

Image Mosaicing with Global Affine Feature Estimation based on Diamond Search

Rosida Vivin Nahari

Faculty of Engineering, Trunojoyo University
Bangkalan, Indonesia

Abstract - Image mosaicing is the process of presenting a complete image from a sequence of smaller images or video frames. Most of existing algorithms only focus on capturing static scenes. This research develops a mosaicing technique for cases where there are moving objects in input frames. There are three stages in this system. The first one is the preprocessing stage, which yields local motion and global affine estimates. The second stage is the registration of image mosaics, and the last one is the integration of image mosaics. Some trials in estimation stage have been conducted and the results show that the diamond technique can be used to initialize a background region with the average PNSR 26.5116 dB and the computation time 8.41 seconds. The selection of global affine features from the background mask region has a great influence on the result of mosaic registration since it can decrease the key point selection error of image joining in the region of moving objects.

Keywords : *image mosaicing, Diamond search, global affine features, moving object*

I. INTRODUCTION

In recent years, image mosaicing have been an active area of research in the fields of computer vision, photography, digital image processing, and computer graphics. The applications of image mosaicing include: the construction of satellite photographs and aerial mapping, photo editing, and various creative works in virtual environments. Image mosaicing is a technique to extend fields of view (FOV) by means of presenting images based on registration method from a sequence of images. In the field of video coding, image mosaicing is a technique to present a video sequence in the scene based format and remove redundant information.

Traditional techniques of image mosaicing were firstly used with static input frames, i.e. it only joins images without considering the motion of moving objects in the frames and the one due to the camera during image capturing [1]. Recent studies have developed dynamic mosaicing technique, in which it does not only consider the improvement of registration in the background, but also consider the image registration of moving objects [15]. The detection of background and foreground regions can be carried out by estimating local and global motions [5].

The feature detection of background region plays important role on the results of image mosaicing [12]. In [15]

studied the feature-based image mosaicing to produce a background model using features obtained from block-based detections using the full search method. The drawback of this method is that it needs long computation time [9]. The development of hierarchical search method improves this computation time, compared to the full search method [3]. However, the hierarchical technique might discard important information from some pixels and therefore the best feature might also be discarded [18]. The trade-off solution between the computation time and the quality of the feature block-based detection for full search is developing patterns from block search to diamond search [18].

The resulting estimated features of a background region obtained from the full search method still contain features which are supposed to be foreground's, due to the global motion of the camera [3]. This error can be minimized by compensating the predicted frames resulting from local motion estimates using global affine technique [6].

This research uses the features obtained from the diamond search method which initialize the global affine stage. Therefore, it is expected that more appropriate features can be obtained for the image mosaic registration stage.

II. IMAGE MOSAICING

Image mosaicing is the process of joining several images to produce a panoramic or larger image [8]. Mosaicing can also be defined as the process of presenting a complete image from a sequence of images or video frames [3]. On the other hand, image stitching is the technique of combining several images sequentially to produce a composite image [11].

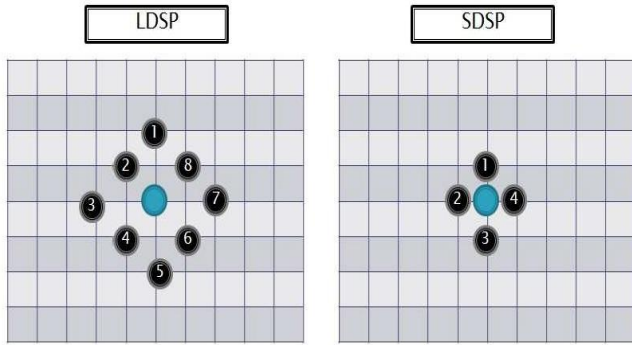


Figure 1. The process of motion vector search using Diamond search

2.1 Local Motion Features

The diamond pattern, according to [18], is a more efficient estimator. Compared to the *Block-based Gradient Descent Algorithm*, the Large Diamond Search Pattern (LDSP) is able to find large motion blocks with few search points while Small Diamond Search Pattern (SDSP) is able to find small motion vectors (stationary ones). Patterns of LDP and SDSP are depicted in Figure 1.

The process shown in Figure 2 is an initial search of motion vectors; blocks denoted by point indices 1-7 will be checked which one has the smallest SAD. Figure 2, illustrates that the 4th point has the minimum value. Therefore, the search process using LDSP is repeated.

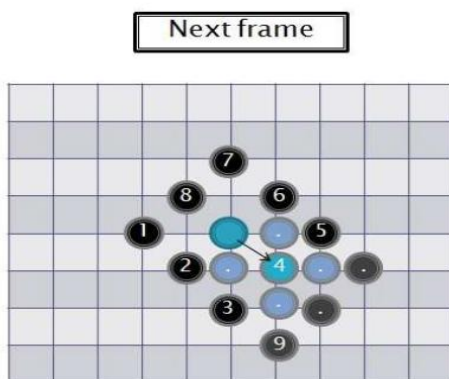


Figure 2. Searching Motion vector with Diamond Search

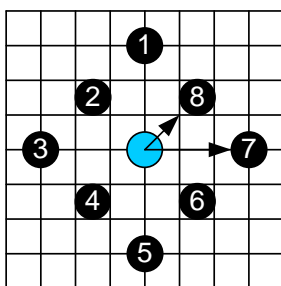


Figure 3. Initial process of Diamond search

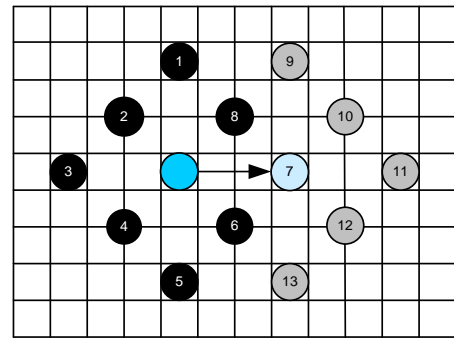


Figure 4. The process of Diamond search per key point

The searching process using diamond pattern is depicted in Figure 3. Point “7” has minimum SAD in the initial step and five new search points, i.e. “9”, “10”, “11”, “12” dan “13”, are formed to be checked in the next steps.

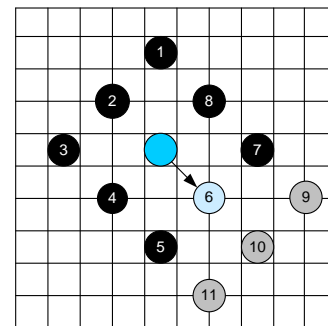


Figure 5. The process of Diamond search per edge point

2.2 Global Affine Features

Features obtained from the local motion estimates are only the motions of moving objects, therefore the moving regions due to camera motions have to estimated first before they are chosen as the key points for image registration. The estimation of global motions can be done using transformation models such as translation, rigid, affine, and projective transformations [3].

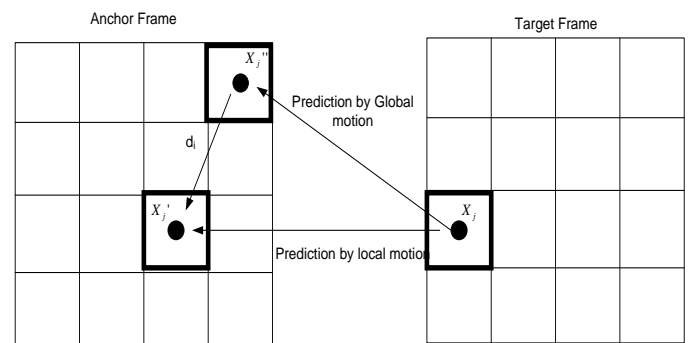


Figure 6. The geometric relationship between motion vector obtained from local motion and from global motion

2.3 Image Registration

The geometric registration of image mosaicing inputs is used to align the position of an image to other images or to transform the coordinate of an image to others' by minimizing the error function of least square method used to find the motion parameters. Image mosaicing consists of three registration steps, i.e. forward homography calculation, bounding box construction, and backward homography calculation. The forward homography matrix is a 2D perspective model with 8 parameters used for mosaicing [14].

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -u_1x_1 & -u_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -u_2x_2 & -u_2y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -u_3x_3 & -u_3y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -u_4x_4 & -u_4y_4 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -v_1x_1 & -v_1y_1 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -v_2x_2 & -v_2y_2 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -v_3x_3 & -v_3y_3 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -v_4x_4 & -v_4y_4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} \quad (1)$$

2.4 Image Mosaic Integration

The steps to construct mosaics using destination scan method are described in the flowchart below.

1. Do processes 2-7, repeat it with condition $x = 1$ up to $x = \mathbf{bw}$
2. Do processes 3-7, repeat it with condition $y = 1$ up to $y = \mathbf{bh}$
3. Find point (u,v) , that is the point (x,y) in the original image which are warped into mosaic image using backward homography \mathbf{H}_{BS} .
4. Determine the value of c , that is the color of (u,v) in the original image. If (u,v) is located outside the original image, then there is no color value entered into c .
5. Repeat step 3 and 4 for destination image using backward homography \mathbf{H}_{BD} .
6. Compute the final color of mosaic image based on the color obtained from the two input images. The final color is denoted by \mathbf{C}_{final}
7. Set the color of mosaic image in point (x,y) to \mathbf{C}_{final}

2.5 Blending

Blending is a stage of combining the colors of the first and the second images in the overlapping region. This research uses weighted average method by considering the average weight of the color intensity, based on the distance between the pixel coordinate and the image center [2]. The weight of each pixel is determined using Eq. (2).

$$w(x, y) = \left(1 - \left| 1 - \frac{2x}{lebar} \right| \right) \left(1 - \left| 1 - \frac{2y}{tinggi} \right| \right) \quad (2)$$

Next, the pixel color intensity in overlapping region is computed using Eq.(3).

$$I(x, y) = \frac{\sum_i w_i(x, y) f_i(x, y)}{\sum_i w_i(x, y)} \quad (3)$$

i : The index of input image having overlapping region.

$I(x, y)$: The pixel intensity of image mosaics in overlapping region.

$f_i(x, y)$: The pixel intensity of the i^{th} image input.

III. RESEACRH METHODS

This section describes the research steps. The first one is the preprocessing process of mosaicing, i.e. to find appropriate features in the background region by constructing background mask, using hexagon search with the improvement of global features through affine transformation. Next, the key points of background mask are detected for the image mosaic registration process.

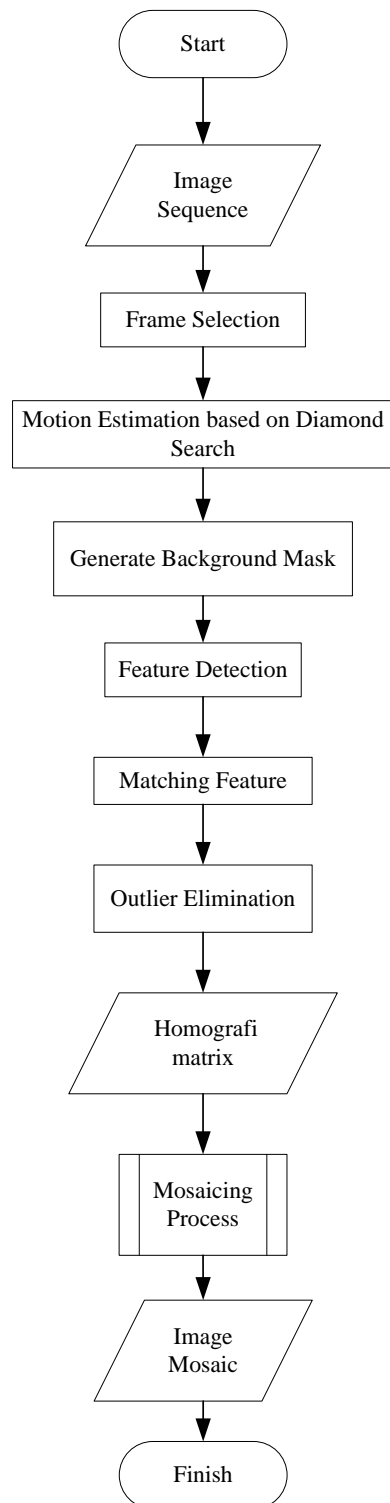


Figure 7. The main design of feature detection for mosaicing process

After obtaining the projection parameters, all images can be warped according to their coordinates. The transformation parameters of each frame are mapped into the reference coordinate frame by combining their transformation matrices. This research chooses the first frame as the reference frame and warps all other images into the coordinate of this first image. Therefore, information concerning camera motions is

useful to construct the dynamic mosaic from background, which is well-integrated and is saved in a single image.

IV. RESULTS AND DISCUSSION

4.1 Research data

This research uses “coastguard” video with QCIF format (176×144) . Not all frames in “coastguard” video are used in testing process. The choice of frames is based on the large similarity measure of frames in order to reduce the computation time since the selected frames can be a representation of other frames.

Then, the background regions of the above frame data are determined to enable the registration stage of image mosaics. The mosaicing process only uses anchor frames. The results show that the average PSNR of diamond search is 26.5116, better than that of EBMA search, which is 26.2596, and that of HBMA search, which is 26.0270.



Figure 8. Input images of “coastguard” sequence: the 5th and 20th frame

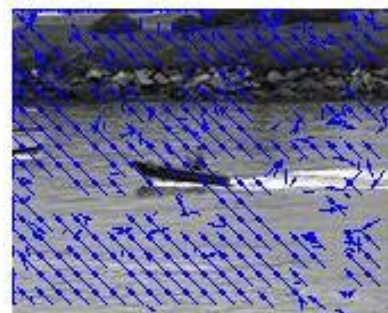


Figure 9. Motion vector of diamond search



Figure 10. Motion prediction using diamond search



Figure 11. Image differences using diamond search in the 5th and 20th frame



Figure 12. Morphology operation for image differences from diamond search



Figure 13. Detection of foreground and background regions

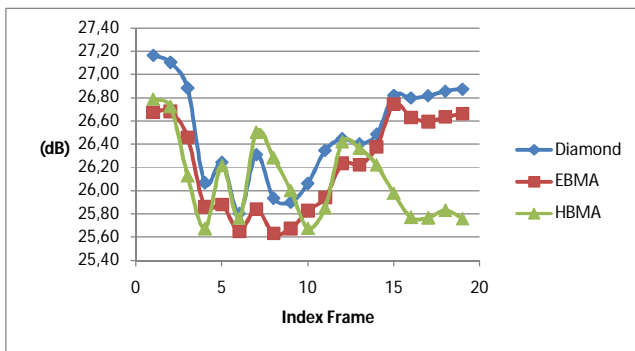


Figure 14. MSE comparison of diamond, EBMA, and HBMA searches

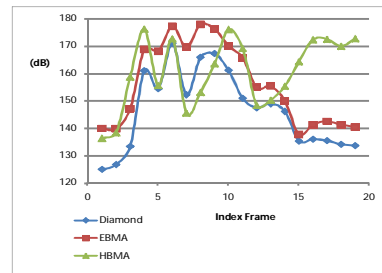


Figure 15. PSNR comparison of diamond, EBMA, and HBMA searches

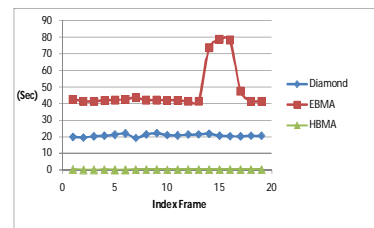


Figure 16. Computation time comparison of diamond, EBMA, and HBMA searches

Hierarchical-based search yields the lowest computation time, that is 0.2631 seconds in average. The next rank belongs to the Diamond technique, which is 8.4122 seconds in average. The EBMA technique needs the longest computation time, namely 47.1528 seconds in average. The pattern based search technique has the computation time 20.8001 seconds in average.

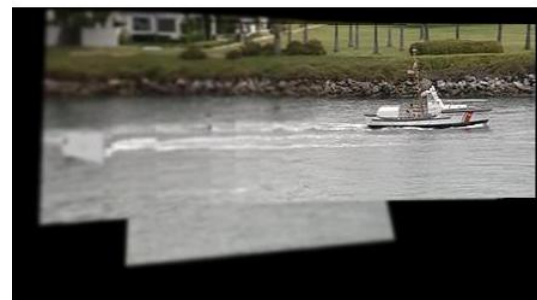


Figure 17. The result of mosaicing 20 frame

V. CONCLUSION

Based on the research results, the following conclusions are obtained:

1. The PSNR of diamond search is 26.5116 in average, better than the PSNRs of the two other methods, i.e. EBMA method yields PSNR of 26.2596 while the lowest quality is obtained from HBMA method with PSNR of 26.0270. Therefore, the use of diamond search is very suitable for obtaining the motion vectors of the motion estimates, as the first initialization for the global motion estimates; moreover, the efficiency of computation time is 8.41 seconds in average.
2. The estimation of global motion after the local motion process by affine transformation is used to correct the foreground detection error due to the camera motion.
3. The use of background mask for feature extraction and mosaicing processes is able to reduce the selection error of the best feature points, used for image registration. On the contrary, the mosaicing results without using background mask sometimes select features which are part of moving foreground.
4. The appropriate chosen of frames is useful to speed up the mosaicing process of image sequences without involving the whole frames.
5. The implementation of motion estimates is proved to be useful in constructing image mosaics.

Further development of results in this research need to be undertaken so that it can become a preprocessing method for other applications. It is suggested that the system can be developed using faster and more accurate segmentation process in obtaining the background mask from a sequence of images. It is expected that the selection of transformation can be developed for more accurate global motion in order to reduce noises due to camera motions.

REFERENCES

- [1] David G. Lowe.(1999).*Object recognition from local scale-invariant features. International Conference on Computer Vision*, Corfu, Greece. pp. 1150-1157.
- [2] Harris, C. and M. Stephen, (1988), *A combined corner and edge detection*. In M. Matthews, M. editor, Proceedings of the 4th ALVEY vision conference, pages 147-151, University of Manchester, England.
- [3] Hsu, C.T and Y.C. Tsan.(2004). Mosaics of video sequences with moving objects. *Signal Processing: Image Communication* 19, pages 81-98.
- [4] Ishfaq, Z. Weiguo, L. Jiancong and L. Ming, (2006). A Fast Adaptive Motion Estimation Algorithm, *IEEE Transactions On Circuits And Systems For Video Technology*, Vol. 16, NO. 3.
- [5] Jorge Badenas, JoseH Miguel Sanchiz, Filiberto Pla. (2001). *Motion-Based Segmentation And Region Tracking In Image Sequences*. Pattern Recognition.
- [6] Kunter, M.(2008).*Advances in Sprite-based Video Coding*. Dissertation. Technische Universitat Berlin.
- [7] Lee, C.-H. and L.-H. Chen,(1997). "A Fast Motion Estimation Algorithm Based on the Block Sum Pyramid," *IEEE Trans. Image Processing*, vol. 6, pp. 1587-1591.
- [8] Levin, A., Assaf Zomet, Shmuel Peleg, and Yair Weiss.(2004). *Seamless Image stitching in the gradient domain*, in European Conference On Computer Vision, vol. 4, pp. 377-389.
- [9] Liu Lei, Wang Zhiliang, Liu Jiwei, Cui Zhaohui, (2009). "Fast Global Motion Estimation," *Proc. of 2nd IEEE Int. Conf. on Broadband Network & Multimedia Technology (IC-BNMT 09)*, pp. 220-225, Beijing.
- [10] Martin A. Fischler and Robert C. Bolles (1981). "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". *Comm. of the ACM* 24 (6): 381-395.
- [11] Nielsen, F.(2000). Randomized Adaptive Algorithms for Mosaicing Systems, *IEICE Transactions of Information & Systems*, vol.E83-D, no.7.
- [12] Rocha, A, Ferreira, R, and Campilho,A, (2000). *Image Mosaicing using Corner Detection*, *SIAR2000-V Ibero-American Symposium on Pattern Recognition*.
- [13] Szeliski, R., (1996). *Video mosaics for virtual environments*. *Computer Graphics Applications* 16 (3): 22-30.
- [14] Szeliski, R and Shum, H,(1997), *Creating Full View Panoramic Image Mosaics and Environment Maps*, In *Proc. of SIGGRAPH*, , pp.251-258
- [15] Shen, H. (2004). *Moving Object Reconstruction on Background Mosaics of Dynamic Video Sequences*.Thesis. School of Computing at the National University of Singapore.
- [16] Vinukonda, P.(2011), *A Study Of The Scale-Invariant Fitur Transform On A Parallel Pipeline*, Jawaharlal Nehru Technological University, Hyderabad.
- [17] Zhang, D. and G. Lu. (2001).*Segmentation of moving objects in image sequence: A review. Circuits, Systems, and Signal Processing*, 20(3):143 - 183.
- [18] Zhu Ce, L.Xiao Lin, And Lap-Pui Chau.(2002).*Hexagon-Based Search Pattern For Fast Block Motion Estimation*. *IEEE Transactions On Circuits And Systems For Video Technology*, Vol. 12, No. 5.
- [19] Zomet, A and S Peleg.(1998).*Applying super-resolution to panoramic mosaics*. In *Workshop on Applications of Computer Vision*, pages 286 - 287.