International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 12 | Dec 2019 www.irjet.net p-ISSN: 2395-0072

An Approach to FPGA based Implementation of Image Mosaicing using Neural Networks

Dhairya Gada[1], Gagan Agarwal[2], Kenil Gadhiya[3], Darshana Sankhe[4]

Students [1] [2] [3], Professor [4], Dept. of Electronics Engineering, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

---

Abstract – *Panoramic images provide an immersive experience to the person viewing the image. The creation of an image mosaic from a sequence of overlapped views is the most widely used method of obtaining a larger view of a scene than that available with a single view. It has been a subject of research for many years, and recently its demand for the application in Digital 360 Camera Systems has increased. However, mosaicing images requires extensive computational power. In this paper we propose an FPGA based approach for Image Mosaicing. To further improve the process of Image Stitching we will be using Neural Networks for the same.*

***Key Words*: Image Stitching, Image Mosaicing, FPGA, Artificial Intelligence, Neural Networks**

1. INTRODUCTION

A Panoramic image is fundamentally a wide-angle depiction of space. And recently many applications of such wide-angle images have emerged. For instance, virtual reality scenarios that are inspired by real world, virtual tourism, surveillance of a wide field etc.

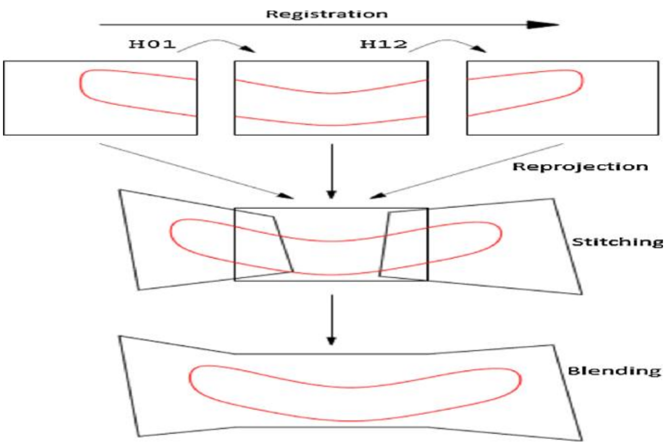

Fig-1: Steps of Image Mosaicing.

Image mosaicing is one of the techniques of image processing which is useful for tiling multiple digital images to create panoramic images. Mosaicing is stitching together of several arbitrarily shaped images to form one large balanced image so that the boundaries between the original images are not seen. As shown in Fig. 1, mosaicing involves various steps of image processing: registration, reprojection, stitching, and blending. [1]

In registration, the overlapping regions are identified. Then, the different images are re-projected on a single coordinate frame. The images are then stitched and the image boundaries are removed by a process called blending.

The registration process requires features to be extracted from the image. The features extracted from an image using conventional image processing techniques are not always relevant and may also include noise. Therefore, a properly trained convolutional neural networks can be used in image mosaicing to extract the feature points that are more accurate.

The entire process of Image Mosaicing is computationally intensive and operating on high resolution images is time consuming [2]. Since image processing operations basically involve working with pixel values that are arranged in a matrix structure, the individual pixel values or multiple group of pixel values in an image can be operated upon parallelly.

This parallelism can be leveraged by using parallel processors or FPGAs. Use of such hardware systems can greatly improve the timing efficiency of the mosaicing process [3,4].

FPGA (Field Programmable Gate Array) is a reconfigurable hardware that can be programmed to realize application specific hardware. Therefore, an image processing algorithm can be implemented on FPGA such that all the logic blocks operate on various sub-parts of the image and the combination of all their outputs deliver the final resulting image.

2. Proposed System Model

The basic structure of the system is elucidated in this section, with an overall system block representation and an explanation for each relevant subsystem involved in the proposed scheme.

The system proposed can be represented by a block diagram as shown below, in which the blocks are devised and named as per the proposed function of each block. The block diagram of the system makes use of 5 primary components:

- Obtaining Input Images
- Pre-Processing
- Image Stitching
- Post processing
- Stitched Output Image

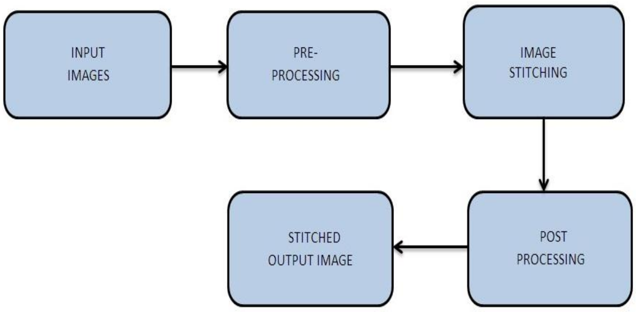

Fig-2: Block Diagram Representation of the System

2.1 Input Images

A set of images are given as input to the system. The images are such that they have around 40% of overlapping sections in which matches would be established to merge the images.

2.2 Pre-Processing

Pre-processing is done for the improvement of the image data that suppresses unwanted distortions or enhances some image features for further processing process. In this stage, the images are transformed to extract features which can then be used as inputs to a neural network.

2.3 Image Stitching

It is the process of combining multiple images with overlapping fields of views to produce a segmented panoramic image. Image stitching enables the combinations of multiple images to create a large picture that is beyond the aspect ratio and resolution of the camera's individual shots. At this stage, the features extracted by the transform will be matched using a trained neural network to identify the overlapping regions.

2.4 Post Processing and Stitched Output Images

In this stage, the matched features are transformed into a continuous image by proper pixel value allotment at the overlapping regions. The final output image is the stitched version of the set of input images with the common sections merged together.

3. System Implementation

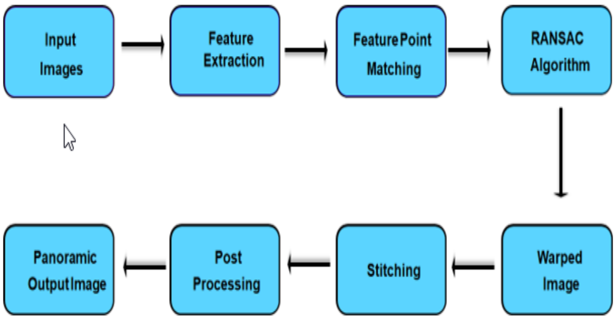

Fig-3: Process Flow

3.1 Input Images

To test the approach of Image mosaicing using Neural Networks, we used a set of 3 images such that they had at least 40% overlap between them.

Fig-4: Input Images

3.2 Feature Extraction

Feature Extraction is a type of reduction in the dimension such that it efficiently represents interesting parts of an image as a compact feature vector. This is done with the help of VGG-16 neural network. The outputs of specific layers of the network are used to create descriptors for the keypoints in the images. [5]

3.3 Feature Matching

Feature Point Matching means finding similar matching points between two images. The descriptors calculated in the previous step are used for this purpose. Euclidean distance is calculated for 2 keypoints using their respective descriptors and the points with the shortest distances are considered as matchpoints which are further used for the calculation of homography matrix

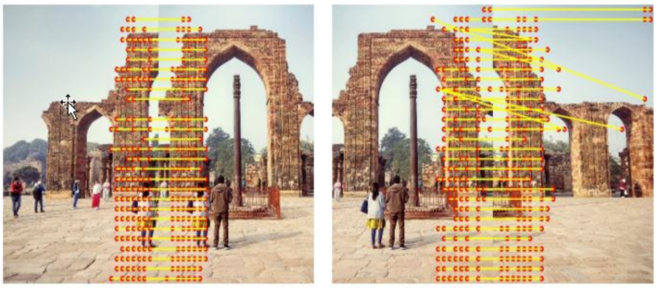

Fig-5: Keypoint Matches

3.4 RANSAC Algorithm

We make use of the RANSAC algorithm [6] to determine the Homography Matrix. Homography gives the relation between the similar keypoints in the two images. Homography matrix of the second image with respect to the first image is used to warp the second image into a global coordinate frame with respect to the first image.

$$\begin{bmatrix} x2 \\ y2 \\ 1 \end{bmatrix} = \begin{bmatrix} h1 & h2 & h3 \\ h4 & h5 & h6 \\ h7 & h8 & 1 \end{bmatrix} \begin{bmatrix} x1 \\ y1 \\ 1 \end{bmatrix}$$

Fig-6: Representation of Homography Matrix

H is the 3 x 3 Homography Matrix where h3 is translation on X-axis, h6 is Translation on Y-axis, h1, h2, h4 and h6 represent Scaling and Rotation and h7 and h8 stand for Horizontal and Vertical Distortion.

Ideally, we can use Direct Linear Transformation on 4 Best Match Points to determine the homography. However, since the matches are not always perfect and there are chances of imperfect matches a robust algorithm such as RANSAC is used.

RANSAC stands for Random Sample Consensus. In RANSAC multiple iterations are performed over a number of randomly selected feature-matched pairs to produce a homography matrix that best-fits all the matched pairs. After a number of iterations, the homography matrix that satisfies the major number of pairs is taken as the final result.

RANSAC involves the following steps:

- Randomly selecting a subset of the data set.
- Fitting a model to the selected subset.
- Determining the number of outliers.

3.5 Warped Image

Image warping is the process of digitally manipulating an image such that any shapes portrayed in the image have been significantly distorted. It is used for correcting image distortion as well as creative purposes (e.g., morphing). The images are multiplied with the homography matrix calculated in the previous step. This converts the individual images from a local coordinate system to a global coordinate system.

Fig-6: Images (b) and (c) in a global coordinate system w.r.t image (a)

3.6 Stitching

In this step, a pixel value is assigned to each coordinate in the global coordinate system thus making a continuous image out of the 3 input images

Fig-7: Images (b) and (c) in a global coordinate

3.7 Post Processing and Output Image

Post processing involves cylindrical warping, straightening, bundle adjustments, blending and cropping. These all processes collectively eliminate the effect of different light exposures and geometries in the input images. The output image thus produced does not have any unwanted input image edges visible, thus producing a seamless panoramic image.

Fig-8: Output Image

4. Proposed Approach to FPGA based implementation

It is observed that the Image Processing Operations along with the computations performed by the Neural Network are iterative in nature and thus time consuming when running these processes on a conventional computer which are sequential in nature [2]. Furthermore, the processing time to obtain the final output increases as the number of images increase.

Thus, we propose to implement this system on an FPGA device that inherently supports parallel computation on the Image Matrices as well as neural network operations.

This solution can be approached by using Vivado HLS [7] to create custom IP blocks for the necessary operations. Followed by the used of Vivado Design Suite [8]to create the complete block diagram of the entire system. This involves defining the port to be used for communication between the FPGA board and the PC and configuring the connections to be used. Vivado converts the block diagram into Verilog and generates a .bit file which is then used to program the FPGA

5. Results

Artificial Intelligence and image processing techniques are thus used to get a mosaiced image from the multiple input images with a certain number of overlapping regions.

6. Conclusions

The concept of image mosaicing provides wider angle images without the use of high-end equipment. However, it is a computationally heavy task and traditional hardware fails to meet deadlines in case of real time mosaicing. Hence, the use of Neural Networks and FPGAs is justified to speed up the process by making use of their inherent parallelism.

The proposed approach finds applications mainly in the fields of enhancing consumer experience and security. For e.g., viewing a hotel room before booking or watching a sporting event. Also, a 360° CCTV would provide much better coverage than a normal one. One possible defence application includes getting live 360° feed from an unmanned vehicle's POV. Such images can be used to map regions and landscapes. Use of 360° videos can revolutionize the entertainment industry by providing a much richer immersive experience to the viewer.

REFERENCES

[1] Soo-Hyun CHO, Yun-Koo CHUNG and Jae Yeon LEE, "Automatic Image Mosaic System Using Image Feature Detection and Taylor Series," in Proc. Of the 7th International Conference on Digital Image Computing: Techniques and Applications, Sydney, Australia,pp.549-556, 2003

[2] P. Athanas & A. Abbott, Addressing the Computational Requirements of Image Processing with a Custom Computing Machine: An Overview, Workshop on Reconfigurable Architectures, IPPS '95, 1-15, 1995

[3] U Ramacher et al, A 53-GOPS Programmable Vision Processor for Processing, Coding-Decoding and Synthesizing of Images, Proc. 27th Eur. Solid-State Circ. Conf., Villach, Austria, 2001, 160-163

[4] T Minami et al, A 300-MOPS Video Signal Processor with a Parallel Architecture, IEEE J. of Solid-State Circ., 26(12), 1991, 1868-1875.F. Yang, M. Paindavoine, "Implementation of an RBF neural network on embedded systems: real-time face tracking and identity verification," in IEEE Transaction on Neural Networks, vol. 14, no. 5, pp.1162-1175, 2003

[5] Z. Yang and T. Dan and Y. Yang, "Multi-Temporal Remote Sensing Image Registration Using Deep Convolutional Features," in IEEE Access, 2018

[6] Martin A. Fischler and Robert C. Bolles "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography"

[7] "Vivado HLS" Vivado High Level Synthesis. [Online]. Available: https://www.xilinx.com/products/design-tools/vivado/integration/esl-design.html. [Accessed: 30-Nov-2019]

[8] "Vivado Design Suite," Vivado Design Suite. [Online]. Available: https://www.xilinx.com/products/design-tools/vivado.html. [Accessed: 30-Nov-2019]