Image Mosaicing with Invariant Features Detection using SIFT

By Jagjit Singh

Lovely Professional University, India

Abstract - There are situations where it is not possible to capture larger views with the given imaging media such as still cameras or video recording machines in a single stretch because of their inherent limitations. So to avoid such conditions a term Image Mosaicing comes into play. This Paper presents a complete system for mosaicing a group of still images with some amount of overlapping between every two successive images. Mainly the idea is to wrap up the overlapping areas within the group of images. Detection for the common area is done using common features by the help of feature extraction from the images. In this paper technique used for the feature extraction is SIFT which is used to extract invariant features which are stable in nature. Invariant features are those features of an image which does not change even after the scaling, rotation, or zooming, change in illumination of the image is done. Multiple level filtering and downsampling are the key factors of the SIFT. So the steps involved are feature detection, matching of stable features, wrapping up of features around those feature locations. Mosaicing part consists of two major part and those are transformation matrix and bilinear interpolation. Mosaiced images are full length images which consist of all the group images.

GJCST-F Classification: I.4.7

IMAGE MOSAICING WITH INVARIANT FEATURES DETECTION USING SIFT

Strictly as per the compliance and regulations of:

Image Mosaicing with Invariant Features Detection using SIFT

Jagjit Singh

Abstract - There are situations where it is not possible to capture larger views with the given imaging media such as still cameras or video recording machines in a single stretch because of their inherent limitations. So to avoid such conditions a term Image Mosaicing comes into play. This Paper presents a complete system for mosaicing a group of still images with some amount of overlapping between every two successive images. Mainly the idea is to wrap up the overlapping areas within the group of images. Detection for the common area is done using common features by the help of feature extraction from the images. In this paper technique used for the feature extraction is SIFT which is used to extract invariant features which are stable in nature. Invariant features are those features of an image which does not change even after the scaling, rotation, or zooming, change in illumination of the image is done. Multiple level filtering and downsampling are the key factors of the SIFT. So the steps involved are feature detection, matching of stable features, wrapping up of features around those feature locations. Mosaicing part consists of two major part and those are transformation matrix and bilinear interpolation. Mosaiced images are full length images which consist of all the group images.

I. Introduction

The basis of the image mosaicing technique is to find the common part in the two images which are going to be mosaiced. Different techniques are available but every technique contains some or the other flaw in it. So after the literature review, the technique best suitable for common invariant feature extraction is SIFT (Scale Invariant Feature Technique). Scale Invariant Feature Transform is an approach for extracting distinctive invariant features from images, and it has been successfully applied to many computer vision problems (e.g. face recognition and object detection) [1]. The features are invariant to image scaling, translation, and rotation, and partially invariant to illumination changes. This approach transforms an image into a large collection of local feature vectors, each of which is invariant to image translation, scaling, and rotation, and partially invariant to illumination changes. The scale-invariant features are efficiently identified by using a staged filtering approach. [2] The first stage identifies key locations in scale space by looking for locations that are maxima or minima of a difference-of-Gaussian function. Each point is used to generate a feature vector that describes the local image region sampled relative to its scale-space coordinate frame. The features achieve partial invariance to local variations, by blurring image gradient locations.

To achieve rotation invariance and a high level of efficiency, we have chosen to select key locations at maxima and minima of a difference of Gaussian function applied in scale space. This can be computed very efficiently by building an image pyramid with resampling between each level.

To generate the next pyramid level, we resample the already smoothed image B using bilinear interpolation with a pixel spacing of 1.5 in each direction. While it may seem more natural to resample with a relative scale of, the only constraint is that sampling be frequent enough to detect peaks. The 1.5 spacing means that each new sample will be a constant linear combination of 4 adjacent pixels. This is efficient to compute and minimizes aliasing artefacts that would arise from changing the resampling coefficients. [2]

Maxima and minima of this scale-space function are determined by comparing each pixel in the pyramid to its neighbors. First, a pixel is compared to its 8 neighbors at the same level of the pyramid. If it is a maxima or minima at this level, then the closest pixel location is calculated at the next lowest level of the pyramid, taking account of the 1.5 times resampling. If the pixel remains higher (or lower) than this closest pixel and its 8 neighbors, then the test is repeated for the level above.
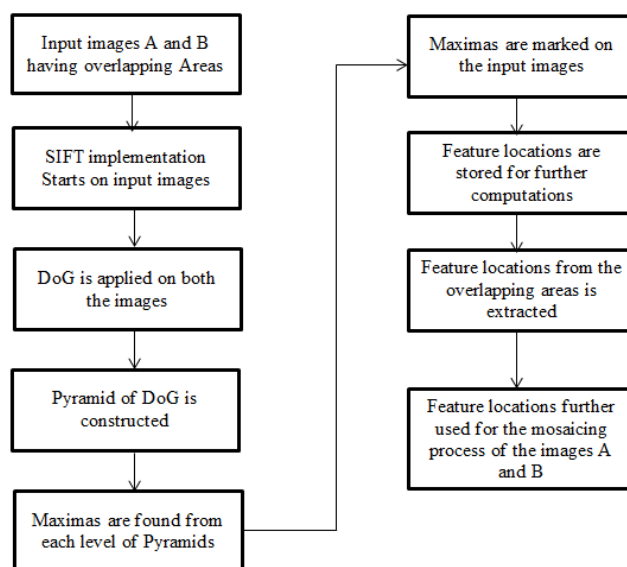

Figure 1 : Basic Implementation of SIFT and Mosaicing

Author : Lovely Professional University, Jalandhar, India.
E-mail : Jagjit.sh@gmail.com

II. Research Methodology

Steps involved in the image mosaicing using SIFT is as follow. Below represents the all the basic steps involved with the help of the basic flow charts.

a) Difference of Gaussian (DoG)

Difference of Gaussians is a feature enhance-ement algorithm that involves the subtraction of one blurred version of an original image from another, less blurred version of the original. In the simple case of grayscale images, the blurred images are obtained by convolving the original grayscale images with Gaussian kernels having differing standard deviations. Blurring an image using a Gaussian kernel suppresses only high-frequency spatial information. Subtracting one image from the other preserves spatial information that lies between the ranges of frequencies that are preserved in the two blurred images. Thus, the difference of Gaussians is a band-pass filter that discards all but a handful of spatial frequencies. Following flow chart further gives the detail description of the major steps involved in the DoG.
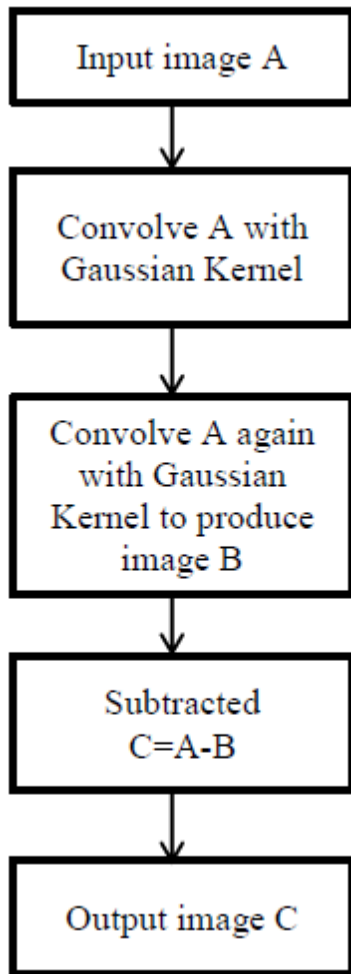
b) Pyramid of DoG

Pyramid or pyramid representation is a type of multi-scale signal representation developed by the computer vision, image processing and signal processing communities, in which an image is subject to repeated smoothing and subsampling. The Difference of Gaussian pyramid is generated from a single input image. The output is a pyramid of several images, each being a unique difference of Gaussian. To generate the pyramid, the input image is repeatedly blurred; the difference between consecutive blur amounts is then output as one Octave of the pyramid. One of the blurred images is down sampled by a factor of two in each direction, and the process occurs again with output in a different size.[3]. In Pyramid of DoG one of the important factors is the Downsampling of the image as to create the next level of the pyramid. In this algorithm the method used for the Downsampling of the images is the Bilinear Interpolation.

In Pyramid of DoG one of the important factor is the Downsampling of the image as to create the next level of the pyramid. In this algorithm the method used for the Downsampling of the images is the Bilinear Interpolation.
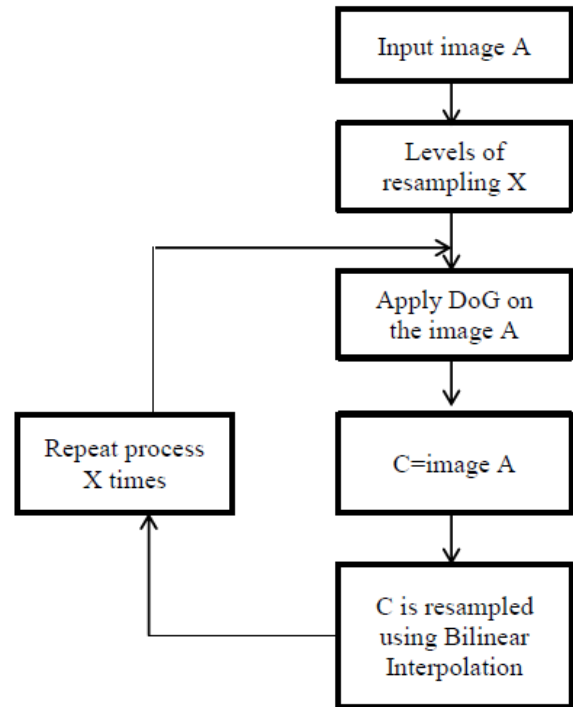

Figure 2 : Flow chart for DoG

Figure 3 : Flow chart of Pyramid of DoG

c) Bilinear Interpolation

Interpolation is a method of constructing new data points within the range of a discrete set of known data points. Bilinear interpolation is an extension of linear interpolation for interpolating functions of two variables. In this project Bilinear Interpolation is used to Downsample the images. It is mainly a combination of

the bilinear interpolation and the Nearest Neighbour Image Scaling approach. Nearest neighbour is the simplest and fastest implementation of image scaling technique. The principle in image scaling is to have a reference image and using this image as the base to construct a new scaled image. The constructed image will be smaller, larger, or equal in size depending on the scaling ratio. When enlarging an image, we are actually introducing empty spaces in the original base picture. From the image in figure, a n image with dimension (w1 = 4, h1 = 4) is to be enlarged to (w2 = 8, h2 = 8). The black pixels represent empty spaces where interpolation is needed, and the complete picture is the result of nearest neighbor interpolation. Scaling algorithm is to find appropiate spot to put the empty spaces inside the original image, and to fill all those spaces with livelier colors. For the nearest neighbor technique, the empty spaces will be replaced with the nearest neighboring pixel, hence the name. This results in a sharp but jaggy image, and if the enlarge scale is two; it would seems each pixel has doubled in size. Shrinking, in the other hand involves reduction of pixels and it means lost of irrecoverable information. In this case scaling algorithm is to find the right pixels to throw away.

Figure 4 : Example for bilinear implementation showing the basic process

d) Maximas Finding

Maximas in the given pyramid of DoG are determined by comparing each pixel in the pyramid to its neighbors. First, a pixel is compared to its 8 neighbors at the same level of the pyramid. If it is a maxima or minima at this level, then the closest pixel location is calculated at the next lowest level of the pyramid, taking account of the 1.5 times resampling.[3] If the pixel remains higher (or lower) than this closest pixel and its 8 neighbors, then the test is repeated for the level above. Since most pixels will be eliminated within a few comparisons, the cost of this detection is small and much lower than that of building the pyramid.
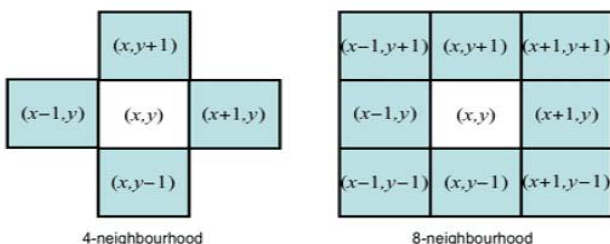

Figure 5 : Figure showing the neighbouhood of a pixel

In this algorithm the maximas are found by simply first comparing the intensity value of a pixel by 8-pixels around that pixel. Neighboring pixels are North, South, West, East, North-east, North-west, South-East, South-West. In the next step the maximas which are extracted from the first scale space are then compared by the maximas extracted from the one level below the scale space. Likewise the maximas are extracted from all the images which are included in the pyramid of DoG.

i. *Mosaicing*

Image mosaics are collection of overlapping images together with coordinate transformations that relate the different image coordinate systems. By applying the appropriate transformations via a warping operation and merging the overlapping regions of warped images, it is possible to construct a single image covering the 25 entire visible area of the scene. This merged single image is the motivation for the term-mosaic. Image mosaicing can be done in a variety of ways. There are many algorithms to do image mosaicing. The algorithm does require effective corner matching. Usually, the algorithms differ in the Image registration process.

After observing the problem domain that is the specifications of the area on which we are working, as simulation of these basic fundamentals is essential for thorough understanding of the algorithm [4]. We chose transformation and Bilinear Interpolation in estimating the missing values.

In case of the image mosaicing, the images specified need not to be necessary that they are aligned as per the images provided. For this fact the proper transformation is performed on each and every image to make it properly aligned as per the first image.

Mainly in dealing with this kind of implementation of algorithm, a common vector is found out. So that the multiplication of that vector with the image aligns it perfectly. Image warping is vector technique used in these algorithms.

III. Results and Conclusions

SIFT implementation is carried out on the test image Test-1.

Figure 6 : Test-1

First step in the SIFT is the DoG of the given Image. On applying the DoG the results are as follow.

Figure 7 : Implementation of DoG

Now the next step in SIFT is the creation of the Pyramid of DoG. Following are the desired output when Pyramid is created from the test image.
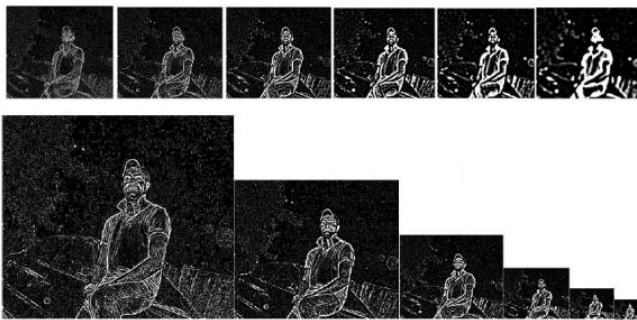

Figure 8 : Pyramid of DoG on the Test image-1

Above are all the desired output that we get after downsampling and then implementing DoG after every Downsampling. The results are quite good and stable for most of the images.

Now the Maxima finding comes after the implementation of the DoG and the Pyramid of the DoG. The outputs are as follow for the maxima finding. First maximas are found within the same space scale and then compaired by the maximas of the one level above it and then with the one level down, so at the end we are left with the maximum number maximas which are quite stable.
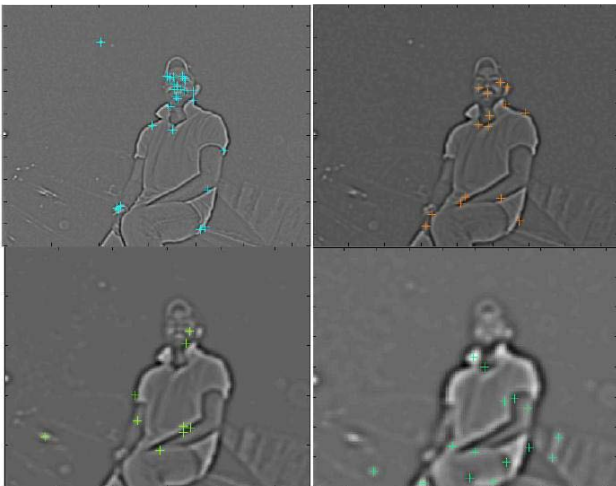

Figure 9 : Maxima finding on the different space scale

After this the marking on the test image is as follow.

Figure 10 : Maxima marking on the original image (Test-1)

Mosaicing part contains the manual detection of the feature locations which are common between the groups of images and then wrapping of the images is done across that point. A new image is generated which is a stich of the two images. The output are as follow.

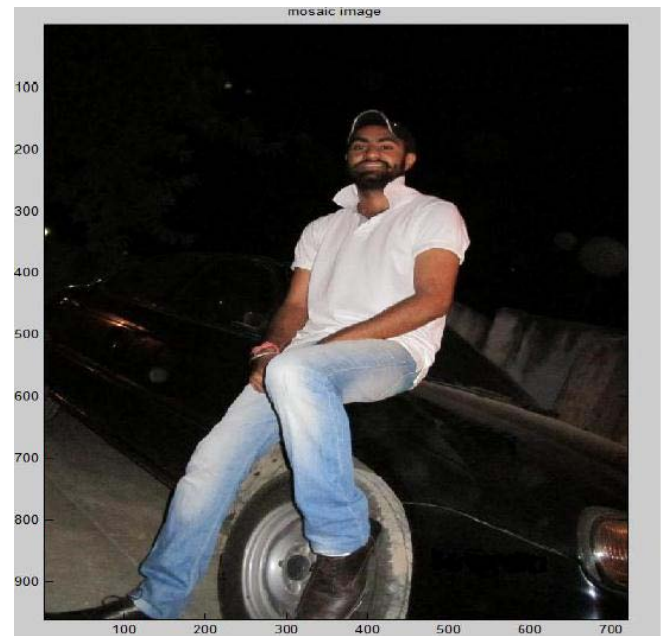*Figure 11 :* Images which are going to be mosaiced

Figure 12 : Resulting image after the mosaicing process (Mosaiced Image)

Above is the final image mosaiced after the whole process of sift and image wrapping.

IV. Conclusion

We presented the implementation of SIFT algorithm for the feature detection for the given image. First the Difference of Gaussian for the test image is found and then pyramid of Difference of Gaussian is created for the desired levels of scale space. DoG is applied and then the Downsampling is done. As we downsample the DoG image ,we get more prominent and less no of features which further helps in less computation and saves time also. Our main aim of using SIFT was that SIFT features are invariant and after implementation we can find that most of the features detected even after the change in the angle and illumination of camera are same if compared with the same image under normal conditions. Maximas that we get for a particular image are constant and invariant to image scaling, translation, and rotation, and partially invariant to illumination changes. The feature locations are picked up from the overlapping areas manually and used for further processing. Mosaicing part which we implemented uses transformation matrix made from the feature locations that are picked up and then inverse mapping is done to create a mosaic with the help of the bilinear interpolation. The mosaics which we are getting are quite good in quality. Illumination and angle of camera does some time affects the mosaic of the images but still the results are worth computing.

References Références Referencias

1. Qi Zhang, Yurong Chen and Yinlong Xu, -SIFT Implementation and Optimization for Multi-Core Systems, Dept. of Computer Science, Univ. of Science and Technology of China.
2. David G. Lowe, ―Object Recognition from Local Scale-Invariant Features', Computer Science Department University of British Columbia Vancouver, B.C., V6T 1Z4, Canada.
3. Ballard Blair and Chris Murphy, ―Difference of Gaussian Scale Space Pyramids for SIFT Feature Detection, Stanford University Department of Computer Science.
4. Soniya Hingu, asst. prof. Narendra. M. patel, image mosaicing using euclidean warping, department of computer engineering, birla vishvakarma vidyalaya, cvm, v.v.nagar.
5. Soo-Hyun CHO , Yun-Koo CHUNG, Jae Yeon LEE,- Automatic Image Mosaic System Using Image Feature Detection and Taylor Series, Proc. VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydney.
6. D. Parks and J. P. Gravel, ―Corner Detection Georgetown University Medical Center.
7. Matthew Brown and David G. Lowe―Automatic Panoramic Image Stitching using Invariant Features, International Journal of Computer Vision 74(1), 59–73, 2007_c 2007 Springer Science + Business Media, LLC. Manufactured in the United States. DOI: 10.1007/s11263-006-0002-3 .
8. Udhav Bhosle, Subhasis Chaudhuri, Sumantra Dutta Roy, ―A Fast Method for Image Mosaicing using Geometric Hashing.
9. Prabhakara Rao G.V, Mahidhar A, ―A Novel Still Image Mosaicing SystemUsing Featureless Registration, Binary Check Stitching and Minimal Blending.
10. Jayachandra Polisetty, Surya Prakash Noolu, Jaya Prakash Varma Champati, P Venkata Srikant Reddy, Ravindranadh Gunnam, ―Image Mosaicing Using Xor Correlation and Fourier Shift Theorem.
11. Sandeep Ghael & Gregory Chew, ―Creating Image Mosaics, EE392J Ð Digital Video Processing (Winter 2000).
12. Kevin E. Loewke, David B. Camarillo, Christopher A. Jobst, and J. Kenneth salisbury, ―"Real-Time Image Mosaicing for Medical Applications".