IJCSN International Journal of Computer Science and Network, Volume 6, Issue 3, June 2017
ISSN (Online) : 2277-5420
www.IJCSN.org
Impact Factor: 1.5

Enhancement in Image Mosaicing using Voronoi and Surf Algorithm

[1] Arya Vijayakumaran Nair; [2] Dr. D. Loganathan; [3] Asha. S

[1] Student, Final Year M.Tech - Cyber Security, MET'S School of Engineering,
Mala, Thrissur, Kerala, India

[2] Professor, Department of CSE, MET'S School of Engineering,
Mala, Thrissur, Kerala, India

[3]Assistant Professor, Department of CSE, MET'S School of Engineering,
Mala, Thrissur, Kerala, India

Abstract - Image mosaicking or image stitching is a process of combining two or more images to create a large panoramic image. The resultant image can also be used for texture mapping of a 3D environment. Consider the case of images taken from a normal camera. We can convert those images in to mosaicked image using mosaicking technique. This paper presents a new method of image mosaicking based on SURF and VORONOI. SURF algorithm is used to extract features from the input images. Then the input images are subdivided in to regions based on VORONOI method. In order to match the regions from VORONOI, ZNCC (Zero Mean Normalized Cross Correlation) method is used and geometric transformation of input images are calculated based on that value the two images are merged together to form mosaicked image. Finally to create a natural-looking mosaics, a quasi homography warp is applied, which balances the perspective distortion against projective distortion in non-overlapping region. This method helps to reduce projection time and execution time. It is faster than traditional mosaicking algorithms.

Keywords - *Image mosaicking, feature extraction, direct methods, SURF, VORONOI, ZNCC.*

1. Introduction

The word Mosaic originates from an old Italian word 'mosaico' which means a picture or pattern produced by arranging together small pieces of stone, tile, glass, etc. A group of images can be stitched together to form a mosaicked image or a panoramic image, process is called Image mosaicking. Images taken from a normal camera can be converted in to mosaicked image using mosaicking.

Consider the case of a large document, sometimes it may not be possible to capture the document in a single shot. In that case part by part scanning of document is needed and those parts can be combined together to form a large document using mosaicking. Applications of image mosaicking includes construction of satellite photographs, video indexing [13], radiography system, motion detection etc.

Mosaicking process include four different steps

1. Feature Extraction
2. Image Registration
3. Homographic Refinement
4. Image Warping and Blending

The most essential step in mosaicking is Feature Extraction. It includes retrieving an image from some source and some features of the images like edges, corners, etc. are extracted and finally those features are matched to get desired result. Next is Image Registration, it is the process of aligning two or more images of same scene taken from different angle. It is the alignment of two images geometrically - the reference and sensed images. It is the most important step in mosaicking. Third step is Homographic refinement, it is the mapping between two images. In Image warping the position of some points in an image is modified and mapping is done from one image to another. This process does not change the values of pixels. Input images are blended together in to a common plane to get the desired result.

There are two types of image mosaicking methods,

1. Direct method
2. Feature based method

In direct method pixels of two images are compared. It minimizes the sum of absolute difference between overlapping pixels. But their range of convergence is limited.

Next is feature based method, here all feature points in two images are compared by using a local descriptor. Advantage of this method is, it is robust against any type of scene movement in an image and it is faster when compared to direct method. Image mosaicking is also called image stitching.

Some of the application areas are as follows

- The mosaic of satellite remote sensing image.

- Meteorological & environmental monitoring.
- Sea-bottom survey and geological survey.

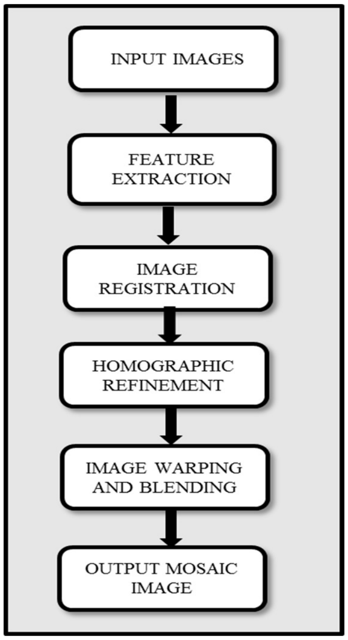

Fig.1. Image mosaicing steps

Fig.2. Input image and feature points

- The 3D rebuilding of objects.
- Video compression, video search browse, and video edit, etc.
- Military reconnaissance and taking evidence.

2. Related Works

Scale Invariant Feature Transform (SIFT), proposed by D.G.Lowe which is a technique to extract feature points. SIFT method can be used to identify an object from an image [3], 3D object recognition [5]. Using nearest-neighbor search shape indexing can be done while using SIFT method [4] and distinctive image features can be identified from scale invariant key points. SIFT algorithm is robust to scale & rotational variation. An efficient and accurate feature point can be extracted from this algorithm. H.Bay et al., [11] proposed Speeded Up Robust Features (SURF), which is 5 times faster than SIFT. This algorithm is rotationally invariant. Smith and J Brady [19] proposed a method for corner detection and structure preserving noise reduction. In this method in order to determine the parts of an image which is closely related to pixels a Non-linear filtering is used. A shape-preserving half projective (SPHP) [2] warp by C.-H. Chang, Y. Sato, and Y.-Y. Chuang, and an adaptive as-natural-as- possible (AANAP) [1] warp by C.-C. Lin, S. U. Pankanti, K. N. Ramamurthy, and A. Y. Aravkin, and warp with global similarity prior (GSP) [16] by Yu-sheng chen and yung-yu chuang address naturalness issues in the non-overlapping region by using a global similarity.

Fig.3. Reference image and feature points

3. Proposed System

A new approach to mosaicking image based on VORONOI regions and SURF is presented. Here SURF algorithm is used to extract interest points or key points from the input images and are matched using the same [12]. Next, each image is divided into several regions using VORONOI diagram. After that VORONOI regions are matched using ZNCC [15] and regions having best correlation scores are selected. By using the seeds of these regions, geometric transformation between the two images are calculated. Finally, the images are overlapped to produce the Mosaicked image and to remove distortion a quasi homography warp is applied.

3.1 Detection of keypoints using SURF

A scale space representation is used to detect the SURF features of an image which is combined with first and second order differential operators. For these operations box filters are used. To identify a selected point as an interest point or not, the determinant of Hessian matrix is used.

$$H(p,\sigma) = \begin{bmatrix} L_{xx}(p,\sigma) & L_{xy}(p,\sigma) \\ L_{yx}(p,\sigma) & L_{yy}(p,\sigma) \end{bmatrix} \quad (1)$$

IJCSN International Journal of Computer Science and Network, Volume 6, Issue 3, June 2017
ISSN (Online) : 2277-5420
www.IJCSN.org
Impact Factor: 1.5

The above equation represents an image I at a point p and scale σ. Lxx(p; σ) is the convolution of Gaussian second order derivative with the image at point with coordinates (x; y).

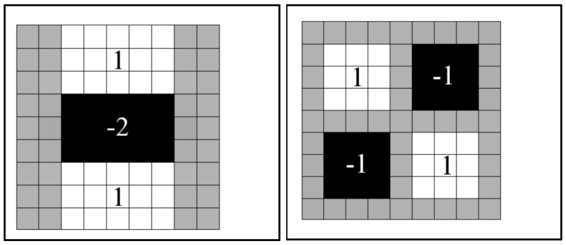

Fig.4. Gaussian second order partial derivatives a) Y direction b) XY direction

To perform efficiently an integral image is used here. Integral image is an intermediate representation for the image and contains the sum of gray scale pixel values of image .The integral image IΣ, can be calculated from image I as

$$I_{\Sigma}(x) = \sum_{i=0}^{i \le x} \sum_{j=0}^{j \le y} I(i, j) \qquad (2)$$

The determinant of the approximated Gaussians is

$$\det(H_{approx}) = D_{xx} D_{yy} - (0.9 D_{xy})^2 \qquad (3)$$

Thus, the interest points, including their scales and locations, are detected in approximate Gaussian scale space (Fig. 2 and Fig. 3). SURF's interest point detector analyse the scale space by increasing the filter size. Image size need not be reduced, but instead of that different scale box filters are applied directly on to the original image (Fig. 5).

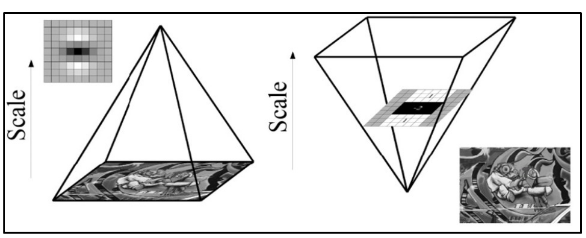

Fig.5. The image size is not reduced and filter size grows at each layer.

For feature description also SURF uses the wavelet responses. To get the SURF feature descriptor a key point's neighbourhood is selected and that region is subdivided in to sub regions and for each sub region the wavelet responses are taken and represented.

The sign of Laplacian which is already computed in the detection is used for underlying interest points. The sign of the Laplacian distinguishes bright blobs on dark backgrounds from the reverse case. In case of matching the features are compared only if they have same type of contrast (based on sign) which allows faster matching (Fig. 6 and Fig. 7).

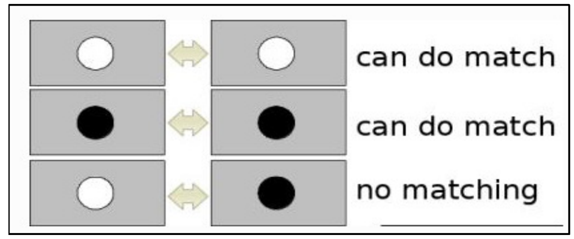

Fig.6. Feature matching

3.2 Division of images in to regions using VORONOI

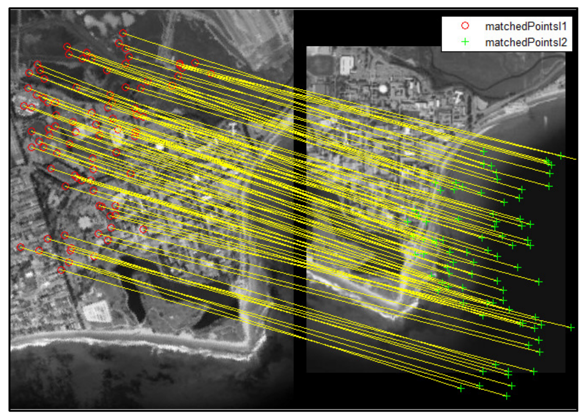

Fig.7. Key points matching

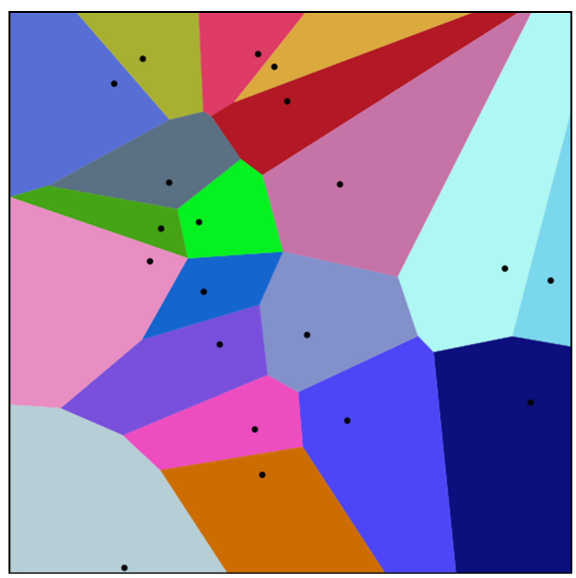

Fig.8. VORONOI diagram

VORONOI diagrams are used to divide an image in to different regions based on set of points [6], [7], [8], [9], [10], [11]. These points are called seeds, generators or

IJCSN International Journal of Computer Science and Network, Volume 6, Issue 3, June 2017
ISSN (Online) : 2277-5420
www.IJCSN.org
Impact Factor: 1.5

sites. For each seed there is a corresponding region consisting of all points closer to that seed than to any other. These regions are called VORONOI cells (Fig 8). Let P be a set of n points in a plane .The VORONOI diagram of P is subdivision of plane in to n cells, one for each site. A point q lies in the cell corresponding to a site $p_i \in$P if Euclidean_Distance (q, p_i) < Euclidean distance (q, p_j), for each $p_i \in$P, j$\neq i$ [14], [18]. Using previously detected control points we apply the VORONOI diagram to divide the two input images in to multiple regions (Fig. 9).

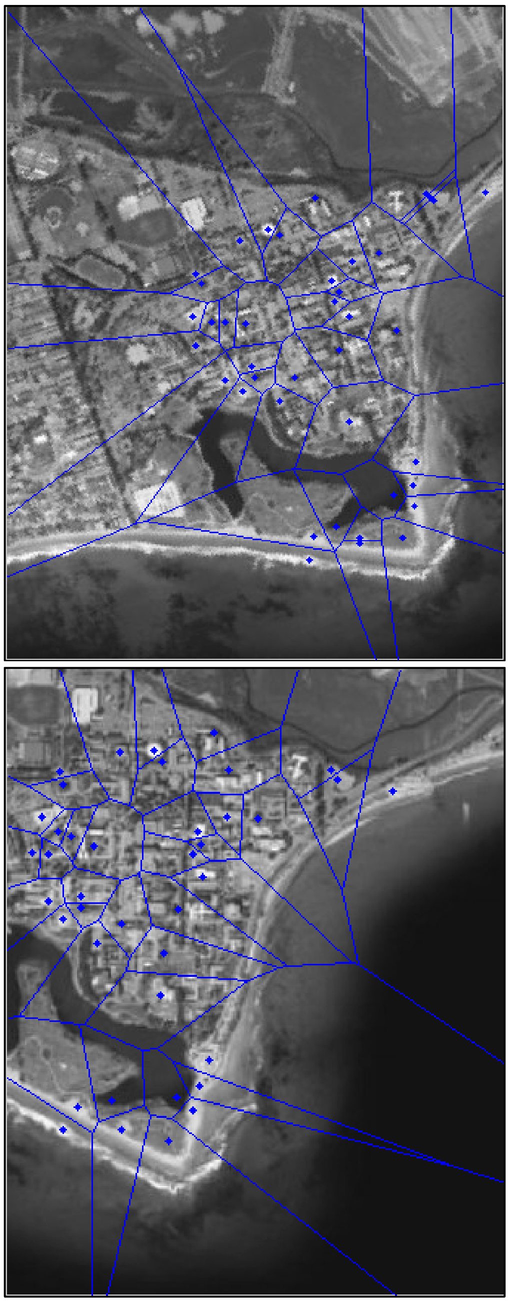

Fig.9. Image division using VORONOI diagram

3.2 VORONOI regions matching

Next step in proposed method is to match those VORONOI regions [21], [22]. For that ZNCC, (Zero mean Normalized Cross Correlation) method is used, formulated as:

$$\mathrm{ZNCC}\left(\mathrm{R}_{\mathrm{i}}^{\mathrm{k}},\mathrm{R}_{\mathrm{j}}^{\mathrm{k}}\right)=\frac{(\mathrm{R}_{\mathrm{i}}^{\mathrm{k}}-\overline{\mathrm{R}_{\mathrm{j}}^{\mathrm{k}}}).(\mathrm{R}_{\mathrm{j}}^{\mathrm{k}}-\overline{\mathrm{R}_{\mathrm{i}}^{\mathrm{k}}})}{\|\mathrm{R}_{\mathrm{i}}^{\mathrm{k}}-\overline{\mathrm{R}_{\mathrm{i}}^{\mathrm{k}}}\|.\|\mathrm{R}_{\mathrm{j}}^{\mathrm{k}}-\overline{\mathrm{R}_{\mathrm{j}}^{\mathrm{k}}}\|} \qquad (4)$$

Here R_i^k is a region of the left image and R_j^k is a region of the right image with the same size. $\overline{R_i^k}$ is the average of the region R_i^k and $\overline{R_j^k}$is the average of the region R_j^k. A thresholdvalue is fixed for the correlation score 0.8 pairs containing correlation score greater than 0.8 are kept in this stage.

3.3 Geometric Transformation

The final step is to calculate the geometric transformation between the two input images based on control points of VORONOI regions previously detected. It is the process of changing the geometry of a raster dataset from one coordinate space to another. For applying the geometric transformation a set of control points are required to register an image on to a projected coordinate system. It is used to find the transformation matrix which maps the greatest number of point pairs between two images. Projective transformation model is applied here, which allows both angular and length distortion, allowing rectangle to be transformed in to an irregular quadrilateral. The control points selected for calculating the transformation will be the generators of VORONOI regions which is having the best correlation scores. Using a geometric transformation T_{ij} both the images are related. T_{ij} is a 3×3 matrix.

$$m_{ki} \sim T_{ij} m_{kj} \qquad (5)$$

m_{ki} is the point k of image I and m_{kj} is the point k of image j and Tij is the geometric transformation of size 3 × 3 which connects image i and j.

Using this transformation matrix the input images can be stitched together. Here the first image is aligned with the reference image so that the pixels representing the same underlying structures are overlaid.

3.4 Quasi-homography warps

This method combines the advantages of homography (line preserving) and similarity (a linear scale function) in the non- overlapping region [1], [17]. Here a forward map H_+ is calculated to get the canvas and a backward map H_+^{-1} is calculated to fill the canvas by bilinear interpolations.

Consider a homography warp H_0

$$f_0(x,y) = \frac{h_1x + h_2y + h_3}{h_7x + h_8y + 1}, \quad (6)$$

$$g_0(x,y) = \frac{h_4x + h_5y + h_6}{h_7x + h_8y + 1}, \quad (7)$$

The horizontal line $y=y_*$ intersects the preimage $x=x_p$ at

$$x_* = \frac{(h_5h_7 - h_6h_7h_8)p + h_2(h_6h_7 - h_4)}{(h_4h_8 - h_5h_7)(ph_7 - h_1)} + \frac{h_3}{ph_7 - h_1} \quad (8)$$

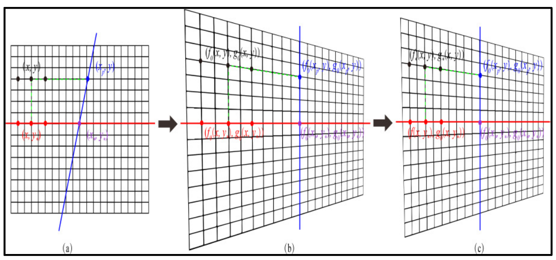

Fig.10. (a) Target image. (b) Reformulation of homography. (c) Derivation of quasi-homography

Taylor's expansion of $f_0(x, y_*)$,

$$f_*(x,y_*) = p + \frac{(h_1h_8 - h_2h_7)y_* + (h_1 - h_3h_7)}{(h_7x_* + h_8y_* + 1)^2}(x - x_*) \quad (9)$$

H_+ can be written as

$$f_+(x,y) = \begin{cases} f_0(x,y), if(x,y_*) \in \mathcal{O}, \\ f_*(x,y), \;\; otherwise, \end{cases} \quad (10)$$

$$g_+(x,y) = \begin{cases} g_0(x,y), if(x,y_*) \in \mathcal{O}, \\ g_*(x,y), \;\; otherwise, \end{cases} \quad (11)$$

Where

$$f_*(x,y) = \frac{w_1x^2y + w_2x^2 + w_3xy + w_4x + w_5y + w_6}{h_7x + h_8y + 1} \quad (12)$$

$$g_*(x,y) = \frac{w_7x^2y + w_8x^2 + w_9xy + w_{10}x + w_{11}y + w_{12}}{h_7x + h_8y + 1} \quad (13)$$

Given $(x', y' \in I')$, then $(x,y) \in I$ is determined by the inverse map H_+^{-1}, and can be calculated using the following

$$H_+^{-1} = \begin{cases} H_0^{-1}, if(x,y_*) \in \mathcal{O}, \\ H_*^{-1}, \;\; otherwise, \end{cases} \quad (14)$$

Where H_*^{-1}

$$x = RootOf(m_1x^2 + m_2x + m_3), \quad (15)$$

$$y = -\frac{h_4x' - h_6h_7x' - h_1y' + h_3h_7y' - h_3h_4 + h_1h_6}{h_4h_8x' - h_5h_7x' - h_1h_8y' + h_2h_7y' - h_2h_4 + h_1h_5} \quad (16)$$

If two images are taken as input, a homography H_0 can be estimated, then a quasi-homography w H_+ can be calculated which extrapolates H_0 in the overlapping region to H_* in the non-overlapping region. For multiple images warping method consist of three stages. In the first stage, a reference image is determined as the standard perspective. Then, a homography is estimated for every two images with overlapping fields of views, and calculate its corresponding quasi-homography. Finally, concatenate all target images in the image plane of the reference image. To blend the images multiband blending is used here.

Fig.11. Mosaicked image

4. Experimental Result

In this section proposed approach and existing approaches are compared. In order to detect the key points SURF algorithm is used, existing systems uses SIFT. SURF consist of Fast-Hessian detector + SURF descriptor and SIFT consist of DOG detector + SIFT descriptor. SURF describes image faster than SIFT by 3 times. SURF is good at handling serious blurring and image rotation. Table 1 compares SURF and SIFT in terms of memory cost, Speed and features detected.

Table. 1 Comparison of SURF and SIFT

Items	SURF	SIFT
Memory cost	64 floats	128 bytes
Speed(Time to detect and describe 4000 features)	2.4 seconds	6 seconds
Features detected in 1024x768 image(Default threshold)	~1000	>3000

The calculation of the geometric transformation by the proposed method is faster in terms of calculation time. In proposed method, the calculation of the transformation is performed only once, but in RANSAC method [20], different iterations are needed to get the required accuracy. So the computational complexity is reduced is this method.

Fig. 12. Comparison of different warps

In projection step a single image is created by combining different parts of the captured scene (I1 and I2) by the projection of two images on a flat surface. Each point of the panorama is a projection of a point of the set of original images taking into account that it may be projected onto several images simultaneously by using the previously calculated transformation. In proposed method VORONOI method is used to match regions to reduce the projection time. Quasi-homography warp used in the proposed system is compared with Homography, SPHP [2] and Auto Stitch. Fig.12 illustrates results of image stitching. In warping using Homography the trees shown in the pictures are enlarged but all the straight lines are preserved here. In SPHP it curves buildings and grounds but without enlarging the region of trees. Due to spherical warps the Auto Stitch produces a fish eye like mosaic. Quasi-homography balances projective distortion against perspective distortion by preserving horizontal lines and vertical lines.

4. Conclusion

An improved method of image mosaic based on VORONOI regions and SURF algorithm is proposed here. In this method the control pints from two input images are extracted using SURF algorithm. VORONOI method is used to divide the images in different regions based on the control points. Then using ZNCC correlation VORONOI region matching is performed. Finally geometric transformation is calculated between the two images, for this the regions with highest correlation scores is selected and the input images are overlapped to get the desired mosaicked. To create a natural-looking mosaics, a quasi homography warp is applied, which balances the perspective distortion against projective distortion in non-overlapping region

References

[1] C.-C. Lin, S. U. Pankanti, K. N. Ramamurthy, and A. Y. Aravkin, "Adaptive as-natural-as-possible image stitching," in Proc. IEEE Conf. Comput. Vis. Pattern Recog., Jun. 2015, pp. 1155–1163.

[2] C.-H. Chang, Y. Sato, and Y.-Y. Chuang, "Shape-preserving halfprojective warps for image stitching," in Proc. IEEE Conf. Comput. Vis.Pattern Recog., May 2014, pp. 3254–3261.

[3] D. G Lowe, "Object recognition from local scale-invariant features‖, International Conference on Computer Vision, vol. 2, pp.1150– 1157,1999 J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[4] D. G. Lowe, ―Distinctive Image Features from Scale-Invariant Keypoints‖, International Journal of Computer Vision, vol. 60, pp. 91- 110, 2004. M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[5] D. G. Lowe, ―Local feature view clustering for 3D object recognition‖, IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, pp.682-688, 2001. K. Elissa, "Title of paper if known," unpublished.

[6] D. T. Lee and A. K. Lin. Generalized Delaunay triangulation for planar graphs. Discrete Comput. Geom., 1:201–217, 1986.

[7] D. T. Lee. On k-nearest neighbor Voronoi diagrams in the plane. IEEE Trans. Comput., C-31:478–487, 1982.

[8] Du, Qiang, Vance Faber, and Max Gunzburger. "Centroidal Voronoi Tessellations: Applications and Algorithms." SIAM Review 21 (1999): 637-676.

[9] F. Aurenhammer and H. Edelsbrunner, "An Optimal Algorithm for Constructing the Weighted Voronoi

Diagram in the Plane." Pattern Recognition 17, 2 (1984): 251-257.

[10] Fang X, Zhu J, Luo B (2012) Image mosaic with relaxed motion. SIViP 6(4):647–667. doi:10.1007/s11760-010-0194-4

[11] Fortune, Steven. "A Sweepline Algorithm for Voronoi Diagrams." Proceedings of the Second Annual ACM Symposium on Computational Geometry Yorktown Heights, New York: Association for Computing Machinery, 1986: 313-322.

[12] H. Bay, et al., ―Speeded up Robust Features‖, Proc. of the 9th European Conf. on Computer Vision, Cambridge, U.K., pp. 404-417, 2006

[13] J. Beis, and D. G. Lowe, ―Shape indexing using approximate nearest-neighbour search in high-dimensional spaces‖, Conference on Computer Vision and Pattern Recognition, Puerto Rico, pp. 1000–1006, 1997..

[14] Ju, Lili, Qiang Du, and Max Gunzburger. "Probabilistic Methods for Centroidal Voronoi Tessellations and Their Parallel Implementations." Parallel Computer 28 (2002):1477-1500.

[15] M. Laraqui, A. Saaidi , A. Mouhib ,M Abarkan(2015) Images matching using voronoï regions propagation. 3D Res 6(3):1–16. doi:10.1007/s13319-015-0056-5

[16] M. Lhuillier , L. Quan (2002) Quasi-dense reconstruction from image sequence. In Computer Vision—ECCV 2002, 125–139. Springer Berlin Heidelberg. doi:10.1007/3-540-47967-8_9

[17] N. I. S. with the Global Similarity Prior, "Yu-sheng chen and yung-yuchuang," in Proc. Eur. Conf. Comput. Vis., Oct. 2016, pp. 186–201.

[18] N.-M. Lˆe. Randomized incremental construction of simple abstract Voronoi diagrams in 3-space. In Proc. 10th Internat. Conf. Fund. Comput. Theory, volume 965 of Lecture Notes Comput. Sci., pages 333–342. Springer-Verlag, 1995.

[19] Okabe, Atsuyuki, Barry Boots, Kokichi Sugihara, and Sung Nok Chiu. Spatial Tessellations: Concepts and Applications of Voronoi Diagrams, 2nd Ed.. Chichester, West Sussex, England: John Wiley & Sons, 1999.

[20] P J Anju , Dr.D.Loganathan, (2016) Image Fusion using NSCT Theory and Wavelet Transform for Medical Diagnosis, International Journal of Computer Science and Information Technologies, Vol. 7 (3) , 1507-1510

[21] S. Peleg and J. Herman, ‖Panoramic mosaics by manifold projection‖ Proc. of IEEE Computer Society conference on Computer Vision and Pattern Recognition, San Juan, pp.338-343, 1997.

[22] Wang X, Ying X, Liu Y-J, Xin S-Q,WangW, Gu X,Mueller-WittigW, He Y (2015) Intrinsic computation of centroidal Voronoi tessellation (CVT) on meshes. Comput Des, 51–61. doi:10.1016/j.cad.2014.08.023

Authors -

Arya Vijayakumaran Nair received the B. Tech degree in Information Technology from M. G University, Kerala, India, in 2012, and currently doing M. Tech in Cyber Security in MET'S School of Engineering, Mala – APJ Abdul Kalam Technological University, Kerala, India.

Dr. D. Loganathan is a Professor and Head of Computer Science and Engineering department in MET'S School of Engineering, Mala, Trissur, Kerala. After his B.E., and M.E degree, he accomplished a doctoral degree from Anna University, Chennai, India. He has more than 20 years of teaching experience and having 8 years of research experience in engineering field. His research interest includes Wireless Communication, Wireless Ad hoc Networks and Image Processing. He has published several research papers in various international journals.

Asha. S is an Assistant professor in MET'S School of Engineering Mala, Thrissur, Kerala. Received the Btech degree in Govt. Engineering College Idukki, Kerala and M.Tech from Anna University. She has more than 4 years of teaching experience. Subject of interest are programming languages such as C, C++, Theory of computation, Compiler design.