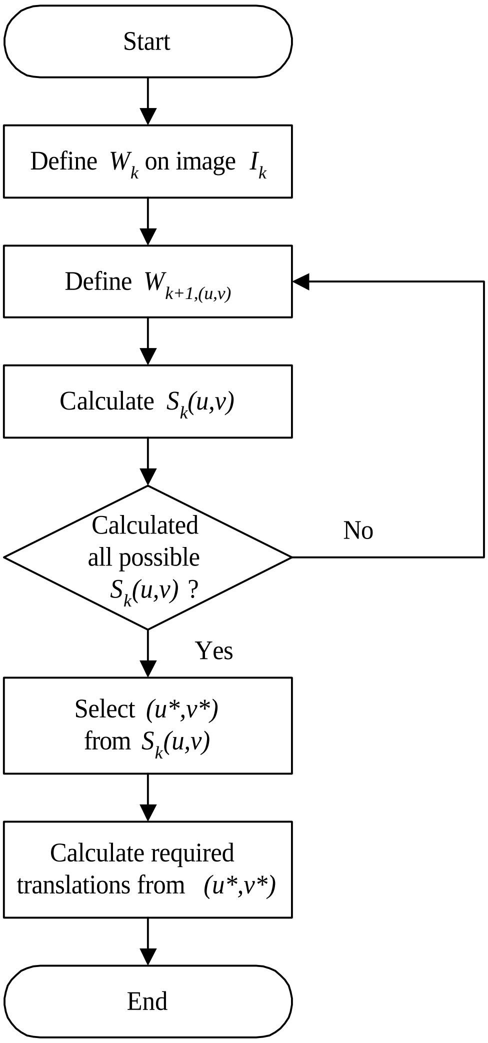


Figure 9: *Flowchart of steps involved in image registration.*

In this method, the similarity measure between the two windows is the sum of the absolute differences of the two windows. Therefore, the similarity measure for position (u,v), is calculated by Eqn. 3,

$$S_k(u,v) = \sum_{i=1}^{m}\sum_{j=1}^{n} |W_k(i,j) - W_{k+1,(u,v)}(i,j)| \quad . \tag{3}$$

Once the similarity measures for all of the possible positions have been calculated, the optimal matching position, denoted by *(u*,v*)*, is chosen by examine the magnitudes of values in S_k. In this method it is intuitive to choose the position which has the minimum similarity measure as the optimal matching position. The reason is that the sum of the absolute differences of two windows has been used as the similarity measure and a smaller difference usually implies a greater degree of similarity between the windows. Therefore, a position *(u*,v*)* is selected, such that the value of the similarity measure at that position is the minimum in all of the calculated similarity measures. That is, at the optimal position *(u*,v*)*, the value of $S_k(u^*,v^*)$, is minimal among $\{S_k(i,j)\}$, as shown in Eqn. 4,

$$S_k(u^*,v^*) = \min_{1 \le i \le H-m,\ 1 \le j \le L-n} \{S_k(i,j)\} \quad . \tag{4}$$

The transformation required to align I_k and I_{k+1} can then be obtained from *(u*,v*)*. The steps involved in this image registration method can be represented by the flowchart in Fig. 9. The images registered by this method are shown in Fig. 10.

Figure 10: *Registered images.*

Figure 11: *Registered image.*

However, at times, this method may return the wrong translations due to various reasons. One of these reasons is that the position where the similarity measure for the windows is globally minimal, has been assumed to provide the translations for the optimal alignment of the images. In the case where the intensities between adjacent images differ significantly, the absolute differences of the averaged intensities may not be a good indication of the similarity of the images. Fig. 11 shows how two adjacent images with visible intensity differences between them are registered by this method. From the images, it is obvious that the image has not been correctly registered in the vertical direction.

3.2 Registration using a different feature set

One approach to avoid the misalignment caused by the differences in the intensity between adjacent images is to choose a feature which is not as dependent upon the intensity as the averaged intensity used in the basic method. A feature that has such property is the set of edges in an image. The set of edges is also a good feature to be used in registering images, since a lot of the information, such as the size, shape and orientation of objects, are conveyed by the edges. The intensity shift between adjacent images also does not affect the edge images significantly.

In this image registration method, the binary edge image of the image is used as the feature set to avoid the misalignment. The similarity measure, search set and the search strategy remain as before so that the effect of varying the feature set can be observed.

To generate the binary edge map, a suitable edge operator needs to be selected [Kle96]. Since our aim is to reduce the effect of the differences in the intensities of the images, a simple edge operator that provides reasonable edge images for the images suffices. The edge image generated by the edge operator is converted into a binary image so that only the significant edges remain in the feature set.

After comparing the binary edge images generated by the Sobel, Prewitt and Kirsch operators, it has been found that the Sobel operators are able to provide the binary edge images more suited to our purpose. Fig. 12 shows the images registered using binary edge images generated by the Sobel operators as the feature set. By comparing Figs. 11 and 12, it can be seen that the correct translations required to align the images have been obtained by using the binary edge image in the feature set. Therefore, using the binary edge image in the feature set can improve the accuracy of image registration by reducing the adverse effect of the intensity differences in adjacent images.

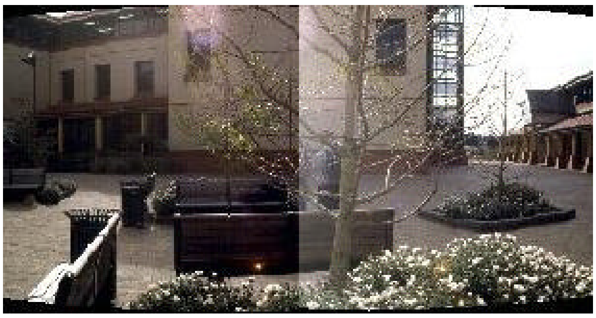


Figure 12: *Registered images.*

3.3 Registration using different similarity measures

Three similarity measures are investigated in this section, the squared differences, correlation product and the standard deviation of the differences of the intensity values in the defined windows. According to the similarity measures used, different criteria are used to determine the optimal matching position.

The first similarity measure is the sum of the square differences of the intensity values in the windows, W_k and W_{k+1} [Shi87]. It has been assumed that lower similarity measures indicate the windows are more alike. According to this assumption, the optimal translations required to align the images are indicated by the position where the similarity measure is the global minimum. Fig. 13 shows the image registered by choosing the position with the minimum value of similarity measures.

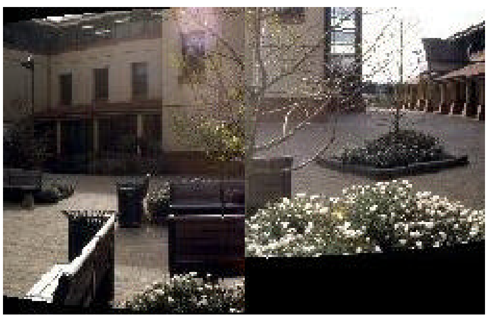


Figure 13: *Image registered using sum of squared differences.*

The second similarity measure is the sum of the correlation product of W_k and W_{k+1}. With this similarity measure, it is assumed that a higher similarity measure indicates a better match. The image registered using the product of the windows as the similarity measure is shown in Fig. 14.

Figure 14: *Image registered using sum of product.*

The third similarity measure to be investigated is the standard deviation of the intensity differences between the windows. This similarity measure is propose to deal with the situation where there is an intensity shift between adjacent images. It has been thought that the intensity shift between adjacent images can result in large differences for the windows at the optimal matching position, thus causing incorrect translations to be returned. However, in the same situation, the standard deviation of the intensity differences for matching windows may be quite small, since the intensity value differs with the corresponding position in W_{k+1} by approximately the same amount in each position of W_k. Therefore, by using the standard deviation of the intensity values as the similarity measure, the image registration method can be made more tolerant to the intensity shifts between adjacent images. In this method, we assume that if two windows are similar, then the intensity differences of the windows should be uniform, resulting in smaller standard deviation. Fig. 15 shows the images registered this using the standard deviation of the intensity differences as the similarity measure.

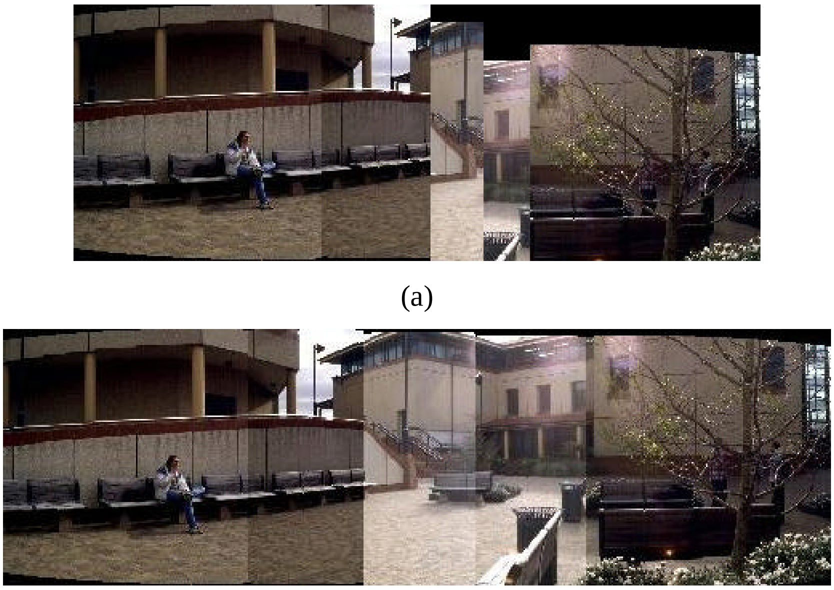
Figure 15: *Images registered using standard deviation of differences.*

Despite the intensity differences between the images, the images have been well aligned, as seen in Fig. 15. Therefore, using the standard deviation of the intensity differences as the similarity measure is one possibility of avoiding misalignment caused by the intensity differences between the images.

3.4 Registration by restricting the search set

For images acquired to generate a panoramic image, the rotations or translations of the camera between the acquisitions of successive image should be quite similar, if not constant. Hence the transformations required to align each pair of these images should also be quite similar, making it possible to estimate the next transformation from the previously obtained transformations.

Let (u_{k-1}^*,v_{k-1}^*) denote the position with the optimal similarity measure for images I_{k-1} and I_k, where images I_{k-1} and I_k, are both from a series of images acquired for the generation of a panoramic image. According to the assumption about the acquisition of images, the position with the optimal similarity measure for images I_k and I_{k+1} is within a neighbourhood of the position with the optimal similarity measure for images I_{k-1} and I_k. Therefore, we can use the previously obtained matching position to define a region, within which the position with the optimal similarity measure for images I_k and I_{k+1} is contained.



(a)

(b)

Figure 16: *Image registered by (a) initial method and (b) restricting the search set.*

From Fig. 16, it can be seen that by restricting the search set during image registration, more accurate alignment can be obtained. Therefore, for image registration in the generation of panoramic images, it is advantageous to improve the accuracy of image registration by restricting the search set.

3.5 Registration with step search strategy

A few step search algorithms have been proposed for estimations of motion field in MPEG images [Tzi94]. The advantage of using a step search algorithm is that the required number of calculations for similarity measures can be reduced dramatically. In this section, we look at an image registration method which uses a 2D binary search to locate the position with the optimal similarity measure, (u_k^*,v_k^*).

In this search strategy, only five positions are evaluated for their respective similarity measures in each iteration of the search. The five positions include a centre point and four points respectively in the north, east, south and west of the centre point. Initially, the centre point, (u_0,v_0), is centred in the region being searched. The distances between the centre point and the four other points are the step sizes. Initially, the vertical step size, $s_{u,0}$, is half the distance between the centre point and the top border, and the horizontal step size, $s_{v,0}$, is half the distance between the centre point and the left border.

The search starts by evaluating the similarity measure for the window located at each of the five initial positions, *i.e.*, (u_0,v_0), $(u_0-s_{u,0},v_0)$, $(u_0,v_0+s_{v,0})$, $(u_0+s_{u,0},v_0)$ and $(u_0,v_0-s_{v,0})$. After obtaining five similarity measures, $S_k(u_0,v_0)$, $S_k(u_0-s_{u,0},v_0)$, $S_k(u_0,v_0+s_{v,0})$, $S_k(u_0+s_{u,0},v_0)$ and $S_k(u_0,v_0-s_{v,0})$, the position with the optimal similarity measure (u_1,v_1) is chosen as the centre point for the next search iteration. The similarity measure is optimised by minimising the sum of the absolute differences between W_k and W_{k+1}. After determining the next centre position,(u_1,v_1), the vertical and horizontal steps, $s_{u,0}$ and $s_{v,0}$, are halved to give the next steps, $s_{u,1}$ and $s_{v,1}$,

When both step sizes reach 1, the similarity measures of eight positions around the centre point are evaluated to determine the position with the optimal similarity measure. Since the step size is reduced by half in each iteration, the search algorithm converges to a solution very quickly. The image registered using the step search algorithm is shown in Fig. 17.

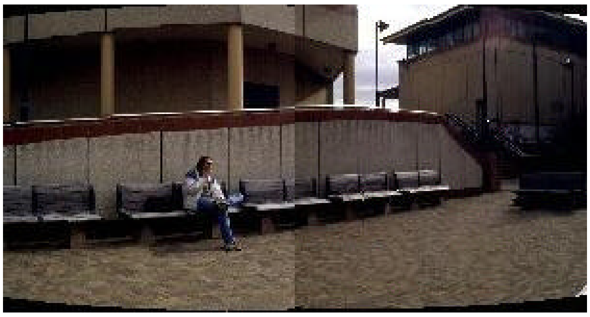


Figure 17: *Image registered using step search.*

In Fig. 17, it can be seen that the images have been well registered by this registration method. Nevertheless, at times, using a step search strategy may miss out the position of the globally optimal similarity measure and return a position with locally optimal similarity measure due to the fact that not all of the positions within the search region are evaluated. Therefore, even though the use of step search algorithms significantly reduces the required number of calculations, it may not be practical to use step search algorithms in the process of image registration. However, it might be worthwhile to investigate the use of step search algorithms in providing an estimation of the optimal position.

3.6 Combination of methods

In this section, we briefly describe a few image registration methods that are constructed by varying the contents of two or more components in the initial image registration method. This is done so that the constructed image registration methods may have the combined benefits of the selected components.

3.6.1 Registration with binary edge image and restricted search set

This image registration method uses the binary edge image as the feature set and a restricted neighbourhood as the search set. The similarity measure is the absolute differences of the pixel values in the binary edge images and the search strategy is an exhaustive search over the defined search set.

With this combination, the chance of misalignment can be reduced by limiting the search to a defined neighbourhood, within which the optimal overlapping position is guaranteed to occur. By using the binary edge image as the feature set, we can increase the image registration method's tolerance to the intensity differences between the images to be registered.

3.6.2 Registration with restricted search set and different similarity measures

The registration methods discussed in this section uses the averaged intensity values as the feature set, but with different similarity measures. The search set is restricted and the search strategy is the exhaustive search over all possible translations.

The similarity measures in this section are calculated from the squared intensity differences, the correlation product of the intensities, and the standard deviation of the intensity differences. By keeping the other components the same, it is possible to observe the effect of varying the similarity measure while limiting the search to a defined neighbourhood.

3.6.3 Registration with restricted search set and step search

To improve the accuracy of the image registration method that uses a step search as the search strategy, we restrict the search set for the method. The feature set is the averaged intensity or the binary edge image. The similarity measure is either the absolute differences of the intensities, or the standard deviation of the intensity differences.

Recall that the a possible cause of error in using a step search algorithm is that the local optima can make the algorithm converge towards the wrong solution. Therefore, a restricted search set is used in this section, to lower the chance of the algorithm converging towards a local optimum away from the position with the globally optimal similarity measure.

To further increase the chance of returning the correct translations for aligning images, we need reliable similarity measures. From the previous sections, it has been found that the binary edge image, combined with the absolute differences of the pixel values, provides a good indication of the similarity between the features being compared. Therefore, in one image registration method, we use the binary edge image as the feature set, the sum of absolute differences as the similarity measure, with a restricted search set and the binary step search algorithm. It has also been found that when the averaged intensity is used in the feature set, the standard deviation of the intensity differences between the compared features is a good similarity measurement. Therefore, an image registration method using the averaged intensity as the feature set, the standard deviation of the intensity differences as the similarity measure, a restricted search set, and the binary step search strategy is also constructed.

3.7 Comparison of methods

In this section, the images registered by every image registration method mentioned in this section are used to evaluate the performance of each method quantitatively.

Each pair of images in the input series of images is manually aligned so that the images appear to be best aligned according to human perception. The translations thus obtained are used as the best possible translations for aligning pairs of images and act as the standard for evaluating the performances of different image registration methods.

Table 1 shows the manually acquired translations for the images and Fig. 18 shows the series of images after they have been joined according to the translations shown in Table 1.

Image no.	1-2	2-3	3-4	4-5	5-6	6-7	7-8
Vert.trans.	-5	3	1	-1	-1	2	2
Horiz. trans.	156	154	160	158	167	159	163

Table 1: *Manually obtained translations.*

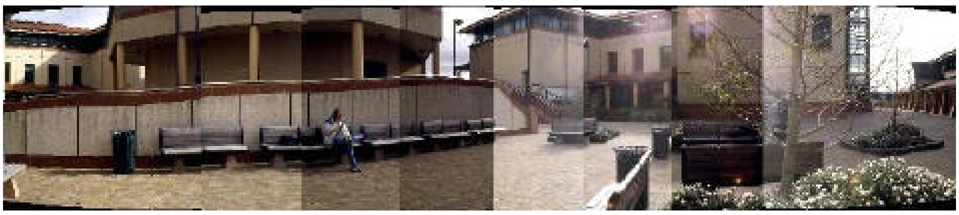


Figure 18: *Manually registered images.*

To simplify the reference to different image registration methods, Table 2 provides the numberings for all of the image registration methods mentioned in Section 3. The methods are referenced to by their numbers hereafter.

Method	Description
1	Basic image registration method (Section 3.1)
2	Uses binary edge image as feature space (Section 3.2)
3	Uses squared intensity difference as similarity measure (Section 3.3)
4	Uses correlation as similarity measure (Section 3.3)
5	Uses standard deviation of intensity differences as similarity measure (Section 3.3)
6	Uses restricted search space (Section 3.4)
7	Uses a step search algorithm as search strategy (Section 3.5)
8	Uses restricted search space and binary edge image as feature space (Section 3.6.1)
9	Uses restricted search space and squared intensity differences as similarity measure (Section 3.6.2)
10	Uses restricted search space and correlation as similarity measure (Section 3.6.2)
11	Uses restricted search space and standard deviation of intensity differences as similarity measure (Section 3.6.2)
12	Uses restricted search space, binary edge image as feature space and a step search algorithm as search strategy (Section 3.6.3)
13	Uses restricted search space, standard deviation of intensity differences as similarity measure and a step search algorithm as search strategy (Section 3.6.3)

Table 2: *Numberings for image registration methods.*

Method	Sum of squared vertical differences	Sum of squared horizontal differences	Sum of squared differences	Performance ranking
1	3664	6696	10360	12
2	44	240	284	2
3	1552	5440	6992	11
4	13892	12196	26088	13
5	26	282	308	4
6	23	700	723	5
7	2870	2779	5649	10
8	10	240	250	1
9	151	708	859	6
10	499	1768	2267	8
11	17	282	299	3
12	77	1925	2002	7
13	166	2142	2308	9

Table 3: *Summary of the differences and ranking of the methods.*

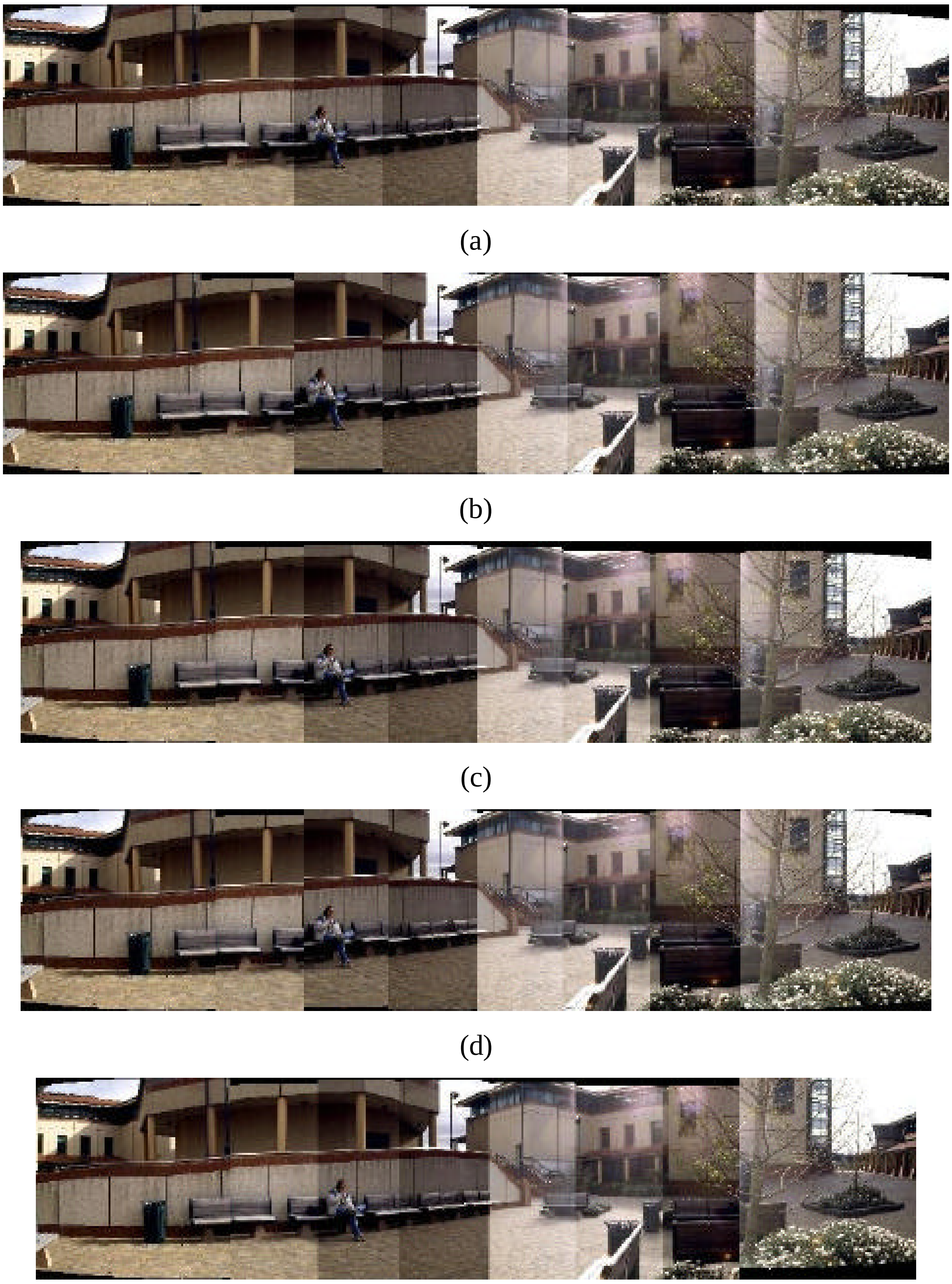
3.7.1 Comparisons

Table 3 shows the sum of the squared differences for translations obtained manually and translations obtained by each method. The methods are ranked according to the magnitudes of the differences. The ranking of 1 is given to the method which returned the translations closest to the best possible translations.

From Table 3, we see that method numbers 8, 2, 11, 5 and 6 are the top five image registration methods. The descriptions of these methods can be found with reference to Table 2.

The images registered by the top five methods are shown in Fig. 19 so that we can compare the visual appearances of the registered images.

According to the evaluation, it has been found that methods 8, 2, 11, 5 and 6 have the best performance out of the image registration methods mentioned in this chapter. The methods have been evaluated based on the differences between the translations obtained by each method and the translations obtained manually.



(a)

(b)

(c)

(d)

(e)

Figure 19: *Images registered by the top five registration methods, (a) method 8, (b) method 2, (c) method 11, (d) method 5 and (e) method 6.*

4 Image merging

Image merging is the process of adjusting the values of pixels in two registered images, such that when the images are joined, the transition from one image to the next is invisible. At the same time, the merged images should preserve the quality of the input images as much as possible.

In an ideal case, the overlapping region of adjacent images should be identical, so that the intensity values of I_k are equal to intensity values of the corresponding position in I_{k+1} for any point *(a,b)* within the overlapping region. However, due to various reasons, including the lighting condition, the geometry of the camera set-up and other reasons mentioned in Section 2, the overlapping regions of adjacent images are almost never the same. Therefore, removing part of the overlapping regions in adjacent images and concatenating the trimmed images often produce images with distinctive *seams*. A seam is the artificial edge produced by the intensity differences of pixels immediately next to where the images are joined.

One approach to remove the seam is to perform the intensity adjustment locally, within a defined neighbourhood of the seam, so that only the intensity values in the neighbourhood are affected by the adjustment [Mil75]. Another approach is to perform a global intensity adjustment on the images to be merged, so that apart from the intensity values within the overlapping regions, intensity values outside the overlapping regions may also need to be adjusted [Mil77]. In this section, image merging by local intensity adjustments are investigated.

One of our objectives is to merge the images so that the seam between images is visually undetectable. The second objective is to preserve the quality of the original images as much as possible. So that the merged image is not seriously degraded by the intensity adjustment required to remove the seam.

Four image merging methods are described and discussed with respect to their behaviours. The images merged by each of these methods are shown and used to evaluate the performances of the methods. The image merging method that has the best performance according to the two objectives is used in our image stitcher for the generations of panoramic images.

4.1 Linear distribution of intensity differences

An image merging method which uses a linear ramp to spread the intensity differences of the pixels which are immediately next to the seam has been proposed by D. L. Milgram for blending pairs of grey level satellite images. This merging method has been adapted to merge colour images in this section.

Let the region in I_k which appears in the final image be bounded by two parameters, l_k and r_k, representing the leftmost and the right most column of I_k which are in the final image. Fig. 20 shows the positions of l_k and r_k.

To merge the images, we must determine the position of the seam. The position of the seam is determined on the size and location of the overlapping regions between I_k and I_{k+1}. The size of the overlapping regions can be obtained from the translations required to align the images and are respectively denoted by

$$I^*{}_k(i,j) = I_k(i, L_k - t_v + j) \quad \text{and}$$
$$I^*{}_{k+1}(i,j) = I_{k+1}(i,j) \quad , \tag{5}$$

for $1 \le j \le t_v$ and $1 \le i \le H$, where t_v is the horizontal translation and H is the height of I_k.

Since images are less likely to be distorted near the centre, it is desirable to have the contributing regions close to the centre of the images to avoid distortions in the merged image. By placing the seam in the centre of the overlapping regions, the regions contributed to the final image are as close to the centres of both images as possible.

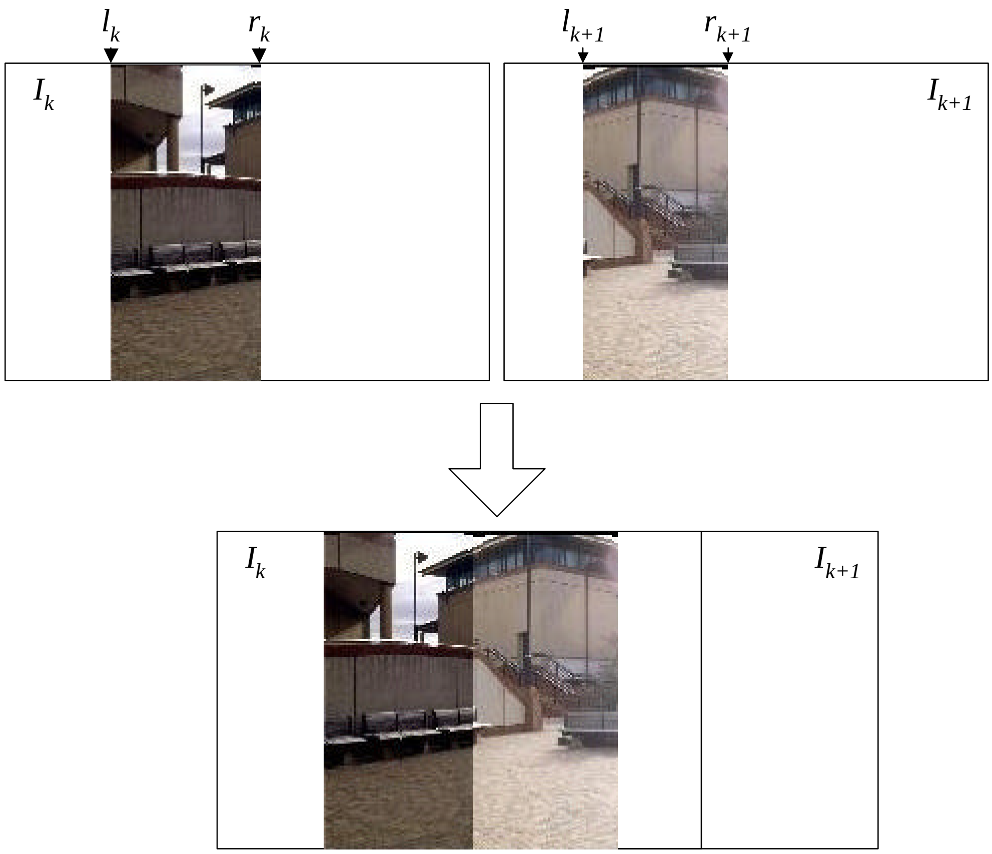


Figure 20: *The contribution of I_k to the final image is bounded by l_k and r_k.*

The seam in the merged image is formed by the r_kth column in I_k and the l_{k+1}th column in I_{k+1}. According to the position of the seam, the right most column in I_k and the left most column in I_{k+1} that appear in the final image are given by Eqn. 6,

$$r_k = L_k - \frac{t_v}{2} \quad \text{and}$$

$$l_{k+1} = \frac{t_v}{2} \quad . \tag{6}$$

The differences in the intensity values across the seam are calculated by

$$E_k(i) = I_k(i, r_k) - I_{k+1}(i, l_{k+1}) \quad , \tag{7}$$

where $1 \le i \le H$.

A neighbourhood, N, is selected so that the intensity differences can be distributed across the seam. The distributed intensity difference, e_k, is given by Eqn. 8,

$$e_k(i) = \frac{1}{2N} E_k(i) \quad . \tag{8}$$

A neighbourhood of half of the width of the overlapping region, *i.e.*, $t_v/2$, has been selected, so that the intensity adjustment is applied over the entire overlapping region. For each pixel in the defined neighbourhood, a weighted intensity difference, e_k, is added to or subtracted from it. So that the intensity values in the merged images are given by

$$I_k(i, r_k - j) = I_k(i, r_k - j) - (N - j)e_k(i) \quad \text{and}$$
$$I_{k+1}(i, l_{k+1} + j) = I_{k+1}(i, l_{k+1} + j) + (N - j)e_k(i) \quad , \tag{9}$$

for $1 \le j \le N$ and $1 \le i \le H$. Note that the weighing function in this method is a linear function inversely proportionally to the distance away from the seam.

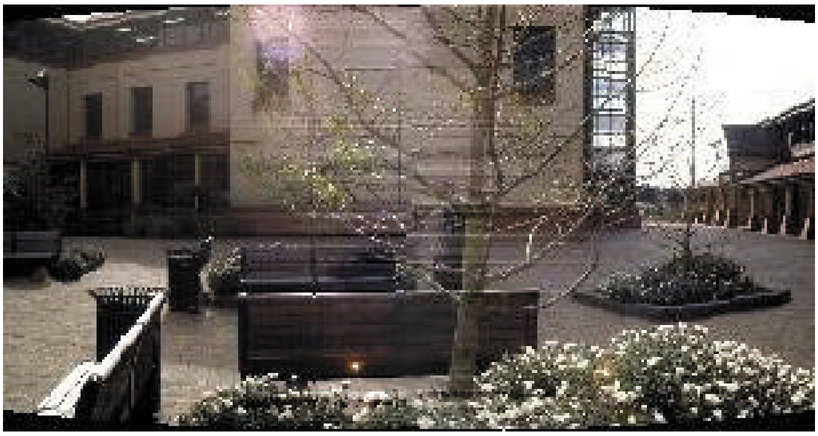


Figure 21: *Image merged by method 1.*

The image shown in Fig. 21 does not have the distinctive seam presented in Fig. 12. However, some horizontal stripes have appeared across the merged region. The horizontal stripes appeared because the intensity values have only been adjusted with respect to the horizontal direction. Therefore, in the merged images, discontinuities of intensities in the vertical direction occur where the values of intensities were uniform in the input images. Furthermore, each intensity difference in E_k is calculated from the intensity values of two pixels. Since the intensity values of the *i*th row are adjusted according to the value of $E_k(i)$, a large fluctuation in the magnitude of $E_k(i)$ causes the whole row to be much brighter or darker than its neighbouring rows.

In the case when there is only an intensity shift between I_k and I_{k+1}, the intensity differences across the seam should be quite uniform. However, in practise, the intensity differences across the seam are seldom uniform. There are often fluctuations in the intensity differences, caused by the position of the seam and the inconsistencies in features across the seam.

Fig. 22 shows the values of E_k for the images merged in Fig. 21. In Fig. 22, the two lines mark the upper and lower 10% of the intensity differences. From the graph of E_k, we can see that there are quite a few sudden leaps in the intensity differences. The star and cross signs respectively mark the positions where the values of E_k are above the upper 10% or below the lower 10% of the intensity differences. The positions where the values of E_k are above the upper 10% or below the lower 10% of the intensity differences are mapped onto the merged image shown in Fig. 23.

From Fig. 23, it can be seen that most of the signs mark where the horizontal stripes occur, showing that the horizontal stripes are caused by the large fluctuations of the intensity differences. If we are able to avoid using the obtained values of E_k at those marked positions, it may be possible to eliminate the occurrence of the horizontal stripes and improve the quality of the merged image.

The advantage of this image merging method is that the detailed features in the input images, such as lines, small objects, edges or corners, remain intact in the merged image if the selected neighbourhood is large enough. A larger N implies that the intensity differences can be spread over a larger area, therefore, the intensity increment or decrement, e_k, between each column is smaller, thus preserving the relative intensity differences. If the neighbourhood is small, the magnitude of e_k is larger, which means that the relative differences of the intensities might not be so well preserved. Large intensity increments may also be visually observable, resulting in a detectable seam between joined images. However, a smaller N means that more pixels retain the intensity values of the original images, thus also preserve the quality of the original images in the merged image.

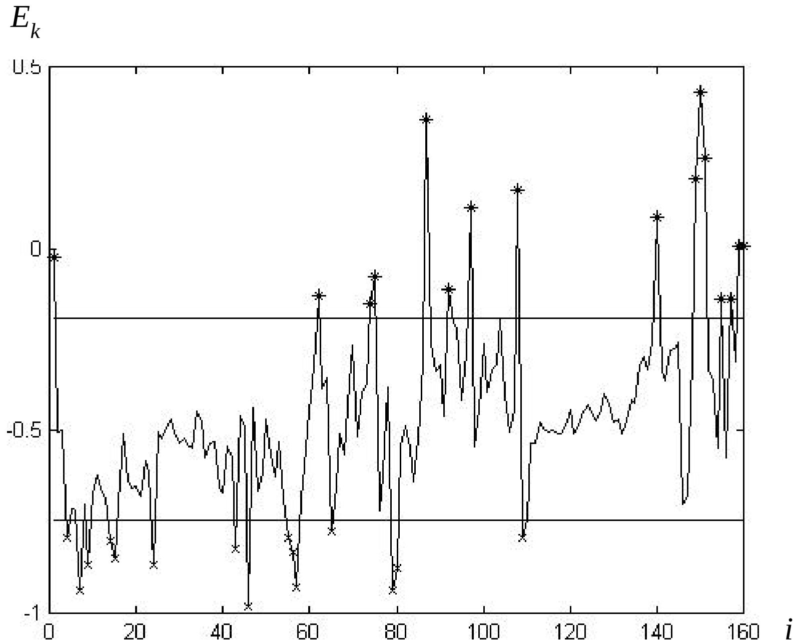


Figure 22: *Graph for the values of E_k.*

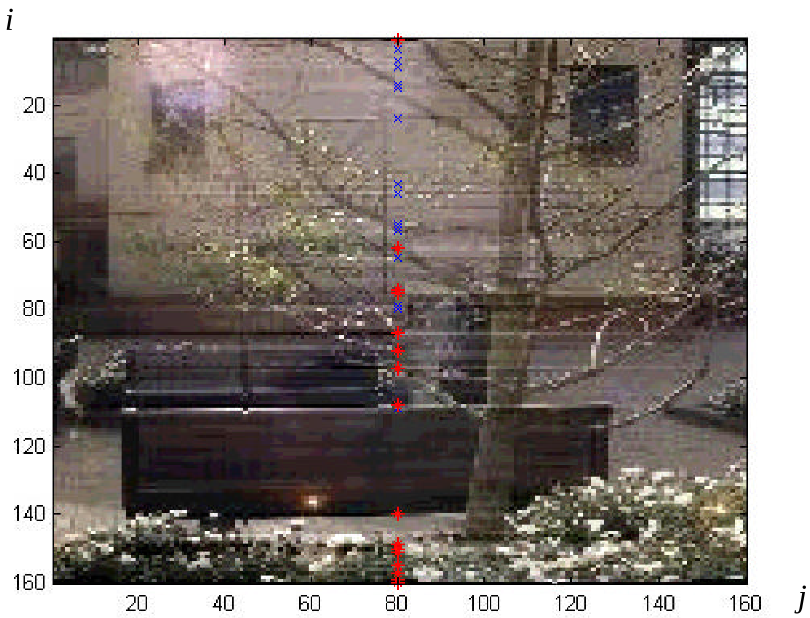


Figure 23: *Merged image with marked horizontal strips.*

4.2 Linear distribution of median intensity differences

In this section, the median of intensity differences is used to avoid selecting the large fluctuations of the intensity differences as discussed in Section 4.1. We define the intensity difference for row i, $E'_k(i)$,

$$E_k^{'}(i) = median\{E_k(i-c), E_k(i-c+1), \ldots, E_k(i), \ldots, E_k(i+c-1), E_k(i+c)\} \quad , \tag{10}$$

as the median of the intensity differences for $2c+1$ intensity values. The image merged by this algorithm is shown in Fig. 24, given that the value of c is 2, which means that each $E'_k(i)$ is the median of five intensity differences.

Figure 24: *Image merged by method 2.*

From Fig. 24, we can see that the seam is invisible under visual observations. The horizontal stripes caused by large fluctuations in the intensity differences are also absent from the merged image. It can also be seen that the details in the input images have been retained in the merged image.

4.3 Intensity adjustment with respect to corresponding pixels in overlapping region

The image merging methods described in Sections 4.1 and 4.2 both use the intensity differences across the seam to calculate the intensity values in the merged images. The advantage with these methods is that the relative intensity differences can be maintained in the merged images. However, with such methods, the intensity calculations for the entire row are based on the intensity difference between two pixels, regardless of the intensity discontinuity caused in the vertical direction. Therefore, it is not possible to completely avoid the appearances of horizontal lines caused by the fluctuations of the intensity differences.

In this section, we take a different approach by gradually adjusting the contributions of the intensities from the input images. The contributions of the intensities are linearly varied across a defined neighbourhood, N, about the seam. The neighbourhood is of the same size as in methods 1 and 2, *i.e.*, N is equal to $t_v/2$. On the left most column of the defined neighbourhood, image I_k contributes 100% of the intensity values, but the percentage contributed by image I_k linearly decreases as we move away from the left most column. The contribution from image I_k is 50% in the centre of the neighbourhood, *i.e.*, immediately on the left hand side of the seam. The contribution from I_k is 0% on the right most column of the neighbourhood. The progression of the percentage contributed by image I_{k+1} is the inverse of image I_k.

Eqn. 11 shows the intensity calculation for this method. The intensity values within the defined neighbourhood of the seam in the merged image are represented by $\bar{I}$,

$$\bar{I}_k(i, r_k + j) = \frac{N-j}{2N} . I_k(i, r_k + j) + \frac{N+j}{2N} I_{k+1}(i, l_{k+1} + j - 1) \quad , \tag{11}$$

where $1 \le i \le H$ and $-N < j < N$.

The image merged by this method is shown in Fig. 25. In Fig. 25, the merged image does not have the undesired horizontal stripes. By visual observation, the merged image also shows no significant loss of detail, as the detailed features in the image can be easily distinguished.

Figure 25: *Image merged by method 3.*

This method works well in cases where the overlapping regions of image I_k and I_{k+1} differ by an intensity shift, *i.e.*, the features in the overlapping regions occupy the same position and have the same size, but the intensity values for the features may differ by a constant amount. However, the relative positions and shapes of the features within the overlapping regions may vary due to object movement, parallax errors in image acquisition, or failure to find the ideal translations between images during image registration. In such cases, calculating the intensity values with respect to the intensity values in the corresponding positions produces an effect similar to exposing the negative film more than once when taking a photograph. Under this effect, the affected objects appear at different positions and/or with different orientations, within a single image. It is possible for the same object appears twice in the merged image in the merging of two images, thus this effect is known as the *double exposure effect*. Fig. 26 shows the double exposure effect produced by this image merging method, where the white arrow indicates the presence of the double exposure effect within the merged image.

Figure 26: *Double exposure effect in the merged image.*

Nevertheless, if the images have been correctly registered and the overlapping regions from the left and the right images differ only by a constant amount, then this method is able to eliminate seams without introducing intensity discontinuities in the vertical directions.

4. 4 Intensity adjustment with respect to median filtered regions

In this method, a low-pass filter is applied on portions of the defined neighbourhood to extract the low frequency components. Since the low frequency components contribute to the overall intensities of the image, we calculate the intensity values in the merged image according to the intensity value in the low-pass filtered images.

The low-pass filter used in this method is the median filter, chosen for its ability to remove small objects, yet preserving the edges on the filtered image. The removal of small objects which occupy different positions or orientations within the overlapping regions helps to reduce the occurrence of the double exposure effect.

The median filter is applied to the portion of the overlapping region where the intensity contributions to the final image are less than 50%. The 50% threshold has been chosen because it is not desirable to have the low-pass filtered images contributing too much into the merged image. Otherwise the merged images will not retain the details from the input images, which are represented by the high frequency components. Therefore, the threshold has been chosen so that the input images are the main contributor at all times, hence preserving the quality of the original images as much as possible. Eqn. 12 calculates the intensities in the merged image. The the median filtered images are denoted by $\tilde{I}$ and the calculated intensity values within the defined neighbourhood of the seam are denoted by $\bar{I}$,

$$\bar{I}_k(i, r_k + j) = \begin{cases} \dfrac{(N-j)\tilde{I}_k(i, r_k + j)}{2N} + \dfrac{N+j}{2N} I_{k+1}(i, l_{k+1} + j - 1), \text{ if } \dfrac{N-j}{2N} \le 0.5 \\ \\ \dfrac{(N-j)I_k(i, r_k + j)}{2N} + \dfrac{N+j}{2N} \tilde{I}_{k+1}(i, l_{k+1} + j - 1), \text{ if } \dfrac{N+j}{2N} < 0.5 \end{cases}, \tag{12}$$

where $1 \le i \le H$ and $-N < j < N$.

In this work, we have experimented with a 7 by 7 median filter. The image merged by this method is shown in Fig. 27, where the seam between the joined images has been successfully removed. The images with the double exposure effect shown in Fig. 26 are merged by this image merging method in Fig. 28 to demonstrate the elimination of the double exposure effect.

Figure 27: *Image merged by method 4.*

Figure 28: *Elimination of double exposure effect by method 4.*

From Fig. 28, we see that it is possible to reduce the double exposure effect by applying the median filter over a selected neighbourhood on each of the images. However, by close inspection of the images merged by methods 3 and 4, it has been found that by using intensity values from both images to calculate the intensity values in the merged images, the intensity differences between adjacent positions are not as well preserved as in methods 1 and 2.

Overall, image merging methods 1 and 2 are sensitive to the intensity differences between images, but are less effected by the differences in the position and orientations of features within the overlapping regions. Methods 3 and 4 are less sensitive to the intensity differences between adjacent images, but are more suitable for situations where the position and orientations of features within the overlapping regions are similar.

4.5 Comparison of methods

The comparison of the image merging methods is done with respect to two criteria. The first criterion is the elimination of the seam, such that there is no visible boundary between joined images in the merged image. The second criterion is the preservation of image quality, by which we mean that features such as edges, corners and other fine details should not be blurred or degraded in any other way.

Since the elimination of the seam is able to be judged visually, it is not necessary to quantitatively evaluate this criterion. In this work, the differences of the *contrast* of the merged and the original images are used to evaluate the quality of the merged images. Contrast is used to mean the intensity differences between neighbouring pixels, or the first order derivative of the intensity values. Eqn. 13 shows the calculation for image contrast , I'_k, of the image I_k in the vertical direction,

$$I'_k(i,j) = I_k(i+1,j) - I_k(i,j) \ , \tag{13}$$

for $1 \le i < H$ and $1 \le j \le L$.

To compare the original and merged images, the area corresponding to the overlapping regions of the original images is extracted and divided into two halves. Fig. 29 illustrate the merged image, $I_{k,k+1}$, and two halves corresponding to the overlapping regions of the original images, A and B.

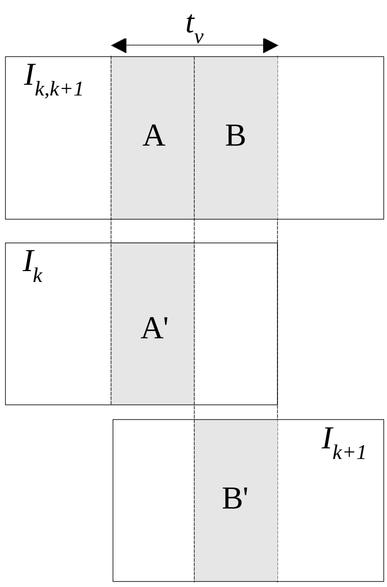


Figure 29: *Regions for comparisons.*

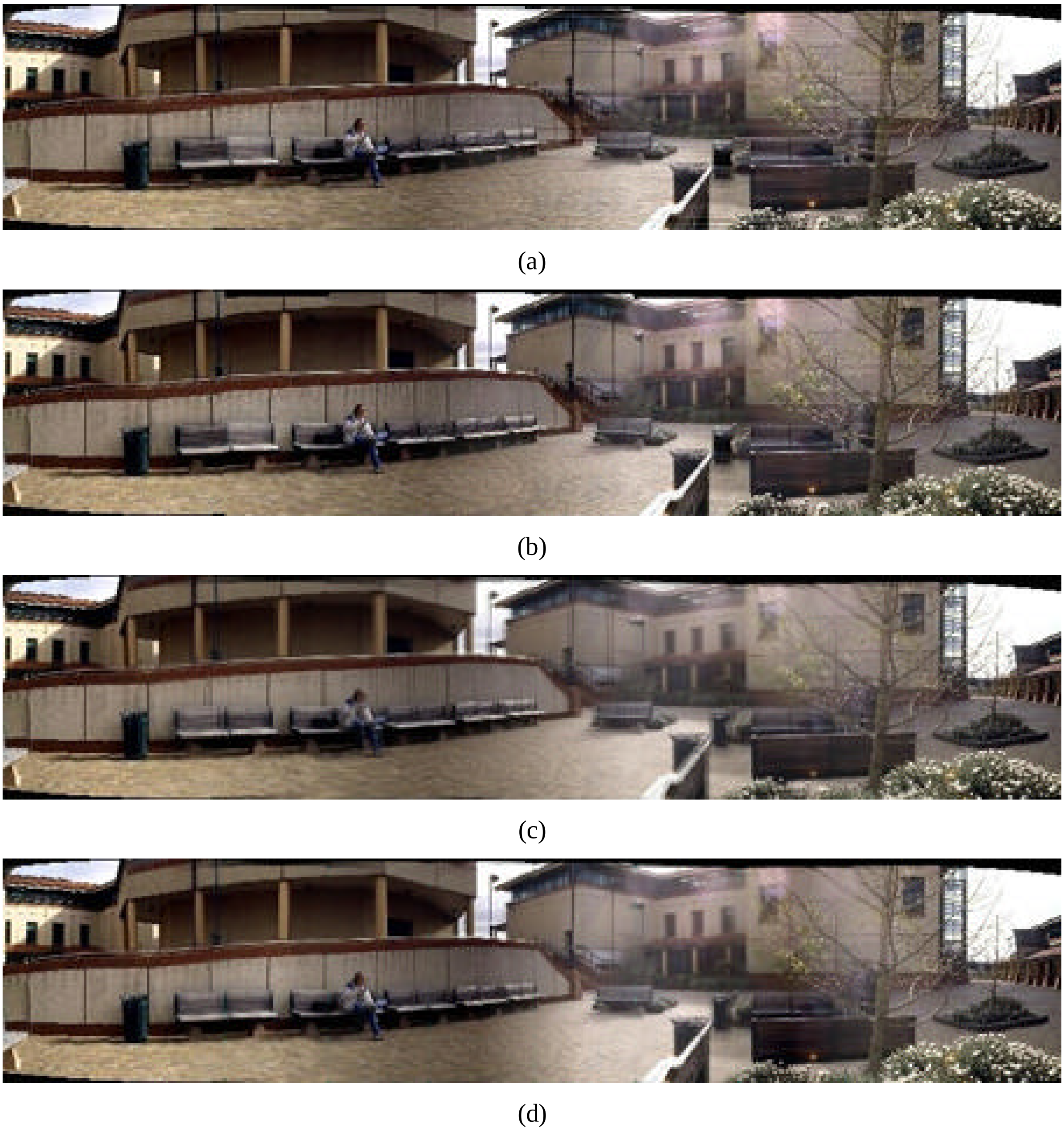
By using the image merging methods described in Sections 4.1 to 4.4, the left half of the overlapping region is mainly contributed by the left input image, *i.e.*, from A’, and the right half of the overlapping region is mainly contributed by the right input image, *i.e.*, from B’. Therefore, the contrast values for regions A and B, are subtracted from the contrast values for regions A’ and B’ respectively to find the differences in the contrast values. The performance measures d_A and d_B, for regions A and B, are given in Eqn. 14,

$$d_A = \sum_{i=1}^{H}\sum_{j=1}^{t_v/2} |I'_k(i, L_k - t_v + j) - I'_{k,k+1}(i, L_k - t_v + j)| \quad \text{and}$$

$$d_B = \sum_{i=1}^{H}\sum_{j=1}^{t_v/2} |I'_{k+1}(i, \frac{t_v}{2} + j) - I'_{k,k+1}(i, L_k - \frac{t_v}{2} + j)| \qquad , \tag{14}$$

where L and H respectively denote the width and height of the images.

Fig. 30 shows a series of eight images merged by the image merging methods discussed in this work. The series of images joined without any merging can be found in Fig. 18. Table 4 contains the sums of the absolute contrast differences between the merged and input images.



(a)

(b)

(c)

(d)

Figure 30: *Images merged by methods 1 to 4, (a) to (d)*

Image no.	Method 1		Method 2		Method 3		Method 4	
	d_A	d_B	d_A	d_B	d_A	d_B	d_A	d_B
1-2	71343	73055	7367	6979	34928	28621	24887	22767
2-3	44770	44795	5645	5432	22230	30749	22036	22882
3-4	51577	51585	5351	5851	21091	23890	18761	19508
4-5	69622	70211	10566	6638	40398	22722	30050	21769
5-6	53650	53509	6112	7370	29372	30109	20622	25959
6-7	54929	54963	9465	9631	25879	30180	19875	24457
7-8	78524	78171	11359	10625	37791	37246	31392	27548

Table 4: *Sum of the contrast differences.*

By using the contrast differences to measure the image quality of the merged images, the image merging method discussed in Section 4.2 has been found to have the best performance.

The reason that the second method has better quality when compared with the other method is that the adjustment to the original intensity levels is kept to a necessary minimum. The amount of adjustment is also proportional to the intensity differences between the joined images. Therefore, large intensity differences between the images indicate that the merged images may not retain the quality of the original images as well as when the intensity differences are small.

Even though the first merging method uses similar calculations for the intensities in the merged image, it is over sensitive to the fluctuations in intensity differences. The fluctuations in intensity differences cause horizontal stripes to occur, thus degrading quality of the image merged by the first method.

The quality of the merged images is dependent upon the positions, sizes and orientations of features within the overlapping regions of the input images for the third and fourth image merging methods. The sensitivity to the differences in features causes an effect resembling a double exposed photograph in the images merged by the third method. However, it is often difficult to control the positions, sizes and orientations of features so that they are exactly the same within the overlapping regions of the input images. Therefore, it is less advantageous for these two methods to be dependent upon such factors.

In the fourth method, the sensitivity to the differences in features has been reduced by the use of a median filtered region in the intensity calculations. However, due to the amount of intensity alterations required, the quality of the merged images is still not as good as that of the second method.

The contrast of the images merged by methods 3 and 4 are also not as high as the image merged by methods 1 and 2. It has been noted that the reduction of contrast may be another form of the double exposure effect where the distances between the double exposed features are very small, one or two pixels, for example.

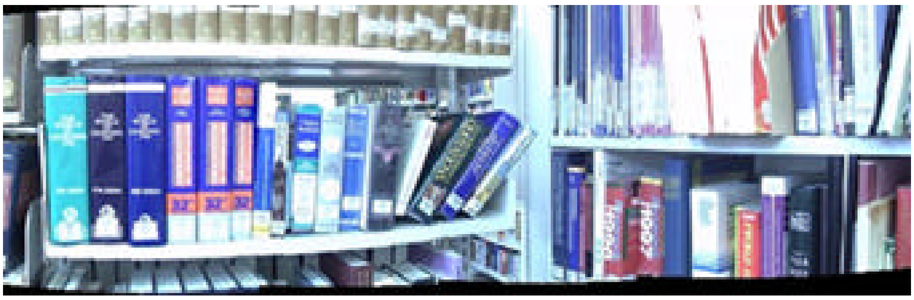
5 Resultant panoramic images

From the evaluations conducted in Sections 3 and 4, the best performing image registration and image merging methods are used to implement our image stitcher. Panoramic images stitched by the implemented image stitcher are shown in the following.

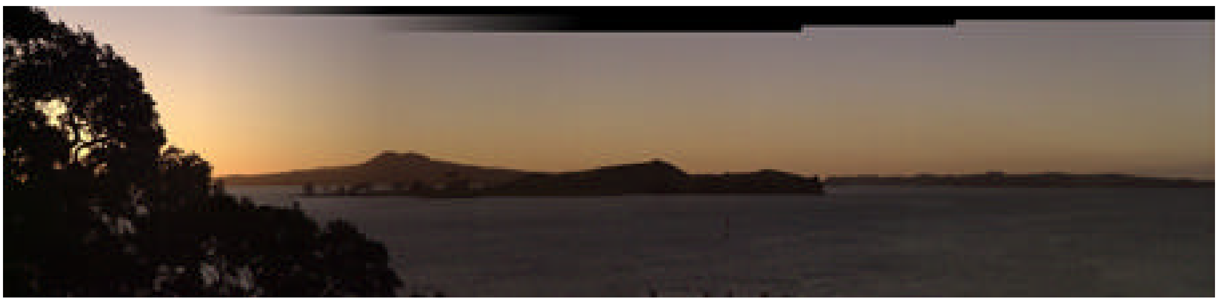
Note that some panoramic images have been divided into two halves to fit onto the page.

(a)

(b)



(c)



(d)

Figure 31: *Resultant panoramic images*

6 Conclusions and future directions

In this work, we have discussed the different set-up's in the acquisition of images for generating panoramic images, the stitching of the acquired images and the resultant panoramic images produced by our stitcher. By doing so, we hope to provide a better understanding of the different stages involved in the generation of panoramic images.

The process of image stitching has been divided into two parts, image registration and image merging. Both parts have been described in detail and different methods have been discussed for achieving the registration and merging of the images. The different methods suggested for registration and merging of the images are evaluated according to the requirements in each part. According to the evaluations, the best performing methods are used to implement our image stitcher. Panoramic images generated by our image stitcher have also been shown.

The image stitcher provides a cost effective and flexible alternative to acquire panoramic images using a panoramic camera. The generated panoramic images can be used for 360 degree interactive panoramic movies, such as QuickTime VR®, or multi-node panoramic movies.

Our stitcher program can be used to stitch images in the generation of image based or hybrid virtual environments. Currently, stitched panoramic images have been used in the generation of interactive movies which enable the user to navigate predetermined paths. However, such panoramic movies have limited user interactions. By implementing an image stitcher into the virtual environment generator, it is possible to generate views as the user travel between defined positions in the virtual environment. This allows more freedom with respect to the viewing of the environment and user navigation. Within selected positions of the panoramic movies, the user is able to have a 360 degree view of the environment.

The panoramic images stitched by a stitcher can also be used in applications where the camera is unable to obtain a full view of the object of interest. The full view of the object can be constructed using the image stitcher using overlapping regional images acquired for the object.

The methods provided in Sections 3 and 4 are possibilities for achieving image registration and merging. Evaluations of these methods provide indications on the behaviour of each method. The information provided in this work may be used to adapt the methods separately to applications other than image stitching, such as the registration of aerial or satellite images, or the merging of unrelated images for special effects.

In image registration, it will be desirable to find a method that can provide the optimal translations between the input images regardless of the adverse factors, such as the intensity shifts between images, unwanted camera movement, similarity of features, lack of features.

The image merging methods discussed in this work all produce certain degrees of degradation in the merged images, such as the striping or the double exposure effects, due to the required intensity adjustment. Therefore, it will be worthwhile to find an image merging method which eliminates the seams between the joined images without causing any degradation in the merged images, or at least reduce the degradation to such a degree that it is not visually detectable.

In conclusion, we hope that our work is able to provide some fundamental ideas to those wishing to further investigate the process of image stitching and the image stitcher in the future.

7 References

[Bar72] D. I. Barnea and H. F. Silverman, "A class of algorithms for fast digital registration", IEEE Trans. Comput., C-21, 1972, pp. 179-186.

[Bra95] R. N. Bracewell, Two dimensional imaging, Prentice-Hall, New Jersey, 1995.

[Bro92] L. G. Brown, "A survey of image registration techniques", Computing Surveys, vol. 24, 1992, no. 4, pp.325-376.

[Che97] C. Chen and R. Klette, “An image stitcher and its application in panoramic movie making”, Proc. DICTA’97, Dec. 1997, pp.101-106.

[Gam88] J. P. Gambotto, “Segmentation and interpretation of infrared image sequences”, vol. 3, Advances in computer vision and image processing, JAI Press, Connecticut, 1988.

[Gon87] R. C. Gonzalez and P. Wintz, Digital image processing, 2nd ed., pp. 425-426, Addison-Wesley, Massachusetts, 1987.

[Kas90] R. Kasturi and M. M. Trivedi, Image analysis applications, pp. 208-216, Marcel Dekker, New York, 1990.

[Kle96] R. Klette and P. Zamperoni, Handbook of image processing operators, Chichester: John Wiley and Sons, 1996.

[Mil75] D. L. Milgram, “Computer methods for creating photomosaics”, IEEE Trans. Comput., vol. C-24, Nov. 1975, pp.1113-1119.

[Mil77] D. L. Milgram, “Adaptive techniques for photomosaicking”, IEEE Trans. Comput. vol. C-26, pp.1175-1180, Nov. 1977.

[Rit96] G. X. Ritter and J. N. Wilson, Handbook of computer vision algorithms in image algebra, pp. 225-239, CRC Press, 1996.

[Ros82] A. Rosenfeld and A. C. Kak, Digital image processing, 2nd ed., vol. 2, Academic Press, 1982.

[Shi87] Y. Shirai, Three-dimensional computer vision, pp.127-129, Springer-Verlag, Berlin, 1987.

[Shu97] H. Shum and R. Szeliski, Panoramic image mosaics, Technical Report, http://www.research.microsoft.com/, 1997.

[Son93] M. Sonka, Vaclav Hlavac and R. Boyle, Image processing, analysis and machine vision, pp.176-179, Chapman & Hall, London, 1993.

[Tzi94] G. Tziritas and C. Labit, Advances in image communication, vol. 4, Motion analysis for image sequence coding, Elsevier, 1994.