ISSN 2229-5518

Image Mosaicing using Feature based technique

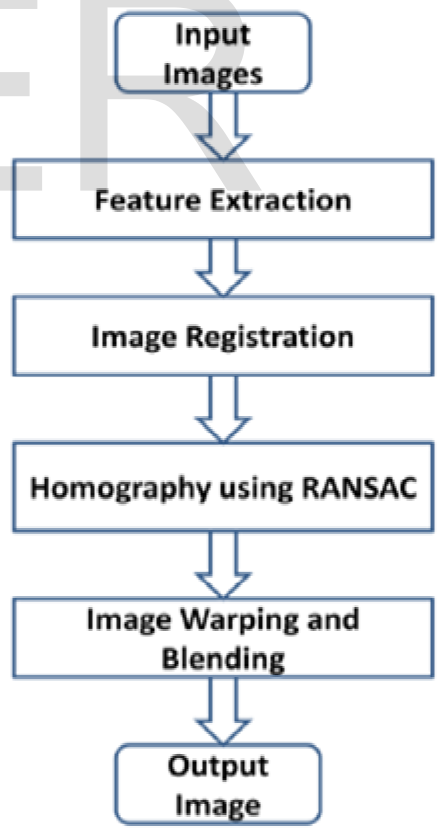
Kali Krishna Kota, Shruti Jonnalagadda

Abstract— The automatic construction of large, high resolution image mosaics is an active research area in the field of photogrammetry, computer vision, image processing and computer graphics. The need for developing a technique which does the process of combining multiple photographic images with overlapping fields of view to produce a segmented panorama or high-resolution image is because the human visual system has a field of view of around 135 x 200 degrees, but a typical camera has a field of view of only 35 x 50 degrees. To overcome the barrier of a bounded field of view, to expand one's visual perceptive arena, Image Mosaicing technique has been developed. It is a field of study in the discipline of Computer Vision whose study is devoted to discovering algorithms, data representations, and computer architectures that embody the principles underlying visual capabilities. It thus is a multi-disciplinary field of science and technology that depends on information taken from the images. It aims at building artificial systems to simulate the seeing sense of living beings. The mosaiced image (stitched image) of various images of the same scene is obtained by following a sequence of steps. The image is first analysed and all the local features(corners) are detected by employing feature detection algorithms. Most efficient algorithm being Scale Invariant Feature Transform(SIFT). These local features detected are matched and the overlapped portion of the images of the same scene is located. These images are then aligned according to the planar coordinate transformations. After appropriate alignment, excess portion of the overlapped region is removed with the help of a procedure known as homography using RANSAC algorithm. Later to obtain a seamless view of the scene captured without the stitched region being visible is obtained by blending both the images. This process flow results in a high resolution well stitched image output.

Key Terms— Feature detection, SIFT, Corners, Registration, Homography, RANSAC, Blending

1 Introduction

Image mosaicing is the alignment of multiple images into larger compositions which represent portions of a 3D scene. The mosaicing method described here is concerned with images which can be registered by means of a planar homography: views of a planar scene from arbitrary viewpoints, or views of a scene taken by a rotating camera. There are three important factors to consider when constructing a mosaic: the estimation of a set of homographies which are consistent over all the views; the choice of reprojection manifold on which the images are composited; and the choice of algorithm for blending the overlapping images. Image Mosaicing has become an active area of research in the fields of photogrammetry, computer vision, image processing, and computer graphics. The various methods adopted for image mosaicing can be broadly classified into direct methods and feature based methods. Direct methods are found to be useful for mosaicing large overlapping regions, small translations and rotations. Feature based methods can usually handle small overlapping regions and in general tend to be more accurate but computationally intensive. Image Mosaicing is basically the stitching of multiple correlated images to generate a larger wide-angle image of a scene. It is a synthetic composition generated from a sequence of images and it can be obtained by understanding geometric relationships between images. The geometric relations are the coordinate systems that relates the different image coordinate syste-ms. The appropriate transformations are applied with the help of a warping operation than by merging the overlapping regions of warped images. This merged single image is the output mosaiced image. Various steps involved in image mosaicing are shown in the form of a flow chart.



ISSN 2229-5518

2 Fundamental methods for Mosaicing

2.1 Direct method

In direct method, all the pixel intensities of the images are compared with each other. It minimizes the sum of absolute differences between overlapping pixels. In this method, each pixel is compared with each other. They are scale and rotation variant. Direct method uses information gathered from the image alignment. It measures the contribution of every pixel in the image. The main disadvantage of direct method is that they have a limited range of convergence. Direct Method uses information from all pixels. It iteratively updates an estimate of homography so that a particular cost function is minimized. Phase-Correlation is also used to estimate the a few parameters of homography.

2.2 Feature based method

In feature-based technique, two images are compared using local descriptors. Image mosaicing based on feature-based techniques, feature extraction, registration, and blending are various steps involved. The main characteristics of feature based method include noise, scale, translation, and rotation invariance. This method is faster and is automated.

2.3 Feature based method over Direct method

Direct method, due to pixel to pixel comparison is very complex and is very costly. It is very slow in comparison to the feature based method. Feature based method has the advantage that it is invariant to a number of characteristics as mentioned above. Also it has the advantage that it is highly automatized due to accurate keypoint descriptors.

3 Steps involved in Image Mosaicing

3.1 Feature Extraction

Feature Extraction is the first step involved in Image Mosaicing. Feature extraction method in image mosaicing extract distinct features(corners) from the images which can be used to match the similarity for estimation of relative transformation between the images. Features are informative and non-redundant data that contain relevant information related to the input data (In this context, related to the images). This kind of relevant information is required to do the intended task i.e. image stitching. The importance of feature extraction is that it helps in describing a given data set accurately with minimal information. In our context, feature extraction is the step where relevant information regarding the input images is extracted with the help of an algorithm. The algorithm is chosen in such a way that the computational time, memory storage is very less and efficient. Features have application dependent definition. In order to represent an image, features like edges, corners, blobs, ridges are defined. The choice of whether an image be represented by edges, corners or blobs is dependent on the amount of accuracy required for our task. The best feature to define an image is a corner. Detailed explanation of all the features of an image and why corner is preferred to the rest of the features is explained in detail in the next section. Features based methods have shown much advantage over direct mosaicing methods in both time and space complexity.

3.1.1 Importance of features

There is no universal or exact definition of what constitutes a feature, and the exact definition often depends on the problem or the type of application. Given that, a feature is defined as an "interesting" part of an image, and features are used as a starting point for many computer vision algorithms. Since features are used as the starting point and main primitives for subsequent algorithms, the overall algorithm will often only be as good as its feature detector. Consequently, the desirable property for a feature detector is repeatability: whether or not the same feature will be detected in two or more different images of the same scene. Importance of local features may also be highlighted because of the below given points:

Locality:
Features are local, so they are robust to occlusion and clutter.
Distinctiveness:
They can differentiate a large database of objects.
Quantity:
Hundreds or thousands of features are found in a single image.
Efficiency:
Real-time performance with the help of local features are achievable.
Generality:
Exploit different types of features in different situations.

3.1.2 Types of Features

EDGES:
Edges are points where there is a boundary (or an edge) between two image regions. In general, an edge can be of almost arbitrary shape, and may include junctions. In practice, edges are usually defined as sets of points in the image which have a strong gradient magnitude. Furthermore, some common algorithms will then chain high gradient points together to form a more complete description of an edge. These algorithms usually place some constraints on the properties of an edge, such as shape, smoothness, and gradient value. Locally, edges have a one-dimensional structure. The disadvantage with selecting edges as interest points or features is they are variant to all the transformations except for brightness i.e. they are invariant to brightness alone.

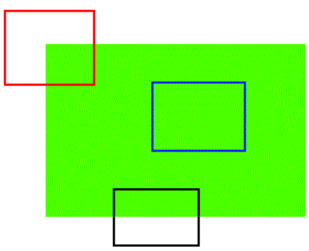
CORNERS:
The terms corners and interest points are used somewhat interchangeably and refer to point-like features in an image, which have a local two dimensional structure. The name "Corner" arose since early algorithms first performed edge detection, and then analyzed the edges to find rapid changes in direction (corners). These algorithms were then developed so that explicit edge detection was no longer required, for instance by looking for high levels of curvature in the image gradient. It was then noticed that the so-called corners were also being detected on parts of the image which were not corners in the traditional sense (for instance a small bright spot on a dark background may be detected). These points are fre-

ISSN 2229-5518
quently known as interest points, but the term "corner" is used by tradition.

BLOBS:
Blobs provide a complementary description of image structures in terms of regions, as opposed to corners that are more point-like. Nevertheless, blob descriptors may often contain a preferred point (a local maximum of an operator response or a center of gravity) which means that many blob detectors may also be regarded as interest point operators. Blob detectors can detect areas in an image which are too smooth to be detected by a corner detector. Consider shrinking an image and then performing corner detection. The detector will respond to points which are sharp in the shrunk image, but may be smooth in the original image. It is at this point that the difference between a corner detector and a blob detector becomes somewhat vague. To a large extent, this distinction can be remedied by including an appropriate notion of scale. Nevertheless, due to their response properties to different types of image structures at different scales, the LoG and DoH blob detectors are also mentioned in the article on corner detection.

RIDGES:
For elongated objects, the notion of ridges is a natural tool. A ridge descriptor computed from a grey-level image can be seen as a generalization of a medial axis. From a practical viewpoint, a ridge can be thought of as a one-dimensional curve that represents an axis of symmetry, and in addition has an attribute of local ridge width associated with each ridge point. Unfortunately, however, it is algorithmically harder to extract ridge features from general classes of grey-level images than edge-, corner- or blob features. Nevertheless, ridge descriptors are frequently used for road extraction in aerial images and for extracting blood vessels in medical images—see ridge detection.

3.1.3 Corners over all the features

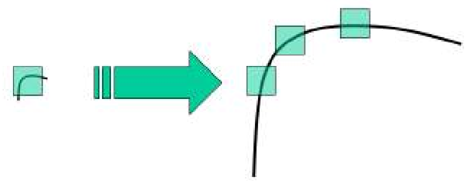


image

In the above image, the blue patch is a flat area, if it is moved all through the central region, we cannot differentiate the difference i.e. wherever you move the blue patch, it looks the same. The black patch shows an edge. If you move it in vertical direction (i.e. along the gradient) it changes, but if it is placed on top of the edge which is on the other side, it looks the same. Now in the case of red patch, a corner, wherever the patch is moved, it looks different. So basically, corners are considered to be good features in an image. In some cases, blobs are also considered good features, but not always. This is why corners are very often used as interest points.
(This illustration has been referred to, from Open CV)

3.1.4 Corner detection Algorithm – SIFT (Scale Invariant Feature Transform)

There are a number of corner detection algorithms such as Harris, SURF, FAST, BRIEF, ORB etc. In this paper we will only talk about SIFT algorithm which is a keypoint/interest point extractor. David Lowe from University of British Columbia, introduced Scale Invariant Feature Transform (SIFT) in his paper, Distinctive Image Features from Scale-Invariant Keypoints. The algorithm was patented by the university later on. The main advantage of this algorithm comes from the fact that it rotational as well as scale invariant.



image

A clear understanding of scale variance can be achieved from the image above. (Illustration from Open CV).
A corner within the small window is flat when it is zoomed in the same window. This is not the case when using SIFT algorithm. As mentioned above, one of its major advantage. Rotational invariance may also be understood on same lines. A corner remains to be a corner even though it is rotated multiple times.
Significance of Scale Invariance from a practical standpoint is:
Images of a scene from different viewpoints with different depths when captured may be of different sizes. In such a case algorithms which are scale variant fail to conclude that all the images of the single scene are same as they fail to match keypoints/ interest points/ corners. SIFT being scale invariant succeeds in concluding that all the images are same.

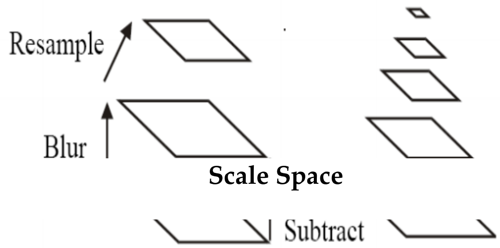
STEPS FOR EXTRACTING INTEREST POINTS/ KEY POINTS

1. Scale Space Representation
2. Key Point Localisation
3. Orientation Assignment
4. Key Point Descriptor
5. Key Point Matching

Scale Space Representation:
Scale Space Representation is done in order to find potential areas in the image where features may be found. The result of this representation will not yield accurate locations of all the features. It is implemented by using a Difference of Gaussian function to identify potential key points. The location of peaks in the scale space determines the possible location of interest

ISSN 2229-5518

point. To understand how it is done, first a scale space is to be formed. This is done in the following way.



Scale Space

age
it is
are
many. To overcome this difficulty a continuous spectrum of scales is formed. After the scale space is formed points are analysed, to conclude if they are interest points or not.

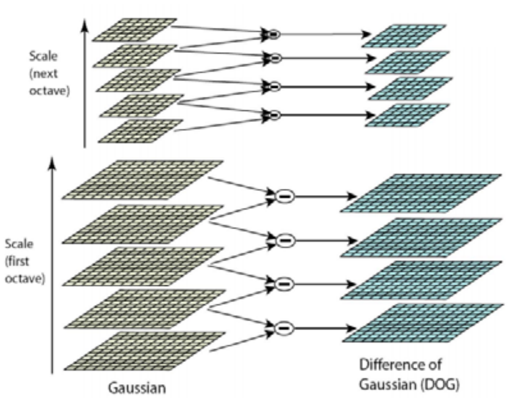
In general Laplacian of Gaussian (LoG) is computed, but for computational complexity reduction Difference of Gaussian (DoG) is used. With the help of the heat equation it can be proven that LoG is equivalent to DoG. It can be seen in the image shown below.

$$\frac{\partial G}{\partial \sigma} = \sigma \Delta^2 G \qquad \text{Heat Equation}$$

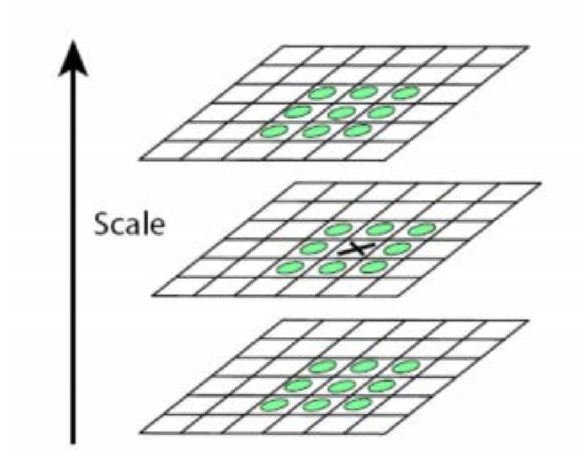
$$\sigma \Delta^2 G = \frac{\partial G}{\partial \sigma} = \frac{G(x,y,k\sigma) - G(x,y,\sigma)}{k\sigma - \sigma}$$

$$G(x,y,k\sigma) - G(x,y,\sigma) \approx (k-1)\sigma^2 \Delta^2 G$$

By computing DoG, rate of change of Gaussian function with respect to scale can be known. DoG Process can be illustrated visually in the image below.



Octaves in the image refer to a set of images with scales σ, k σ, k^2* σ etc. with a fixed resolution. For instance, Octave-1 may contain 5 scales with a resolution of 500x500, Octave-2 is obtained by subsampling Octave-1. In this way each and every column is scanned for the detection of a keypoint. After DoG is computed, local maxima in the scale space is a potential interest point. Emphasis is on the word 'potential' because, all the local maxima are not stable interest points. In order to be certain that a particular maxima is an accurate interest point, we take a 3x3 neighbourhood around the point. This neighbourhood is also considered in the scales above and below the possible interest point. The pixel intensity of the possible interest point is compared with neighbouring 26 points. If it is either maxima, or a minima, it is a SIFT interest point.



Scale Space Peak detection

All the interest points are detected in the same way. Sometimes there is a possibility that, detected keypoints are not repeatable i.e. scaling of that keypoint gives a different result. In order to eliminate such kind of unstable keypoints, Keypoint Localization is done.

KeyPoint Localization:

As mentioned above, KeyPoint localization is done for more accurate and stable interest points. KeyPoint localization is done in two steps:

1.Initial Outlier Rejection:

In this step, Taylor series of DoG function is computed and maxima/minima is obtained. If the maxima/minima ≥ Vth (Threshold Value), interest point is very much stable and is accepted, otherwise it is rejected.

After this level of interest point rejection, total number of keypoints/interest points obtained will come down.

1. 2.Further Outlier Rejection:

Sometimes, there is a possibility that edges are confused be confused to be corners. To eliminate such possibilities, further outlier rejection is done. In order to do so, DoG surface is assumed to be a surface. Principal curvature(PC) of the surface is computed and compared with a threshold value. If PC ‹ threshold, interest point is accepted otherwise it is rejected.

Mathematical Illustration:

ISSN 2229-5518

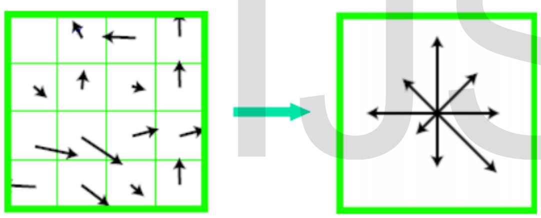
$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \qquad Tr(H) = D_{xx} + D_{yy} = \lambda_1 + \lambda_2$$

$$Det(H) = D_{xx}D_{yy} - (D_{xy})^2 = \lambda$$

$$\frac{Tr(H)^2}{Det(H)} = \frac{(r+1)^2}{r} \qquad r = \frac{\lambda_1}{\lambda_2}$$

Keypoint is eliminated if $r > 10$ as it may be a potential edge.

Orientation Assignment:
Orientation Assignment is the step where orientation of the SIFT interest point is to be computed. Orientation assignment will help in achieving rotational invariance. This is done by evaluating orientation of an interest point which is most dominating and align all the other interest points along it. First Derivative of each pixel is found which gives information about magnitude and direction also. Orientation of all the points around the interest point is computed and a histogram of 36 bins is formed. This histogram gives information regarding the number of pixels having same orientation angle. Each bin is 10° and thus we have 36 bins. Accurate weights with respect to magnitudes are assigned. Most dominant orientation (with regard to magnitude) are assigned to the interest point.



Histogram

Local Image Descriptors:

SIFT is very popular mainly because of its descriptors. A SIFT descriptor is more useful than a detector as it describes the keypoint with respect to various parameters which further are used to match interest points. To describe a keypoint, a window is drawn around the interest point. Information such as intensity, change in illumination etc. is extracted and is stored. But this is a low level description method. Accurate description is given by computing relative orientation and magnitude in a 16x16 neighbourhood at key point. This 16x16 neighbourhood is divided in 4x4 blocks. So we have a total of 16 blocks. Histogram of each block is computed which is 8 dimensional vector. All the 16 histograms are concatenated resulting in a 128 dimensional vector.

Key Point Matching:

Key points obtained are matched against a database of relevant images. Euclidean distance between vectors is found and nearest neighbour is declared as the one with minimum Euclidean distance. But in some cases, the second closest-match may be very near to the first. It may happen due to noise or some other reasons. In that case, ratio of closest-distance to second-closest distance is taken. If it is greater than 0.8, they are rejected. It eliminates around 90% of false matches while discards only 5% correct matches.

3.2 Image Registration

Image registration is the central task in image mosaicing which is the process of establishing correspondence between images that require to be stitched. Registration can be done in two ways. Direct registration and Feature based Registration. Since we are dealing with feature based technique of mosaicing the latter is discussed. Registration basically refers to the geometric alignment of the image set. These different sets of data may consist of two or more digital images taken of a single scene from different sensors at different times or from different viewpoints. In image registration the geometric correspondence between the images is established so that they may be transformed, compared and analyzed in a common reference frame. This is of practical importance in many fields, including remote sensing, computer vision, medical imaging. To register images, we need to determine geometric transformation that aligns images with respect to the reference image. So to conclude, the main purpose of image registration is to create geometric correspondence between images in order to compare images and apply necessary steps to stitch images appropriately. A common approach to register images is by computing homography. Details about Homography are dealt with in the next section.

3.3 Estimating Homography

Homography is a process which relates pixel coordinates of images to be stitched. Homography application to each pixel results in a new image which is a warped version of the original image.
This step basically determines the type of transformation that is required to be applied to register the images based on the correspondences obtained previously. Homography describes how a planar surface transforms when imaged through pin-hole cameras that have a different position and orientation in space i.e. it deals with linear transformation in projective space. The homography matrix, H describing the translation, rotation and scaling operations is a 3x3 matrix with 9 parameters. The points p′ and p in projective space are related (upto a scale factor) as follows

$$p' \sim Hp$$

$$p' \sim \begin{pmatrix} \theta_1 & \theta_4 & \theta_7 \\ \theta_2 & \theta_5 & \theta_8 \\ \theta_3 & \theta_6 & \theta_9 \end{pmatrix} p$$

Given a set of corresponding points p and p′ the problem of finding the homography matrix is equivalent to solving a system of linear equations. So, to determine 8 parameters a minimum of 4 pairs of correspondences are required. However, in reality large number of correspondences, several of which will be outliers. The presence of outliers will lead to an erroneous Homography matrix. So, outliers are to be removed.

3.4 RANSAC (Method to remove Outliers)

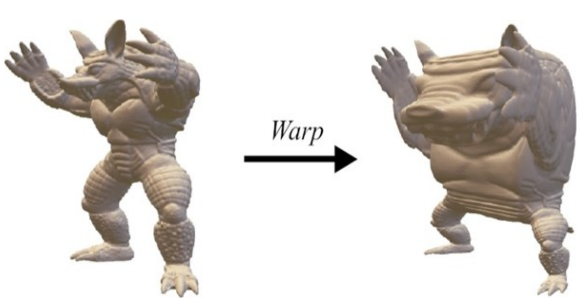
The presence of outliers during computation of Homography matrix will lead to an erroneous result. These outliers are removed using RANSAC. It separates inliers and outliers.
RANdom SAmple Consensus is an algorithm which solves different problems in the context of outliers.
This algorithm can be implemented in 3 steps:

1. Sample data points
2. Compute required parameters to fit a model.
3. Decide if the set of sampled points are inliers with the help of a preset threshold of the model
4. Repeat the process for different of sampled data points.

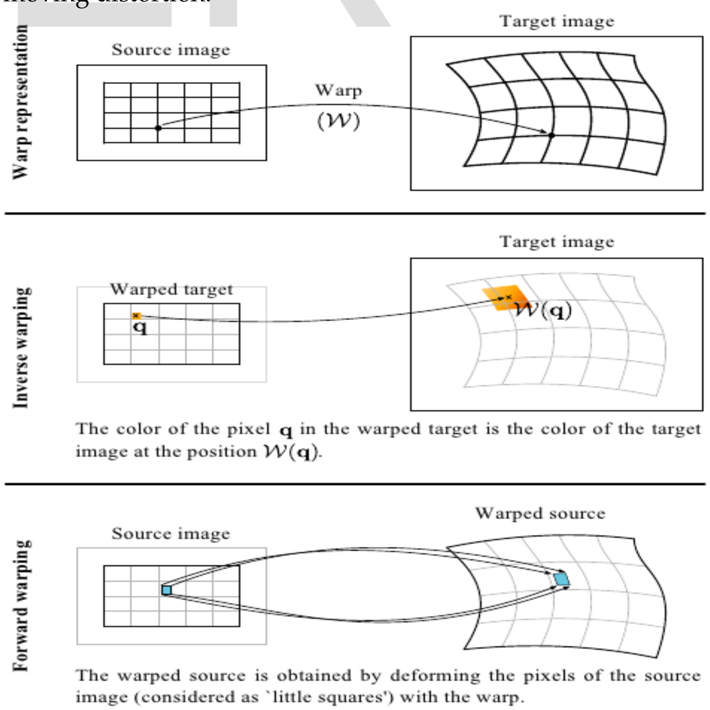
First, data points are sampled from the large set of putative correspondences. These points are solved for the parameters of the model. The result is verified in such a way that, sampled data points fit a predefined threshold. If the fraction of the number of inliers over the total number points in the set exceeds a predefined threshold, re-estimate the model parameters using all the identified inliers and discard. If that does not happen, the above procedure is repeated N times. The number of iterations, N, is chosen in such a way that, it is high enough to ensure the probability p (usually set to 0.99) that at least one of the sets of random samples does not include an outlier. Let u represent the probability that any selected data point is an inlier 1 and v = 1 − u the probability of observing an outlier. N iterations of the minimum number of points denoted m are required,
where $1 - p = (1 - u\ m)\ N$ (1) and thus with some manipulation, $N = \log(1 - p) \log(1 - (1 - v)m)$.

3.4 Image Warping

Image Warping is the process of digitally manipulating an image such that any shapes portrayed in the image have been significantly distorted. Warping may be used for correcting image distortion.



While an image can be transformed in various ways, pure warping means that points are mapped to points without changing the colours. This can be based mathematically on any function from part of the plane to the plane. If the function is injective the original can be reconstructed. If the function is a bijection any image can be inversely transformed. The last step is to warp and blend all the input images to an output composite mosaic. First we need to make out the output mosaic size by computing the range of warped image coordinates for each input image. As described earlier we can easily do this by mapping four corners of each source image forward and computing the minimum x, minimum y, maximum x and maximum y coordinates to determine the size of the output image. Finally, x-offset and y-offset values specifying the offset of the reference image origin relative to the output panorama needs to be calculated. The next step is to use the inverse warping as described above for mapping the pixels from each input image to the plane defined by the reference image. Information from multiple images of the same scene. These images may be captured from different sensors, acquired at different times, or having different spatial and spectral characteristics. The object of the image fusion is to retain the most desirable characteristics of each image. With the availability of multi-sensor data in many fields, image fusion has been receiving increasing attention in the researches for wide spectrums of applications. Warping may also be defined as a process wherein any distortion of the images of the same scene are removed accurately without any relevant information being removed or according to the requirement distortions may be added. It efficiently removes undesired information which results in a distortion of the image mosaic. It is an effective means of retaining the image quality by either adding or removing distortion.

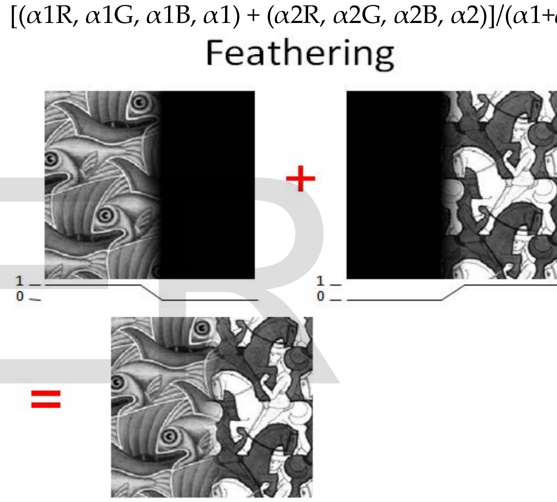


ISSN 2229-5518

3.5 Image Blending

Once we have registered all of the input images with respect to each other, we need to decide how to produce the final stitched image. This Blending procedure involves selecting a final compositing surface, for e.g., a flat surface or a cylindrical surface. Finally, a decision must be made on how to blend them in an order to create a mosaiced output. The first step to be taken is to decide on the way the final image should be represented. If only a few images are stitched together, a natural approach is to select one of the images as the reference and to then warp all of the other images into the reference coordinate system. The resulting composite is sometimes called a flat panorama. Since the projection onto the final surface is still a perspective projection, hence straight lines remain straight. There are many different projective layouts on which image stitching can be used, such as rectilinear projection, where the stitched image is viewed on a two dimensional plane intersecting the panosphere in a single point. Lines that are straight in reality are shown as straight regardless of their directions on the image. One case of rectilinear projection is the use of cube faces with cubic mapping for panorama viewing. It also shows the cylindrical projection where the stitched image shows a 360° horizontal field of view and a limited vertical field of view. Panoramas in this projection are meant to be viewed as though the image is wrapped into a cylinder and viewed from within. To build a cylindrical panorama, a sequence of images is taken by a camera mounted on a levelled tripod. If the camera focal length or field of view is known, each perspective image can be warped into cylindrical coordinates. Two types of cylindrical warping are forward warping and inverse warping. In forward warping, the source image is mapped onto cylindrical surface, but it can have holes in the destination image (because some pixels may never get mapped there). Therefore, we use inverse mapping where each pixel in the destination image is mapped to the source image. Since the mapping is unlikely to be exactly on the pixel values, bilinear interpolation is used to calculate the colours at the destination pixels. Once the source pixels have been mapped onto the final composite surface, the second step is to blend them in order to create the output mosaic or a panorama. If all of the images are in perfect registration and identically exposed, this is an easy problem (any pixel combination will do). There are many different pixels blending methods used in image stitching, such as feathering image blending, gradient domain and Image Pyramid blending. Feathering image blending is a technique used in computer graphics software to smooth or blur the edges of a feature; it is the simplest approach, in which the pixel values in the blended regions are, weighted average from the two overlapping images. We generally use alpha factor often called alpha channel having the value 1 at the centre pixel and becomes 0 after decreasing linearly to the border pixels. Where atleast two images overlap occurs in an output mosaic we will use the alpha values as follows to compute the colour at a pixel. Suppose there are two images, $I1$, $I2$, overlapping in the output image; each pixel (x, y) in image Ii (i=1,2) is represented as $Ii(x, y) = (\alpha iR, \alpha iG, \alpha iB, \alpha i)$ where (R,G,B) are the colour values at the pixel. We will compute the pixel value of (x, y) in the stitched output image with the help of the formula shown below:

$$[(\alpha 1R, \alpha 1G, \alpha 1B, \alpha 1) + (\alpha 2R, \alpha 2G, \alpha 2B, \alpha 2)]/(\alpha 1+\alpha 2).$$



Sometimes this simple approach doesn't work well (for example in the presence of exposure differences). But if all the images were taken at the same time and using high quality tripods, therefore, this simple algorithm produces excellent results. An alternative approach to multi-band image blending is to perform the operations in the gradient domain. Here, instead of working with the initial color values, the image gradients from each source image are copied; in a second pass, an image that best matches these gradients is reconstructed. Copying gradients directly from the source images after seam placement is just one approach to gradient domain blending. Another important approach of image blending is Image Pyramid blending; the image pyramid is actually a representation of the image by a set of the different frequency-band images (i.e. Hierarchical representation of an image at different resolution). Image pyramid provides many useful properties for many applications, such as noise reduction, image analysis, image enhancement, etc. Laplacian pyramid is an algorithm using Gaussian to blend the image while keeping the significant feature in the meantime. It downsizes the image into different levels (sizes) with Gaussian. Later, it expands the Gaussian in to the lower lever and subtracts from the image in that lever to acquire the Laplacian image. This Laplacian Pyramid is the true useful member of the image pyramid. Each layer of this pyramid is the band-pass image. We also see that even after its frequencies are calculated, the local features of the image are still present. The final step is to blend the pixels colours in the overlapped region.

4 Conclusion

Image mosaicing is a powerful tool for generating a larger view of the scene. The main focus of this report was to give a better understanding of a technique called Image Mosaicing using feature based technique i.e., by using SIFT algorithm to detect keypoints also known as features of input images. Images are then registered accordingly and then homography is computed using RANSAC algorithm, later images are blended together in order to give a seamless view.This technique can be utilized for the better interpretation of the scene in the fields like medial imagery etc where angle of image capturing is constrained by other real life limitation. A good and efficient methodology is used for image mosaicing but still further improvements can be done in the specialized sub-parts. Some of the sub-parts like image blending, image registration or interpolation can be more optimized and can be made more efficient.

5 Future Scope

With the advancement in the field of Computer Vision, the use of Image Mosaicing technique has been improved. Also due to the requirement of a greater field of view image mosacing has been proven to be effective.

This paper on Image Mosaicing using feature based technique puts together the knowledge of fundamental concepts required to apply Image Mosaicing. Along with all the fundamental aspects it also shows a detailed study on all the efficient intermediate steps. With this level of exposure to Image Mosaicing study, it can be extended further to the following given areas.

1.Application to digital video: Modern video imagery is, almost without exception, subject to some form of lossy compression. The current crop of compression methods, such as Motion-JPEG and MPEG, are based on block-wise transform encoding. The image is divided into typically 8x8 pixel blocks and each block is projected into an orthogonal basis, such as the Discrete Cosine Transform. Transform components with coefficients below a threshold are dropped, and the remaining coefficients are quantized, packed into space-efficient data structures, and finally compressed using a loss-less method. Decompressing/reconstructing the image is simply the inverse process, but the information lost in discarding and quantizing the original transform coefficients means that the decompressed image will not be identical to the original. Globally, the difference is hardly noticeable, but at the level of the individual blocks, the difference can be severe. The problem is illustrated in Figure, which shows an image before and after JPEG compression. Although globally the images appear very similar, the close-ups reveal that high-frequency information has been lost. Consequently, a major source of degradation in modern video images is due to lossy-compression. However, the compression/decompression operations could be thought of as providing an explicit generative model describing the transformation of the original high-resolution image into the degraded, decompressed image. Unlike the generative model used throughout this thesis, whose parameters involve the blur and sampling rate, the parameters of this modern model would depend on knowing which transform components were dropped in each block, and on the level of quantization applied. These values maybe recovered from the compressed image file itself. Such a generative model could be used as a direct replacement for the one described here, allowing super-resolution restoration from multiple transform-encoded images. [11]

Figure showing drawbacks of compression

3D Image Mosaicing :

In the proposed theory of Image Mosaicing approach for a large scale 3D-scene modelling, the computation of matching is efficiently distributed in 3 stages, Camera pose estimation, Image Mosaicing and 3D image reconstruction. [11]

References

[1] J.S. Bridle, "Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition," Neurocomputing—Algorithms, Architectures and Applications, F. Fogelman-Soulie and J. Herault, eds., NATO ASI Series F68, Berlin: Springer-Verlag, pp. 227-236, 1989. (Book style with paper title and editor)

[2] Photogrammetry – SIFT features and ransac (https://www.youtube.com/watch?v=oT9c_LlFBqs)

[3] SIFT - https://www.youtube.com/watch?v=NPcMS49V5hg

[4] R.C. Gonzales R.E. Woods. Digital Image Processing. Addison-Wesley Publishing Company, 1993

[5] A Survey on Image Mosaicing Techniques- International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 2, February 2013- Hemlatha Joshi(CSIT)

[6] Image Stitching based on Feature Extraction Techniques: A Survey: International Journal of Computer Applications (0975 – 8887) Volume 99– No.6, August 2014-Hazem Elbakry, Mohammed Elmogy, Ebtsam Adel

ISSN 2229-5518
[7] Vimal Singh Bind, P. R. (2013). A Robust Technique for Feature-based Image Mosaicing using Image Fusion.

[8] M.A. Fischler and R.C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 24(6):381–395, 1981.

[9] International Journal of Computer Applications (0975 – 8887) Volume 99– No.6, August 2014

[10] http://www.cse.yorku.ca/~kosta/CompVis_Notes/ransac.pdf

[11] Image Mosaicing and Super Resolution by David Peter Capel